

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Apremilast (Otezla®)

Amgen GmbH

Modul 4 C

Erwachsene Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom (BS) assoziiert sind und für die eine systemische Therapie infrage kommt

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	10
Abkürzungsverzeichnis	11
4 Modul 4 – allgemeine Informationen	15
4.1 Zusammenfassung der Inhalte von Modul 4.....	16
4.2 Methodik.....	23
4.2.1 Fragestellung.....	23
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	24
4.2.3 Informationsbeschaffung.....	26
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	27
4.2.3.2 Bibliografische Literaturrecherche.....	27
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken.....	28
4.2.3.4 Suche auf der Internetseite des G-BA.....	30
4.2.3.5 Selektion relevanter Studien.....	31
4.2.4 Bewertung der Aussagekraft der Nachweise.....	31
4.2.5 Informationssynthese und -analyse.....	33
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	33
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	33
4.2.5.3 Meta-Analysen.....	46
4.2.5.4 Sensitivitätsanalysen.....	47
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	47
4.2.5.6 Indirekte Vergleiche.....	51
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	54
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	54
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	54
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	54
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	56
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken.....	58
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA.....	59
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	60
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	61
4.3.1.2.1 Studiendesign und Studienpopulationen.....	61
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	68
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	69
4.3.1.3.1 Endpunkt – RCT.....	70
4.3.1.3.1.1 Vollständige Remission der oralen Aphthen (komplettes Ansprechen) - RCT.....	73

4.3.1.3.1.2	Schmerzhaftigkeit oraler Aphthen mittels VAS - RCT.....	89
4.3.1.3.1.3	Veränderung des Behçet's Syndrome Activity Score (BSAS) zu Woche 12 - RCT	97
4.3.1.3.1.4	Veränderung des Behçet's Disease Current Activity Form (BDCAF) zu Woche 12 – RCT	100
4.3.1.3.1.5	Veränderung des Behçet's disease Quality of Life (BD-QoL) zu Woche 12 – RCT	106
4.3.1.3.1.6	Verbesserung des 36-Item Short Form Health Survey, Version 2 (SF-36v2) um $\geq 2,5$ Punkte zu Woche 12 – RCT.....	110
4.3.1.3.1.7	Sicherheit (unerwünschte Ereignisse, UE) - RCT	115
4.3.1.3.2	Subgruppenanalysen – RCT.....	124
4.3.1.3.2.1	Subgruppenanalysen für den Endpunkt Veränderung der Anzahl oralen Aphthen	131
4.3.1.3.2.2	Subgruppenanalysen für den Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission	133
4.3.1.3.2.3	Subgruppenanalysen für den Endpunkt Schmerzhaftigkeit oraler Aphthen mittels VAS.....	136
4.3.1.3.2.4	Subgruppenanalysen für den Endpunkt Veränderung des BDCAF zu Woche 12.....	139
4.3.1.3.2.5	Subgruppenanalysen für den Endpunkt Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12.....	144
4.3.1.4	Liste der eingeschlossenen Studien - RCT	148
4.3.2	Weitere Unterlagen.....	148
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	148
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	149
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	149
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	150
4.3.2.1.1.3	Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken	153
4.3.2.1.1.4	Studien aus der Suche auf der Internetseite des G-BA.....	154
4.3.2.1.1.5	Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich	155
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	157
4.3.2.1.2.1	Studiendesign und Studienpopulationen	157
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	167
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	167
4.3.2.1.3.1	Endpunkt – indirekte Vergleiche aus RCT.....	168
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	170
4.3.2.1.4	Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT...	170
4.3.2.2	Nicht randomisierte vergleichende Studien.....	170
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	170
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	171
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	172
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	172

4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	173
4.3.2.2.4	Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien	173
4.3.2.3	Weitere Untersuchungen.....	173
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	174
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	174
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	175
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	175
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	176
4.3.2.3.4	Liste der eingeschlossenen Studien – weitere Untersuchungen.....	176
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	176
4.4.1	Beurteilung der Aussagekraft der Nachweise	176
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	179
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	186
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	186
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	186
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	187
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	187
4.5.4	Verwendung von Surrogatendpunkten	187
4.6	Referenzliste.....	188
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		195
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken.....		207
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		214
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken).....		218
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		229
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		243

Tabellenverzeichnis

	Seite
Tabelle 4-1: Ausmaß des Zusatznutzens auf Endpunktebene	17
Tabelle 4-2: Übersicht über Ein- und Ausschlusskriterien von Studien zur Bewertung von Apremilast	25
Tabelle 4-3: Übersicht über Ein- und Ausschlusskriterien von Studien zur Bewertung der zweckmäßigen Vergleichstherapien (zVT)	26
Tabelle 4-4: Demografische Charakterisierung der Patientenpopulation	34
Tabelle 4-5: Krankheitsspezifische Charakterisierung der Patientenpopulation	36
Tabelle 4-6: Patientenrelevante Endpunkte zum Nachweis des Zusatznutzens von Apremilast	37
Tabelle 4-7: A priori festgelegte Subgruppenanalysen der Studie RELIEF	48
Tabelle 4-8: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	55
Tabelle 4-9: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel	56
Tabelle 4-10: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel	58
Tabelle 4-11: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel	59
Tabelle 4-12: Studienpool – RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-13: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	62
Tabelle 4-14: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel	63
Tabelle 4-15: Demografische Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	63
Tabelle 4-16: Krankheitsspezifische Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	65
Tabelle 4-17: Dosistitrationsschema der Apremilast-Gruppe (Gruppe 1)	67
Tabelle 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel	69
Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	70
Tabelle 4-20: Operationalisierung von Vollständige Remission der oralen Aphthen (komplettes Ansprechen)	73
Tabelle 4-21: Bewertung des Verzerrungspotenzials für Vollständige Remission der oralen Aphthen (komplettes Ansprechen) in RCT mit dem zu bewertenden Arzneimittel	74

Tabelle 4-22: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	75
Tabelle 4-23: Sensitivitätsanalyse: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 (NRI) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	76
Tabelle 4-24: Ergebnisse für die Zeit bis zur vollständigen Remission der oralen Aphthen aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	76
Tabelle 4-25: Ergebnisse für Veränderung der Anzahl oraler Aphthen zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	78
Tabelle 4-26: Ergebnisse für Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	78
Tabelle 4-27: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population).....	80
Tabelle 4-28: Ergebnisse für Veränderung der Anzahl oraler Aphthen in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	81
Tabelle 4-29: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der genitalen Aphthen (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	83
Tabelle 4-30: Ergebnisse für die Veränderung der Anzahl genitaler Aphthen (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population).....	84
Tabelle 4-31: Ergebnisse für neue oder sich verschlimmernde Hautläsionen im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	85
Tabelle 4-32: Ergebnisse für neue oder sich verschlimmernde Arthritis im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	86
Tabelle 4-33: Ergebnisse für neue oder sich verschlimmernde gastrointestinale Manifestationen im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population).....	86
Tabelle 4-34: Ergebnisse für neue oder sich verschlimmernde Manifestationen des ZNS im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	87
Tabelle 4-35: Operationalisierung von Schmerzhaftigkeit oraler Aphthen mittels VAS	89
Tabelle 4-36: Bewertung des Verzerrungspotenzials für Schmerzhaftigkeit oraler Aphthen mittels VAS in RCT mit dem zu bewertenden Arzneimittel	90
Tabelle 4-37: Rücklaufquoten der VAS von Baseline bis Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	91

Tabelle 4-38: Ergebnisse für Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	91
Tabelle 4-39: Ergebnisse für die Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	92
Tabelle 4-40: Ergebnisse für die Veränderung der Schmerzhaftigkeit der oralen Aphthen mittels VAS (MMRM) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	94
Tabelle 4-41: Ergebnisse für die Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	95
Tabelle 4-42: Operationalisierung von Veränderung des BSAS zu Woche 12	97
Tabelle 4-43: Bewertung des Verzerrungspotenzials für Veränderung des BSAS zu Woche 12 in RCT mit dem zu bewertenden Arzneimittel	97
Tabelle 4-44: Rücklaufquoten des BSAS-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	98
Tabelle 4-45: Ergebnisse für Veränderung des BSAS zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	99
Tabelle 4-46: Ergebnisse für Veränderung des BSAS in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population).....	99
Tabelle 4-47: Operationalisierung von Veränderung des BDCAF zu Woche 12	100
Tabelle 4-48: Bewertung des Verzerrungspotenzials für Veränderung des BDCAF zu Woche 12 in RCT mit dem zu bewertenden Arzneimittel	101
Tabelle 4-49: Rücklaufquoten des BDCAF-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	102
Tabelle 4-50: Ergebnisse für Veränderung des BDCAF zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	103
Tabelle 4-51: Ergebnisse für die Veränderung des BDCAF in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	104
Tabelle 4-52: Operationalisierung von Veränderung des BD-QoL zu Woche 12	106
Tabelle 4-53: Bewertung des Verzerrungspotenzials für Veränderung des BD-QoL zu Woche 12 in RCT mit dem zu bewertenden Arzneimittel	107
Tabelle 4-54: Rücklaufquoten des BD-QoL-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	108
Tabelle 4-55: Ergebnisse für Veränderung des BD-QoL zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	108
Tabelle 4-56: Ergebnisse für Veränderung des BD-QoL in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population).....	109
Tabelle 4-57: Operationalisierung von Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12.....	110

Tabelle 4-58: Bewertung des Verzerrungspotenzials für Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12 mit dem zu bewertenden Arzneimittel.....	111
Tabelle 4-59: Rücklaufquoten des SF-36-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	112
Tabelle 4-60: Ergebnisse für Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12 (NRI) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population).....	113
Tabelle 4-61: Operationalisierung von Sicherheit	115
Tabelle 4-62: Bewertung des Verzerrungspotenzials für Sicherheit in RCT mit dem zu bewertenden Arzneimittel	116
Tabelle 4-63: Ergebnisse für Sicherheit zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	117
Tabelle 4-64: Ergebnisse der UE von besonderem Interesse (Diarrhoe) zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	118
Tabelle 4-65: Ergebnisse der UE bei mindestens 10 % der Patienten in einem Studienarm nach SOC und PT zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)	119
Tabelle 4-66: Ergebnisse der UE, die zum Therapieabbruch führten, nach SOC und PT zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)	120
Tabelle 4-67: Ergebnisse für Sicherheit von Woche 12 bis Woche 64 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (AAT-Population).....	121
Tabelle 4-68: Ergebnisse der UE bei mindestens 10 % der Patienten in einem Studienarm nach SOC und PT von Woche 12 bis Woche 64 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (AAT-Population).....	122
Tabelle 4-69: Ergebnisse der UE, die von Woche 12 bis Woche 64 zum Therapieabbruch führten, nach SOC und PT aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (AAT-Population)	123
Tabelle 4-70: Matrix der durchgeführten Subgruppenanalysen.....	126
Tabelle 4-71: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für RELIEF und Subgruppenmerkmale	127
Tabelle 4-72: Subgruppenanalysen für Endpunkt Veränderung der Anzahl oraler Aphthen zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	131
Tabelle 4-73: Subgruppenanalysen für Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	133
Tabelle 4-74: Subgruppenanalysen für Endpunkt Schmerzhaftigkeit oraler Aphthen mittels VAS (MMRM) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	136
Tabelle 4-75: Subgruppenanalysen für Endpunkt Veränderung des BDCAF zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	139

Tabelle 4-76: Subgruppenanalysen für Endpunkt Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12 (NRI) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)	144
Tabelle 4-77: Übersicht der eingeschlossenen Studie RELIEF	148
Tabelle 4-78: Liste der Studien des pharmazeutischen Unternehmers – RCT für den adjustierten indirekten Vergleich	150
Tabelle 4-79: Studien des pharmazeutischen Unternehmers, die nicht für den indirekten Vergleich herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel	150
Tabelle 4-80: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich.....	153
Tabelle 4-81: Relevante Studien aus der Suche auf der Internetseite des G BA – RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich.....	155
Tabelle 4-82: Studienpool – RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich.....	156
Tabelle 4-83: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	158
Tabelle 4-84: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	160
Tabelle 4-85: Ein- und Ausschlusskriterien der RCT für den indirekten Vergleich.....	161
Tabelle 4-86: Relevante Unterschiede in den Baseline Charakteristika zwischen RELIEF und Melikoglu 2005	163
Tabelle 4-87: Relevante Unterschiede in den Baseline Charakteristika zwischen RELIEF und Hamuryudan 1998	165
Tabelle 4-88: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	167
Tabelle 4-89: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	167
Tabelle 4-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	168
Tabelle 4-91: Operationalisierung von <Endpunkt xxx>.....	168
Tabelle 4-92: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	169
Tabelle 4-93: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	169
Tabelle 4-94: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	171
Tabelle 4-95: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien	172
Tabelle 4-96: Operationalisierung von <Endpunkt xxx>.....	172
Tabelle 4-97: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	172

Tabelle 4-98: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen	175
Tabelle 4-99: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	175
Tabelle 4-100: Ausmaß des Zusatznutzens auf Endpunktebene.....	180
Tabelle 4-101: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	186
Tabelle 4-102 (Anhang): Studiendesign und -methodik für Studie RELIEF.....	229
Tabelle 4-103 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie CC-10004-BCT-002.....	244

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	57
Abbildung 2: Kaplan-Meier-Kurven für Zeit bis zur vollständigen Remission der oralen Aphthen (komplettes Ansprechen) bis Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	77
Abbildung 3: Kaplan-Meier-Kurven für Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission bis Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel.....	79
Abbildung 4: Zeitlicher Verlauf der Veränderung der Anzahl oraler Aphthen (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	82
Abbildung 5: Zeitlicher Verlauf für Veränderung der Schmerzhaftigkeit der oraler Aphthen mittels VAS bis Woche 12 (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel.....	94
Abbildung 6: Zeitlicher Verlauf für die Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	96
Abbildung 7: Zeitlicher Verlauf für Veränderung des BSAS (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel.....	100
Abbildung 8: Zeitlicher Verlauf für Veränderung des BDCAF – Indexwert (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	105
Abbildung 9: Zeitlicher Verlauf für Veränderung des BDCAF – Patientenwahrnehmung (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel.....	105
Abbildung 10: Zeitlicher Verlauf für Veränderung des BDCAF – Untersucherwahrnehmung (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel.....	105
Abbildung 11: Zeitlicher Verlauf für Veränderung des BD-QoL (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel	109
Abbildung 12: Kaplan-Meier-Kurven für den Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission, Subgruppe Alkoholkonsum: Ja, aus Studie RELIEF mit dem zu bewertenden Arzneimittel.....	134
Abbildung 13: Kaplan-Meier-Kurven für den Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission, Subgruppe Alkoholkonsum: Nein, aus Studie RELIEF mit dem zu bewertenden Arzneimittel.....	134
Abbildung 14: Flussdiagramm der bibliografischen Literaturrecherche - Suche nach randomisierten, kontrollierten Studien für die zVT	152
Abbildung 15: Flow-Chart zur Studie REFLIEF	242

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AAT	Mit Apremilast behandelt (Apremilast-as-treated)
AESI	Unerwünschte(s) Ereignis(se) von besonderem Interesse (adverse event of special interest)
ALT	Alanin-Aminotransferase
AMIS	Arzneimittelinformationssystem
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse (analysis of covariance)
APR	Apremilast
AST	Aspartat-Aminotransferase
AUC	Fläche unter der Kurve (area under the curve)
AUC _{w0-12}	AUC zu Woche 12
AWG	Anwendungsgebiet
BD	Behçet-Syndrom (Behçet's Disease)
BDCA	Behçet's Disease Current Activity
BDCAF	Behçet's Disease Current Activity Form
BDCAI	Behçet's Disease Current Activity Index
BD-QoL	Behçet's Disease Quality of Life
bid	Zweimal täglich (bis in die)
BMI	Body Mass Index
BMJV	Bundesministerium der Justiz und für Verbraucherschutz
BS	Behçet-Syndrom (Behçet-Syndrome)
BSAS	Behçet's Syndrome Activity Score
CENTRAL	Cochrane Central Register of Controlled Trials
CMH-Test	Cochran-Mantel-Haenszel-Test
CONSORT	Consolidated Standards of Reporting Trials
CTCAE	Allgemeine Terminologie und Merkmale unerwünschter Ereignisse (Common Terminology Criteria for Adverse Events)
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
EBM	Evidenzbasierte Medizin (evidence based medicine)
EKG	Elektrokardiogramm

Abkürzung	Bedeutung
EMA	Europäische Arzneimittel-Agentur (European Medicines Agency)
EMBASE	Excerpta Medica dataBASE
EU	Europäische Union (European Union)
EU-CTR	EU Clinical Trials Register
EULAR	European League Against Rheumatism
GCP	Gute klinische Praxis (Good Clinical Practice)
G-BA	Gemeinsamer Bundesausschuss
HIV	Humanes Immundefizienz Virus
HPRA	Health Products Regulatory Authority
HR	Hazard Ratio
ICTRP	International Clinical Trials Registry Platform
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interaktive Response-Technologie (Interactive Response Technology)
ISG	International Study Group for Behçet's Disease
ITT	Intention-to-Treat
KI	Konfidenzintervall
LOCF	Letzte verfügbare Beobachtung (last observation carried forward)
LS-MW	Kleinste-Quadrate-Mittelwert (least square-mean value)
LS-MWD	Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference)
MCID	Minimale klinisch bedeutsame Differenz (minimal clinically important difference)
MCS	Psychische Summenskala (mental component summary)
MEDLINE	Medical Literature Analysis and Retrieval System Online
MedDRA	Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung (Medical Dictionary for Regulatory Activities)
MMRM	Gemischtes Modell mit Messwiederholungen (mixed-effects model for repeated measures)
MSU	Monosodium urate
MTC	Mixed Treatment Comparison
MW	Mittelwert

Abkürzung	Bedeutung
NRI	Non-Responder-Imputation
NSAR	Nichtsteroidale Antirheumatika
OC	Vorliegende Beobachtungen (observed cases)
OR	Odds Ratio
p.o.	Oral (per os)
PCS	Psychische Summenskala (physical component summary)
PGA	Physician's Global Assessment
PICO	Population, Intervention, Comparison, Outcome
PP	Per-Protocol
PRO	Patientenberichteter Endpunkt (patient-reported outcome)
PT	Bevorzugter Begriff (preferred term)
RCT	Randomisierte kontrollierte Studie (randomized controlled trial)
RD	Risikodifferenz (risk difference)
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SD	Standardabweichung (standard deviation)
SE	Standardfehler (standard error)
SF-36v2 / SF-36	36-Item Short Form Health Survey, Version 2
SGB	Sozialgesetzbuch
SGB V	Sozialgesetzbuch Fünftes Buch
SMD	Standardisierte Mittelwertdifferenz
SMQ	Standardised MedDRA Queries
SOC	Systemorganklasse (System Organ Class)
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegende(s) unerwünschte(s) Ereignis(se)
TNF	Tumornekrosefaktor
TNF- α	Tumornekrosefaktor-alpha
UE	Unerwünschte(s) Ereignis(se)
ULN	Obergrenze des Normalwertes (upper limit of normal)
USA	Vereinigte Staaten von Amerika (United States of America)

Abkürzung	Bedeutung
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
VAS	Visuelle Analogskala
WBC	Weißer Blutkörperchen (white blood cell)
WHO	Weltgesundheitsorganisation (World Health Organization)
ZNS	Zentralnervensystem
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

Wie sind medizinischer Nutzen und Zusatznutzen von Apremilast bei erwachsenen Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt, im Vergleich zur zweckmäßigen Vergleichstherapie (zVT) „Therapie nach Maßgabe des Arztes“ zu bewerten?

Datenquellen

Zur Beantwortung der Fragestellung wird die Zulassungsstudie RELIEF, eine doppelblinde, randomisierte, Placebo-kontrollierte Studie (randomized controlled trial, RCT), herangezogen. Die Bewertung erfolgt unter Berücksichtigung patientenrelevanter Endpunkte in Bezug auf Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Sicherheit. Ein indirekter Vergleich war nach Prüfung nicht durchführbar. Es gibt neben Apremilast keine weiteren Arzneimittel, die länderübergreifend einheitlich im Anwendungsgebiet (AWG) zugelassen sind. Die Studie RELIEF stellt somit die bestmögliche Evidenz zur Ableitung des Zusatznutzens dar.

Ein-/Ausschlusskriterien für Studien

Es wurde eine Suche in Studienregistern und eine bibliografische Literaturrecherche durchgeführt, um neben der Zulassungsstudie RELIEF weitere relevante Studien mit Apremilast bei erwachsenen Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt, zu identifizieren. Die Ein-/Ausschlusskriterien dieser Recherche sind in Tabelle 4-2 dargestellt. Es wurden keine weiteren relevanten Studien identifiziert.

Um relevante Studien für einen potenziellen indirekten Vergleich zu identifizieren, wurde eine weitere Suche für die zVT „Therapie nach Maßgabe des Arztes“ durchgeführt (Tabelle 4-3). Die Suchstrategie war analog zu der von Apremilast. Es konnten keine Studien identifiziert werden, die sich für einen indirekten Vergleich eignen.

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

Zur Bewertung der Aussagekraft der Nachweise wurden Angaben zu Verzerrungsaspekten sowohl auf Studienebene als auch auf Endpunktebene dokumentiert und bewertet (Anhang 4-F). Im Anschluss an die Beurteilung des Verzerrungspotenzials erfolgte die Klassifizierung der Wahrscheinlichkeit des Zusatznutzens für jeden patientenrelevanten Endpunkt gemäß den Anforderungen im Methodenpapier des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (IQWiG 2017).

Die Charakteristika der in die Bewertung eingeschlossenen Zulassungsstudie RELIEF wurden anhand der Items 2b bis 14b des CONSORT-Statements separat dargestellt (Anhang 4-E).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Für die Bestimmung des medizinischen Zusatznutzens von Apremilast wurde die zweiarmige RCT RELIEF herangezogen (Evidenzstufe Ib).

Nachfolgend werden in Tabelle 4-1 alle Ergebnisse zum medizinischen Nutzen und Zusatznutzen unter Angabe des Ausmaßes des endpunktspezifischen Zusatznutzens für die im Dossier berücksichtigten Endpunkte dargestellt.

Tabelle 4-1: Ausmaß des Zusatznutzens auf Endpunktebene

Endpunkte	Studie RELIEF (Apremilast vs. Placebo)	Ausmaß des Zusatznutzens
Morbidität		
Vollständige Remission der oralen Aphthen	Vollständiges Ansprechen Ergebnisse: 56,7 % vs. 25,2 % RR [95 %-KI]: 2,25 [1,55; 3,27] p-Wert < 0,0001	beträchtlich
Zeit bis zur vollständigen Remission der oralen Aphthen (Median in Wochen)	Ergebnis: 2,14 vs. 8,14 HR [95 %-KI]: 2,40 [1,69; 3,40] p-Wert < 0,0001	
Schmerzhaftigkeit oraler Aphthen mittels VAS	Ergebnis (LS-MW): -39,04 vs. -18,77 SMD [95 %-KI]: -1,90 [-2,24; -1,57]	
Veränderung BSAS	Ergebnis (LS-MW): -17,35 vs. -5,41 SMD [95 %-KI]: -0,76 [-1,05; -0,48]	
Veränderung BDCAF	Patientenwahrnehmung: Ergebnis (LS-MW): -1,68 vs. -0,67 SMD [95 %-KI]: -0,64 [-0,92; -0,36]	
	Untersucherwahrnehmung: Ergebnis (LS-MW): -1,64 vs. -0,69 SMD [95 %-KI]: -0,65 [-0,93; -0,37]	

Endpunkte	Studie RELIEF (Apremilast vs. Placebo)	Ausmaß des Zusatznutzens
Gesundheitsbezogene Lebensqualität		
Veränderung BD-QoL	Ergebnis (LS-MW): -3,47 vs. -0,52 SMD [95 %-KI]: -0,51 [-0,79; -0,24]	beträchtlich
Verbesserung der gesundheitsbezogenen Lebensqualität anhand SF-36	Körperliche Rollenfunktion: Ergebnisse: 43,3 % vs. 24,3 % RR [95 %-KI]: 1,78 [1,19; 2,67] p-Wert = 0,0042	
	Körperliche Schmerzen Ergebnisse: 58,6 % vs. 42,7 % RR [95 %-KI]: 1,37 [1,04; 1,80] p-Wert = 0,0226	
	Allgemeine Gesundheitswahrnehmung: Ergebnisse: 47,1 % vs. 30,1 % RR [95 %-KI]: 1,57 [1,11; 2,24] p-Wert = 0,0322	
	Vitalität: Ergebnisse: 60,6 % vs. 41,8 % RR [95 %-KI]: 1,45 [1,11; 1,90] p-Wert = 0,0069	
	Soziale Funktionsfähigkeit: Ergebnisse: 51,9 % vs. 26,2 % RR [95 %-KI]: 1,99 [1,37; 2,88] p-Wert = 0,0002	
	Emotionale Rollenfunktion: Ergebnisse: 46,1 % vs. 32,0 % RR [95 %-KI]: 1,43 [1,01; 2,02] p-Wert = 0,0405	
	Psychisches Wohlbefinden: Ergebnisse: 54,8 % vs. 39,8 % RR [95 %-KI]: 1,42 [1,06; 1,89] p-Wert = 0,0322	
	Psychische Summenskala: Ergebnisse: 54,8% vs. 33,0 % RR [95 %-KI]: 1,72 [1,26; 2,36] p-Wert = 0,0016	

Endpunkte	Studie RELIEF (Apremilast vs. Placebo)	Ausmaß des Zusatznutzens
Sicherheit		
UE gesamt	Ergebnisse: 78,8 % vs. 71,8 % RR [95 %-KI]: 1,10 [0,94; 1,28] p-Wert = 0,2436	kein geringerer oder höherer Schaden
SUE	Ergebnisse: 2,9 % vs. 3,9 % RR [95 %-KI]: 0,74 [0,17; 3,24] p-Wert = 0,6917	
Therapieabbruch aufgrund von UE	Ergebnisse: 2,9 % vs. 4,8 % RR [95 %-KI]: 0,59 [0,15; 2,42] p-Wert = 0,4633	
Schwere UE (CTCAE \geq 3)	Ergebnisse: 5,8 % vs. 5,8 % RR [95 %-KI]: 0,99 [0,33; 2,97] p-Wert = 0,9863	
UE von besonderem Interesse: Diarrhoe	Gesamtrate: Ergebnisse: 31,7 % vs. 17,5 % RR [95 %-KI]: 1,82 [1,10; 3,01] p-Wert = 0,0176	
UE bei mindestens 10 % der Patienten in einem Studienarm nach SOC und PT zu Woche 12	SOC: Erkrankungen des Gastrointestinaltrakts PT: Diarrhoe Ergebnisse: 41,3 % vs. 20,4 % RR [95 %-KI]: 2,03 [1,30; 3,16] p-Wert = 0,0011	
BDCAF: Behçet's Disease Current Activity Form; BD-QoL: Behçet's disease quality of life; BSAS: Behçet's Syndrome Activity Score; CTCAE: Common Terminology Criteria for Adverse Events; HR: Hazard Ratio; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); n.b.: nicht berechenbar; PT: Bevorzugter Begriff (preferred term); RR: Relatives Risiko; SF-36: 36-Item Short Form Health Survey, Version 2; SMD: Standardisierte Mittelwertdifferenz; SOC: Systemorganklasse (System Organ Class); SUE: Schwerwiegende unerwünschte/s Ereignis/se; UE: Unerwünschte/s Ereignis/se; VAS: Visuelle Analogskala		

Mortalität

In der Studie RELIEF traten keine Todesfälle auf.

Morbidität

Vollständige Remission oraler Aphthen

Unter Apremilast erreichten mehr als doppelt so viele Patienten eine vollständige Remission der oralen Aphthen im Vergleich zu Placebo (56,7 % vs. 25,2 %). Das relative Risiko (RR) für eine vollständige Remission der oralen Aphthen zwischen Apremilast und Placebo betrug 2,25 (95 %-Konfidenzintervall (KI): [1,55; 3,27]; $p < 0,0001$). Die Zeit bis zum Erreichen der vollständigen Remission wurde im Apremilast-Arm im Median nach etwa zwei Wochen erreicht. Bei Patienten, die Placebo erhielten, trat dies erst nach etwa acht Wochen ein. Der

Unterschied entspricht einem HR von 2,40 [95 %-KI: 1,69; 3,40]. Bei über 30 % der Patienten bildeten sich zudem nach vollständiger Remission bis Woche 12 keine neuen oralen Aphthen.

In einer zusätzlichen 52-wöchigen Extensionsphase wurden diese positiven Effekte von Apremilast bestätigt. Die zum Ende der Placebo-kontrollierten Phase (Woche 12) erzielte Remission der oralen Aphthen blieb für Patienten, die weiterhin Apremilast erhielten, bis zum Behandlungsende (Woche 64) erhalten. Patienten, die in der Placebo-kontrollierten Phase Placebo erhielten, erhielten stattdessen unter Beibehaltung der Verblindung in der Extensionsphase ebenfalls Apremilast. Für dieses Patientenkollektiv wurde bis Woche 64 eine mit dem Apremilast-Arm vergleichbare Remission der oralen Aphthen berichtet. Dies verdeutlicht den raschen Wirkeintritt und die langanhaltende Wirksamkeit von Apremilast. Nach dem Absetzen der Behandlung mit Apremilast zu Woche 64 trat bei allen Patienten, unabhängig davon, ob diese in der Placebo-kontrollierten Phase mit Apremilast oder Placebo behandelt wurden, ein Wiederauftreten von oralen Aphthen ein, wodurch die Kausalität mit dem Behandlungserfolg unter Apremilast unterstrichen wird.

Schmerzhaftigkeit der oralen Aphthen

Die patientenberichtete Schmerzhaftigkeit der oralen Aphthen mittels visueller Analogskala (VAS) ging unter Apremilast im Vergleich zu Placebo um das 1,9-fache zurück (SMD = -1,90; 95 %-KI: [-2,24; -1,57]). In der Extensionsphase hielt dieser Effekt bis zum Behandlungsende in Woche 64 an. Patienten aus der Placebo-Gruppe, die nach Ende der Placebo-kontrollierten Phase statt Placebo Apremilast erhielten, zeigten eine vergleichbare Reduktion der Schmerzhaftigkeit der oralen Aphthen, wodurch die Kausalität mit dem Behandlungserfolg unter Apremilast unterstrichen wird.

Behçet's Syndrome Activity Score (BSAS) und Behçet's Disease Current Activity Form (BDCAF)

Die Ergebnisse der Symptom- und Schmerzerfassung werden durch die Ergebnisse der Fragebögen Behçet's Syndrome Activity Score (BSAS) und Behçet's Disease Current Activity Form (BDCAF) gestützt. Beide Fragebögen sind valide und krankheitsspezifische Instrumente zur Beurteilung der manifestationsübergreifenden, klinischen Aktivität des Behçet-Syndroms, die weitverbreitet eingesetzt werden. Die Ergebnisse bestätigen die vom Patienten empfundene, rasch einsetzende und starke Verbesserung der Krankheitssymptomatik durch Apremilast im gesamten Behandlungsverlauf. Der validierte patientenberichtete Bogen BSAS zeigte eine Verringerung der Krankheitssymptomatik unter Apremilast (SMD = -0,76; 95 %-KI: [-1,05; -0,48]). Der validierte Bogen BDCAF bestätigte einheitlich in allen Skalen eine Verringerung der Krankheitsaktivität (Index: SMD = -0,30; 95 %-KI: [-0,57; -0,02]; Patientenwahrnehmung: SMD = -0,64; 95 %-KI: [-0,92; -0,36]; Untersucherwahrnehmung: SMD = -0,65; 95 %-KI: [-0,93; -0,37]).

Die zum Ende der Placebo-kontrollierten Phase beobachtete Verbesserung in den drei Einzelkomponenten der Krankheitsaktivität anhand des BDCAF-Fragebogens in der Apremilast-Gruppe hielt bis zum Ende der Extensionsphase an. In der Gruppe der Patienten, die in der Placebo-kontrollierten Phase Placebo und in der Extensionsphase Apremilast

erhielten, verbesserte sich die Krankheitsaktivität zu Woche 64 ebenfalls in allen drei Einzelkomponenten und lag jeweils auf vergleichbarem Niveau zur Apremilast-Gruppe.

Gesamtschau Morbidität

Patienten, die am Behçet-Syndrom leiden, sind besonders von den Ausprägungen der oralen Aphthen und den damit verbundenen Schmerzen betroffen, da im Verlauf der Erkrankung bei etwa 98 % der Patienten orale Aphthen auftreten (Altenburg 2012; Davatchi 2014). Die signifikante Remission der oralen Aphthen sowie der ebenfalls signifikante Rückgang des mit den oralen Aphthen verbundenen Schmerzempfindens zugunsten von Apremilast verdeutlichen die hohe therapeutische Wirksamkeit von Apremilast im AWG und den medizinischen Zusatznutzen. Die generelle Krankheitsaktivität durch das Behçet-Syndrom hat einen maßgeblichen Einfluss auf das Wohlbefinden der Patienten. Die validierten Fragebögen BSAS und BDCAF wurden durch Patienten und Untersucher unabhängig erhoben und zeigten einen signifikanten positiven Effekt von Apremilast auf die generelle Krankheitsaktivität des Behçet-Syndroms.

Gemäß AM-NutzenV § 5 Absatz 7 liegt somit ein **beträchtlicher** Zusatznutzen hinsichtlich der Morbidität vor.

Gesundheitsbezogene Lebensqualität

Behçet's disease quality of life (BD-QoL)

Die gesundheitsbezogene Lebensqualität anhand des krankheitsspezifischen Fragebogens „Behçet's disease quality of life“ (BD-QoL) verbesserte sich deutlich unter Apremilast (SMD = -0,51; 95 %-KI: [0,79; 0,24]). Wie auch das Ansprechen hinsichtlich der Symptome des Behçet-Syndroms bleibt auch das erzielte positive Ergebnis am Ende der Placebo-kontrollierten Phase bei Patienten, die Apremilast erhielten, bis zum Behandlungsende in Woche 64 auf einem vergleichbaren Niveau. Patienten, die nach Woche 12 von Placebo auf Apremilast wechselten, erreichten in der Extensionsphase ein vergleichbares Ansprechen wie die Patienten in der Apremilast-Gruppe.

36-Item Short Form Health Survey, Version 2 (SF-36v2)

Bei den Ergebnissen des generischen Fragebogen „36-Item Short Form Health Survey“ (SF-36) erreichten Patienten unter Apremilast in der Mehrheit der Subskalen eine patientenrelevante Verbesserung um $\geq 2,5$ Punkte. Dieser Wert wurde je nach Subskala bei 36,5 % (körperliche Funktionsfähigkeit) bis zu 60,6 % der Patienten (Vitalität) erreicht. Bei sieben von acht einzelnen Funktionsskalen sowie der Summenskala der psychischen Komponente (mental component summary, MCS) wurden signifikante Vorteile zugunsten von Apremilast festgestellt. Besonders auffällig ist das sehr hohe Ansprechen im Bereich der sozialen Aspekte des Fragebogens, d. h., der Unterschied zwischen Apremilast und Placebo ist besonders ausgeprägt bei der sozialen Funktionsfähigkeit (RR: 1,99 [95 %-KI: 1,37; 2,88]) und der psychischen Summenskala (RR: 1,72 [95 %-KI: 1,26; 2,36]).

Gesamtschau gesundheitsbezogene Lebensqualität

Die Symptomatik und Krankheitsaktivität des Behçet-Syndroms beeinflussen unmittelbar die gesundheitsbezogene Lebensqualität der betroffenen Patienten (Alpsoy 2016; Fabiani 2017). Neben einer Linderung der Symptomatik und den Ausprägungen der Manifestationen ist somit die Verbesserung der Lebensqualität der Patienten ein wichtiges Behandlungsziel (Hatemi 2018). Die anhand des BD-QoL Fragebogens erfasste krankheitsspezifische gesundheitsbezogene Lebensqualität verbesserte sich bis zum Ende der Placebo-kontrollierten Studienphase wesentlich. Diese Ergebnisse werden durch den generischen Fragebogen SF-36 gestützt. Bei sieben der acht Funktionsskalen sowie der psychischen Summenskala wurde eine statistisch signifikante und patientenrelevante Verbesserung der gesundheitsbezogenen Lebensqualität unter Apremilast gegenüber Placebo berichtet.

Die patientenrelevanten Effekte in der Kategorie gesundheitsbezogene Lebensqualität wurden in einer zusätzlichen 52-wöchigen Extensionsphase bestätigt: Sowohl bei Patienten, die statt Placebo eine Apremilast-Behandlung erhielten als auch bei zuvor mit Apremilast behandelten Patienten.

Gemäß AM-NutzenV § 5 Absatz 7 liegt somit ein **beträchtlicher** Zusatznutzen hinsichtlich der gesundheitsbezogenen Lebensqualität vor.

Sicherheit

Das am häufigsten berichtete unerwünschte Ereignis (UE) in der Placebo-kontrollierten Phase der Studie war Diarrhoe (41,3 % Apremilast vs. 20,7 % Placebo). In keiner der Fälle wurde die Diarrhoe als schwer (CTCAE ≥ 3) oder schwerwiegend berichtet. In allen weiteren einzelnen UE nach Systemorganklasse (System Organ Class, SOC), nach bevorzugtem Begriff (preferred term, PT) und von besonderem Interesse ergaben sich keine Unterschiede zwischen den beiden Studienarmen. Zusammengefasste UE der Kategorien gesamt, schwer und schwerwiegend sowie jene, die zum Therapieabbruch führten, wiesen vergleichbare Werte auf. Insbesondere da es sich um einen Placebovergleich handelt, liegt ein sehr ausgeglichenes Sicherheitsprofil von Apremilast vor.

Gesamtschau Sicherheit

Insgesamt ist damit im Hinblick auf die Nutzendimension Sicherheit **kein geringerer oder höherer Schaden** belegt.

Weitere Aspekte

Subgruppenanalysen

In keiner der Subgruppenanalysen konnten Belege für eine Modifikation des Behandlungseffektes bzgl. der Mortalität, Morbidität, Sicherheit und gesundheitsbezogene Lebensqualität festgestellt werden (siehe Abschnitt 4.3.1.3.2). Für die Ableitung des Zusatznutzens werden daher die Ergebnisse der Gesamtpopulation herangezogen.

Übertragbarkeit der Ergebnisse

Die Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext ist gegeben (siehe Abschnitt 4.3.1.2.1).

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Insgesamt ergibt sich auf Basis der Kriterien in § 5 Absatz 7 der AM-NutzenV unter Berücksichtigung des Schweregrades der Erkrankung für Apremilast im Vergleich zur zVT „Therapie nach Maßgabe des Arztes“ ein **Hinweis für einen beträchtlichen Zusatznutzen**.

International und auch auf europäischer Ebene gibt es außer Apremilast kein weiteres Arzneimittel, das zur Behandlung von oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind, zugelassen ist. Bei Patienten, für die eine systemische Therapie infrage kommt, zeigte die Studie RELIEF, dass Apremilast:

- signifikant häufiger und früher zu einer vollständigen Remission der oralen Aphthen führt
- die mit der Krankheit und ihrer Ausprägung verbundenen Schmerzen und Krankheitsaktivität signifikant verringert
- die gesundheitsbezogene Lebensqualität verbessert
- ein sehr gutes Sicherheitsprofil aufweist

Diese positiven Effekte konnten zudem bei einem Wechsel zu Apremilast oder bei Fortführung der Therapie über weitere 52 Wochen bestätigt bzw. beibehalten werden.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie

- Endpunkte
- Studientypen

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

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

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

Wie sind medizinischer Nutzen und Zusatznutzen von Apremilast bei erwachsenen Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt, im Vergleich zur zVT „Therapie nach Maßgabe des Arztes“ zu bewerten? Zur Beantwortung der Fragestellung wird die Zulassungsstudie RELIEF, eine zweiarmige doppelblinde, randomisierte, Placebo-kontrollierte Studie, herangezogen. Die Bewertung erfolgt unter Berücksichtigung patientenrelevanter Endpunkte in Bezug auf Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Sicherheit.

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.

Es wurden Ein- und Ausschlusskriterien definiert, um diejenigen Studien zu identifizieren, die zur Beantwortung der in Abschnitt 4.2.1 dargelegten Fragestellung geeignet sind. Die Selektionskriterien berücksichtigen die Kategorien Patientenpopulation, Intervention, Vergleichstherapie, Endpunkte, Studientyp und -dauer, Sprache, Publikationstyp sowie Studienstatus (Tabelle 4-2, Tabelle 4-3).

Tabelle 4-2: Übersicht über Ein- und Ausschlusskriterien von Studien zur Bewertung von Apremilast

Kriterium		Einschlusskriterium	Ausschlusskriterium
1	Patientenpopulation	Erwachsene Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt	Patientenpopulationen, die nicht dem zugelassenen Anwendungsgebiet entsprechen
2	Intervention	Apremilast gemäß Fachinformation: 30 mg oral, zweimal pro Tag (im Abstand von 12 Stunden)	Abweichende Intervention oder Dosierung
3	Vergleichstherapie	Keine Einschränkung	Keine Einschränkung
4	Endpunkte	Mortalität Morbidität Gesundheitsbezogene Lebensqualität Sicherheit	Patientenrelevante Endpunkte nicht erfasst
5	Studientyp	RCT Meta-Analyse von RCT Systematische Übersichtsarbeiten (Reviews) zu RCT	Nicht randomisierte, nicht kontrollierte Studien Nicht-klinische Studien Case Reports Case Series
6	Studiendauer	Keine Einschränkung	Keine Einschränkung
7	Sprache	Englisch oder Deutsch	Andere als Englisch oder Deutsch
8	Publikationstyp	Vollpublikation oder Bericht verfügbar	Keine Primärpublikation, Review-Artikel (narratives Review), Abstract, Poster
9	Studienstatus	Abgeschlossene Studien, für die Ergebnisse vorliegen	Noch nicht begonnene, laufende oder abgeschlossene Studien, für die noch keine Ergebnisse vorliegen
CONSORT: Consolidated Standards of Reported Trials; RCT: Randomisierte kontrollierte Studie (randomized controlled trial)			

Um für einen potenziellen indirekten Vergleich relevante Studien zu identifizieren, wurde für die zVT „Therapie nach Maßgabe des Arztes“ eine weitere systematische Literaturrecherche durchgeführt. Die Kriterien waren analog zur Suche für Apremilast.

Tabelle 4-3: Übersicht über Ein- und Ausschlusskriterien von Studien zur Bewertung der zweckmäßigen Vergleichstherapien (zVT)

Kriterium		Einschlusskriterium	Ausschlusskriterium
1	Patientenpopulation	Erwachsene Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt	Patientenpopulationen, die nicht dem zugelassenen Anwendungsgebiet entsprechen
2	Intervention	Therapie nach Maßgabe des Arztes mit einem der folgenden Wirkstoffe in der fachinformationskonformen Dosierung: Azathioprin, Prednison, Prednisolon, Thalidomid, Interferon-alfa, TNF- α -Inhibitoren, Ciclosporin, Dapson	Abweichende Intervention oder Dosierung
3	Vergleichstherapie	Keine Einschränkung	Keine Einschränkung
4	Endpunkte	Mortalität Morbidität Gesundheitsbezogene Lebensqualität Sicherheit	Patientenrelevante Endpunkte nicht erfasst
5	Studientyp	RCT Meta-Analyse von RCT systematische Übersichtsarbeiten (Reviews) zu RCT	Nicht randomisierte, nicht kontrollierte Studien Nicht-klinische Studien Case Reports Case Series
6	Studiendauer	Keine Einschränkung	Keine Einschränkung
7	Sprache	Englisch oder Deutsch	Andere als Englisch oder Deutsch
8	Publikationstyp	Vollpublikation oder Bericht verfügbar	Keine Primärpublikation, Review-Artikel (narratives Review), Abstract, Poster
9	Studienstatus	Abgeschlossene Studien, für die Ergebnisse vorliegen	Noch nicht begonnene, laufende oder abgeschlossene Studien, für die noch keine Ergebnisse vorliegen
CONSORT: Consolidated Standards of Reported Trials; RCT: Randomisierte kontrollierte Studie (randomized controlled trial); TNF- α : Tumornekrosefaktor-alpha			

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

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

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

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

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

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

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

Die bibliografische Literaturrecherche nach RCT mit dem zu bewertenden Arzneimittel im vorliegenden AWG wurde am 20. März 2020 gemäß der Vorgaben in den folgenden Datenbanken, jeweils über die Suchoberfläche Ovid, durchgeführt:

- Medical Literature Analysis and Retrieval System Online (MEDLINE) (inklusive „in-process & other non-indexed citations“)
- Excerpta Medica dataBASE (EMBASE)
- Cochrane Central Register of Controlled Trials (CENTRAL)

Die Suchstrategie wurde an die jeweilige Datenbank angepasst. Für MEDLINE und EMBASE wurde der nach Verfo empfohlene Suchfilter nach Wong verwendet (Wong 2006). Alle Suchvorgänge wurden nach den Population, Intervention, Comparison, Outcome (PICO)-Kriterien in den drei Blöcken „Indikation“, „Intervention“ und „Studiendesign“ durchgeführt. Für die Datenbank „CENTRAL“ wurde auf den Block „Studientyp“ verzichtet, da die Datenbank den Block „Studientyp“ bereits beinhaltet. Berücksichtigt wurden Publikationen in englischer und deutscher Sprache.

Die systematische bibliografische Literaturrecherche nach RCT zur zVT „Therapie nach Maßgabe des Arztes“ erfolgte am 20. März 2020 ebenfalls in den zuvor genannten Datenbanken. Dabei wurden die Wirkstoffe Azathioprin, Prednison, Prednisolon, Thalidomid, Interferon-alfa, TNF- α -Inhibitoren, Ciclosporin und Dapson erfasst. Die Suchstrategie entsprach dem zu bewertenden Arzneimittel.

Im Anhang 4-A sind die Suchstrategien detailliert und replizierbar dargestellt.

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 im vorliegenden AWG sowie mit den zuvor genannten zVT wurde zwischen 20. März 2020 und 15. April 2020 in den folgenden Studienregistern und Studienergebnisdatenbanken durchgeführt:

- [Clinicaltrials.gov](https://www.clinicaltrials.gov) (<https://www.clinicaltrials.gov>)
- EU Clinical Trials Register (EU-CTR) (<https://www.clinicaltrialsregister.eu>)
- International Clinical Trials Registry Platform (ICTRP) Search Portal (WHO) (<http://apps.who.int/0000000000/Default.aspx>)
- Clinical Data (Suchportal der European Medicines Agency, EMA) (<https://clinicaldata.ema.europa.de/web/cdp/home>)

- Arzneimittelinformationssystem (AMIS), Datenbank des Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI)
(<https://www.dimdi.de/dynamic/de/arzneimittel/arzneimittel-recherchieren/>)

Für die Suche kam für jedes Studienregister bzw. jede Studienergebnisdatenbank eine angepasste Suchstrategie zum Einsatz. Die Recherche für Clinical Data und AMIS beschränkte sich entsprechend den Vorgaben auf die Suche nach Einträgen mit Ergebnisberichten, die bereits anderweitig (z. B. über die bibliografische Literaturrecherche und Studienregisterrecherche) identifiziert wurden. Die Dokumentation der Suchstrategie war in diesen beiden Datenbanken nicht erforderlich. Es wurden abgeschlossene, abgebrochene und laufende Studien erfasst. Die Recherche wurde auf die Indikation mit den Stichwörtern „Behcet“ oder „Behçet“ begrenzt, um ausschließlich Studien im vorliegenden AWG von Apremilast zu identifizieren.

Im Anhang 4-B sind die Suchstrategien detailliert und replizierbar dargestellt.

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 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.

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

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 nach RCT mit dem zu bewertenden Arzneimittel im vorliegenden AWG sowie mit den zVT wurde am 30. März 2020 entsprechend den Vorgaben auf der Website des Gemeinsamen Bundesausschusses (G-BA) (<https://www.g-ba.de/informationen/nutzenbewertung/>) durchgeführt. Dabei wurde nach Nutzenbewertungen entsprechend der (frühen) Nutzenbewertung nach § 35a SGB V 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 Bewertung der in der bibliografischen Literaturrecherche identifizierten Publikationen erfolgte zunächst auf Titel- bzw. Abstract- und anschließend auf Volltextebene anhand der in Abschnitt 4.2.2 erläuterten Ein- und Ausschlusskriterien (Tabelle 4-2, Tabelle 4-3). Zwei Personen nahmen die Bewertung unabhängig voneinander vor, dabei wurden eventuelle Abweichungen in der Bewertung bis zu einer Konsensfindung diskutiert. Der Selektionsprozess wurde elektronisch dokumentiert.

Die Identifikation relevanter Studien aus Studienregistern und Studienergebnisdatenbanken folgte der gleichen Prozedur.

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.

Zur Bewertung der Aussagekraft der Nachweise wurden Angaben zu Verzerrungsaspekten sowohl auf Studienebene als auch auf Endpunktebene mithilfe des Bewertungsbogens in Anhang 4-F gemäß den oben beschriebenen Vorgaben des G-BA extrahiert und dokumentiert. Nach der Beurteilung des Verzerrungspotenzials auf Studien- und Endpunktebene zur Graduierung des

Ausmaßes der Ergebnissicherheit (s. o.) erfolgte die Klassifizierung der Wahrscheinlichkeit des Zusatznutzens für jeden patientenrelevanten Endpunkt gemäß den Anforderungen im IQWiG-Methodenpapier (IQWiG 2017).

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.

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 Charakteristika der in die Bewertung eingeschlossenen RCT RELIEF wurden anhand der Items 2b bis 14b des CONSORT-Statements separat dargestellt (Anhang 4-E). Studienziele und -design, Patientenpopulation, Interventionen und Zielkriterien sind in den Items 2b bis 6, Fallzahlbestimmung, Randomisierung, Verblindung und statistische Methoden in den Items 7 bis 12 und Patientenfluss, inklusive Flow-Chart, ferner Informationen zu Aufnahme und Behandlungsdauer der Patienten sowie zum Studienende in den Items 13 bis 14b abgebildet.

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

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.

Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

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

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

Patientencharakteristika

Die demografischen und krankheitsspezifischen Patientencharakteristika zu Baseline der Studie RELIEF sind in Tabelle 4-4 und Tabelle 4-5 dargestellt:

Tabelle 4-4: Demografische Charakterisierung der Patientenpopulation

Charakteristikum	Ausprägung
Alter (Jahre)	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
Alterskategorie 1 (Jahre), n (%)	<ul style="list-style-type: none"> • < 65 • ≥ 65
Alterskategorie 2 (Jahre), n (%)	<ul style="list-style-type: none"> • < 40 • ≥ 40 bis < 65 • ≥ 65
Geschlecht, n (%)	<ul style="list-style-type: none"> • Männlich • Weiblich
Ethnie, n (%)	<ul style="list-style-type: none"> • Ureinwohner Amerikas oder Alaskas • Asiatisch • Schwarz oder Afroamerikaner • Hawaiianer oder Einwohner anderer pazifischer Inseln • Weiß • Nicht erfasst / nicht berichtet
Region ^a n (%)	<ul style="list-style-type: none"> • Europa^b • Nordamerika^c • Asien^d • Rest der Welt^e
Körpergewicht (kg)	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
Gewichtskategorie (kg), n (%)	<ul style="list-style-type: none"> • < 50

Charakteristikum	Ausprägung
	<ul style="list-style-type: none"> • ≥ 50 bis < 60 • ≥ 60 bis < 70 • ≥ 70 bis < 85 • ≥ 85
Körpergröße (cm)	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
BMI ^f (kg/m ²), n (%)	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
BMI-Kategorie (kg/m ²), n (%) ^f	<ul style="list-style-type: none"> • $< 18,5$ • $\geq 18,5$ bis < 25 • ≥ 25 bis < 30 • ≥ 30
Alkoholkonsum, n (%)	<ul style="list-style-type: none"> • Ja <ul style="list-style-type: none"> < 1 Drink/Woche 1 bis 14 Drinks/Woche > 14 Drinks/Woche • Nein
Tabakkonsum, n (%) ^g	<ul style="list-style-type: none"> • Aktiver Raucher • Ehemaliger Raucher • Nichtraucher
<p>a: Die weitere Stratifizierungsmöglichkeit des Charakteristikums „Region“ in die Kategorien „Japan“ und „Nicht-Japan“ bleibt unberücksichtigt, da weniger als 20 % der Patienten aus Japan stammten. Die japanischen Patienten sind in der für das Nutzendossier berücksichtigten Regionsstratifizierung Kategorie „Asien“ erfasst. Die nicht-japanischen Patienten sind innerhalb der drei anderen Kategorien erfasst.</p> <p>b: Frankreich, Deutschland, Griechenland und Italien.</p> <p>c: USA.</p> <p>d: Japan und Südkorea.</p> <p>e: Libanon, Israel und Türkei.</p> <p>f: Berechnung auf Basis des Körpergewichts vor der ersten Dosis der Studienmedikation und der Körpergröße zu Screening.</p> <p>g: Patienten mit dem Status „ehemaliger Raucher“ wurden in der Auswertung als „aktive Raucher“ gezählt.</p> <p>BMI: Body Mass Index; k.A.: Keine Angabe; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ausprägung; SD: Standardabweichung (standard deviation); USA: Vereinigte Staaten von Amerika (United States of America)</p>	

Tabelle 4-5: Krankheitsspezifische Charakterisierung der Patientenpopulation

Charakteristikum	Ausprägung
Erkrankungsdauer des Behçet-Syndroms (Jahre)	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
Erkrankungsdauer nach Behçet-Syndrom-Kategorie (Jahre), n (%)	<ul style="list-style-type: none"> • ≤ 2 • > 2 bis ≤ 6 • > 6 bis < 10 • ≥ 10 • fehlend
Anzahl oraler Aphthen	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
Anzahl oraler Aphthen, n (%)	<ul style="list-style-type: none"> • ≥ 2 bis ≤ 5 • > 5 bis ≤ 10 • > 10
Schmerzhaftigkeit der oralen Aphthen mittels VAS	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
BSAS	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
BDCAF Indexwert	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
BD-QoL	<ul style="list-style-type: none"> • N • MW (SD) • Median (Min, Max)
<i>Vorangegangene Therapien</i>	
Vorbehandlung mit Immunsuppressiva, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Vorbehandlung mit Colchicin, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Vorbehandlung mit oralen Kortikosteroiden, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Vorbehandlung mit topischen Kortikosteroiden, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Vorbehandlung mit Biologika ^a , n (%)	<ul style="list-style-type: none"> • Ja • Nein
Vorbehandlung mit NSAR, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Vorbehandlung mit anderen Analgetika oder Anästhetika als NSAR	<ul style="list-style-type: none"> • Ja • Nein
<p>a: Vorherige Biologika-Therapie erlaubt für andere Indikationen, einschließlich anderer Manifestationen der BD.</p> <p>BD: Behçet-Syndrom (Behçet's disease); BDCAF: BD Current Activity Form; BD-QoL: BD Quality of Life; BSAS: Behçet's Syndrome Activity Score; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ausprägung; NSAR: Nichtsteroidale Antirheumatika; SD: Standardabweichung (standard deviation); VAS: Visuelle Analogskala</p>	

Patientenrelevante Endpunkte

Entsprechend § 2 Abs. 3 der AM-NutzenV besteht der Nutzen eines Arzneimittels in einem patientenrelevanten therapeutischen Effekt (BMJV 2019). Als patientenrelevant erachten G-BA und IQWiG übereinstimmend die Kategorien Mortalität, Morbidität, Lebensqualität und Sicherheit (G-BA 2020; IQWiG 2017). Dabei gilt es zu beurteilen, wie ein Patient fühlt, wie er seine Funktionen und Aktivitäten wahrnehmen kann und ob bzw. wie lange er überlebt (IQWiG 2017).

Die zur Nutzenbewertung herangezogenen patientenrelevanten Endpunkte sind in Tabelle 4-6 abgebildet. Eine Beschreibung der Patientenrelevanz, Operationalisierung und Validität findet sich im anschließenden Fließtext.

Tabelle 4-6: Patientenrelevante Endpunkte zum Nachweis des Zusatznutzens von Apremilast

Studienendpunkt	Mortalität	Morbidität	Gesundheitsbezogene Lebensqualität	Sicherheit
Todesfälle	■			
Vollständige Remission der oralen Aphthen (komplettes Ansprechen)		■		
Schmerzhaftigkeit oraler Aphthen mittels VAS		■		
Veränderung des BSAS zu Woche 12		■		
Veränderung des BDCAF zu Woche 12		■		
Veränderung des BD-QoL zu Woche 12			■	
Verbesserung des SF-36v2 um $\geq 2,5$ Punkte zu Woche 12			■	
Sicherheit <ul style="list-style-type: none"> • UE gesamt • SUE • Schwere UE • Therapieabbruch aufgrund von UE • UE von besonderem Interesse 				■

BD: Behçet-Syndrom (Behçet's disease); BDCAF: BD Current Activity Form; BD-QoL: BD Quality of Life; BSAS: Behçet's Syndrome Activity Score; SF-36: 36-Item Short Form Health Survey, Version 2; SUE: Schwerwiegende(s) unerwünschte(s) Ereignis(se); UE: Unerwünschte(s) Ereignis(se); VAS: Visuelle Analogskala

Mortalität

Die Mortalität wurde im Rahmen der Studie RELIEF über UE, die zum Tode führten, erfasst. Auf das Ergebnis dieser Nutzenbewertung haben die Ergebnisse zur Mortalität keinen Einfluss, da im Rahmen der Studie keine Todesfälle auftraten.

Morbidität

Vollständige Remission der oralen Aphthen (komplettes Ansprechen)

Patienten, die am Behçet-Syndrom leiden, sind besonders von den Ausprägungen der oralen Aphthen und den damit verbundenen Schmerzen betroffen, da im Verlauf der Erkrankung bei etwa 98 % der Patienten orale Aphthen auftreten (Altenburg 2012; Davatchi 2014). Die Darstellung des Endpunkts „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ erfolgt über den Anteil der Patienten mit vollständiger Remission der oralen Aphthen. Zur Stützung der Ergebnisse werden die folgenden Analysen zusätzlich und ergänzend dargestellt:

- Zeit bis zur vollständigen Remission der oralen Aphthen
- Anzahl oraler Aphthen
- Dauer der vollständigen Remission der oralen Aphthen

Patientenrelevanz

Orale Aphthen repräsentieren sowohl die häufigste Erstmanifestation als auch die häufigste klinische Manifestation des Behçet-Syndroms (Alpsoy 2007; Altenburg 2012; Sibley 2014; Zouboulis 1997). Charakteristisch für orale Aphthen sind Schmerzen sowie Einschränkungen bei Nahrungsaufnahme, sprachlicher Kommunikation und körperlicher Aktivität. Daraus resultiert für die Patienten eine maßgebliche Beeinträchtigung des Gesundheitsstatus und der Lebensqualität (Canpolat 2011; Gilworth 2004; Mumcu 2006). Die Läsionen treten rezidivierend auf, wobei ihre Anzahl pro Schub variiert. In den meisten Fällen sind mindestens zwei Läsionen gleichzeitig vorhanden (Davatchi 2010). Die Patienten empfinden ihre Erkrankung als umso belastender, je mehr Läsionen auftreten. Eine vollständige Remission der oralen Aphthen oder zumindest eine Verringerung ihrer Anzahl ist demzufolge mit einem unmittelbaren Nutzen für den Patienten assoziiert. Dies bestätigt die Patientenrelevanz des Endpunkts.

Operationalisierung

Die Zeit bis zur vollständigen Remission, die Zeit bis zum Rezidiv und die Anzahl der oralen Aphthen wurden im Studienverlauf vom Prüfarzt durch visuelles Auszählen erfasst.

Die statistische Analyse basierte auf der ITT-Population. Diese umfasste alle Patienten, bei denen mindestens die einmalige Behandlung mit der Studienmedikation erfolgte.

Der Endpunkt „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ wurde wie folgt operationalisiert:

- Anteil der Patienten mit vollständiger Remission der oralen Aphthen

Als vollständige Remission wird der Status definiert, in dem zum Untersuchungszeitpunkt keine oralen Aphthen erkennbar sind.

Zu Woche 12 fehlende Werte wurden mittels letzter verfügbarer Beobachtung (Last Observation Carried Forward, LOCF) imputiert (Hauptanalyse). Zusätzlich wurden zu Woche 12 fehlende Werte als Non-Responder imputiert (Sensitivitätsanalyse).

- Zeit bis zur vollständigen Remission

Für die Analyse des supportiv dargestellten Endpunkts „Zeit bis zur vollständigen Remission“ wurden alle Patienten berücksichtigt, die mindestens eine vollständige Remission im Zeitraum von Baseline bis einschließlich Woche 12 aufwiesen. Patienten ohne Ergebnis nach Baseline wurden zum Zeitpunkt der ersten Einnahme der Studienmedikation zensiert.

- Anzahl oraler Aphthen

Für den supportiv dargestellten Endpunkt „Anzahl oraler Aphthen“ wurden die Ergebnisse als Veränderung zu Woche 12 gegenüber Baseline dargestellt. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert.

- Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission

In die statistische Analyse des supportiven Endpunkts „Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission“ wurden diejenigen Patienten der ITT-Population eingeschlossen, die bis Woche 12 eine vollständige Remission der oralen Aphthen mit anschließendem Rezidiv zeigten. Als Zeit wurde die Dauer zwischen der ersten vollständigen Remission und dem ersten Wiederauftreten der oralen Aphthen definiert. Patienten ohne dokumentiertes Rezidiv zwischen Baseline und Woche 12 wurden zum Zeitpunkt der letzten Erfassung der oralen Aphthen zensiert. Patienten, bei denen nach erstmaligem Erreichen einer vollständigen Remission keine weiteren Ergebnisse vorlagen, wurden zum Zeitpunkt der dokumentierten vollständigen Remission zensiert.

Validität

Die in die Studiendurchführung involvierten Prüfarzte waren ausschließlich qualifizierte, mit dem Behçet-Syndrom vertraute Experten, die in der Lage waren, Ausprägung und Verlauf der oralen Aphthen über eine visuelle Auszählung zuverlässig zu erfassen. Die vollständige Remission der oralen Aphthen sowie ihr Wiederauftreten nach vollständiger Remission lassen sich darüber hinaus objektiv auf visuellem Wege durch den Prüfarzt erfassen. Die Validität des Endpunkts „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ kann daher bestätigt werden.

Weitere Manifestationen des Behçet-Syndroms

Zusätzlich zu den oralen Aphthen wurden weitere relevante Manifestationen des Behçet-Syndroms in der Studie RELIEF erfasst und dokumentiert. Im Dossier werden diese über den Studienverlauf hinweg deskriptiv dargestellt.

Folgende weitere Manifestationen des Behçet-Syndrom werden dargestellt:

- Genitale Aphthen
- Hautläsionen
- Arthritis
- Uveitis
- Gastrointestinale Manifestationen
- Vaskuläre Manifestationen
- Manifestationen des Zentralnervensystems

Da die Studie RELIEF nicht darauf ausgerichtet war, die weiteren Manifestationen ausgewogen zu erfassen und zu bewerten, ist eine interferenzstatistische Auswertung nicht zielführend und wird nicht zur Ableitung des Zusatznutzens herangezogen. Zudem mussten die weiteren schweren Manifestationen bei Baseline in einem nicht-behandlungsbedürftigen Zustand sein. Daher erfolgt ausschließlich eine deskriptive Darstellung.

Schmerzhaftigkeit oraler Aphthen

Patientenrelevanz

Das Auftreten oraler Aphthen wird häufig von lokalen Schmerzen begleitet. Es handelt sich damit um ein charakteristisches Begleitsymptom und spiegelt den Schweregrad der Erkrankung unmittelbar wider. Schmerz jeglicher Ursache führt bei den betroffenen Personen zu einer wesentlichen Verschlechterung der gesundheitsbezogenen Lebensqualität (Bellach 2000; Ellert 2004). Diese Assoziation wurde auch für die im Rahmen des Behçet-Syndroms auftretenden oralen Aphthen bestätigt (Bodur 2006; Moses Alder 2008; Mumcu 2009). Dies wiederum unterstreicht, dass eine Verringerung ihrer Schmerzhaftigkeit vom Patienten als unmittelbarer Nutzen empfunden wird. Die Patientenrelevanz für den Endpunkt ist somit gegeben.

Operationalisierung

Die Schmerzhaftigkeit der oralen Aphthen wurde zu allen Untersuchungszeitpunkten vom Patienten mittels VAS beurteilt (siehe Abschnitt 4.3.1.3.1.2). Die VAS umfasst eine gerade Linie von 100 mm Länge. Auf dieser Linie markiert der Patient sein aktuelles Schmerzempfinden zwischen linkem bzw. unterem Ausgangswert (keine Schmerzen = 0 mm) und rechtem bzw. oberem Maximalwert (unerträgliche Schmerzen = 100 mm).

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten Patienten berücksichtigte, die mindestens eine Dosis der Studienmedikation erhalten und die VAS zu Baseline komplettiert hatten. Die Auswertung erfolgte mit drei Analyseansätzen. Im ersten

Ansatz werden die Ergebnisse als Veränderung zu Woche 12 gegenüber Baseline dargestellt, wobei zu Woche 12 fehlende Werte mittels LOCF imputiert wurden. Im zweiten Ansatz erfolgte eine weitere Analyse mittels gemischtem Modell mit Messwiederholungen (Mixed-effects model for repeated measures, MMRM). Im dritten Ansatz erfolgt die Ergebnisdarstellung als Veränderung gegenüber Baseline zu den einzelnen Erhebungszeitpunkten bis Woche 12 mittels vorliegender Beobachtungen (Observed cases, OC).

Validität

Die VAS ist eine valide und verlässliche Methode zur Selbstbewertung des aktuellen Schmerzempfindens durch den Patienten mit einer hohen Sensitivität (Schomacher 2008). Dies stützt die Validität des Endpunkts „Schmerzhaftigkeit oraler Aphthen“.

Veränderung des Behçet's Syndrome Activity Score (BSAS) zu Woche 12

Patientenrelevanz

Der BSAS stellt ein Messinstrument zur Bewertung der klinischen Aktivität des Behçet-Syndroms durch den Patienten dar. Patientenberichtete Endpunkte (patient-reported outcomes, PRO) zum Krankheitsgeschehen und dessen Änderungen durch eine Arzneimittelintervention spiegeln die subjektive Einschätzung des Patienten unmittelbar wider und sind somit patientenrelevant. Als krankheitsspezifischer Fragebogen adressiert der BSAS alle im Rahmen des Behçet-Syndroms möglichen Organbeteiligungen. Dies unterstreicht seine Bedeutung im Rahmen der Erkrankung und die Patientenrelevanz des Endpunkts „Veränderung des BSAS zu Woche 12“. Weiterhin werden innerhalb des Scores die oralen Aphthen, ein wesentliches Symptom der Erkrankung, abgedeckt.

Operationalisierung

Der BSAS beinhaltet zehn Fragen zur Bewertung der Krankheitsaktivität in den letzten vier Wochen. Erfasst werden orale und genitale Aphthen, Hautläsionen, ferner gastrointestinale, neurologische, vaskuläre, Gelenk- und Augenbeteiligungen sowie die aktuelle allgemeine Krankheitsaktivität. Jede Frage wird mit einer Punktzahl zwischen 0 und 10 bewertet, der Gesamtscore liegt somit zwischen 0 und 100. Ein höherer Wert bedeutet eine höhere Krankheitsaktivität und entspricht damit einem schlechteren Zustand (Yilmaz 2013).

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten Patienten berücksichtigte, die mindestens eine Dosis der Studienmedikation erhalten und den BSAS zu Baseline komplettiert hatten. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline dargestellt. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert.

Validität

Der BSAS-Fragebogen wurde in englischer Sprache entwickelt und in weitere Sprachen übersetzt und validiert (Choi 2016; Yilmaz 2013). Zur Verwendung in der Studie RELIEF wurde er auch ins Deutsche übersetzt. Die Zertifizierung der deutschen Sprachversion erfolgte durch eine unabhängige dritte Partei (Amgen 2014c).

Veränderung des Behçet's Disease Current Activity Form (BDCAF) zu Woche 12

Patientenrelevanz

Der BDCAF ist ein Fragebogen zur Beurteilung der aktuellen Aktivität des Behçet-Syndroms, der sowohl vom Patienten als auch vom Prüfarzt und voneinander unabhängig komplettiert wird. Als krankheitsspezifischer Fragebogen adressiert der BDCAF alle im Rahmen des Behçet-Syndroms möglichen Organbeteiligungen. Die oralen Aphthen, ein wesentliches Symptom der Erkrankung, werden ebenfalls abgedeckt.

Operationalisierung

Der BDCAF-Fragebogen besteht aus drei unabhängigen Komponenten zur Krankheitsaktivität in den letzten vier Wochen, dem Behçet's Disease Current Activity Indexwert (BDCAI-Score) sowie der Wahrnehmung der Krankheitsaktivität durch den Patienten und der Gesamtwahrnehmung der Krankheitsaktivität durch den Untersucher.

Der BDCAI beinhaltet insgesamt 12 Fragen zur Krankheitsaktivität in den letzten vier Wochen, aufgeteilt in Fragen zu:

- Kopfschmerz,
- oralen Aphthen
- genitalen Aphthen,
- Hautläsionen
- Gelenkbeteiligung
- gastrointestinalen Symptomen
- Beteiligung der Augen
- Beteiligung des Nervensystems
- großen Gefäßentzündungen

Der Gesamtscore des BDCAI liegt zwischen 0 und 12 (Choi 2015; Lawton 2004; Yilmaz 2013). Ein höherer Wert ist mit einer höheren Krankheitsaktivität assoziiert.

Die Wahrnehmung der Krankheitsaktivität in den letzten vier Wochen durch den Patienten und den Untersucher erfolgt jeweils mit Hilfe einer siebenstufigen Likert-Skala von sehr schlecht bis sehr gut (Punktwert von 1 bis 7) (Choi 2015; Lawton 2004).

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten Patienten berücksichtigte, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Score verfügbar war. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline dargestellt. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert.

Validität

Der BDCAF ist ein häufig eingesetztes, verlässliches Messinstrument zur Beurteilung und Dokumentation klinischer Aktivität bei Patienten mit dem Behçet-Syndrom. Er wurde in englischer Sprache entwickelt (Bhakta 1999) und in verschiedene Sprachen übersetzt und validiert (Choi 2015; Hamuryudan 1999; Harzallah 2013; Neves 2007).

Gesundheitsbezogene Lebensqualität

Veränderung des Behçet's Disease Quality of Life (BD-QoL) zu Woche 12

Patientenrelevanz

Entsprechend der AM-NutzenV und der Verfahrensordnung des G-BA ist die Verbesserung der gesundheitsbezogenen Lebensqualität als patientenrelevanter Endpunkt anzusehen (BMJV 2019; G-BA 2020). PRO zu gesundheitsbezogener Lebensqualität und ihren Änderungen geben das subjektive Befinden des Patienten unmittelbar wieder. Der im Rahmen der Studie RELIEF verwendete Fragebogen, BD-QoL, erfüllt als krankheitsspezifischer Fragebogen die Anforderungen an eine verlässliche Erfassung der PRO bei Patienten mit dem Behçet-Syndrom.

Operationalisierung

Der BD-QoL beinhaltet 30 Aussagen zum Einfluss des Behçet-Syndroms auf Aktivitäten des täglichen Lebens und die gesundheitsbezogene Lebensqualität, die der Patient mit „wahr“ (Punktwert = 1) oder „nicht wahr“ (Punktwert = 0) kommentiert. Der Gesamtscore ist eine einfache ungewichtete Summe der Einzelergebnisse und liegt zwischen 0 und 30; ein höherer Wert markiert eine stärkere subjektive Beeinträchtigung.

Für die statistische Analyse wurde die Intention-to-treat (ITT)-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Gabe der Studienmedikation erhalten und den Fragebogen zu Baseline komplettiert hatten. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert.

Validität

Der BD-QoL stellt ein valides Messinstrument dar, das entwickelt wurde, um die Auswirkungen des Behçet-Syndroms auf die Lebensqualität des Patienten zu bewerten (Gilworth 2004). Der Einsatz des Bogens zur Erhebung der Lebensqualität wurde validiert (Touma 2011; Yi 2008). Zur Verwendung in der Studie RELIEF wurde der BD-QoL aus der englischen Originalsprache u. a. in die deutsche und türkische Sprache übersetzt. Die Zertifizierung der beiden Sprachversionen erfolgte durch eine dritte unabhängige Partei (Amgen 2014a, 2014b).

Verbesserung des 36-Item Short Form Health Survey, Version 2 (SF-36) um $\geq 2,5$ Punkte zu Woche 12

Patientenrelevanz

Beim SF-36 handelt es sich um einen generischen Fragebogen zur Selbstbeurteilung der gesundheitsbezogenen Lebensqualität durch den Patienten (Ware, JR 1992; Ware, JR 2000). Der SF-36 zeigte bei Patienten mit dem Behçet-Syndrom einen Zusammenhang zwischen

Krankheitsaktivität und -schwere auf der einen und der Beeinträchtigung der gesundheitsbezogenen Lebensqualität auf der anderen Seite (Canpolat 2011; Ertam 2009; Melikoğlu 2014; Mumcu 2006). Der SF-36 ist weiterhin ein valides Messinstrument, welches in vielen anderen Erkrankungen (z. B. Rheumatoide Arthritis und Psoriasis) eingesetzt wird.

Operationalisierung

Der SF-36 bildet die gesundheitsbezogene Lebensqualität über zwei Hauptdimensionen ab, die körperliche und die psychische Summenskala, die ihrerseits in jeweils vier Subskalen gegliedert sind. Zu den Subskalen der körperlichen Summenskala (physical component summary, PCS) gehören körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen und allgemeine Gesundheitswahrnehmung. Die psychische Summenskala untergliedert sich in die Subskalen für Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden. Zwei bis zehn Fragen pro Subskala ergeben insgesamt 36 Fragen, von denen eine Frage zur Gesundheitsveränderung in den vergangenen 12 Monaten in der Auswertung unberücksichtigt bleibt (Ware, JR 1992; Ware, JR 2000). Alle Fragen werden mit Hilfe einer drei- bis sechsstufigen Likert-Skala beantwortet und die erhaltenen Werte für jede Subskala in einen Bereich zwischen 0 (schlechteste Lebensqualität) und 100 (beste Lebensqualität) transformiert (Morfeld 2008; Ware, JR 2000). Die Scores der beiden Summenskalen werden auf einen Mittelwert von 50 und eine Standardabweichung von 10 transformiert. Höhere Werte bedeuten eine bessere gesundheitsbezogene Lebensqualität.

Die statistische Analyse basierte auf denjenigen Patienten der ITT-Population, die zu Woche 12 eine Verbesserung um mindestens 2,5 Punkte in jeder Skala aufwiesen. Patienten mit fehlenden Werten zu Woche 12 wurden als Non-Responder eingestuft.

Validität

Der SF-36 ist ein weltweit eingesetztes generisches PRO-Instrument zur Messung der gesundheitsbezogenen Lebensqualität (Ware, JR 1992; Ware, JR 2000). Er wurde in mehr als 100 Sprachen übersetzt (Methode der Rückübersetzung) und validiert, darunter auch in Deutsch (Bullinger 1995). Die Version 2 des SF-36 ist ebenfalls in über 100 Übersetzungen verfügbar; die deutsche Version 2 wurde erstmals im Jahr 2002 eingesetzt (Ellert 2013; Morfeld 2005). Auch der G-BA berücksichtigt den SF-36-Fragebogen regelhaft in seinen Nutzenbewertungsverfahren zur Einstufung des Zusatznutzens, beispielsweise bei Guselkumab (Anwendungsgebiet Plaque Psoriasis) (G-BA 2018).

Sicherheit

Unerwünschte Ereignisse (UE)

Patientenrelevanz

UE als Maß für die Verträglichkeit einer Therapie (therapiebedingte Morbidität) werden generell als patientenrelevanter Endpunkt betrachtet, da sie sich unmittelbar auf den Gesundheitszustand eines Patienten auswirken (BMJV 2019; G-BA 2020; IQWiG 2017).

Operationalisierung

Folgende sicherheitsrelevanten Endpunkte wurden berücksichtigt:

- UE gesamt
- Schwerwiegende UE (SUE)
- Schwere UE
- Therapieabbruch aufgrund von UE
- UE von besonderem Interesse
 - Diarrhoe
 - Suizidgedanken und suizidales Verhalten
 - Maligne Erkrankungen
 - Schwere kardiale Komplikation
 - Opportunistische Infektionen

Es wurden alle UE berücksichtigt, die vom Tag der ersten Einnahme bis maximal 28 Tage nach der letzten Einnahme der Studienmedikation auftraten.

Zur Beurteilung des Schweregrades wurden die UE nach der Ausprägung „leicht“, „mäßig“ oder „schwer“ klassifiziert. Die Kodierung der UE basierte auf der Version 20.0 des Medizinischen Wörterbuchs für Aktivitäten im Rahmen der Arzneimittelzulassung (Medical Dictionary for Regulatory Activities, MedDRA). Sie erfolgte hierarchisch von der entsprechenden SOC zum PT (ICH 2011, 2017).

Die statistische Analyse der Sicherheitsendpunkte erfolgte für die Sicherheitspopulation, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation (as-treated) erhalten hatten. Zum einen wird die Gesamtrate der sicherheitsrelevanten Endpunkte zu Woche 12 dargestellt und zum anderen diejenigen UE mit SOC und PT, die bei mindestens 10 % der Patienten sowie bei mindestens zehn Patienten und mindestens 1 % der Patienten in einer Behandlungsgruppe bis Woche 12 auftraten.

Validität

Zur Gewährleistung der Patientensicherheit stand ein unabhängiges externes Datenüberwachungskomitee bereit, das die Einhaltung der guten klinischen Praxis (Good Clinical Practice, GCP)-Anforderungen im Studienverlauf in regelmäßigen Abständen überwachte (Amgen 2017). Eine subjektive Beurteilung der UE durch den Prüfarzt kann somit weitgehend ausgeschlossen und der Endpunkt als valide angesehen werden.

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.

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.

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 erforderlich.

Für die Nutzenbewertung von Apremilast im vorliegenden AWG wird ausschließlich die Studie RELIEF herangezogen.

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.

Nicht zutreffend.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

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

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

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

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

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

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

Subgruppenanalysen waren in der Studie RELIEF für folgende Effektmodifikatoren a priori festgelegt (Tabelle 4-7):

Tabelle 4-7: A priori festgelegte Subgruppenanalysen der Studie RELIEF

Subgruppenmerkmal	Ausprägung
Geschlecht	<ul style="list-style-type: none"> • Männlich • Weiblich
Alter (Jahre)	<ul style="list-style-type: none"> • < 40 • ≥ 40 bis < 65 • ≥ 65
Anzahl oraler Aphthen zu Baseline	<ul style="list-style-type: none"> • ≥ 2 bis ≤ 5 • > 5 bis ≤ 10 • > 10
Geografische Region ^a	<ul style="list-style-type: none"> • Europa • Nordamerika • Asien • Rest der Welt
Ethnie	<ul style="list-style-type: none"> • Weiß • Nicht weiß

Subgruppenmerkmal	Ausprägung
Gewicht zu Baseline (kg)	<ul style="list-style-type: none"> • < 50 • ≥ 50 bis < 60 • ≥ 60 bis < 70 • ≥ 70 bis < 85 • ≥ 85
BMI zu Baseline (kg/m ²)	<ul style="list-style-type: none"> • < 18,5 • ≥ 18,5 bis < 25 • ≥ 25 bis < 30 • ≥ 30
Alkoholkonsum	<ul style="list-style-type: none"> • Ja • Nein
Tabakkonsum	<ul style="list-style-type: none"> • Aktiver Raucher • Ehemaliger Raucher • Nichtraucher
Erkrankungsdauer (Jahre)	<ul style="list-style-type: none"> • ≤ 2 • > 2 bis ≤ 6 • > 6 bis ≤ 10 • > 10
Vorherige Colchicin-Einnahme	<ul style="list-style-type: none"> • Ja • Nein
Vorherige Einnahme oraler Kortikosteroide	<ul style="list-style-type: none"> • Ja • Nein
<p>a: Die weitere Stratifizierungsmöglichkeit des Charakteristikums „Region“ in die Kategorien „Japan“ und „Nicht-Japan“ bleibt unberücksichtigt, da weniger als 20 % der Patienten aus Japan stammten. Die japanischen Patienten sind in der für das Nutzendossier berücksichtigten Regionsstratifizierung Kategorie „Asien“ erfasst. Die nicht-japanischen Patienten sind innerhalb der drei anderen Kategorien erfasst.</p> <p>BMI: Body Mass Index</p> <p>Quelle: Amgen 2019</p>	

Begründung für die Trennpunkte bei der Kategorisierung von quantitativen Merkmalen

Alter

Die definierten Trennpunkte orientieren sich zum einen am Durchschnittsalter der Patienten zum Zeitpunkt der Erstmanifestation des Behçet-Syndroms und zum anderen an den für die erwachsene Bevölkerung üblichen Bereichen (Alpsoy 2007; Altenburg 2012; Zouboulis 1997). Das Behçet-Syndrom manifestiert sich meistens im dritten oder vierten Lebensjahrzehnt (20 bis 40 Jahre). Selten tritt es im Kindesalter (bis 15 Jahre) oder bei Erwachsenen jenseits des fünfzigsten Lebensjahres auf (Davatchi 2014; Greco 2018; Yazici 1984).

Die im Dossier dargestellten Trennpunkte für „Alter“ wurden gemäß statistischem Analyseplan der Studie RELIEF a priori festgelegt.

Anzahl oraler Aphthen zu Baseline

Die definierten Trennpunkte orientieren sich an Angaben in der Literatur, die von zwei oder mehr oralen Läsionen pro Rezidiv bzw. von einem gesicherten Behçet-Syndrom bei

Vorhandensein von mehr als sechs oralen Aphthen zur selben Zeit berichten (Davatchi 2010; Main 1992).

Die im Dossier dargestellten Trennpunkte für „Anzahl oraler Aphthen zu Baseline“ wurden gemäß statistischem Analyseplan der Studie RELIEF a priori festgelegt.

Gewicht zu Baseline

Die definierten Trennpunkte markieren die für erwachsene Personen üblichen Bereiche.

Die im Dossier dargestellten Trennpunkte für „Gewicht zu Baseline“ wurden gemäß statistischem Analyseplan der Studie RELIEF a priori festgelegt.

BMI zu Baseline

Die definierten Trennpunkte orientieren sich an den von der WHO festgelegten Kategorien für erwachsene Personen mit Untergewicht bis Normalgewicht, Übergewicht und Adipositas (WHO 2019).

Die im Dossier dargestellten Trennpunkte für „BMI zu Baseline“ wurden gemäß statistischem Analyseplan der Studie RELIEF a priori festgelegt.

Erkrankungsdauer

Die definierten Trennpunkte orientieren sich an der Spezifität des Behçet-Syndroms, bei dem es sich um eine chronische Erkrankung handelt, die im Zeitverlauf intermittierend auftreten kann.

Die im Dossier dargestellten Trennpunkte für „Erkrankungsdauer“ wurden gemäß statistischem Analyseplan der Studie RELIEF a priori festgelegt.

Darstellung der Subgruppenanalysen

Subgruppenanalysen waren ausschließlich für den primären Endpunkt der Studie RELIEF, die Fläche unter der Kurve (area under the curve, AUC) von Baseline zu Woche 12 (AUC_{W0-12}) für die Anzahl der oralen Aphthen, a priori definiert. Für die im Rahmen des Dossiers betrachteten Endpunkte zu Morbidität, gesundheitsbezogener Lebensqualität und Sicherheit (siehe Abschnitt 4.2.5.2) erfolgten die Subgruppenanalysen post-hoc.

Die Subgruppenanalysen der Wirksamkeitsendpunkte basierten auf den Daten der ITT-Population. Für die Subgruppenanalysen der Sicherheitsendpunkte wurden die Daten der Sicherheitspopulation herangezogen.

Die Prüfung auf Effektmodifikation für die binären Endpunkte erfolgte mittels Cochran-Mantel-Haenszel-Test für die Responder-Analysen und anhand eines stratifizierten Cox-Regressionsmodells für die Time-to-Event-Analysen. Für die kontinuierlichen Endpunkte wurde eine Kovarianzanalyse (analysis of covariance, ANCOVA) durchgeführt mit Behandlung, Geschlecht und Region als Faktoren und dem Baseline-Wert als Kovariable.

Im Rahmen der Subgruppenanalysen wurden mögliche Effektmodifikationen bezüglich der Effektschätzer mit Hilfe des Interaktions-p-Wertes untersucht und tabellarisch dargestellt. Bei Vorliegen einer Interaktion ($p < 0,05$) wurden die Ergebnisse für die jeweiligen Subgruppen separat beschrieben.

Die Quantifizierung der Unterschiede zwischen den Behandlungen in jeder Subgruppe erfolgte für die binären Endpunkte über die Effektschätzer RR, Odds Ratio (OR), Risikodifferenz (RD) und HR mit den dazugehörigen 95 %-KI und den p-Werten. Für die kontinuierlichen Endpunkte wurden die Kleinste-Quadrate-Mittelwertdifferenz (least-square-Mittelwertdifferenz, LS-MWD) und die standardisierte Mittelwertdifferenz (SMD) nach Hedges' g mit den dazugehörigen 95 %-KI und den p-Werten berechnet.

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

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.

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).

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).*

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

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

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

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

- *Art und Umfang von Sensitivitätsanalysen.*

Es wurde nach Studien mit einem der als zVT infrage kommenden Wirkstoffe für einen adjustierten indirekten Vergleich gesucht.

Aus den Treffern der systematischen Recherche wurde der generierte Studienpool anhand folgender Merkmale auf die Durchführbarkeit eines adjustierten indirekten Vergleichs geprüft:

- Vergleichbarkeit der Patientenpopulation hinsichtlich:
 - Ein- und Ausschlusskriterien
 - Baseline-Charakteristika
- Vergleichbarkeit der Intervention (Brückenkomparator):
 - Effekte in den Placebo-Gruppen
- Vergleichbarkeit der Operationalisierung der patientenrelevanten Endpunkte

Fehlende Informationen führten zum Ausschluss der jeweiligen Studie. Auf Vergleichbarkeit der Interventionen wird nur getestet, wenn die Effektschätzer in der Placebo-Gruppe der Apremilast-Studie RELIEF mit der Placebo-Gruppe der jeweiligen Studie mit einem der als zVT infrage kommenden Wirkstoffe vergleichbar sind.

Sollten die Voraussetzungen erfüllt sein, wird das Verfahren von Bucher et al. verwendet (Bucher 1997).

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 4-8: 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
RELIEF	ja	ja	abgeschlossen	Placebo-kontrollierte Doppelblindphase: 12 Wochen aktive Behandlungsphase: 52 Wochen	Apremilast 30 mg bid Placebo
BCT-001	ja	ja	abgeschlossen	Placebo-kontrollierte Doppelblindphase: 12 Wochen aktive Behandlungsphase: 12 Wochen	Apremilast 30 mg bid Placebo
bid: zweimal täglich (bis in die)					

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

Die Informationen in Tabelle 4-8 entsprechen dem Stand vom 30. März 2020.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-8 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-9: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
BCT-001 ^a	Ausschlussgrund 1 ("Patientenpopulationen, die nicht dem zugelassenen Anwendungsgebiet entsprechen")
a: Medikamentöse Vorbehandlung des Behçet-Syndroms nicht verpflichtend. Damit kann keine Aussage darüber getroffen werden, ob Patienten für eine systemische Therapie infrage kommen.	

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.

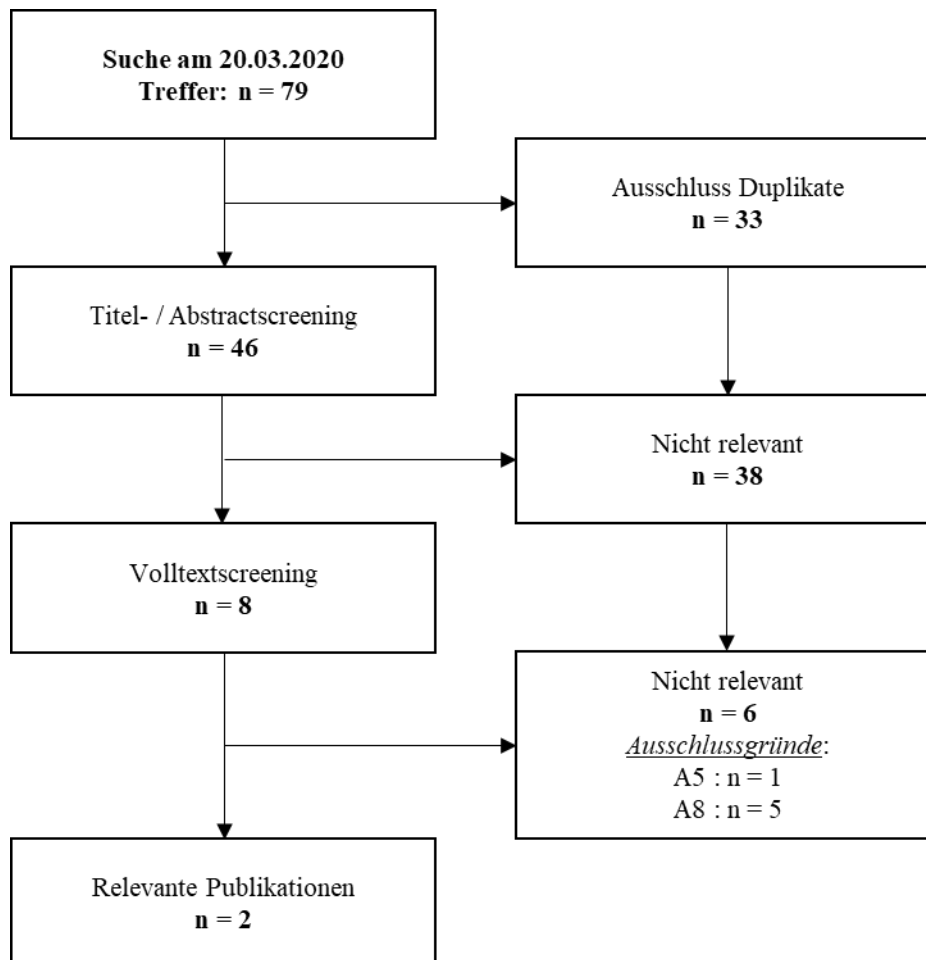


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Zur Identifizierung relevanter Primärpublikationen für RCT mit dem zu bewertenden Arzneimittel wurde am 20. März 2020 eine systematische bibliografische Literaturrecherche durchgeführt. Das vollständige Suchprofil befindet sich in Anhang 4-A.

Zwei Personen nahmen die Bewertung anhand der in Tabelle 4-2 definierten Ein- und Ausschlusskriterien unabhängig voneinander vor, dabei wurden eventuelle Abweichungen in der Bewertung bis zu einer Konsensfindung diskutiert. Der Selektionsprozess wurde elektronisch dokumentiert. Die bibliografische Literaturrecherche zu RCT mit dem zu bewertenden Arzneimittel Apremilast im vorliegenden AWG ergab 79 Treffer, darunter 33 Duplikate. Von den verbliebenen 46 Treffern wurden 38 nach Titel- / Abstractscreening und sechs nach anschließendem Volltextscreening begründet ausgeschlossen. Somit konnten zwei für die Nutzenbewertung relevante Publikationen identifiziert werden.

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 4-8) 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 4-10: 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)
RELIEF BCT-002 NCT02307513 2014-002108-25	clinicaltrials.gov (ClinicalTrials.gov 2014a) EU Clinical Trials Register (EU-CTR 2014)	ja	ja	abgeschlossen
BCT-001 NCT00866359 2008-002722-11 ^b	clinicaltrials.gov (ClinicalTrials.gov 2014b) WHO ICTRP (ICTRP 2014)	ja	ja	abgeschlossen
<p>a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>b: Ausschluss der Studie, da medikamentöse Vorbehandlung des Behçet-Syndroms nicht verpflichtend.</p> <p>EU-CTR: European Union Clinical Trials Register; ICTRP: International Clinical Trials Registry Platform; WHO: World Health Organization</p>				

Die Studie BCT-001 wurde aufgrund der nicht verpflichtenden medikamentösen Vorbehandlung des Behçet-Syndroms als nicht relevant ausgeschlossen. Daher kann für diese Studie nicht angenommen werden, dass sie für Patienten, für die eine systemische Therapie infrage kommt, repräsentativ ist und somit die Zielpopulation der vorliegenden Nutzenbewertung widerspiegelt. Dies war in der Studie BCT-001 bei Patienten ohne

medikamentöse Vorbehandlung nicht der Fall, da die internationalen Leitlinien bei diesen Patienten in der Erstlinienbehandlung oraler Aphthen im Rahmen des Behçet-Syndroms topische Kortikosteroide empfehlen (Hatemi 2018). Die Studie RELIEF hingegen bildet die Patientenpopulation ab, da hier alle Patienten erst nach mindestens einer nicht-biologischen medikamentösen Vorbehandlung und bei Eignung für eine systemische Therapie für die Studie infrage kamen.

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.

Stand der Informationen in Tabelle 4-10 ist der 30. März 2020.

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 4-8) 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 4-11: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

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

Geben Sie an, welchen Stand die Information in Tabelle 4-11 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Informationen in Tabelle 4-11 ist der 30. März 2020.

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 4-12: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienberichte (ja/nein [Zitat])	Register-einträge ^c (ja/nein [Zitat])	Publikation und sonstige Quellen ^d (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebo-kontrolliert						
RELIEF	ja	ja	nein	ja (Amgen 2019)	ja clinicaltrials.gov (ClinicalTrials.gov 2014a) EU Clinical Trials Register (EU-CTR 2014)	ja (Hatemi 2019)
<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>EU-CTR: European Union Clinical Trials Register</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 4-13: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
RELIEF	RCT, doppelblind, multizentrisch, Placebo-kontrolliert, parallel	Erwachsene Patienten (≥ 18 Jahre) mit diagnostiziertem Behçet-Syndrom anhand der ISG Kriterien, mindestens eine nicht-biologische Vorbehandlung des Behçet-Syndroms	Gruppe 1: Apremilast (n = 104) Gruppe 2: Placebo (n = 103)	Screening-Phase: bis zu sechs Wochen Placebo-kontrollierte Doppelblindphase: 12 Wochen Aktive Behandlungsphase: 52 Wochen (beide Gruppen Apremilast) Nachbeobachtung: vier Wochen Datenschnitte nach Komplettierung der Woche 12 sowie nach Woche 28 (Wirksamkeit) und nach Behandlung von 100 Patienten über ein Jahr (Sicherheit)	53 Studienzentren in zehn Ländern: Asien (Japan, Südkorea) Europa (Frankreich, Deutschland, Griechenland, Italien) USA Rest der Welt (Libanon, Israel, Türkei) 12/2014 – 9/2018	<u>Primärer Endpunkt:</u> AUC für die Anzahl oraler Aphthen <u>Sekundäre Endpunkte:</u> Anzahl, Ansprechrare und Schmerzhaftigkeit der oralen Aphthen, Zeit bis zu vollständiger Remission, Zeit bis zum Rezidiv, Krankheitsaktivität über PRO (mittels BSAS und BDCAF) und gesundheitsbezogener Lebensqualität (mittels BD-QoL und SF-36), Sicherheit (UE)
<p>AUC: Fläche unter der Kurve (area under the curve); BDCAF: Behçet’s Disease Current Activity Form; BD-QoL: Behçet’s disease Quality of Life; BSAS: Behçet’s Syndrome Activity Score; ISG: International Study Group for Behçet’s Disease; PRO: Patientenberichtete Endpunkte (patient-reported outcome); RCT: randomisierte kontrollierte Studie (randomized controlled trial); SF-36: 36-Item Short Form Health Survey, Version 2; UE: Unerwünschte/s Ereignis/se; USA: Vereinigte Staaten von Amerika (United States of America)</p> <p>Quelle: Amgen 2019</p>						

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

Studie	Apremilast	Placebo	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
RELIEF	Apremilast 30 mg p.o. bid im Abstand von ca. 12 Stunden für einen Zeitraum von 64 Wochen	Placebo p.o. bid im Abstand von ca. 12 Stunden für einen Zeitraum von 12 Wochen im Anschluss daran: Apremilast 30 mg p.o. bid im Abstand von ca. 12 Stunden für einen Zeitraum von 52 Wochen	Die Patienten waren mit mindestens einer nicht-biologischen Therapie des Behçet-Syndroms, z. B. mit topischen Kortikosteroiden oder systemischen Wirkstoffen, vorbehandelt und für eine (weitere) systemische Therapie zur Behandlung oraler Aphthen geeignet. Eine Dosisänderung von Apremilast im Studienverlauf war nicht vorgesehen.
bid: zweimal täglich (bis in die); p.o.: oral (per os) Quelle: Amgen 2019			

Tabelle 4-15: Demografische Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Charakteristikum	Apremilast (N = 104)	Placebo (N = 103)	Gesamt (N = 207)
Alter (Jahre)			
n	104	103	207
MW (SD)	39,4 (12,1)	40,6 (12,7)	40,0 (12,4)
Median (Min; Max)	37 (20; 67)	40 (19;72)	40 (19; 72)
Alterskategorie 1 (Jahre), n (%)			
< 65	101 (97,1)	99 (96,1)	200 (96,6)
≥ 65	3 (2,9)	4 (3,9)	7 (3,4)
Alterskategorie 2 (Jahre), n (%)			
< 40	58 (55,8)	45 (43,7)	103 (49,8)
40 bis < 65	43 (41,3)	54 (52,4)	97 (46,9)
≥ 65	3 (2,9)	4 (3,9)	7 (3,4)
Geschlecht, n (%)			
Männlich	40 (38,5)	40 (38,8)	80 (38,6)
Weiblich	64 (61,5)	63 (61,2)	127 (61,4)
Ethnie, n (%)			
Ureinwohner Amerikas oder Alaskas	0	1 (1,0)	1 (0,5)
Asiatisch	32 (30,8)	30 (29,1)	62 (30,0)
Schwarz oder Afroamerikaner	1 (1,0)	0	1 (0,5)
Hawaiianer oder Einwohner anderer pazifischer Inseln	0	1 (1,0)	1 (0,5)
Weiß	69 (66,3)	68 (66,0)	137 (66,2)
Nicht erfasst / nicht berichtet	2 (1,9)	3 (2,9)	5 (2,4)

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

Charakteristikum	Apremilast (N = 104)	Placebo (N = 103)	Gesamt (N = 207)
Region, n (%)			
Europa ^a	25 (24,0)	27 (26,2)	52 (25,1)
Nordamerika ^b	14 (13,5)	11 (10,7)	25 (12,1)
Asien ^c	32 (30,8)	29 (28,2)	61 (29,5)
Rest der Welt ^d	33 (31,7)	36 (35,0)	69 (33,3)
Körpergewicht (kg)			
n	104	103	207
MW (SD)	66,4 (16,03)	70,0 (16,45)	68,2 (16,30)
Median (Min; Max)	63,25 (40,0; 117,8)	68,00 (39,6; 120,0)	65,00 (39,6; 120,0)
Gewichtskategorie (kg), n (%)			
< 50	14 (13,5)	10 (9,7)	24 (11,6)
≥ 50 bis < 60	23 (22,1)	19 (18,4)	42 (20,3)
≥ 60 bis < 70	28 (26,9)	26 (25,5)	54 (26,1)
≥ 70 bis < 85	29 (27,9)	29 (28,2)	58 (28,0)
≥ 85	10 (9,6)	19 (18,4)	29 (14,0)
Körpergröße (cm)			
n	104	103	207
MW (SD)	165,4 (10,71)	165,3 (9,66)	165,4 (10,18)
Median (Min; Max)	165 (144,7; 201,0)	166 (145,0; 188,0)	165 (144,7; 201,0)
BMI ^e (kg/m ²), n (%)			
n	104	103	207
MW (SD)	24,1 (4,50)	25,59 (5,794)	24,84 (5,224)
Median (Min; Max)	23,6 (16,7; 40,6)	23,94 (17,0; 45,0)	23,72 (16,7; 45,0)
BMI-Kategorie (kg/m ²), n (%) ^e			
< 18,5	7 (6,7)	3 (2,9)	10 (4,8)
≥ 18,5 bis < 25	60 (57,7)	51 (49,5)	111 (53,6)
≥ 25 bis < 30	29 (27,9)	29 (28,2)	58 (28,0)
≥ 30	8 (7,7)	20 (19,4)	28 (13,5)
Alkoholkonsum, n (%)			
Ja	35 (33,7)	31 (30,1)	66 (31,9)
< 1 Drink/Woche	19 (18,3)	19 (18,4)	38 (18,4)
1 bis 14 Drinks/Woche	16 (15,4)	12 (11,7)	28 (13,5)
> 14 Drinks/Woche	0	0	0
Nein	69 (66,3)	72 (69,9)	141 (68,1)
Tabakkonsum, n (%) ^f			
Aktiver Raucher	20 (19,2)	12 (11,7)	32 (15,5)
Ehemaliger Raucher	16 (15,4)	22 (21,4)	38 (18,4)
Nichtraucher	68 (65,4)	69 (67,0)	137 (66,2)
<p>a: Frankreich, Deutschland, Griechenland und Italien. b: Vereinigte Staaten von Amerika. c: Japan und Südkorea. d: Libanon, Israel und Türkei. e: Berechnung auf Basis des Körpergewichts vor der ersten Dosis der Studienmedikation und der Körpergröße zu Screening. f: Patienten mit Status sowohl aktiver als auch ehemaliger Raucher wurden zu aktiven Rauchern gezählt. BMI: Body Mass Index; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten mit Ausprägung; SD: Standardabweichung (standard deviation) Quelle: Amgen 2019</p>			

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

Charakteristikum	Apremilast (N = 104)	Placebo (N = 103)	Gesamt (N = 207)
Dauer des Behçet-Syndroms (Jahre) n MW (SD) Median (Min; Max)	103 6,7 (7,4) 4,10 (0,0; 39,0)	102 6,9 (8,0) 4,27 (0,12; 44,8)	205 6,8 (7,7) 4,27 (0,02; 44,8)
Dauer des Behçet-Syndroms (Jahre) in Kategorien, n (%)			
≤ 2	32 (30,8)	34 (33,0)	66 (31,9)
> 2 bis ≤ 6	30 (28,8)	25 (24,3)	55 (26,6)
> 6 bis < 10	15 (14,4)	16 (15,5)	31 (15,0)
≥ 10	26 (25,0)	27 (26,2)	53 (25,6)
fehlend	1 (1,0)	1 (1,0)	2 (1,0)
Anzahl oraler Aphthen n MW (SD) Median (Min; Max)	104 4,2 (3,7) 3 (2; 32)	103 3,9 (2,7) 3 (2; 19)	207 4,1 (3,2) 3 (2; 32)
Anzahl oraler Aphthen, n (%)			
> 2 bis ≤ 5	91 (87,5)	89 (86,4)	180 (87,0)
> 5 bis < 10	10 (9,6)	11 (10,7)	21 (10,1)
≥ 10	3 (2,9)	3 (2,9)	6 (2,9)
Schmerzhaftigkeit der oralen Aphthen mittels VAS n MW (SD) Median (Min; Max)	103 61,2 (27,6) 67,0 (0; 100)	102 60,8 (26,9) 64,5 (1; 100)	205 61,0 (27,2) 66,0 (0; 100)
BSAS n MW (SD) Median (Min; Max)	104 42,8 (16,2) 41,00 (8,0; 81,5)	103 44,3 (16,9) 43,00 (9,0; 90,0)	207 43,5 (16,5) 42,00 (8,0; 90,0)
BDCAF Indexwert n MW (SD) Median (Min; Max)	104 3,7 (1,6) 4,0 (1; 7)	102 3,6 (1,7) 4,0 (1; 9)	206 3,7 (1,6) 4,0 (1; 9)
BD-QoL n MW (SD) Median (Min; Max)	104 10,2 (8,23) 8,00 (0,0; 28,0)	103 11,2 (8,2) 12,00 (0,0; 29,0)	207 10,7 (8,2) 10,00 (0,0; 29,0)
<i>Vorangegangene Therapien</i>			
Vorbehandlung mit Immunsuppressiva, n (%)			
Ja	14 (13,5)	14 (13,6)	28 (13,5)
Nein	90 (86,5)	89 (86,4)	179 (86,5)
Vorbehandlung mit Colchicin ^a , n (%)			
Ja	52 (50,0)	57 (55,3)	109 (52,7)
Nein	52 (50,0)	46 (44,7)	98 (47,3)
Vorbehandlung mit oralen Kortikosteroiden, n (%)			
Ja	17 (16,3)	15 (14,6)	32 (15,5)
Nein	87 (83,7)	88 (85,4)	175 (84,5)

Charakteristikum	Apremilast (N = 104)	Placebo (N = 103)	Gesamt (N = 207)
Vorbehandlung mit topischen Kortikosteroiden, n (%)			
Ja	13 (12,5)	16 (15,5)	29 (14,0)
Nein	91 (87,5)	91 (87,5)	178 (86,0)
Vorbehandlung mit Biologika ^b , n (%)			
Ja	2 (1,9)	3 (2,9)	5 (2,4)
Nein	102 (98,1)	100 (97,1)	202 (97,6)
Vorbehandlung mit NSAR, n (%)			
Ja	27 (26,0)	42(40,8)	69 (33,3)
Nein	77 (74,0)	61 (59,2)	138 (66,7)
Vorbehandlung mit anderen Analgetika oder Anästhetika als NSAR, n (%)			
Ja	16 (15,4)	21 (20,4)	37 (17,9)
Nein	88 (84,6)	82 (79,6)	170 (82,1)

a: Ein Patient in der Placebo-Gruppe erhielt Colchimax und wurde in die Analyse für die Vormedikation mit Colchicin einbezogen.

b: Vorherige Biologika-Therapie erlaubt für andere Indikationen, inklusive andere Ausprägung der BD.

BD: Behçet-Syndrom (Behçet's disease); BDCAF: BD Current Activity Form; BD-QoL: BD Quality of Life; bid: zweimal täglich (bis in die); BSAS: Behçet's Syndrome Activity Score; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten mit Ausprägung; NSAR: Nichtsteroidale Antirheumatika; SD: Standardabweichung (standard deviation); VAS: Visuelle Analogskala

Quelle: Amgen 2019

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.

Bei der Studie RELIEF handelt es sich um eine multinationale, multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte Phase-III-Studie. Es wurden erwachsene Patienten im Alter ab 18 Jahren mit Behçet-Syndrom untersucht. Die Diagnose erfolgte anhand der Kriterien der International Study Group (ISG) for Behçet's Disease. Die Patienten mussten mindestens eine nicht-biologische Vortherapie erhalten haben und zu Studienbeginn mindestens zwei orale Aphthen aufweisen, deren Ausmaß eine (weitere) topische Behandlung als ungeeignet oder unzureichend ausschloss. Die Studie RELIEF ist die bisher größte Studie zur Behandlung oraler Aphthen bei Patienten mit Behçet-Syndrom.

Insgesamt wurden 207 Patienten in die Studie eingeschlossen. Die Zuteilung der Patienten in eine der beiden Behandlungsgruppen erfolgte randomisiert im Verhältnis 1:1. Bei 104 Patienten wurde Apremilast 30 mg p.o. bid (60 mg Tagesdosierung) angewendet, bei 103 Patienten Placebo p.o. bid. Die zweimal tägliche Anwendung erfolgte in beiden Gruppen im Abstand von etwa 12 Stunden. Apremilast wurde in der ersten Woche schrittweise von 10 mg über 20 mg auf 30 mg je Anwendung titriert. Hierzu erhielten die Patienten der Apremilast-Gruppe (Gruppe 1) pro Einnahmezeitpunkt eine Tablette Apremilast in einer Dosierung von 10 mg, 20 mg oder 30 mg. Das exakte Titrationsschema ist in Tabelle 4-17 dargestellt.

Tabelle 4-17: Dosistitrationsschema der Apremilast-Gruppe (Gruppe 1)

Tag 1	Tag 2		Tag 3		Tag 4		Tag 5		Ab Tag 6	
Morgens	Morgens	Abends	Morgens	Abends	Morgens	Abends	Morgens	Abends	Morgens	Abends
10 mg	10 mg	10 mg	10 mg	20 mg	20 mg	20 mg	20 mg	30 mg	30 mg	30 mg

Die Placebo-Gruppe (Gruppe 2) erhielt die identische Anzahl an Placebo-Tabletten je Tag. Das Aussehen der Placebo-Tabletten war zu den Apremilast-Tabletten identisch.

Patienten der Apremilast-Gruppe waren den Patienten der Placebo-Gruppe hinsichtlich der demografischen und krankheitsspezifischen Patientencharakteristika vergleichbar (Tabelle 4-15; Tabelle 4-16). Das mediane Alter der Gesamtpopulation betrug 40,0 Jahre; der Anteil der Frauen fiel mit 61,4 % etwa doppelt so hoch aus wie der der Männer. Das Behçet-Syndrom dauerte im Median seit 4,3 Jahren an. Am häufigsten waren Patienten der Gruppe zuzuordnen, die seit zwei oder weniger Jahren erkrankt war (31,4 %). Alle Patienten mussten bei Studienbeginn mindestens zwei orale Aphthen aufweisen. Zu Studienbeginn wiesen 87,0 % zwischen zwei und fünf orale Aphthen auf. Im Median wiesen die Patienten drei orale Aphthen auf. Alle Patienten waren im Laufe ihrer Krankheitsgeschichte des Behçet-Syndroms gemäß Protokoll vorbehandelt. Mit 52,7 % bzw. 33,3 % waren Colchicin bzw. NSAR die am häufigsten eingesetzten Therapeutika. Weniger als 20 % der Patienten hatten Immunsuppressiva, orale oder topische Kortikosteroide sowie andere Analgetika / Anästhetika als NSAR als Vormedikation erhalten. Fünf Patienten (2,4 %) waren mit Biologika für andere Indikationen oder andere Ausprägungen des Behçet-Syndroms vorbehandelt.

Die Studie umfasste vier Phasen: eine bis zu sechswöchige Screeningphase, eine initiale 12-wöchige doppelblinde, Placebo-kontrollierte Behandlungsphase, eine anschließende 52-wöchige aktive Behandlungsphase / Extensionsphase sowie eine vierwöchige Nachbeobachtungsphase zur Überwachung des Sicherheitsprofils.

Im Anschluss an die 12-wöchige doppelblinde, Placebo-kontrollierte Behandlungsphase wurden alle Patienten für 52 Wochen ausschließlich mit Apremilast 30 mg p.o. bid weiterbehandelt. Zur Aufrechterhaltung der Verblindung der initialen 12-wöchigen Behandlungsphase blieb die Studienmedikation auch während der anschließenden 52-wöchigen aktiven Behandlungsphase verblindet. Zu Woche 12 wurde ausschließlich bei den Patienten der

Gruppe, die in der Placebo-kontrollierten Phase Placebo erhielten, Apremilast mit 10 mg beginnend über 20 mg auf 30 mg titriert, wie vorstehend für die Apremilast-Gruppe beschrieben (Tabelle 4-17). Die Patienten der Apremilast-Gruppe erhielten weiterhin 30 mg p.o. bid des Wirkstoffs.

Die Apremilast-Gruppe erreichte gegenüber der Placebo-Gruppe zu Woche 12 statistisch signifikant überlegene Ergebnisse für die Wirksamkeitsendpunkte „Vollständige Remission der oralen Aphthen“, „Schmerzhaftigkeit der oralen Aphthen mittels VAS“ sowie die krankheitsspezifische Morbidität anhand der „BSAS“- und „BDCAF“-Fragebögen und die gesundheitsbezogene Lebensqualität anhand der „BD-QoL“- und „SF-36“-Fragebögen. Das Sicherheitsprofil war mit Ausnahme der in der Apremilast-Gruppe statistisch signifikant häufiger aufgetretenen, als UE von besonderem Interesse vordefinierten Diarrhoe in beiden Behandlungsgruppen vergleichbar.

Übertragbarkeit auf den deutschen Versorgungskontext

In die Studie RELIEF wurden Patienten aus Europa und den USA (37,2 %), Asien (29,5 %) und Rest der Welt (33,3 %) eingeschlossen. Dabei lebten die meisten Patienten in Ländern, die im Hinblick auf ihre Gesellschaftssysteme, Kultur und Ethnologie zu den Ländern der westlichen Welt zählen. Wie in Modul 3.2 bereits erläutert, tritt das Behçet-Syndrom besonders häufig bei Personen auf, die den Bevölkerungsgruppen entlang der ehemaligen Seidenstraße entstammen. In Europa sind dies vornehmlich Personen, die in der Türkei leben und türkischstämmige Migranten. In Deutschland leben nach aktuellen Angaben von 2018 etwa 2.769.000 türkischstämmige Einwohner. Mit sieben Studienteilnehmern, die aus Deutschland stammen und 54 Studienteilnehmern, die aus der Türkei stammen, entstammt ein repräsentativer Teil der Studienpopulation (N = 205) entweder aus einer Bevölkerungsgruppe oder einer Regionen, die dem im Versorgungsalltag erwarteten Patienten in Deutschland entspricht. Es ist somit davon auszugehen, dass die Studienpopulation den deutschen Versorgungskontext abbildet.

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			
RELIEF	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Zuteilung der Patienten in einen der beiden Studienarme, Apremilast oder Placebo, für die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase erfolgte randomisiert mit Hilfe einer interaktiven Response-Technologie (IRT). Eine adäquate Erzeugung der Randomisierungssequenz und eine verdeckte Gruppenzuteilung waren damit ebenso gewährleistet wie die Verblindung der Patienten und der behandelnden Personen. Im Studienverlauf erfolgte die Zuteilung kodierter Kits der Studienmedikation entsprechend der von der IRT vergebenen Randomisierungsnummer durch den Krankenhausapotheker oder autorisiertes Personal am Prüfzentrum. In den Studienunterlagen finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es liegen auch keine weiteren, das Ergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial als niedrig eingestuft wird.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

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

Studie	Mortalität	Morbidität	Gesundheitsbezogene Lebensqualität	Sicherheit
Endpunkt	Todesfälle	Vollständige Remission der oralen Aphthen (komplettes Ansprechen)	Veränderung des BD-QoL zu Woche 12	UE gesamt
		Schmerzhaftigkeit oraler Aphthen mittels VAS	Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12	SUE
		Veränderung des BSAS zu Woche 12		Schwere UE
		Veränderung des BDCAF zu Woche 12		Therapieabbruch aufgrund von UE
				UE von besonderem Interesse
RELIEF	ja	ja	ja	ja
BD: Behçet-Syndrom (Behçet's disease); BDCAF: BD Current Activity Form; BD-QoL: BD Quality of Life; BSAS: Behçet's Syndrome Activity Score; SF-36: 36-Item Short Form Health Survey, Version 2; SUE: Schwerwiegende(s) unerwünschte(s) Ereignis(se); UE: Unerwünschte(s) Ereignis(se); VAS: Visuelle Analogskala				

4.3.1.3.1 Endpunkt – 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 Vollständige Remission der oralen Aphthen (komplettes Ansprechen) - 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 von Vollständige Remission der oralen Aphthen (komplettes Ansprechen)

Studie	Operationalisierung
RELIEF	<p>Der Endpunkt „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ wird über den Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 erfasst.</p> <p>Zusätzlich werden die Zeit bis zur vollständigen Remission, die Anzahl und die Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission dargestellt.</p> <p>Zur Stützung der Ergebnisse zu Woche 12 werden die Daten aus der Extensionsphase (ab Woche 12 bis Woche 68) für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen und die Anzahl oraler Aphthen zusätzlich dargestellt.</p> <p><u>Anteil der Patienten mit vollständiger Remission der oralen Aphthen</u></p> <p>Der Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 wurde vom Prüfarzt durch visuelles Auszählen der oralen Aphthen erfasst. Für die statistische Analyse wurde die ITT-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten hatten. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert (Hauptanalyse). Zusätzlich wurden zu Woche 12 fehlende Werte als Non-Responder imputiert (Sensitivitätsanalyse). In die Ergebnisdarstellung werden, neben Anzahl und prozentualem Anteil der Patienten mit Ereignis bzw. Imputation, RR, OR und RD mit dem jeweiligen 95 %-KI sowie der p-Wert einbezogen. Die Ergebnisse ab Woche 12 bis Woche 68 werden anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p> <p><u>Zeit bis zur vollständigen Remission der oralen Aphthen</u></p> <p>Die Zeit bis zur vollständigen Remission der oralen Aphthen wurde vom Prüfarzt im Studienverlauf auf Basis des visuellen Auszählens der oralen Aphthen erfasst. In der statistischen Analyse wurden diejenigen Patienten der ITT-Population berücksichtigt, die mindestens eine vollständige Remission der oralen Aphthen im Zeitraum von Baseline bis einschließlich Woche 12 zeigten. Patienten ohne Post-Baseline-Erfassung der oralen Aphthen wurden zum Zeitpunkt der ersten Einnahme der Studienmedikation zensiert. Zur Ergebnisdarstellung werden Median und HR, jeweils mit 95 %-KI, sowie der p-Wert herangezogen.</p> <p><u>Anzahl oraler Aphthen</u></p> <p>Die Anzahl der oralen Aphthen wurde vom Prüfarzt zu allen Untersuchungszeitpunkten durch visuelles Auszählen erfasst.</p> <p>Für die statistische Analyse wurde die ITT-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten hatten. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline anhand von MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie SMD mit 95 %-KI dargestellt. Die Ergebnisse ab Woche 12 bis Woche 68 werden anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p>

Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission

Die Zeit bis zum Rezidiv nach vollständiger Remission wurde vom Prüfarzt im Studienverlauf auf Basis des visuellen Auszählens der oralen Aphthen erfasst.

Die statistische Analyse basierte auf denjenigen Patienten der ITT-Population, die bis Woche 12 eine vollständige Remission der oralen Aphthen mit anschließendem Rezidiv aufwiesen. Als Zeit wurde die Dauer zwischen der ersten vollständigen Remission und dem ersten Wiederauftreten der oralen Aphthen definiert. Patienten ohne dokumentiertes Rezidiv zwischen Baseline und Woche 12 wurden zum Zeitpunkt der letzten Erfassung der oralen Aphthen zensiert. Patienten ohne Ergebnis nach Erreichen der vollständigen Remission wurden zum Zeitpunkt der dokumentierten vollständigen Remission zensiert. Zur Ergebnisdarstellung werden Median und HR, jeweils mit 95 %-KI, sowie der p-Wert herangezogen.

HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz; MW: Mittelwert; N: Anzahl der Patienten; OC: Vorliegende Beobachtungen (observed cases); OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz

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 Vollständige Remission der oralen Aphthen (komplettes Ansprechen) 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
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig
a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die gesamte ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten. Es fanden sich keine Aspekte in den

Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand adäquater Methoden erhoben wurde. Die in die Studiendurchführung involvierten Prüfarzte waren ausschließlich qualifizierte, mit dem Behçet-Syndrom vertraute Experten. Sie waren damit in der Lage, Ausprägung und Verlauf der oralen Aphthen über eine visuelle Auszählung zuverlässig zu erfassen. Die vollständige Remission der oralen Aphthen sowie ihr Wiederauftreten nach vollständiger Remission lassen sich darüber hinaus objektiv auf visuellem Wege durch den Prüfarzt erfassen. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ in der Studie RELIEF als niedrig eingestuft wird.

Stellen Sie die Ergebnisse für den Endpunkt „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Ergebnisse der Hauptanalyse des Endpunkts „Vollständige Remission der oralen Aphthen (komplettes Ansprechen)“ über den Anteil der Patienten abgebildet, die zu Woche 12 eine vollständige Remission erreichten. Darüber hinaus erfolgt die Darstellung des Endpunkts anhand der Ergebnisse der Zusatzanalysen für die Zeit bis zur vollständigen Remission der oralen Aphthen, der Veränderung ihrer Anzahl sowie der Zeit bis zum Rezidiv nach vollständiger Remission.

Tabelle 4-22: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo			
N	n1 (%)	n2 (%)	N	n1 (%)	n2 (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a
104	59 (56,7)	8 (7,7)	103	26 (25,2)	22 (21,4)	2,25 [1,55; 3,27]	3,90 [2,16; 7,03]	0,31 [0,19; 0,44]	< 0,0001

a: p-Wert des Cochran-Mantel-Haenszel-Tests.
 ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); N: Anzahl der Patienten; n1: Anzahl der Patienten mit Ereignis; n2: Anzahl der Patienten mit Imputation (LOCF); OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko
 Quelle: Amgen 2020a

In der Apremilast-Gruppe zählten 59 Patienten (56,7 %) zu den Respondern hinsichtlich einer vollständigen Remission der oralen Aphthen zu Woche 12. Im Vergleich dazu zählten in der Placebo-Gruppe 26 Patienten (25,2 %) zu den Respondern. Mit einem p-Wert von < 0,0001 liegt ein statistisch hochsignifikantes Ergebnis vor. Dies verdeutlicht die Überlegenheit von

Apremilast in der Verbesserung des Leitsymptoms und der Schwere des Behçet-Syndroms (Tabelle 4-22).

Tabelle 4-23: Sensitivitätsanalyse: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 (NRI) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo			
N	n1 (%)	n2 (%)	N	n1 (%)	n2 (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a
104	55 (52,9)	8 (7,7)	103	23 (22,3)	23 (22,3)	2,37 [1,58; 3,54]	3,92 [2,15; 7,18]	0,31 [0,18; 0,43]	< 0,0001

a: p-Wert des Cochran-Mantel-Haenszel-Tests.
 ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); N: Anzahl der Patienten; n1: Anzahl der Patienten mit Ereignis; n2: Anzahl der Patienten mit Imputation (Non-Responder); NRI: Non-Responder-Imputation; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko
 Quelle: Amgen 2020a

Die Ergebnisse der Sensitivitätsanalyse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen zu Woche 12 mit der Non-Responder-Imputation sind konsistent zu den Ergebnissen mit LOCF-Imputation (Tabelle 4-22, Tabelle 4-23). Die Robustheit der Analyse mittels LOCF konnte somit gezeigt werden.

Tabelle 4-24: Ergebnisse für die Zeit bis zur vollständigen Remission der oralen Aphthen aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	n (%)	Median ^a (Wochen) [95 %-KI]	N	n (%)	Median ^a (Wochen) [95 %-KI]	HR ^b [95 %-KI]	p-Wert ^c
104	84 (80,77)	2,14 [2,00; 4,00]	103	54 (52,43)	8,14 [4,71; n.b.]	2,40 [1,69; 3,40]	< 0,0001

a: Der Median wurde auf Basis der Kaplan-Meier-Methode ermittelt.
 b: Das HR mit dem dazugehörigen 95 %-KI basiert auf dem stratifizierten Cox-Modell (Kovariablen: Anzahl der oralen Aphthen zu Baseline, Stratifizierungsfaktoren: Geschlecht und Region).
 c: Der p-Wert basiert auf dem stratifizierten Log-rank-Test mit Geschlecht und Region als Stratifizierungsfaktoren.
 HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; n.b.: nicht berechenbar
 Quellen: Amgen 2018a, 2020a

Eine vollständige Remission der oralen Aphthen zu jedem möglichen Zeitpunkt zwischen Baseline und Woche 12 erreichten 84 Patienten (80,77 %) der Apremilast-Gruppe und 54 Patienten (52,43 %) der Placebo-Gruppe. Die vollständige Remission trat bei den Patienten

der Apremilast-Gruppe im Median nach 2,14 Wochen ein und damit deutlich früher als bei den Patienten der Placebo-Gruppe nach 8,14 Wochen (Tabelle 4-24). Das Ergebnis ist statistisch signifikant ($p < 0,0001$). Das frühere Eintreten der vollständigen Remission wird bereits anhand der Ergebnisse in Woche 2 besonders deutlich (Abbildung 2). In der Apremilast-Gruppe erreichten 44 Patienten (42,3 %) zu diesem Zeitpunkt eine vollständige Remission, im Vergleich zu 12 Patienten (11,7 %) der Placebo-Gruppe. Bei 31 Patienten (29,8 %) der Apremilast-Gruppe hielt die vollständige Remission von Woche 6 bis mindestens Woche 12 an. In der Placebo-Gruppe hielt im Vergleich dazu bei nur fünf Patienten (4,9 %) die vollständige Remission von Woche 6 bis Woche 12 an (Amgen 2019). Die Abbildung 2 verdeutlicht die klinisch relevante Wirksamkeit von Apremilast im Hinblick auf die Anzahl der Patienten mit vollständiger Remission der oralen Aphthen sowie das rasch einsetzende und anhaltende Ansprechen.

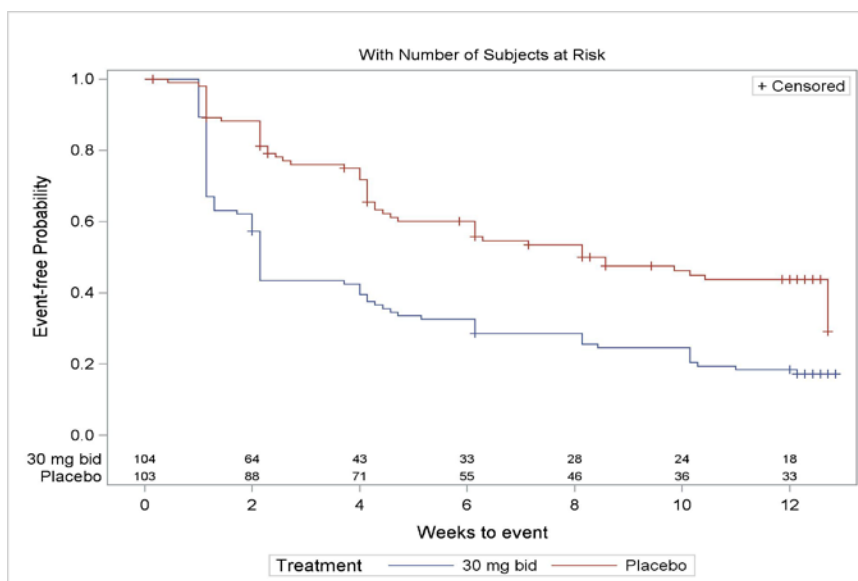


Abbildung 2: Kaplan-Meier-Kurven für Zeit bis zur vollständigen Remission der oralen Aphthen (komplettes Ansprechen) bis Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Tabelle 4-25: Ergebnisse für Veränderung der Anzahl oraler Aphthen zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
104	4,17 (3,65)	-2,74 (0,23)	103	3,93 (2,70)	-1,50 (0,23)	-1,24 [-1,79; -0,69] <0,0001	-0,61 [-0,89; -0,33]

a: p-Wert der ANCOVA.
ANCOVA: Kovarianzanalyse (analysis of covariance); KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz
Quelle: Amgen 2020a

Die Ergebnisse des mittleren Rückgangs der Anzahl oraler Aphthen von Baseline im Vergleich zu Woche 12 zeigten einen statistisch signifikanten Vorteil ($p < 0,0001$) von Apremilast gegenüber Placebo. Da das KI der SMD vollständig außerhalb des Irrelevanzbereichs von -0,2 bis +0,2 lag, sind die Unterschiede zwischen den beiden Gruppen zudem als klinisch relevant anzusehen (Tabelle 4-25).

Tabelle 4-26: Ergebnisse für Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	n (%)	Median ^a (Wochen) [95 %-KI]	N	n (%)	Median ^a (Wochen) [95 %-KI]	HR ^b [95 %-KI]	p-Wert ^c
83	57 (68,67)	4,57 [3,14; 6,14]	53	43 (81,13)	2,29 [2,14; 4,14]	0,61 [0,41; 0,92]	0,0112

a: Der Median wurde auf Basis der Kaplan-Meier-Methode ermittelt.
b: Das HR mit dem dazugehörigen 95 %-KI basiert auf dem stratifizierten Cox-Modell (Stratifizierungsfaktoren: Geschlecht und Region).
c: Der p-Wert basiert auf dem stratifizierten Log-rank-Test mit Geschlecht und Region als Stratifizierungsfaktoren.
HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis
Quellen: Amgen 2018a, 2020a

In der Apremilast-Gruppe bildeten sich bei 31,33 % der Patienten nach vollständiger Remission bis Woche 12 keine neuen oralen Aphthen; in der Placebo-Gruppe lag der Anteil bei 18,87 %. Die mediane Zeit bis zum Rezidiv lag in der Apremilast-Gruppe bei 4,57 Wochen und war

damit statistisch signifikant länger als in der Placebo-Gruppe mit 2,29 Wochen (Tabelle 4-26). In jedem zweiwöchigen Untersuchungsintervall zwischen Woche 2 bis Woche 12 entwickelten weniger Patienten in der Apremilast-Gruppe ein Rezidiv als in der Placebo-Gruppe. Die Überlegenheit von Apremilast gegenüber Placebo konnte somit auch für die Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission bestätigt werden (Abbildung 3).

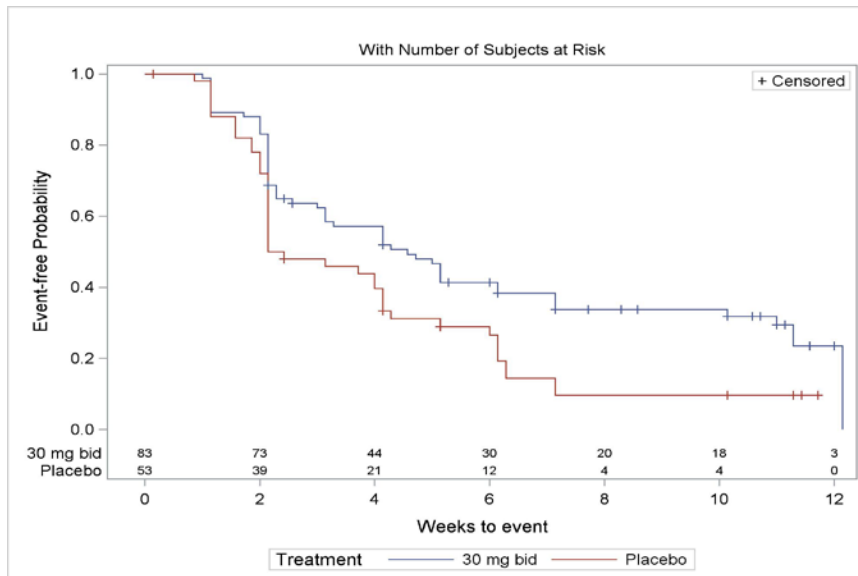


Abbildung 3: Kaplan-Meier-Kurven für Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission bis Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Extensionsphase (ab Woche 12 bis Woche 68)

Nachfolgend werden die Ergebnisse des Endpunkts „Vollständige Remission der oralen Aphthen“ supportiv für die Extensionsphase ab Woche 12 bis einschließlich Woche 68 dargestellt. Zur Darstellung werden der Anteil der Patienten mit vollständiger Remission der oralen Aphthen und die Anzahl oraler Aphthen herangezogen. Für die Analysen der Wirksamkeitseindpunkte wurden ausschließlich diejenigen Patienten berücksichtigt, die nach Beendigung der 12-wöchigen Placebo-kontrollierten Behandlungsphase in die 52-wöchige Extensionsphase wechselten. Da die Patienten der Placebo-Gruppe ab Woche 12 ebenfalls Apremilast erhielten, wird die Gruppe in der Extensionsphase mit Placebo/Apremilast-Gruppe bezeichnet.

Tabelle 4-27: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der oralen Aphthen in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast		Placebo/Apremilast		Apremilast vs. Placebo/Apremilast			
N	n (%)	N	n (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a
Woche 12							
97	55 (56,7)	82	23 (28,1)	2,03 [1,38; 2,99]	3,40 [1,81; 6,37]	0,29 [0,15; 0,43]	0,0001
Woche 16							
95	50 (52,6)	83	53 (63,9)	0,83 [0,65; 1,06]	0,62 [0,34; 1,14]	-0,11 [-0,25; 0,03]	0,0921
Woche 28							
92	57 (62,0)	78	36 (59,0)	1,05 [0,82; 1,34]	1,13 [0,61; 2,10]	0,03 [-0,12; 0,18]	0,6823
Woche 40							
85	50 (58,8)	73	51 (69,9)	0,85 [0,67; 1,06]	0,62 [0,32; 1,21]	-0,11 [-0,26; 0,04]	0,1646
Woche 52							
79	47 (59,5)	70	41 (58,6)	1,05 [0,79; 1,38]	1,03 [0,53; 1,99]	0,01 [-0,15; 0,17]	0,8479
Woche 64							
75	40 (53,3)	67	43 (64,2)	0,85 [0,65; 1,11]	0,66 [0,33; 1,31]	-0,10 [-0,26; 0,06]	0,2219
Woche 68 (Nachbeobachtung)							
85	23 (27,1)	82	19 (23,2)	1,16 [0,69; 1,97]	1,22 [0,60; 2,46]	0,03 [-0,10; 0,17]	0,6097

a: p-Wert des Cochran-Mantel-Haenszel-Tests.
ITT: Intention-to-treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis;
OC: Vorliegende Beobachtungen (observed cases); OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko
Quelle: Amgen 2020a

Der zum Ende der Placebo-kontrollierten Behandlungsphase zu Woche 12 beobachtete statistisch signifikante Unterschied der vollständigen Remission der oralen Aphthen zwischen den beiden Gruppen zugunsten von Apremilast nivellierte sich ab Woche 16, nachdem die Patienten der ursprünglichen Placebo-Gruppe über vier Wochen ebenfalls Apremilast erhalten hatten. Der Anteil der Patienten mit vollständiger Remission der oralen Aphthen fiel ab Woche 16 bis zum Behandlungsende in Woche 64 in beiden Gruppen mit über 50 % vergleichbar hoch aus mit leichten Vorteilen für die Placebo/Apremilast-Gruppe. Zu Woche 68, d. h. vier Wochen ohne Einnahme von Apremilast, reduzierte sich die Responderrate in beiden Gruppen um die Hälfte auf 27,1 % bzw. 23,2 % (Tabelle 4-27).

Tabelle 4-28: Ergebnisse für Veränderung der Anzahl oraler Aphthen in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo/Apremilast			Apremilast vs. Placebo/Apremilast	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Woche 12							
97	3,87 (2,42)	-2,60 (0,23)	82	3,57 (1,87)	-1,65 (0,24)	-0,95 [-1,51; -0,39] 0,0010	-0,50 [-0,80; -0,20]
Woche 16							
95	3,86 (2,44)	-2,77 (0,15)	83	3,57 (1,86)	-2,97 (0,16)	0,20 [-0,17; 0,57] 0,2859	0,16 [-0,13; 0,46]
Woche 28							
92	3,90 (2,46)	-2,78 (0,20)	78	3,56 (1,90)	-2,87 (0,21)	0,09 [-0,40; 0,58] 0,7155	0,06 [-0,25; 0,36]
Woche 40							
85	3,82 (2,23)	-2,92 (0,19)	73	3,62 (1,95)	-3,15 (0,20)	0,23 [-0,24; 0,70] 0,3353	0,15 [-0,16; 0,47]
Woche 52							
79	3,81 (2,30)	-2,99 (0,32)	70	3,64 (1,98)	-2,70 (0,33)	-0,29 [-1,06; 0,48] 0,4574	-0,12 [-0,44; 0,20]
Woche 64							
75	3,60 (1,40)	-2,37 (0,28)	67	3,63 (2,01)	-2,97 (0,28)	0,60 [-0,07; 1,27] 0,0771	0,30 [-0,03; 0,63]
Woche 68 (Nachbeobachtung)							
85	4,26 (3,97)	-1,70 (0,24)	82	3,72 (2,02)	-2,02 (0,24)	0,32 [-0,27; 0,91] 0,2899	0,16 [-0,14; 0,47]
a: p-Wert der ANCOVA. ANCOVA: Kovarianzanalyse (analysis of covariance); KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz Quelle: Amgen 2020a							

Die in der Apremilast-Gruppe zu Woche 12 beobachtete Reduzierung der Anzahl oraler Aphthen hielt im Verlauf der Extensionsphase bis Woche 64 in gleicher Ausprägung an. In der

Placebo/Apremilast-Gruppe verringerte sich die Anzahl der oralen Aphthen vier Wochen nach Beginn der Behandlung mit Apremilast zu Woche 16 deutlich gegenüber Woche 12. Das Ausmaß des Rückgangs der oralen Aphthen blieb auch in der Placebo/Apremilast-Gruppe bis zum Behandlungsende in Woche 64 konstant und befand sich auf einem der Apremilast-Gruppe vergleichbaren Niveau. Zu Woche 68, vier Wochen nach Absetzen von Apremilast, zeigte die Anzahl der oralen Aphthen in beiden Gruppen eine wieder ansteigende Tendenz (Tabelle 4-28; Abbildung 4).

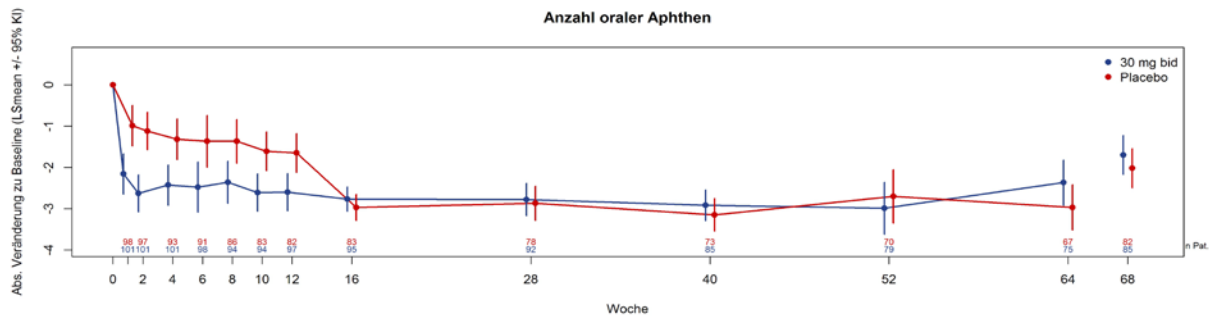


Abbildung 4: Zeitlicher Verlauf der Veränderung der Anzahl oraler Aphthen (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Weitere Manifestationen des Behçet-Syndroms

Nachfolgend werden die Endpunktergebnisse weiterer im Rahmen des Behçet-Syndroms aufgetretenen Manifestationen deskriptiv und supportiv zum primären Wirksamkeitsendpunkt „Vollständige Remission der oralen Aphthen“ dargestellt.

Die Selektionskriterien der Studie RELIEF schlossen die Teilnahme von Patienten mit anderen, schweren Manifestationen des Behçet-Syndroms als den oralen Aphthen aus, sofern diese aufgrund der Schwere der Krankheitsaktivität eine immunsuppressive Therapie erforderten. Zusätzliche Vorgaben oder Einschränkungen hinsichtlich der Krankheitssymptomatik zu Baseline existieren nicht. Weitere schwere Manifestationen des Behçet-Syndroms konnten somit zu Baseline vorliegen, waren aber nicht obligatorisch. Aus diesem Grund erfolgt eine rein deskriptive Darstellung der Symptome.

Vollständige Remission genitaler Aphthen

Tabelle 4-29: Ergebnisse für den Anteil der Patienten mit vollständiger Remission der genitalen Aphthen (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Apremilast N = 17		Placebo ^a N = 17	
	n1	n2 (%)	n1	n2 (%)
Woche 1	17	8 (47,1)	16	1 (6,3)
Woche 2	17	8 (47,1)	16	8 (50,0)
Woche 4	17	10 (58,8)	13	7 (53,8)
Woche 6	16	12 (75,0)	13	6 (46,2)
Woche 8	15	11 (73,3)	11	8 (72,7)
Woche 10	15	9 (60,0)	9	7 (77,8)
Woche 12	16	12 (75,0)	9	7 (77,8)
Woche 16	15	11 (73,3)	8	8 (100,0)
Woche 28	15	13 (86,7)	8	7 (87,5)
Woche 40	13	12 (92,3)	7	5 (71,4)
Woche 52	11	11 (100,0)	7	5 (71,4)
Woche 64	11	11 (100,0)	7	5 (71,4)
Woche 68 (Nachbeobachtung)	12	11 (91,7)	12	6 (50,0)

a: Ab Woche 12 Apremilast (Placebo/Apremilast-Gruppe).
ITT: Intention-to-treat; N: Anzahl der Patienten; n1: Anzahl der Patienten zum Zeitpunkt; n2: Anzahl der Patienten mit Ereignis; OC: Vorliegende Beobachtungen (observed cases)
Quelle: Amgen 2018a

Der Anteil der Patienten mit vollständiger Remission der genitalen Aphthen zu Woche 12 fiel in beiden Behandlungsgruppen mit 75 % (Apremilast) bzw. 77,8 % (Placebo) vergleichbar aus.

In der Extensionsphase hielt die Remissionsrate in der Apremilast-Gruppe an, stieg von Woche 16 bis Woche 52 kontinuierlich auf 100 % an, während die Remissionsrate in der Placebo/Apremilast-Gruppe von 100 % zu Woche 16 bis Woche 40 auf 71,4 % zurückging. In der Nachbeobachtungsphase zu Woche 68 wurde in beiden Behandlungsgruppen ein leichter Rückgang der vollständigen Remission der genitalen Aphthen registriert. Aufgrund der in beiden Gruppen sehr niedrigen Patientenzahl ist eine Interpretation der Daten erschwert (Tabelle 4-29).

Anzahl genitaler Aphthen

Tabelle 4-30: Ergebnisse für die Veränderung der Anzahl genitaler Aphthen (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Apremilast N = 17		Placebo ^a N = 17	
	n	MW (SD)	n	MW (SD)
Baseline	17	2,9 (2,91)	17	2,6 (2,00)
Woche 1	17	2,5 (4,84)	16	2,1 (1,82)
Woche 2	17	1,6 (2,37)	16	1,1 (1,61)
Woche 4	17	1,7 (2,91)	13	1,3 (2,14)
Woche 6	16	0,4 (1,03)	13	1,2 (1,95)
Woche 8	15	0,9 (2,15)	11	0,4 (0,67)
Woche 10	15	0,6 (0,83)	9	0,2 (0,44)
Woche 12	16	0,5 (0,97)	9	0,3 (0,71)
Woche 16	15	0,4 (0,74)	8	0,0 (0,00)
Woche 28	15	0,1 (0,35)	8	0,4 (1,06)
Woche 40	13	0,2 (0,55)	7	0,6 (1,13)
Woche 52	11	0,0 (0,00)	7	0,9 (1,57)
Woche 64	11	0,0 (0,00)	7	0,4 (0,79)
Woche 68 (Nachbeobachtung)	12	0,1 (0,29)	12	1,1 (1,78)

a: Ab Woche 12 Apremilast (Placebo/Apremilast-Gruppe).
ITT: Intention-to-treat; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis;
OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation)
Quelle: Amgen 2018a

Insgesamt 34 Patienten (jeweils 17 in beiden Gruppen) wiesen zu Baseline genitale Aphthen auf. Der Rückgang der genitalen Aphthen von Baseline zu Woche 12 ließ keine relevanten Unterschiede zwischen der Apremilast- und der Placebo-Gruppe erkennen. In der Extensionsphase von Woche 12 bis Woche 64 zeigte sich ein vergleichbares Verhalten wie bei der vollständigen Remission, in der Apremilast-Gruppe war die Anzahl der genitalen Aphthen bis Woche 52 auf N = 0 rückläufig, während sie in der Placebo/Apremilast-Gruppe von N = 0 in Woche 16 wieder anstieg. In der Nachbeobachtungsphase zu Woche 68 ohne Medikation war ein leichter Anstieg der Anzahl genitaler Aphthen in beiden Studiengruppen erkennbar (Tabelle 4-30).

Ein günstiger Behandlungseffekt zeigte sich zu Woche 12 für die Patienten ohne genitale Aphthen zu Baseline. In der Apremilast-Gruppe entwickelte nur ein Patient zwei genitale Aphthen, verglichen mit sieben Patienten in der Placebo-Gruppe (vier Patienten mit einer, zwei Patienten mit zwei und ein Patient mit drei genitalen Aphthen) (Amgen 2018a, 2019). Zwischen Woche 12 und Woche 64 lag der Anteil der während Behandlung neu aufgetretenen genitalen Aphthen zwischen eins und vier in der Apremilast-Gruppe und zwischen zwei und drei in der Placebo/Apremilast-Gruppe (Amgen 2018a).

Neue oder sich verschlimmernde Manifestationen des Behçet-Syndroms

Tabelle 4-31: Ergebnisse für neue oder sich verschlimmernde Hautläsionen im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Apremilast		Placebo ^a	
	N	n (%)	N	n (%)
Woche 1	102	13 (12,7)	98	23 (23,5)
Woche 2	101	18 (17,8)	98	22 (22,4)
Woche 4	101	19 (18,8)	94	27 (28,7)
Woche 6	97	13 (13,4)	92	24 (26,1)
Woche 8	95	15 (15,8)	88	22 (25,0)
Woche 10	94	13 (13,8)	83	19 (22,9)
Woche 12	97	17 (17,5)	82	23 (28,0)
Woche 16	95	14 (14,7)	83	11 (13,3)
Woche 28	92	8 (8,7)	78	17 (21,8)
Woche 40	85	14 (16,5)	73	14 (19,2)
Woche 52	80	17 (21,3)	70	9 (12,9)
Woche 64	76	19 (25,0)	68	13 (19,1)
Woche 68 (Nachbeobachtung)	85	21 (24,7)	82	29 (35,4)

a: Ab Woche 12 Apremilast (Placebo/Apremilast-Gruppe).
 ITT: Intention-to-treat; N: Anzahl Patienten; n: Anzahl der Patienten mit Ereignis; OC: Vorliegende Beobachtungen (observed cases)
 Quelle: Amgen 2018a

Tabelle 4-32: Ergebnisse für neue oder sich verschlimmernde Arthritis im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Apremilast		Placebo ^a	
	N	n (%) ^b	N	n (%) ^c
Woche 1	102	1 (1,0)	98	9 (9,2)
Woche 2	101	0 (0)	98	9 (9,2)
Woche 4	101	2 (2,0)	94	5 (5,3)
Woche 6	97	2 (2,1)	92	1 (1,1)
Woche 8	95	2 (2,1)	88	8 (9,1)
Woche 10	94	2 (2,1)	83	4 (4,8)
Woche 12	97	4 (4,1)	82	5 (6,1)
Woche 16	95	2 (2,1)	83	3 (3,6)
Woche 28	92	0 (0)	78	2 (2,6)
Woche 40	85	1 (1,2)	73	3 (4,1)
Woche 52	80	1 (1,3)	70	5 (7,1)
Woche 64	76	4 (5,3)	68	3 (4,4)
Woche 68 (Nachbeobachtung)	85	3 (3,5)	82	7 (8,5)

a: Ab Woche 12 Apremilast (Placebo/Apremilast-Gruppe).
b: Aktive Arthritis zu Baseline bei 9 Patienten (8,7 %).
c: Aktive Arthritis zu Baseline bei 10 Patienten (9,7 %).
ITT: Intention-to-treat; N: Anzahl Patienten; n: Anzahl der Patienten mit Ereignis; OC: Vorliegende Beobachtungen (observed cases)
Quelle: Amgen 2018a

Tabelle 4-33: Ergebnisse für neue oder sich verschlimmernde gastrointestinale Manifestationen im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Apremilast		Placebo ^a	
	N	n (%)	N	n (%)
Woche 1	102	7 (6,9)	98	6 (6,1)
Woche 2	101	12 (11,9)	98	3 (3,1)
Woche 4	101	6 (5,9)	94	4 (4,3)

Zeitpunkt	Apremilast		Placebo ^a	
	N	n (%)	N	n (%)
Woche 6	97	5 (5,2)	92	1 (1,1)
Woche 8	95	6 (6,3)	88	4 (4,5)
Woche 10	94	5 (5,3)	83	4 (4,8)
Woche 12	97	6 (6,2)	82	5 (6,1)
Woche 16	95	4 (4,2)	83	10 (12,0)
Woche 28	92	6 (6,5)	78	5 (6,4)
Woche 40	85	5 (5,9)	73	2 (2,7)
Woche 52	80	2 (2,5)	70	2 (2,9)
Woche 64	76	1 (1,3)	68	1 (1,5)
Woche 68 (Nachbeobachtung)	85	1 (1,2)	82	5 (6,1)

a: Ab Woche 12 Apremilast (Placebo/Apremilast-Gruppe).

ITT: Intention-to-treat; N: Anzahl Patienten; n: Anzahl der Patienten mit Ereignis; OC: Vorliegende Beobachtungen (observed cases)

Quelle: Amgen 2018a

Tabelle 4-34: Ergebnisse für neue oder sich verschlimmernde Manifestationen des ZNS im Rahmen des Behçet-Syndroms (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Apremilast		Placebo ^a	
	N	n (%)	N	n (%)
Woche 1	102	1 (1,0)	98	0 (0)
Woche 2	101	1 (1,0)	98	2 (2,0)
Woche 4	101	5 (5,0)	94	1 (1,1)
Woche 6	97	2 (2,1)	92	2 (2,2)
Woche 8	95	2 (2,1)	88	2 (2,3)
Woche 10	94	0 (0)	83	2 (2,4)
Woche 12	97	2 (2,1)	82	2 (2,4)
Woche 16	95	0 (0)	83	0 (0)
Woche 28	92	0 (0)	78	1 (1,1)
Woche 40	85	1 (1,2)	73	3 (4,1)

Woche 52	80	1 (1,3)	70	0 (0)
Woche 64	76	1 (1,3)	68	1 (1,5)
Woche 68 (Nachbeobachtung)	85	1 (1,2)	82	3 (3,7)
a: Ab Woche 12 Apremilast (Placebo/Apremilast-Gruppe). ITT: Intention-to-treat; N: Anzahl Patienten; n: Anzahl der Patienten mit Ereignis; OC: Vorliegende Beobachtungen (observed cases); ZNS: Zentralnervensystem Quelle: Amgen 2018a				

Zu den weiteren schweren Manifestationen des Behçet-Syndroms, die im Verlauf der Studie RELIEF als neu aufgetreten oder sich verschlimmernd erfasst wurden, zählten Hautläsionen, Arthritis, Uveitis sowie Manifestationen des gastrointestinalen, vaskulären und Zentralnervensystems.

Die in beiden Gruppen häufigsten neuen oder sich verschlimmernden schweren Manifestationen des Behçet-Syndroms betrafen Hautläsionen (etwa 20 % im gesamten Behandlungsverlauf). Der Anteil der Patienten mit neuen oder sich verschlimmernden Hautläsionen lag in der Apremilast-Gruppe von Baseline bis Woche 12 niedriger als in der Placebo-Gruppe. In der Extensionsphase variierten die Werte in beiden mit Apremilast behandelten Gruppen, waren jedoch weitestgehend vergleichbar (Tabelle 4-31).

Der Anteil an Patienten mit einer neu aufgetretenen oder sich verschlimmernden Arthritis fiel in beiden Gruppen im Studienverlauf mit bis zu 9,2 % in der Placebo-Gruppe und 4,2 % in der Apremilast-Gruppe innerhalb der Placebo-kontrollierten Phase in beiden Gruppen niedrig aus. In der Extensionsphase blieb der Anteil der Patienten mit einer neu aufgetretenen oder sich verschlimmernden Arthritis weiterhin gering und lag in der Placebo/Apremilast-Gruppe minimal höher als in der Apremilast-Gruppe (Tabelle 4-32).

Manifestationen des gastrointestinalen oder Zentralnervensystems traten ebenfalls bei weniger als 10 % der Patienten neu auf oder verschlimmerten sich im Behandlungsverlauf (Ausnahme: gastrointestinale Manifestationen zu Woche 2 bei 11,9 % der Patienten in der Apremilast-Gruppe und zu Woche 16 bei 12,0 % der Patienten in der Placebo/Apremilast-Gruppe). In beiden Gruppen zeigten sich in etwa vergleichbare Werte (Tabelle 4-33 und Tabelle 4-34). Da im Rahmen der Studie zur Differentialdiagnose keine Endoskopie durchgeführt werden musste, ist der Bezug zwischen gastrointestinalen Beschwerden und Behçet-Syndrom nicht endgültig gesichert. Somit sind die Aussagen zu den Manifestationen des Gastrointestinalsystems mit einer zusätzlich erhöhten Unsicherheit behaftet.

Eine neu aufgetretene oder sich verschlimmernde Uveitis wurde lediglich von vier Patienten bis Woche 12 in der Placebo-Gruppe und von einem Patienten zu Woche 64 in der Placebo/Apremilast-Gruppe berichtet.

Kein Patient entwickelte im Behandlungsverlauf neue oder sich verschlimmernde vaskuläre Manifestationen (Amgen 2018a, 2019).

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 wurden keine Meta-Analysen durchgeführt.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Vollständige Remission der oralen Aphthen“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

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 Schmerzhaftigkeit oraler Aphthen mittels 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 4-35: Operationalisierung von Schmerzhaftigkeit oraler Aphthen mittels VAS

Studie	Operationalisierung
RELIEF	<p>Die VAS stellt eine Skala zur Beurteilung des aktuellen Schmerzempfindens durch den Patienten dar. Auf einer geraden Linie von 100 mm Länge markiert der Patient sein aktuelles Schmerzempfinden zwischen linkem / unterem Ausgangswert (keine Schmerzen = 0 mm) und rechtem / oberem Maximalwert (unerträgliche Schmerzen = 100 mm).</p> <p>Die Skala wurde zu Baseline und allen weiteren prädefinierten Untersuchungszeitpunkten erhoben. Für die statistische Analyse wurde die ITT-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten und die VAS zu Baseline beantwortet hatten. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline anhand von MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie SMD mit 95 %-KI dargestellt. In gleicher Weise werden fehlende Werte mittels MMRM berücksichtigt.</p> <p>Zusätzlich werden die Ergebnisse von Woche 1 bis Woche 12 anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p> <p>Zur Stützung der Ergebnisse zu Woche 12 werden die Daten aus der Extensionsphase (ab Woche 12 bis Woche 68) ebenfalls anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p>

ITT: Intention-to-treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least squares mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); LOCF: Letzte verfügbare Beobachtung (last observation carried forward); MMRM: Gemischtes Modell mit Messwiederholungen (mixed-effects model for repeated measures); MW: Mittelwert; OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz; VAS: Visuelle Analogskala

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

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

Tabelle 4-36: Bewertung des Verzerrungspotenzials für Schmerzhaftigkeit oraler Aphthen mittels VAS 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
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig

a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Schmerzhaftigkeit oraler Aphthen mittels VAS“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war. In beiden Behandlungsgruppen wurden jeweils 99 % der Patienten in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei deutlich weniger als 5 %. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand objektiver und adäquater Methoden erhoben wurde. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Schmerzhaftigkeit oraler Aphthen mittels VAS“ in der Studie RELIEF als niedrig eingestuft wird.

Stellen Sie die Ergebnisse für den Endpunkt „Schmerzhaftigkeit oraler Aphthen mittels VAS“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Rücklaufquoten der VAS sowie die Ergebnisse für den Endpunkt „Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS“ dargestellt.

Tabelle 4-37: Rücklaufquoten der VAS von Baseline bis Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Apremilast		Placebo	
	N	n (%)	N	n (%)
Baseline	104	103 (99,0)	103	102 (99,0)
Woche 1	104	95 (91,3)	103	95 (92,2)
Woche 2	104	97 (93,3)	103	96 (93,2)
Woche 4	104	99 (95,2)	103	91 (88,3)
Woche 6	104	97 (93,3)	103	90 (87,4)
Woche 8	104	92 (88,5)	103	85 (82,5) ^a
Woche 10	104	93 (89,4)	103	82 (79,6)
Woche 12	104	95 (91,3)	103	81 (78,6)

a: In Post-hoc-Analysen 86 (83,5 %) Patienten berücksichtigt, davon ein Patient ohne Baseline-Erhebung.
N: Anzahl Patienten der ITT-Population; n: Anzahl der Patienten mit Baseline- und Post-Baseline-Erhebung;
VAS: Visuelle Analogskala
Quelle: Amgen 2020a

Zu Baseline wurde die VAS in beiden Behandlungsgruppen von 99,0 % der Patienten komplettiert. Im Verlauf der 12-wöchigen Placebo-kontrollierten Behandlungsphase fehlten Post-Baseline-Erhebungen von bis zu 12 Patienten (11,5 %) in der Apremilast-Gruppe und 22 Patienten (21,4 %) in der Placebo-Gruppe. Die Rücklaufquoten der VAS lagen zu allen Erhebungszeitpunkten bei über 70 %. Zu Woche 12 betragen sie in der Apremilast-Gruppe 91,3 % und in der Placebo-Gruppe 78,6 % (Tabelle 4-37).

Tabelle 4-38: Ergebnisse für Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
103	61,18 (27,55)	-40,71 (3,34)	102	60,79 (26,92)	-15,89 (3,31)	-24,82 [-32,80; -16,84] <0,0001	-0,85 [-1,14; -0,57]

Einheit der MW = Millimeter

a: p-Wert der ANCOVA, Veränderung gegenüber Baseline (Kovariable: Baseline-Wert, Faktoren: Behandlungsgruppe, Geschlecht und Region).

ANCOVA: Kovarianzanalyse (analysis of covariance); ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz; VAS: Visuelle Analogskala

Quellen: Amgen 2019, 2020a

Die Schmerzhaftigkeit der oralen Aphthen mittels VAS fiel zu Baseline mit 61,18 mm in der Apremilast-Gruppe und 60,79 mm in der Placebo-Gruppe vergleichbar aus. Zu Woche 12 verzeichnete die Placebo-Gruppe eine Reduktion des aktuellen Schmerzempfindens mittels VAS um 15,89 mm, der über der in der Literatur beschriebenen minimalen klinisch bedeutsamen Differenz (minimal clinically important difference, MCID) von 10 mm lag (Dworkin 2008). Mit -40,71 mm erreichte die Apremilast-Gruppe eine signifikante sowie zweieinhalbfach so hohe Verbesserung wie die Placebo-Gruppe, was einem Vierfachen der MCID entsprach. Insgesamt zeigten die Veränderungen der Schmerzhaftigkeit der oralen Aphthen zu Woche 12 gegenüber Baseline zwischen den beiden Behandlungsgruppen eine statistisch signifikante und klinisch relevante Verbesserung zugunsten von Apremilast (Tabelle 4-38). Die Überlegenheit von Apremilast im Vergleich zu Placebo wird somit auch für diesen patientenrelevanten Endpunkt deutlich.

Tabelle 4-39: Ergebnisse für die Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Veränderung zu Woche 1							
95	60,39 (27,70)	-27,02 (3,19)	95	61,93 (26,26)	-15,59 (3,12)	-11,43 [-19,06; -3,80] < 0,0035	-0,43 [-0,71; -0,14]
Veränderung zu Woche 2							
97	61,73 (27,70)	-42,02 (3,03)	96	62,07 (26,50)	-19,19 (2,99)	-22,83 [-30,06; -15,61] < 0,0001	-0,89 [-1,19; -0,60]
Veränderung zu Woche 4							
99	61,22 (27,80)	-39,10 (3,32)	91	60,96 (25,87)	-15,08 (3,31)	-24,02 [-31,94; -16,10] < 0,0001	-0,87 [-1,16; -0,57]

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Veränderung zu Woche 6							
97	60,87 (27,81)	-39,15 (3,21)	90	60,48 (25,77)	-17,39 (3,30)	-21,76 [-29,53; -13,98] < 0,0001	-0,80 [-1,10; -0,51]
Veränderung zu Woche 8							
92	59,93 (28,31)	-40,82 (3,31)	85	60,21 (26,55)	-20,34 (3,40)	-20,47 [-28,46; -12,49] < 0,0001	-0,76 [-1,06; -0,45]
Veränderung zu Woche 10							
93	59,23 (27,78)	-41,48 (3,06)	82	59,34 (26,69)	-22,35 (3,18)	-19,13 [-26,58; -11,68] < 0,0001	-0,76 [-1,07; -0,46]
Veränderung zu Woche 12							
95	60,37 (28,01)	-41,52 (3,45)	81	59,56 (26,78)	-18,79 (3,61)	-22,73 [-31,14; -14,32] < 0,0001	-0,80 [-1,11; -0,50]
Einheit der MW = Millimeter							
a: p-Wert der ANCOVA.							
ANCOVA: Kovarianzanalyse (analysis of covariance); ITT: Intention-to-treat; KI: Konfidenzintervall;							
LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten;							
OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation);							
SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz; VAS: Visuelle Analogskala							
Quellen: Amgen 2018a, 2020a							

Die Veränderung der Schmerzhaftigkeit der oralen Aphthen zeigte zu allen Untersuchungszeitpunkten von Woche 1 bis Woche 12 eine statistisch signifikante Verbesserung zugunsten der Apremilast-Gruppe. Von Woche 2 bis Woche 12 lässt sich darüber hinaus auf Basis der SMD ein klinisch relevanter Vorteil zugunsten von Apremilast ableiten (Tabelle 4-39), den die Abbildung 5 grafisch veranschaulicht.

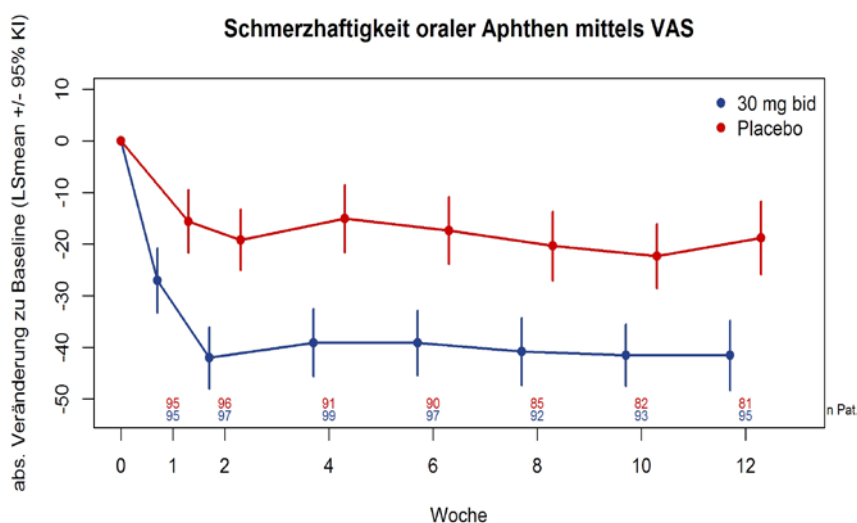


Abbildung 5: Zeitlicher Verlauf für Veränderung der Schmerzhaftigkeit der oraler Aphthen mittels VAS bis Woche 12 (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel
Quelle: Amgen 2020a

Tabelle 4-40: Ergebnisse für die Veränderung der Schmerzhaftigkeit der oralen Aphthen mittels VAS (MMRM) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
102	61,02 (27,63)	-39,04 (1,46)	101	61,18 (26,77)	-18,77 (1,47)	-20,27 [-23,19; -17,35] <0,0001	-1,90 [-2,24; -1,57]

Einheit der MW = Millimeter
a: p-Wert des MMRM, Veränderung gegenüber Baseline (Kovariable: Baseline-Wert, Faktoren: Behandlungsgruppe, Geschlecht und Region).
ITT: Intention-to-treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MMRM: Gemischtes Modell mit Messwiederholungen (mixed-effects model for repeated measures); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz; VAS: Visuelle Analogskala
Quelle: Amgen 2020a

Die Veränderung der Schmerzhaftigkeit der oralen Aphthen anhand des MMRM zeigte eine statistisch signifikante und klinisch relevante Verbesserung unter Apremilast im Vergleich zu Placebo (Tabelle 4-40). Die Ergebnisse der Analysemethoden, Imputation mittels LOCF, OC und MMRM, sind somit konsistent und verdeutlichen die Überlegenheit von Apremilast.

Extensionsphase (ab Woche 12 bis Woche 68)

Nachfolgend werden die Ergebnisse des Endpunkts „Schmerzhaftigkeit oraler Aphthen mittels VAS“ supportiv für die Extensionsphase ab Woche 12 bis Woche 68 dargestellt.

Tabelle 4-41: Ergebnisse für die Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo/Apremilast			Apremilast vs. Placebo/Apremilast	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Woche 12							
95	60,37 (28,01)	-41,52 (3,45)	81	59,56 (26,78)	-18,79 (3,61)	-22,73 [-31,14; -14,32] <0,0001	-0,80 [-1,11; -0,50]
Woche 16							
94	60,65 (28,03)	-42,36 (2,55)	82	59,23 (26,64)	-46,28 (2,66)	3,92 [-2,28; 10,13] 0,2133	0,19 [-0,11; 0,48]
Woche 28							
91	60,11 (28,23)	-40,90 (3,05)	77	59,53 (26,60)	-40,07 (3,22)	-0,83 [-8,34; 6,68] 0,8272	-0,03 [-0,34; 0,27]
Woche 40							
84	61,12 (27,82)	-43,63 (3,03)	73	59,30 (26,08)	-41,45 (3,18)	-2,18 [-9,57; 5,22] 0,5612	-0,09 [-0,41; 0,22]
Woche 52							
78	60,76 (28,43)	-38,97 (3,45)	70	58,59 (26,33)	-36,89 (3,48)	-2,08 [-10,28; 6,13] 0,6179	-0,08 [-0,40; 0,24]
Woche 64							
75	60,83 (28,18)	-35,21 (3,73)	68	59,03 (26,58)	-43,06 (3,75)	7,85 [-1,08; 16,78] 0,0844	0,29 [-0,04; 0,62]

Apremilast			Placebo/Apremilast			Apremilast vs. Placebo/Apremilast	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Woche 68 (Nachbeobachtung)							
84	61,92 (26,95)	-21,39 (3,99)	81	61,09 (26,40)	-21,92 (3,96)	0,53 [-9,30; 10,36] 0,9153	0,02 [-0,29; 0,32]

a: p-Wert der ANCOVA.
ANCOVA: Kovarianzanalyse (analysis of covariance); KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz; VAS: Visuelle Analogskala
Quelle: Amgen 2020a

Die Veränderung der Anzahl oraler Aphthen im Behandlungsverlauf spiegelt sich in dem von den Patienten mittels VAS berichteten Schmerzempfinden wider. Die deutliche Verbesserung der Schmerzhaftigkeit der oralen Aphthen in der Apremilast-Gruppe zu Woche 12 hielt bis zum Behandlungsende in Woche 64 an. In der Placebo/Apremilast-Gruppe verbesserte sich das Schmerzempfinden nach vierwöchiger Einnahme von Apremilast auf ein vergleichbares Niveau, wie zuvor in der Apremilast-Gruppe beobachtet. Dies bestätigte sich im Verlauf der gesamten Extensionsphase. Innerhalb von vier Wochen nach Behandlungsende wurde eine in beiden Gruppen vergleichbare wieder angestiegene Schmerzhaftigkeit der oralen Aphthen berichtet (Tabelle 4-41 und Abbildung 6).

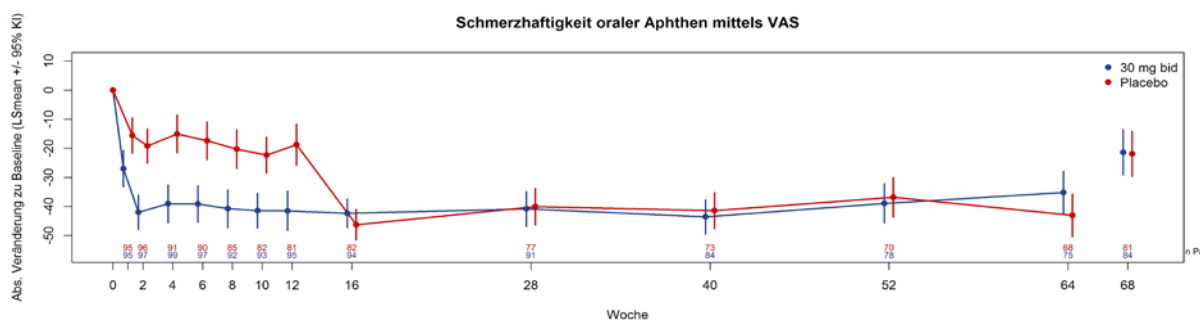


Abbildung 6: Zeitlicher Verlauf für die Veränderung der Schmerzhaftigkeit oraler Aphthen mittels VAS (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Schmerzhaftigkeit oraler Aphthen mittels VAS“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.3 Veränderung des Behçet's Syndrome Activity Score (BSAS) zu Woche 12 - RCT

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

Tabelle 4-42: Operationalisierung von Veränderung des BSAS zu Woche 12

Studie	Operationalisierung
RELIEF	<p>Der BSAS-Fragebogen beinhaltet zehn Fragen zur Bewertung der klinischen Aktivität des Behçet-Syndroms in den letzten vier Wochen durch den Patienten. Erfasst werden orale und genitale Aphthen, Hautläsionen, ferner gastrointestinale, neurologische, vaskuläre und Augenbeteiligungen sowie die aktuelle allgemeine Krankheitsaktivität. Jede Frage wird mit einer Punktzahl zwischen 0 und 10 bewertet, der Gesamtscore liegt somit zwischen 0 und 100. Ein höherer Wert bedeutet eine höhere Krankheitsaktivität und entspricht damit einem schlechteren Zustand.</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline, Woche 12 und Woche 64. Für die statistische Analyse wurde die ITT-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten und den Fragebogen zu Baseline komplettiert hatten. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline anhand von MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie SMD mit 95 %-KI dargestellt. Zur Stützung der Ergebnisse zu Woche 12 werden die Daten aus der Extensionsphase zu Woche 64 anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p>
<p>BSAS: Behçet's Syndrome Activity Score; ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz; MW: Mittelwert; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz</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 4-43: Bewertung des Verzerrungspotenzials für Veränderung des BSAS zu Woche 12 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
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig
<p>a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.</p>						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Veränderung des BSAS zu Woche 12“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die gesamte ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand objektiver und adäquater Methoden erhoben wurde. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Veränderung des BSAS zu Woche 12“ in der Studie RELIEF als niedrig eingestuft wird.

Stellen Sie die Ergebnisse für den Endpunkt „Veränderung des BSAS zu Woche 12“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Ergebnisse für den Endpunkt „Veränderung des BSAS zu Woche 12“ einschließlich der Rücklaufquoten des Fragebogens dargestellt.

Tabelle 4-44: Rücklaufquoten des BSAS-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Apremilast		Placebo	
	N	n (%)	N	n (%)
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)

BSAS: Behçet's Syndrome Activity Score; N: Anzahl Patienten der ITT-Population; n: Anzahl der Patienten mit Baseline- und Post-Baseline-Erhebung
Quelle: Amgen 2020a

Zu Baseline lag für alle Patienten ein vollständiger Rücklauf mit validen Ergebnissen vor. Zu Woche 12 fehlten die Rückläufe von acht Patienten (7,7 %) in der Apremilast-Gruppe und 21 Patienten (20,4 %) in der Placebo-Gruppe. Die Rücklaufquoten des BSAS lagen damit zu beiden Erhebungszeitpunkten bei über 70 % (Tabelle 4-44).

Tabelle 4-45: Ergebnisse für Veränderung des BSAS zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
104	42,75 (16,22)	-17,35 (1,80)	103	44,30 (16,86)	-5,41 (1,78)	-11,94 [-16,20; -7,67] p < 0,0001	-0,76 [-1,05; -0,48]

a: p-Wert der ANCOVA, Veränderung gegenüber Baseline (Kovariable: Baseline-Wert, Faktoren: Behandlungsgruppe, Geschlecht und Region).
ANCOVA: Kovarianzanalyse (analysis of covariance); BSAS: Behçet's Syndrome Activity Score;
ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrat-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrat-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten;
SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz
Quellen: Amgen 2019, 2020a

Die Krankheitsaktivität anhand des BSAS war zu Baseline in beiden Behandlungsgruppen vergleichbar. Mit einem Score von 42,75 in der Apremilast-Gruppe und 44,30 in der Placebo-Gruppe lag sie im mittleren Bereich des Gesamtscores von 100. Zu Woche 12 wurde eine statistisch signifikante und klinisch relevante Verbesserung zugunsten der Apremilast-Gruppe verzeichnet (Tabelle 4-45).

Extensionsphase (ab Woche 12 bis Woche 68)

Nachfolgend werden die Ergebnisse des Endpunkts „Veränderung des BSAS“ supportiv für die Extensionsphase zu Woche 12 und Woche 64 dargestellt.

Tabelle 4-46: Ergebnisse für Veränderung des BSAS in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo/Apremilast			Apremilast vs. Placebo/Apremilast	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Woche 12							
96	41,58 (15,23)	-18,95 (1,80)	82	42,90 (17,07)	-8,11 (1,87)	-10,85 [-15,18; -6,52] < 0,0001	-0,74 [-1,04; -0,44]
Woche 64							
75	41,94 (15,62)	-16,72 (2,41)	68	42,76 (17,27)	-15,81 (2,41)	-0,91 [-6,64; 4,82] 0,7535	-0,05 [-0,38; 0,28]

a: p-Wert der ANCOVA.

ANCOVA: Kovarianzanalyse (analysis of covariance); BSAS: Behçet's Syndrome Activity Score;
 KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten;
 OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation);
 SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz

Quelle: Amgen 2020a

Die deutliche Verbesserung der Krankheitsaktivität anhand des BSAS-Fragebogens in der Apremilast-Gruppe zu Woche 12, bestätigte sich in Woche 64. Ein vergleichbarer Rückgang der Krankheitsaktivität wurde von den Patienten der Placebo/Apremilast-Gruppe nach 52 Wochen Behandlung mit Apremilast berichtet (Tabelle 4-46 und Abbildung 7).

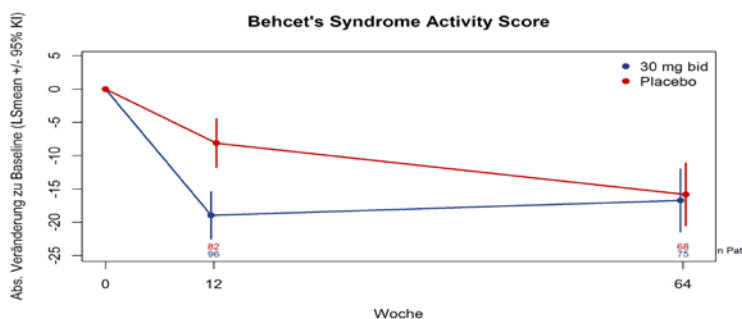


Abbildung 7: Zeitlicher Verlauf für Veränderung des BSAS (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Veränderung des BSAS“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.4 Veränderung des Behçet's Disease Current Activity Form (BDCAF) zu Woche 12 – RCT

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

Tabelle 4-47: Operationalisierung von Veränderung des BDCAF zu Woche 12

Studie	Operationalisierung
RELIEF	<p>Der BDCAF ist ein krankheitsspezifischer Fragebogen zur Beurteilung der Krankheitsaktivität in den letzten vier Wochen. Er besteht aus drei unabhängigen Komponenten, dem BDCA-Indexwert (BDCAI-Score) sowie der Wahrnehmung der Krankheitsaktivität durch den Patienten und der Gesamtwahrnehmung der Krankheitsaktivität durch den Untersucher. Der BDCAI beinhaltet insgesamt 12 Fragen, aufgeteilt in Fragen zu:</p> <ul style="list-style-type: none"> • Kopfschmerz • oralen Aphthen

<ul style="list-style-type: none"> • genitalen Aphthen • Hautläsionen • Gelenkbeteiligung • gastrointestinalen Symptomen • Beteiligung der Augen • Beteiligung des Nervensystems • großen Gefäßentzündungen <p>Der Gesamtscore des BDCAI liegt zwischen 0 und 12; ein höherer Wert ist mit einer höheren Krankheitsaktivität assoziiert.</p> <p>Die Wahrnehmung der Krankheitsaktivität durch den Patienten und den Untersucher erfolgt jeweils mittels einer siebenstufigen Likert-Skala von sehr gut bis sehr schlecht (Punktwert von 1 bis 7).</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline, Woche 12 und Woche 64. Für die statistische Analyse wurde die ITT-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Wert zu Baseline verfügbar war. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline anhand von MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie SMD mit 95 %-KI dargestellt. Zur Stützung der Ergebnisse zu Woche 12 werden die Daten aus der Extensionsphase zu Woche 64 anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p>
<p>BDCAF: Behçet’s Disease Current Activity Form; BDCAI: Behçet’s Disease Current Activity Indexwert; ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz</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 4-48: Bewertung des Verzerrungspotenzials für Veränderung des BDCAF zu Woche 12 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
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig
a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Veränderung des BDCAF zu Woche 12“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war. In der Apremilast-Gruppe wurden alle Patienten, in der Placebo-Gruppe 99 % der Patienten in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei weniger als 5 %. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand objektiver und adäquater Methoden erhoben wurde. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Veränderung des BDCAF zu Woche 12“ in der Studie RELIEF als niedrig eingestuft wird.

Stellen Sie die Ergebnisse für den Endpunkt „Veränderung des BDCAF zu Woche 12“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Ergebnisse für den Endpunkt „Veränderung des BDCAF zu Woche 12“ einschließlich der Rücklaufquoten des Fragebogens dargestellt.

Tabelle 4-49: Rücklaufquoten des BDCAF-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Endpunkt Zeitpunkt	Apremilast		Placebo	
	N	n (%)	N	n (%)
Indexwert				
Baseline	104	104 (100,0)	103	102 (99,0)
Woche 12	104	97 (93,3)	103	81 (78,6) ^a
Patientenwahrnehmung				
Baseline	104	104 (100,0)	103	102 (99,0)
Woche 12	104	97 (93,3)	103	81 (78,6) ^a
Untersucherwahrnehmung				
Baseline	104	104 (100,0)	103	102 (99,0)
Woche 12	104	97 (93,3)	103	81 (78,6) ^a
a: Nur Patienten mit Baseline- und Post-Baseline-Wert berücksichtigt. BDCAF: Behçet's Disease Current Activity Form; N: Anzahl Patienten der ITT-Population; n: Anzahl Patienten mit Erhebung Quellen: Amgen 2018a, 2020a				

Zu Baseline lag für alle Patienten in der Apremilast-Gruppe und für 102 Patienten (99,0 %) der Placebo-Gruppe ein vollständiger Rücklauf vor. Zu Woche 12 fehlten die Rückläufe von sieben Patienten (6,7 %) in der Apremilast-Gruppe und 20 Patienten (19,4 %) in der Placebo-Gruppe. Die Daten des Patienten aus der Placebo-Gruppe ohne Baseline-Erhebung blieben in der Auswertung unberücksichtigt. Die Rücklaufquoten des BDCAF lagen damit zu beiden Erhebungszeitpunkten bei über 70 % (Tabelle 4-49).

Tabelle 4-50: Ergebnisse für Veränderung des BDCAF zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Indexwert							
104	3,70 (1,58)	-0,92 (0,20)	102	3,63 (1,67)	-0,41 (0,20)	-0,52 [-0,99; -0,04] 0,0335	-0,30 [-0,57; -0,02]
Patientenwahrnehmung							
104	4,65 (1,39)	-1,68 (0,18)	102	4,78 (1,33)	-0,67 (0,18)	-1,01 [-1,44; -0,58] < 0,0001	-0,64 [-0,92; -0,36]
Untersucherwahrnehmung							
104	4,47 (1,21)	-1,64 (0,17)	102	4,45 (1,12)	-0,69 (0,17)	-0,95 [-1,35; -0,55] < 0,0001	-0,65 [-0,93; -0,37]
<p>a: p-Wert der ANCOVA, Veränderung gegenüber Baseline (Kovariable: Baseline-Wert, Faktoren: Behandlungsgruppe, Geschlecht und Region).</p> <p>ANCOVA: Kovarianzanalyse (analysis of covariance); BDCAF: Behçet's Disease Current Activity Form; ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz</p> <p>Quelle: Amgen 2020a</p>							

Die mittels BDCAF erfasste Krankheitsaktivität war zu Baseline für die drei Einzelkomponenten des Fragebogens zwischen den beiden Behandlungsgruppen jeweils vergleichbar. Aus dem Indexwert resultierte insgesamt eine etwas geringere Krankheitsaktivität als aus den Beurteilungen durch Patienten und Untersucher.

Die zu Woche 12 dokumentierten Werte zeigten auf Basis von Patientenwahrnehmung und Untersucherwahrnehmung für beide Behandlungsgruppen gegenüber Baseline vergleichbare Veränderungen, die statistisch signifikant und klinisch relevant zugunsten von Apremilast ausfielen. Für den Indexwert wurden zu Woche 12 in der Apremilast-Gruppe statistisch

signifikante, aber klinisch nicht relevante Verbesserungen gegenüber der Placebo-Gruppe berichtet (Tabelle 4-50).

Extensionsphase (ab Woche 12 bis Woche 68)

Nachfolgend werden die Ergebnisse des Endpunkts „Veränderung des BDCAF“ supportiv für die Extensionsphase zu Woche 12 und Woche 64 dargestellt.

Tabelle 4-51: Ergebnisse für die Veränderung des BDCAF in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo/Apremilast			Apremilast vs. Placebo/Apremilast	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Indexwert - Woche 12							
97	3,66 (1,61)	-1,01 (0,20)	81	3,58 (1,67)	-0,68 (0,21)	-0,33 [-0,82; 0,15] 0,1746	-0,20 [-0,50; 0,09]
Indexwert - Woche 64							
75	3,53 (1,56)	-1,40 (0,23)	65	3,69 (1,72)	-1,26 (0,24)	-0,14 [-0,70; 0,43] 0,6342	-0,08 [-0,41; 0,25]
Patientenwahrnehmung - Woche 12							
97	4,62 (1,41)	-1,71 (0,18)	81	4,67 (1,25)	-0,98 (0,19)	-0,73 [-1,18; -0,29] 0,0015	-0,48 [-0,78; -0,18]
Patientenwahrnehmung - Woche 64							
75	4,69 (1,36)	-1,91 (0,21)	65	4,69 (1,32)	-1,70 (0,21)	-0,21 [-0,71; 0,29] 0,4042	-0,14 [-0,47; 0,19]
Untersucherwahrnehmung - Woche 12							
97	4,42 (1,18)	-1,70 (0,17)	81	4,35 (0,96)	-0,93 (0,18)	-0,77 [-1,18; -0,36] 0,0003	-0,55 [-0,85; -0,25]
Untersucherwahrnehmung - Woche 64							
75	4,44 (1,14)	-1,78 (0,19)	65	4,38 (0,98)	-1,83 (0,19)	0,04 [-0,41; 0,49] 0,8495	0,03 [-0,30; 0,36]
a: p-Wert der ANCOVA. ANCOVA: Kovarianzanalyse (analysis of covariance); BDCAF: Behçet's Disease Current Activity Form; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz Quelle: Amgen 2020a							

Die zu Woche 12 beobachtete Verbesserung in den drei Einzelkomponenten der Krankheitsaktivität anhand des BDCAF-Fragebogens in der Apremilast-Gruppe bestätigte sich zu Woche 64. In der Placebo/Apremilast-Gruppe verbesserte sich die Krankheitsaktivität zu Woche 64 ebenfalls in allen drei Einzelkomponenten und lag jeweils auf vergleichbarem Niveau zur Apremilast-Gruppe (Tabelle 4-51, Abbildung 8, Abbildung 9 und Abbildung 10).

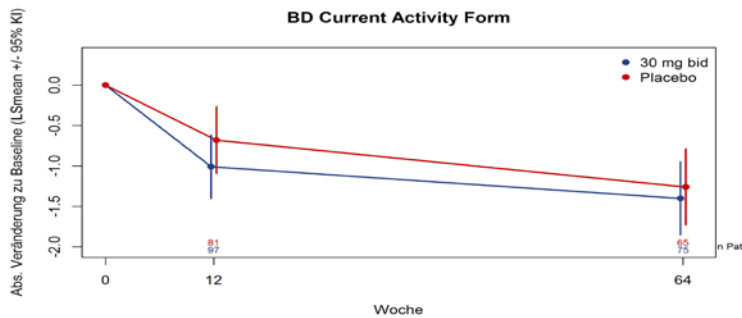


Abbildung 8: Zeitlicher Verlauf für Veränderung des BDCAF – Indexwert (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

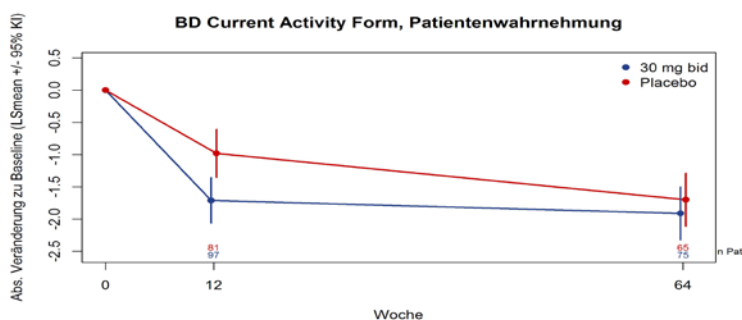


Abbildung 9: Zeitlicher Verlauf für Veränderung des BDCAF – Patientenwahrnehmung (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

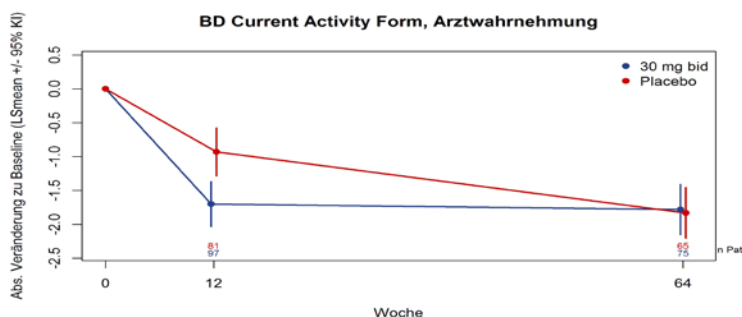


Abbildung 10: Zeitlicher Verlauf für Veränderung des BDCAF – Untersucherwahrnehmung (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Veränderung des BDCAF“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.5 Veränderung des Behçet's disease Quality of Life (BD-QoL) zu Woche 12 – RCT

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

Tabelle 4-52: Operationalisierung von Veränderung des BD-QoL zu Woche 12

Studie	Operationalisierung
RELIEF	<p>Der BD-QoL wurde als krankheitsspezifisches Messinstrument zur Erfassung der gesundheitsbezogenen Lebensqualität von Patienten mit dem Behçet-Syndrom eingesetzt. Der Fragebogen beinhaltet 30 Aussagen zum Einfluss des Behçet-Syndroms auf Aktivitäten des täglichen Lebens und die gesundheitsbezogene Lebensqualität, die der Patient mit „wahr“ (Punktwert = 1) oder „nicht wahr“ (Punktwert = 0) beantwortet. Der Gesamtscore liegt zwischen 0 und 30; ein höherer Wert markiert eine stärkere subjektive Beeinträchtigung.</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline, Woche 12 und Woche 64. Für die statistische Analyse wurde die ITT-Population herangezogen, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Wert zu Baseline verfügbar war. Zu Woche 12 fehlende Werte wurden mittels LOCF imputiert. Die Ergebnisse werden als Veränderung zu Woche 12 gegenüber Baseline anhand von MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie SMD mit 95 %-KI dargestellt. Zur Stützung der Ergebnisse zu Woche 12 werden die Daten aus der Extensionsphase zu Woche 64 anhand der vorstehenden Verfahren aus der OC-Analyse abgebildet.</p>
<p>BD-QoL: Behçet's disease quality of life; ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz</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 4-53: Bewertung des Verzerrungspotenzials für Veränderung des BD-QoL zu Woche 12 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
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig
a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Veränderung des BD-QoL zu Woche 12“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die gesamte ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand valider und adäquater Methoden erhoben wurde. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Veränderung des BD-QoL zu Woche 12“ in der Studie RELIEF als niedrig eingestuft wird.

Stellen Sie die Ergebnisse für den Endpunkt „Veränderung des BD-QoL zu Woche 12“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Ergebnisse für den Endpunkt „Veränderung des BD-QoL zu Woche 12“ einschließlich der Rücklaufquoten des Fragebogens dargestellt.

Tabelle 4-54: Rücklaufquoten des BD-QoL-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Apremilast		Placebo	
	N	n (%)	N	n (%)
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)

BD-QoL: Behçet's disease quality of life; N: Anzahl Patienten der ITT-Population; n: Anzahl Patienten mit Baseline- und Post-Baseline-Erhebung
Quelle: Amgen 2020a

Zu Baseline lag für alle Patienten ein vollständiger Rücklauf vor. Zu Woche 12 fehlten die Rückläufe von acht Patienten (7,7 %) in der Apremilast-Gruppe und 21 Patienten (20,4 %) in der Placebo-Gruppe. Die Rücklaufquoten des BD-QoL lagen zu beiden Erhebungszeitpunkten bei über 70 %. Zu Woche 12 betragen sie in der Apremilast-Gruppe 92,3 % und in der Placebo-Gruppe 79,6 % (Tabelle 4-54).

Tabelle 4-55: Ergebnisse für Veränderung des BD-QoL zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
104	10,22 (8,24)	-3,47 (0,67)	103	11,24 (8,16)	-0,52 (0,66)	-2,95 [-4,53; -1,38] 0,0003	-0,51 [-0,79; -0,24]

a: p-Wert der ANCOVA, Veränderung gegenüber Baseline (Kovariable: Baseline-Wert, Faktoren: Behandlungsgruppe, Geschlecht und Region).
ANCOVA: Kovarianzanalyse (analysis of covariance); BD-QoL: Behçet's disease quality of life;
ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten;
SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz
Quelle: Amgen 2020a

Zu Baseline lag die Beeinträchtigung der Lebensqualität auf Basis des BD-QoL in beiden Behandlungsgruppen auf einem vergleichbaren Niveau: 10,22 in der Apremilast-Gruppe und 11,24 in der Placebo-Gruppe. Zu Woche 12 berichteten die Patienten in der Apremilast-Gruppe eine statistisch signifikante und klinisch relevante Verbesserung der Lebensqualität gegenüber den Patienten der Placebo-Gruppe (Tabelle 4-55).

Extensionsphase (ab Woche 12 bis Woche 68)

Nachfolgend werden die Ergebnisse des Endpunkts „Veränderung des BD-QoL“ supportiv für die Extensionsphase zu Woche 12 und Woche 64 dargestellt.

Tabelle 4-56: Ergebnisse für Veränderung des BD-QoL in der Extensionsphase (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo/Apremilast			Apremilast vs. Placebo/Apremilast	
N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]
Woche 12							
96	9,85 (7,99)	-4,01 (0,74)	82	10,83 (8,23)	-1,02 (0,76)	-3,00 [-4,75; -1,24] 0,0009	-0,50 [-0,80; -0,20]
Woche 64							
75	9,72 (7,74)	-3,72 (0,82)	68	10,78 (8,38)	-3,10 (0,81)	-0,61 [-2,53; 1,31] 0,5291	-0,11 [-0,43; 0,22]

a: p-Wert der ANCOVA.
 ANCOVA: Kovarianzanalyse (analysis of covariance); BD-QoL: Behçet's disease quality of life;
 KI: Konfidenzintervall; LS-MW: Kleinste-Quadrat-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrat-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten;
 OC: Vorliegende Beobachtungen (observed cases); SD: Standardabweichung (standard deviation);
 SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz
 Quelle: Amgen 2020a

Die von den Patienten in der Apremilast-Gruppe zu Woche 12 berichtete Verbesserung der Lebensqualität anhand des krankheitsspezifischen BD-QoL-Fragebogens wurde zu Woche 64 bestätigt. In der Placebo/Apremilast-Gruppe verbesserte sich die Lebensqualität von Woche 12 zu Woche 64 annähernd auf das Niveau der Apremilast-Gruppe (Tabelle 4-56 und Abbildung 11).



Abbildung 11: Zeitlicher Verlauf für Veränderung des BD-QoL (OC) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Quelle: Amgen 2020a

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Veränderung des BD-QoL“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.6 Verbesserung des 36-Item Short Form Health Survey, Version 2 (SF-36v2) um $\geq 2,5$ Punkte zu Woche 12 – RCT

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

Tabelle 4-57: Operationalisierung von Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12

Studie	Operationalisierung
RELIEF	<p>Der generische SF-36-Fragebogen bildet die gesundheitsbezogene Lebensqualität über zwei Hauptdimensionen ab, die körperliche Summenskala (PCS) und die psychische Summenskala (MCS), die sich aus jeweils vier Subskalen zusammensetzen.</p> <p>Die PCS wird aus den folgenden Subskalen gebildet:</p> <ul style="list-style-type: none"> • Körperliche Funktionsfähigkeit • Körperliche Rollenfunktion • Körperliche Schmerzen • Allgemeine Gesundheitswahrnehmung <p>Die MCS wird aus folgenden Subskalen gebildet:</p> <ul style="list-style-type: none"> • Vitalität • Soziale Funktionsfähigkeit • Emotionale Rollenfunktion • Psychisches Wohlbefinden. <p>Zwei bis zehn Fragen pro Subskala ergeben insgesamt 36 Fragen, von denen eine Frage zur Gesundheitsveränderung in den vergangenen 12 Monaten in der Auswertung unberücksichtigt bleibt. Alle Fragen werden mit Hilfe einer drei- bis sechsstufigen Likert-Skala beantwortet und die erhaltenen Werte für jede Subskala in einen Bereich zwischen 0 (schlechteste Lebensqualität) und 100 (beste Lebensqualität) transformiert. Die PCS- und MCS-Scores werden auf einen Mittelwert von 50 und eine Standardabweichung von 10 transformiert. Höhere Werte bedeuten eine bessere gesundheitsbezogene Lebensqualität.</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline und zu Woche 12. Die Analyse der Verbesserung um mindestens 2,5 Punkte in jeder Skala basierte auf der ITT-Population, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Wert zu Baseline verfügbar war. Patienten mit fehlenden Werten zu Woche 12 wurden als Non-Responder eingestuft. In die Ergebnisdarstellung werden, neben Anzahl und prozentuaalem Anteil der Responder, RR, OR und RD mit dem jeweiligen 95 %-KI sowie dem p-Wert einbezogen.</p>
<p>ITT: Intention-to-treat; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); MCS: Psychische Summenskala (mental component summary); OR: Odds Ratio; PCS: Körperliche Summenskala (physical component summary); RD: Risikodifferenz; RR: Relatives Risiko; SF-36: 36-Item Short Form Health Survey, Version 2</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 4-58: Bewertung des Verzerrungspotenzials für Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig
a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die gesamte ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand objektiver und adäquater Methoden erhoben wurde. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ in der Studie RELIEF als niedrig eingestuft wird.

Stellen Sie die Ergebnisse für den Endpunkt „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Rücklaufquoten des SF-36-Fragebogens sowie die Ergebnisse für den Endpunkt „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ für die beiden Summenskalen (PCS und MCS) und die acht Subskalen dargestellt.

Eine Verbesserung der PCS und MCS um jeweils mindestens 2,5 Punkte wurde als MCID definiert (Amgen 2019).

Tabelle 4-59: Rücklaufquoten des SF-36-Fragebogens zu Baseline und Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel

Dimension Zeitpunkt	Apremilast		Placebo	
	N	n (%)	N	n (%)
Körperliche Funktionsfähigkeit				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Körperliche Rollenfunktion				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Körperliche Schmerzen				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Allgemeine Gesundheitswahrnehmung				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Vitalität				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Soziale Funktionsfähigkeit				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Emotionale Rollenfunktion				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
Psychisches Wohlbefinden				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
PCS				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)

Dimension Zeitpunkt	Apremilast		Placebo	
	N	n (%)	N	n (%)
MCS				
Baseline	104	104 (100,0)	103	103 (100,0)
Woche 12	104	96 (92,3)	103	82 (79,6)
MCS: Psychische Summenskala (mental component summary); N: Anzahl Patienten der ITT-Population; n: Anzahl Patienten mit Baseline- und Post-Baseline-Erhebung; PCS: Körperliche Summenskala (physical component summary) Quelle: Amgen 2020a				

Zu Baseline lag für alle Patienten ein vollständiger Rücklauf vor. Zu Woche 12 fehlten die Rückläufe von acht Patienten (7,7 %) in der Apremilast-Gruppe und 21 Patienten (20,4 %) in der Placebo-Gruppe. Die Rücklaufquoten des SF-36 lagen zu beiden Erhebungszeitpunkten bei über 70 %. Zu Woche 12 betrug sie für die acht Subskalen, einschließlich der beiden Summenskalen PCS und MCS, in der Apremilast-Gruppe 92,3 % und in der Placebo-Gruppe 79,6 % (Tabelle 4-59).

Tabelle 4-60: Ergebnisse für Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12 (NRI) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Apremilast			Placebo			Apremilast vs. Placebo			
N	n1 (%)	n2 (%)	N	n1 (%)	n2 (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p- Wert ^a
Körperliche Funktionsfähigkeit									
104	38 (36,5)	8 (7,7)	103	26 (25,2)	21 (20,4)	1,42 [0,94; 2,14]	1,72 [0,94; 3,16]	0,11 [-0,02; 0,23]	0,0804
Körperliche Rollenfunktion									
104	45 (43,3)	8 (7,7)	103	25 (24,3)	21 (20,4)	1,78 [1,19; 2,67]	2,40 [1,32; 4,36]	0,19 [0,06; 0,31]	0,0042
Körperliche Schmerzen									
104	61 (58,6)	8 (7,7)	103	44 (42,7)	21 (20,4)	1,37 [1,04; 1,80]	1,91 [1,10; 3,34]	0,16 [0,02; 0,29]	0,0226
Allgemeine Gesundheitswahrnehmung									
104	49 (47,1)	8 (7,7)	103	31 (30,1)	21 (20,4)	1,57 [1,11; 2,24]	2,08 [1,17; 3,71]	0,16 [0,03; 0,29]	0,0322
Vitalität									
104	63 (60,6)	8 (7,7)	103	43 (41,8)	21 (20,4)	1,45 [1,11; 1,90]	2,18 [1,24; 3,85]	0,18 [0,05; 0,31]	0,0069
Soziale Funktionsfähigkeit									
104	54 (51,9)	8 (7,7)	103	27 (26,2)	21 (20,4)	1,99 [1,37; 2,88]	3,07 [1,70; 5,52]	0,25 [0,12; 0,38]	0,0002

Apremilast			Placebo			Apremilast vs. Placebo			
N	n1 (%)	n2 (%)	N	n1 (%)	n2 (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p- Wert ^a
Emotionale Rollenfunktion									
104	48 (46,1)	8 (7,7)	103	33 (32,0)	21 (20,4)	1,43 [1,01; 2,02]	1,83 [1,03; 3,24]	0,13 [0,00; 0,27]	0,0405
Psychisches Wohlbefinden									
104	57 (54,8)	8 (7,7)	103	41 (39,8)	21 (20,4)	1,42 [1,06; 1,89]	1,84 [1,05; 3,20]	0,14 [0,01; 0,28]	0,0322
PCS									
104	50 (48,1)	8 (7,7)	103	36 (35,0)	21 (20,4)	1,38 [0,99; 1,91]	1,73 [0,98; 3,04]	0,12 [-0,01; 0,26]	0,0606
MCS									
104	57 (54,8)	8 (7,7)	103	34 (33,0)	21 (20,4)	1,72 [1,26; 2,36]	2,52 [1,42; 4,47]	0,21 [0,07; 0,34]	0,0016
a: p-Wert des Cochran-Mantel-Haenszel-Tests. ITT: Intention-to-treat; KI: Konfidenzintervall; MCS: Psychische Summenskala (mental component summary); N: Anzahl der Patienten; n1: Anzahl der Patienten mit Ereignis; n2: Anzahl der Patienten mit Imputation; NRI: Non-Responder-Imputation; OR: Odds Ratio; PCS: Körperliche Summenskala (physical component summary); RD: Risikodifferenz; RR: Relatives Risiko Quelle: Amgen 2020a									

Anhand der einzelnen Skalen des SF-36 ließen sich, mit Ausnahme der körperlichen Funktionsfähigkeit und der PCS, statistisch signifikante Unterschiede zeigen. Insgesamt fielen alle Unterschiede gleichgerichtet und zugunsten von Apremilast aus.

Der Anteil der Responder (Patienten ohne Imputation) lag in der Apremilast-Gruppe zwischen 36,5 % (körperliche Funktionsfähigkeit) und 60,6 % (Vitalität) und in der Placebo-Gruppe zwischen 24,3 % (körperliche Rollenfunktion) und 42,7 % (körperliche Schmerzen). Die größten Unterschiede zwischen den beiden Behandlungsgruppen wurden für die Subskala soziale Funktionsfähigkeit berichtet und die geringsten für die Subskala körperliche Funktionsfähigkeit (Tabelle 4-60).

Insgesamt lässt sich hiermit eine Verbesserung der Lebensqualität bei Anwendung von Apremilast ableiten.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.7 Sicherheit (unerwünschte Ereignisse, UE) - RCT

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

Tabelle 4-61: Operationalisierung von Sicherheit

Studie	Operationalisierung
RELIEF	<p>Es werden alle UE berücksichtigt, die vom Tag der ersten Einnahme bis maximal 28 Tage nach der letzten Einnahme der Studienmedikation auftraten. Alle UE wurden nach der MedDRA-Version 20.0 kodiert. Die Beurteilung des Schweregrades erfolgte nach der Ausprägung „leicht“, „mäßig“ oder „schwer“.</p> <p>Folgende sicherheitsrelevanten Endpunkte wurden erfasst:</p> <ul style="list-style-type: none"> • UE gesamt • SUE • Schwere UE • Therapieabbruch aufgrund von UE • UE von besonderem Interesse <p>Die statistische Analyse erfolgte auf Basis der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation (as-treated) erhalten hatten. Die Analyse umfasst die Gesamtrate der sicherheitsrelevanten Endpunkte zu Woche 12 sowie diejenigen UE mittels SOC und PT, die bei mindestens 10 % der Patienten und die bei mindestens zehn Patienten UND mindestens 1 % der Patienten in einer Behandlungsgruppe bis Woche 12 auftraten. Zur Ergebnisdarstellung werden, neben Anzahl und prozentualem Anteil der Patienten mit mindestens einem Ereignis pro Endpunkt, RR, OR und RD mit dem jeweiligen 95 %-KI und dem p-Wert herangezogen. Zusätzlich werden die UE, die zum Therapieabbruch führten, auf SOC / PT-Ebene deskriptiv abgebildet.</p> <p>Zur Stützung der Ergebnisse zu Woche 12 werden die Daten aus der Extensionsphase zu Woche 64 und Woche 68 für die AAT-Population mit Anzahl und prozentualem Anteil der Patienten mit mindestens einem Ereignis pro Endpunkt zusätzlich dargestellt.</p>
<p>AAT: Mit Apremilast behandelt (Apremilast-as-treated); KI: Konfidenzintervall; MedDRA: Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung (Medical Dictionary for Regulatory Activities); OR: Odds Ratio; PT: Bevorzugter Begriff (preferred term); RD: Risikodifferenz; RR: Relatives Risiko; SOC: Systemorganklasse (System Organ Class); SUE: Schwerwiegende/s unerwünschte/s Ereignis/se; UE: Unerwünschte/s Ereignis/se</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 4-62: Bewertung des Verzerrungspotenzials für Sicherheit 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
RELIEF ^a	niedrig	ja	ja	ja	ja	niedrig
a: Die Bewertung des Verzerrungspotenzials des Endpunkts bezieht sich ausschließlich auf die initiale 12-wöchige Placebo-kontrollierte Behandlungsphase.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das niedrige Verzerrungspotenzial auf Studienebene wird für den Endpunkt „Sicherheit“ bestätigt. Eine Beeinflussung des Ergebnisses ist aufgrund der doppelten Verblindung weder durch die Patienten noch durch die behandelnden bzw. weiterbehandelnden Personen zu erwarten. Die Analyse erfolgte für die gesamte Sicherheitspopulation, die alle mit Studienmedikation behandelten Patienten (as-treated) einschloss. Die Sicherheitspopulation stimmte mit der ITT-Population vollständig überein, da alle behandelten Patienten die zugewiesene Studienmedikation erhielten. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand objektiver und adäquater Methoden erhoben wurde. Es liegen weiterhin keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts „Sicherheit“ in der Studie RELIEF als niedrig eingestuft wird.

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

Placebo-kontrollierte Phase (Baseline bis Woche 12)

Nachfolgend werden die Ergebnisse des Endpunkts „Sicherheit“ für die Gesamtrate der UE, die UE von besonderem Interesse, die UE bei mindestens 10 % der Patienten in einem Studienarm sowie die UE, die zum Studienabbruch führten, zu Woche 12 dargestellt.

Tabelle 4-63: Ergebnisse für Sicherheit zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Apremilast		Placebo		Apremilast vs. Placebo			
N	n (%)	N	n (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a
UE gesamt							
104	82 (78,8)	103	74 (71,8)	1,10 [0,94; 1,28]	1,46 [0,77; 2,76]	0,07 [-0,05; 0,19]	0,2436
SUE							
104	3 (2,9)	103	4 (3,9)	0,74 [0,17; 3,24]	0,74 [0,16; 3,37]	-0,01 [-0,06; 0,04]	0,6917
Therapieabbruch aufgrund von UE							
104	3 (2,9)	103	5 (4,8)	0,59 [0,15; 2,42]	0,58 [0,14; 2,50]	-0,02 [-0,07; 0,03]	0,4633
Schwere UE							
104	6 (5,8)	103	6 (5,8)	0,99 [0,33; 2,97]	0,99 [0,31; 3,18]	-0,00 [-0,06; 0,06]	0,9863
Jeder Patient wird für die zutreffende UE-Kategorie einmal gezählt.							
a: p-Wert des Cochran-Mantel-Haenszel-Tests.							
KI: Konfidenzintervall; N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; SUE: schwerwiegende/s unerwünschte/s Ereignis/se; UE: Unerwünschte/s Ereignis/se							
Quellen: Amgen 2019, 2020a							

Die im Verlauf der initialen 12-wöchigen Placebo-kontrollierten Behandlungsphase nach der ersten Einnahme der Studienmedikation erfassten UE waren zwischen beiden Gruppen vergleichbar. Für 82 Patienten (78,8 %) in der Apremilast-Gruppe und 74 Patienten (71,8 %) in der Placebo-Gruppe wurde mindestens ein UE erfasst. Die Anzahl an SUE, schweren UE sowie Therapieabbruch aufgrund von UE, fiel in beiden Gruppen niedrig aus.

Bei drei Patienten (2,9 %) in der Apremilast-Gruppe und vier Patienten (3,9 %) in der Placebo-Gruppe kam es bis Woche 12 zu mindestens einem SUE. Die Anzahl der Patienten mit SUE lag somit in beiden Behandlungsgruppen unter 5 %. Bei den in der Apremilast-Gruppe aufgetretenen SUE handelte es sich um Behçet-Syndrom, Migräne und Weichteilverletzung. Die Patienten der Placebo-Gruppe berichteten jeweils ein Ereignis von akuter febriler neutrophiler Dermatose, infektiöser Diarrhoe, Erythema multiforme, Genitalinfektion, genitaler Pilzinfektion, oralem Geschwür und Hautläsion als schwerwiegend. Kein SUE wurde bei mehr als einem Patienten beobachtet (Amgen 2019).

Schwere UE traten bei sechs der 82 Patienten (7,3 %) in der Apremilast-Gruppe und sechs der 74 Patienten (8,1 %) in der Placebo-Gruppe mit mindestens einem UE auf. Der Anteil der nicht schweren UE lag somit bei 92,7 % in der Apremilast-Gruppe und 91,9 % in der Placebo-Gruppe

(Tabelle 4-63). Die einzelnen UE wurden bei weniger als 5 % der Patienten als schwere UE beurteilt. Die höchsten Anzahlen an Patienten mit schweren UE wurden in der Apremilast-Gruppe für Übelkeit bei drei Patienten (2,9 %), für Kopfschmerz bei zwei Patienten (1,9 %) sowie für Erbrechen bei einem Patienten (1,0 %) festgestellt. In der Placebo-Gruppe berichtete ein Patient (1,0 %) eine als schweres UE eingestufte Diarrhoe (Amgen 2019).

Zusammenfassend sind die UE zwischen den beiden Behandlungsgruppen vergleichbar. Es wurden wenige schwere UE, SUE und UE, die zu einem Therapieabbruch führten, im Rahmen der Studie berichtet; statistisch signifikante Unterschiede lagen dabei nicht vor. Auch eine Häufung von spezifischen schweren UE, SUE und UE, die zu einem Therapieabbruch führten, lag nicht vor. Dies verdeutlicht die gute Verträglichkeit von Apremilast.

Tabelle 4-64: Ergebnisse der UE von besonderem Interesse (Diarrhoe) zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Apremilast		Placebo		Apremilast vs. Placebo			
N	n (%)	N	n (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a
Gesamtrate Diarrhoe^b							
104	33 (31,7)	103	18 (17,5)	1,82 [1,10; 3,01]	2,19 [1,14; 4,23]	0,14 [0,03; 0,26]	0,0176
Schwere Diarrhoe^b							
104	0 (0,0)	103	1 (1,0)	0,33 [0,01; 8,01] ^c	0,33 [0,01; 8,12] ^c	-0,01 [-0,03; 0,01] ^c	0,3150
Schwerwiegende Diarrhoe^b							
104	0 (0,0)	103	0 (0,0)	0,99 [0,02; 49,45] ^c	0,99 [0,02; 50,39] ^c	-0,00 [-0,00; 0,00] ^c	n.b.
a: p-Wert des Cochran-Mantel-Haenszel-Tests. b: Diarrhoe als UE von besonderem Interesse ist entsprechend Protokoll definiert als mindestens zwei wässrige oder flüssige Stuhlgänge an einem Tag. c: Modell ohne Stratifikationsfaktoren. KI: Konfidenzintervall; N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; n.b.: nicht berechenbar; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; UE: Unerwünschte/s Ereignis/se Quellen: Amgen 2019, 2020a							

Diarrhoe ist ein bekanntes, häufig auftretendes UE bei Behandlung mit Phosphodiesterase 4-Inhibitoren, zu denen Apremilast zählt (Amgen 2019). Als a priori definiertes UE von besonderem Interesse trat sie bei 33 Patienten (31,7 %) in der Apremilast-Gruppe und 18 Patienten (17,5 %) in der Placebo-Gruppe auf. Lediglich einer der 18 Patienten in der Placebo-Gruppe (5,6 %) bezeichnete seine Diarrhoe als schwer; damit lag der Anteil der als nicht schwer bewerteten Diarrhoe bei 100 % (Apremilast-Gruppe) und 94,4 % (Placebo-Gruppe). Kein Diarrhoe-Ereignis wurde als schwerwiegend eingestuft (Tabelle 4-64).

Die übrigen a priori definierten UE von besonderem Interesse, Suizidgedanken und suizidales Verhalten, maligne Erkrankungen, schwere kardiale Komplikationen und opportunistische Infektionen, wurden in keiner der beiden Gruppen bis Woche 12 beobachtet (Amgen 2019).

Tabelle 4-65: Ergebnisse der UE bei mindestens 10 % der Patienten in einem Studienarm nach SOC und PT zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

SOC PT	Apremilast		Placebo		Apremilast vs. Placebo			
	N	n (%)	N	n (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a
Infektionen und parasitäre Erkrankungen								
Oberer Atemweg	104	12 (11,5)	103	5 (4,9)	2,38 [0,87; 6,51]	2,56 [0,87; 7,54]	0,07 [-0,01; 0,14]	0,0806
Erkrankungen des Nervensystems								
Kopfschmerz	104	15 (14,4)	103	11 (10,7)	1,35 [0,65; 2,80]	1,41 [0,61; 3,24]	0,04 [-0,05; 0,13]	0,4176
Erkrankungen des Gastrointestinaltrakts								
Diarrhoe	104	43 (41,3)	103	21 (20,4)	2,03 [1,30; 3,16]	2,75 [1,48; 5,11]	0,21 [0,09; 0,33]	0,0011
Übelkeit	104	20 (19,2)	103	11 (10,7)	1,80 [0,91; 3,57]	1,99 [0,90; 4,40]	0,09 [-0,01; 0,18]	0,0855
Patienten mit multiplen UE in einer SOC werden für diese SOC einmal gezählt. Jeder Patient wird für das zutreffende UE (PT) einmal gezählt. a: p-Wert des Cochran-Mantel-Haenszel-Tests. KI: Konfidenzintervall; N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; PT: Bevorzugter Begriff (preferred term); RD: Risikodifferenz; RR: Relatives Risiko; SOC: Systemorganklasse (System Organ Class); UE: Unerwünschte/s Ereignis/se Quellen: Amgen 2018a, 2019, 2020a								

UE, die bei mindestens 10 % der Patienten, unabhängig vom Schweregrad, bis Woche 12 in einer der beiden Behandlungsgruppen auftraten, umfassten obere Atemwegsinfektion, Kopfschmerz, Diarrhoe und Übelkeit. Das von 43 Patienten (41,3 %) in der Apremilast-Gruppe und 21 Patienten (20,4 %) in der Placebo-Gruppe am häufigsten berichtete UE war Diarrhoe (Tabelle 4-65). Die Abweichung der Zahlen bei der Diarrhoe nach SOC / PT und der Diarrhoe als UE von besonderem Interesse ergibt sich daraus, dass Letztere nur jene Ereignisse erfasst, bei der es zu mindestens zwei wässrigen oder flüssigen Stuhlgängen pro Tag kam. Es traten keine UE bei mindestens zehn Patienten UND mindestens 1 % der Patienten in einem Studienarm auf (Amgen 2019).

Tabelle 4-66: Ergebnisse der UE, die zum Therapieabbruch führten, nach SOC und PT zu Woche 12 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

SOC PT	Apremilast		Placebo	
	N	n (%)	N	n (%)
Erkrankungen des Nervensystems				
Kopfschmerz	104	1 (1,0)	103	1 (1,0)
Lethargie	104	0	103	1 (1,0)
Gefäßerkrankungen				
Behçet-Syndrom	104	1 (1,0)	103	0
Erkrankungen der Atemwege, des Brustraums und Mediastinums				
Husten	104	0	103	1 (1,0)
Erkrankungen des Gastrointestinaltrakts				
Oberbauchschmerzen	104	1 (1,0)	103	0
Übelkeit	104	1 (1,0)	103	0
Erbrechen	104	1 (1,0)	103	0
Diarrhoe	104	0	103	1 (1,0)
Orale Ulzera	104	0	103	1 (1,0)
Erkrankungen der Haut und des Unterhautzellgewebes				
Akute febrile neutrophile Dermatose	104	0	103	1 (1,0)
Pemphigus	104	0	103	1 (1,0)
Hautläsion	104	0	103	1 (1,0)
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen				
Muskuloskelettale Brustschmerzen	104	0	103	1 (1,0)
N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugter Begriff (preferred term); SOC: Systemorganklasse (System Organ Class); UE: Unerwünschte/s Ereignis/se Quellen: Amgen 2018a, 2020a				

Der Anteil der Patienten mit UE, die zum Therapieabbruch führten, fiel mit drei Patienten (2,9 %) in der Apremilast-Gruppe und fünf Patienten (4,9 %) in der Placebo-Gruppe insgesamt niedrig und vergleichbar aus. Kein UE, das zum Therapieabbruch führte, trat in den beiden Behandlungsgruppen bei mehr als einem Patienten auf (Tabelle 4-66).

Zusammenfassend wurde Diarrhoe als UE von besonderem Interesse und als UE nach SOC / PT von statistisch signifikant mehr Patienten in der Apremilast-Gruppe berichtet als in der Placebo-Gruppe. Die übrigen UE nach SOC / PT, die bei mehr als 10 % der Patienten auftraten, ließen keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsgruppen erkennen.

Extensionsphase (ab Woche 12 bis Woche 68)

Nachfolgend werden die Ergebnisse des Endpunkts „Sicherheit“ supportiv für die Gesamtrate der Sicherheit, die UE von besonderem Interesse, die UE bei mindestens 10 % der Patienten in einem Studienarm sowie die UE, die zum Studienabbruch führten, für die Extensionsphase von Woche 12 bis Woche 64 dargestellt. Für die Analysen der Sicherheitsendpunkte wurden alle mit Apremilast behandelten (Apremilast-as-treated, AAT) Patienten (AAT-Population) berücksichtigt.

Tabelle 4-67: Ergebnisse für Sicherheit von Woche 12 bis Woche 64 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (AAT-Population)

Ereignis Zeitpunkt	Apremilast		Placebo/Apremilast	
	N	n (%)	N	n (%)
UE gesamt				
Woche 64	104	77 (74,0)	83	70 (84,3)
SUE				
Woche 64	104	8 (7,7)	83	6 (7,2)
Therapieabbruch aufgrund von UE				
Woche 64	104	9 (8,6)	83	3 (3,6)
Schwere UE				
Woche 64	104	13 (12,5)	83	4 (4,8)
AAT: Mit Apremilast behandelt (Apremilast-as-treated); N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; SUE: schwerwiegende/s unerwünschte/s Ereignis/se; UE: Unerwünschte/s Ereignis/se Quellen: Amgen 2018a, 2020a				

Die Gesamtzahl an UE, SUE und zum Therapieabbruch führenden UE in der Apremilast-Gruppe fiel zwischen dem Ende der 12-wöchigen Placebo-kontrollierten Behandlungsphase und der Extensionsphase vergleichbar aus. Die Anzahl an Patienten mit schweren UE lag zu Woche 64 etwas höher als zu Woche 12, allerdings wurden nur zwei UE, Diarrhoe und Übelkeit, bei zwei Patienten (1,9 %) in der Apremilast-Gruppe als schwer bezeichnet. Alle weiteren schweren UE traten bei nur einem Patienten (1 %) auf (Amgen 2020a). Das Sicherheitsprofil in der Placebo/Apremilast-Gruppe unterschied sich in der Extensionsphase nicht nennenswert von der der Apremilast-Gruppe. Etwas geringere Raten wurden für UE, die zum Therapieabbruch führten und schwere UE registriert. Insgesamt wurde kein UE von mehr als einem Patienten in der Placebo/Apremilast-Gruppe als schwer eingestuft (Tabelle 4-67) (Amgen 2020a).

Tabelle 4-68: Ergebnisse der UE bei mindestens 10 % der Patienten in einem Studienarm nach SOC und PT von Woche 12 bis Woche 64 aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (AAT-Population)

SOC PT	Apremilast		Placebo/Apremilast	
	N	n (%)	N	n (%)
Erkrankungen des Nervensystems				
Kopfschmerz				
Woche 64	104	14 (13,5)	83	14 (16,9)
Erkrankungen des Gastrointestinaltrakts				
Diarrhoe				
Woche 64	104	26 (25,0)	83	23 (27,7)
Übelkeit				
Woche 64	104	9 (8,6)	83	11 (13,2)
<p>a: Diarrhoe als UE von besonderem Interesse ist entsprechend Protokoll definiert als mindestens zwei wässrige oder flüssige Stuhlgänge an einem Tag.</p> <p>AAT: Mit Apremilast behandelt (Apremilast-as-treated); N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugter Begriff (preferred term); SOC: Systemorganklasse (System Organ Class); UE: Unerwünschte/s Ereignis/se</p> <p>Quellen: Amgen 2018a, 2020a</p>				

Der während der Extensionsphase aufgetretene Kopfschmerz innerhalb der Apremilast-Gruppe trat in gleicher Ausprägung auf wie während der ersten 12 Wochen Behandlung. Die Rate an Diarrhoe- und Übelkeit-Ereignissen war in der Extensionsphase deutlich rückläufig mit 26 % bzw. 8,6 % gegenüber 41,3 % bzw. 19,2 % zu Woche 12. Übelkeit wurde mit 14,5 % von einem etwas höheren Teil der Patienten in der Placebo/Apremilast-Gruppe berichtet als in der Apremilast-Gruppe. Hinsichtlich der übrigen UE waren die Patientenzahlen zwischen den beiden Gruppen vergleichbar (Tabelle 4-68).

Tabelle 4-69: Ergebnisse der UE, die von Woche 12 bis Woche 64 zum Therapieabbruch führten, nach SOC und PT aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (AAT-Population)

SOC PT	Apremilast		Placebo/Apremilast	
	N	n (%)	N	n (%)
Erkrankungen des Nervensystems				
Spannungskopfschmerz	104	1 (1,0)	83	0
Gefäßerkrankungen				
Behçet-Syndrom	104	0	83	1 (1,2)
Arterienthrombose	104	1 (1,0)	83	0
Infektionen und parasitäre Erkrankungen				
Atypische mykobakterielle Infektion	104	1 (1,0)	83	0
Herpes zoster	104	1 (1,0)	83	0
Lymphknotentuberkulose	104	1 (1,0)	83	0
Neuritis vestibularis	104	0	83	1 (1,2)
Erkrankungen des Gastrointestinaltrakts				
Abdominalschmerz	104	0	83	1 (1,2)
Übelkeit	104	1 (1,0)	83	0
Pankreatitis akut	104	1 (1,0)	83	0
Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen)				
Brustkrebs	104	1 (1,0)	83	0
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort				
Thoraxschmerzen nicht kardialen Ursprungs	104	1 (1,0)	83	0
Augenerkrankungen				
Skleritis	104	1 (1,0)	83	0
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen				
Myositis	104	1 (1,0)	83	0
AAT: Mit Apremilast behandelt (Apremilast-as-treated); N: Anzahl der Patienten der Sicherheitspopulation; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugter Begriff (preferred term); SOC: Systemorganklasse (System Organ Class); UE: Unerwünschte/s Ereignis/se				
Quellen: Amgen 2018a, 2020a				

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt „Sicherheit“ auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

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.
- 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.

¹⁶ unbesetzt

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 4-70: Matrix der durchgeführten Subgruppenanalysen

Endpunkt Studie	Geschlecht	Alter	Orale Aphthen zu Baseline	Region	Ethnie	Gewicht zu Baseline	BMI zu Baseline	Alkohol-konsum	Tabak-konsum	Erkrankungs-dauer	Vormedikation	
											Colchicin	Orale Kortikosteroide
Anteil der Patienten mit vollständiger Remission der oralen Aphthen (komplettes Ansprechen)												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Zeit bis zur vollständigen Remission der oralen Aphthen												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Anzahl der oralen Aphthen												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Dauer der vollständigen Remission der oralen Aphthen												
RELIEF	○	○	n.d.	○	○	○	○	○	○	○	○	○
Schmerzhaftigkeit der oralen Aphthen mittels VAS												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Veränderung des BSAS zu Woche 12												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Veränderung des BDCAF zu Woche 12												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Veränderung des BD-QoL zu Woche 12												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12												
RELIEF	○	○	n.d.	○	○	○	○	○	○	○	○	○
Sicherheit (Gesamtrate UE)												
RELIEF	○	○	○	○	○	○	○	○	○	○	○	○

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

○: Post-hoc durchgeführte Subgruppenanalyse

BD: Behçet-Syndrom (Behçet’s disease); BDCAF: BD Current Activity Form; BD-QoL: BD Quality of Life; BMI: Body Mass Index; BSAS: Behçet’s Syndrome Activity Score; n.d.: Nicht durchgeführt; SF-36: 36-Item Short Form Health Survey, Version 2; UE: Unerwünschte/s Ereignis/se; VAS: Visuelle Analogskala

Stellen Sie anschließend in Tabelle 4-71 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ($p < 0,05$) Interaktionsterme.

Tabelle 4-71: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für RELIEF und Subgruppenmerkmale

Endpunkt Studie	Geschlecht	Alter	Orale Aphthen zu Baseline	Region	Ethnie	Gewicht zu Baseline	BMI zu Baseline	Alkoholkonsum	Tabakkonsum	Erkrankungsdauer	Vormedikation	
											Colchicin	Orale Kortikosteroide
Anteil der Patienten mit vollständiger Remission der oralen Aphthen (komplettes Ansprechen)												
RELIEF	p=0,9229	p=0,0549	p=0,4927	p=0,1864	p=0,8530	p=0,7615	p=0,2752	p=0,8137	p=0,6136	p=0,3744	p=0,0999	p=0,3155
Zeit bis zur vollständigen Remission der oralen Aphthen												
RELIEF	p=0,9171	p=0,9445	p=0,4186	p=0,2161	p=0,1869	p=0,4625	p=0,6658	p=0,6573	p=0,5227	p=0,0529	p=0,6828	p=0,2486
Anzahl der oralen Aphthen												
RELIEF	p=0,8461	p=0,0811	p=0,0132	p=0,4169	p=0,3407	p=0,6576	p=0,7631	p=0,4778	p=0,2291	p=0,1495	p=0,1721	p=0,4207
Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission												
RELIEF	p=0,6216	p=0,7382	k.A.	p=0,7117	p=0,9452	p=0,8667	p=0,1931	p=0,0284	p=0,6085	p=0,0782	p=0,8401	p=0,1271
Schmerzhaftigkeit der oralen Aphthen mittels VAS (LOCF)												
RELIEF	p=0,8168	p=0,4386	p=0,9654	p=0,2723	p=0,4041	p=0,6638	p=0,4997	p=0,1641	p=0,2664	p=0,8084	p=0,7702	p=0,7528
Schmerzhaftigkeit der oralen Aphthen mittels VAS (MMRM)												
RELIEF	p=0,9565	p=n.b.	p=0,1147	p=0,3876	p=0,9778	p=0,9676	p=0,5686	p=0,0058	p=0,0780	p=0,1692	p=0,0254	p=0,0049
Veränderung des BSAS zu Woche 12												
RELIEF	p=0,7988	p=0,6249	p=0,9208	p=0,7250	p=0,6958	p=0,5135	p=0,4691	p=0,0558	p=0,0635	p=0,4621	p=0,2445	p=0,2654

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

Endpunkt Studie	Geschlecht	Alter	Orale Aphthen zu Baseline	Region	Ethnie	Gewicht zu Baseline	BMI zu Baseline	Alkoholkonsum	Tabakkonsum	Erkrankungsdauer	Vormedikation	
											Colchicin	Orale Kortikosteroide
Veränderung des BDCAF zu Woche 12 - Indexwert												
RELIEF	p=0,6214	p=0,3614	p=0,9222	p=0,1334	p=0,8876	p=0,1483	p=0,4689	p=0,6146	p=0,3034	p=0,0943	p=0,0289	p=0,1817
Veränderung des BDCAF zu Woche 12 - Patientenwahrnehmung												
RELIEF	p=0,4300	p=0,1106	p=0,9732	p=0,6616	p=0,5536	p=0,0643	p=0,0349	p=0,9637	p=0,2580	p=0,0406	p=0,0955	p=0,5803
Veränderung des BDCAF zu Woche 12 - Untersucherwahrnehmung												
RELIEF	p=0,1480	p=0,0279	p=0,7217	p=0,4509	p=0,4724	p=0,3895	p=0,1250	p=0,7292	p=0,3865	p=0,0434	p=0,0686	p=0,5699
Veränderung des BD-QoL zu Woche 12												
RELIEF	p=0,1143	p=0,5708	p=0,3834	p=0,2470	p=0,4106	p=0,6110	p=0,5749	p=0,1258	p=0,1135	p=0,7903	p=0,2662	p=0,2270
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12 (Körperliche Funktionsfähigkeit)												
RELIEF	p=0,8813	p=n.b.	p=0,5115	p=0,2923	p=0,4242	p=0,1252	p=0,0459	p=0,1442	p=0,2279	p=0,8075	p=0,1083	p=0,2448
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12 (Körperliche Rollenfunktion)												
RELIEF	p=0,9554	p=n.b.	p=n.b.	p=0,4582	p=0,5319	p=0,8669	p=0,4108	p=0,2708	p=0,0212	p=0,8632	p=0,0389	p=7985
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12 (Körperliche Schmerzen)												
RELIEF	p=0,8890	p=n.b.	p=0,2051	p=0,6951	p=0,7369	p=0,9344	p=0,3116	p=0,7408	p=0,2834	p=0,5726	p=0,3559	p=0,3234
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12 (Allgemeine Gesundheitswahrnehmung)												
RELIEF	p=0,9610	p=n.b.	p=n.b.	p=0,9453	p=0,7982	p=0,7728	p=n.b.	p=0,1187	p=0,7460	p=0,2305	p=0,6685	p=0,9657
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12 (Vitalität)												
RELIEF	p=0,940	p=0,6568	p=0,9424	p=0,6501	p=0,7298	p=0,4104	p=0,2355	p=0,7396	p=0,0209	p=0,4726	p=0,0176	p=0,2298
Verbesserung des SF-36 um ≥ 2,5 Punkte zu Woche 12 (Soziale Funktionsfähigkeit)												
RELIEF	p=0,9694	p=n.b.	p=0,2482	p=0,6945	p=0,8497	p=0,3863	p=0,8957	p=0,5143	p=0,4362	p=0,3588	p=0,1799	p=0,2764

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

Endpunkt Studie	Geschlecht	Alter	Orale Aphthen zu Baseline	Region	Ethnie	Gewicht zu Baseline	BMI zu Baseline	Alkoholkonsum	Tabakkonsum	Erkrankungsdauer	Vormedikation	
											Colchicin	Orale Kortikosteroide
Verbesserung des SF-36 um \geq 2,5 Punkte zu Woche 12 (Emotionale Rollenfunktion)												
RELIEF	p=0,6921	p=0,2036	p=0,9209	p=0,9300	p=0,5140	p=0,9235	p=0,9301	p=0,3549	p=0,1901	p=0,7132	p=0,1409	p=0,8227
Verbesserung des SF-36 um \geq 2,5 Punkte zu Woche 12 (Psychisches Wohlbefinden)												
RELIEF	p=0,3318	p=0,7883	p=0,4981	p=0,5278	p=0,2007	p=0,8751	p=0,4506	p=0,9624	p=0,2906	p=0,3839	p=0,1585	p=0,0992
Verbesserung des SF-36 um \geq 2,5 Punkte zu Woche 12 (PCS)												
RELIEF	p=0,3374	p=n.b.	p=0,2220	p=0,5101	p=0,8265	p=0,9478	p=0,1581	p=0,0473	p=0,6463	p=0,7395	p=0,2669	p=0,7973
Verbesserung des SF-36 um \geq 2,5 Punkte zu Woche 12 (MCS)												
RELIEF	p=0,5211	p=n.b.	p=0,4489	p=0,5807	p=0,1464	p=0,8940	p=0,7363	p=0,7360	p=0,6125	p=0,9471	p=0,0031	p=0,7349
Sicherheit (Gesamtrate UE)												
RELIEF	p=0,8607	p=0,5418	p=n.b.	p=0,9660	p=0,9120	p=0,4685	p=n.b.	p=0,5498	p=0,7053	p=0,4849	p=0,1138	p=0,4341
Sicherheit (SUE)^a												
RELIEF	p=0,2525	p=n.b.	p=n.b.	p=0,9826	p=0,5450	p=0,9451	p=0,6200	p=0,1837	p=0,3775	p=n.b.	p=0,5543	p=0,1779
Sicherheit (schwere UE)^a												
RELIEF	p=0,2185	p=n.b.	p=0,1959	p=n.b.	p=n.b.	p=n.b.	p=n.b.	p=0,5728	p=0,6835	p=0,2331	p=0,8483	p=0,2403
Sicherheit (UE, die zum Therapieabbruch führten)^a												
RELIEF	p=0,6648	p=n.b.	p=0,5328	p=0,1822	p=0,1929	p=n.b.	p=n.b.	p=0,1080	p=0,2313	p=n.b.	p=0,7361	p=1,0000
<p>a: Weniger als zehn Ereignisse in einer der Subgruppen aufgetreten.</p> <p>BD: Behçet-Syndrom (Behçet's disease); BDCAF: BD Current Activity Form; BD-QoL: BD Quality of Life; BMI: Body Mass Index; BSAS: Behçet's Syndrome Activity Score; k.A.: Keine Angabe; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); MCS: Psychische Summenskala (mental component summary); MMRM: Gemischtes Modell mit Messwiederholungen (mixed-effects model for repeated measures); n.b.: Nicht berechenbar; PCS: Körperliche Summenskala (physical component summary); SF-36: 36-Item Short Form Health Survey, Version 2; SUE: Schwerwiegend/s unerwünschte/s Ereignis/se; UE: Unerwünschte/s Ereignis/se; VAS: Visuelle Analogskala</p> <p>Quelle: Amgen 2020a</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.

4.3.1.3.2.1 Subgruppenanalysen für den Endpunkt Veränderung der Anzahl oraler Aphthen

Tabelle 4-72: Subgruppenanalysen für Endpunkt Veränderung der Anzahl oraler Aphthen zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktionstests
Alle Patienten	104	4,17 (3,65)	-2,74 (0,23)	103	3,93 (2,70)	-1,50 (0,23)	-1,24 [-1,79; -0,69] <0,0001	-0,61 [-0,89; -0,33]	Nicht zutreffend
Anzahl oraler Aphthen zu Baseline									
≥ 2 bis ≤ 5	91	3,30 (0,90)	-2,26 (0,18)	89	3,10 (0,90)	-1,28 (0,18)	-0,98 [-1,48; -0,48] 0,0002	-0,57 [-0,87; 0,28]	0,0132
> 5 bis ≤ 10	10	6,80 (0,80)	-4,66 (0,93)	11	7,30 (1,30)	-2,40 (0,89)	-2,26 [-4,99; 0,48] 0,1006	-0,73 [-1,61; 0,16]	
> 10	3	22,30 (8,70)	-13,63 (3,73)	3	16,30 (2,50)	-9,70 (3,73)	-3,92 [-21,83; 13,98] 0,5357	-0,46 [-2,08; 1,17]	
<p>a: p-Wert der ANCOVA. ANCOVA: Kovarianzanalyse (analysis of covariance); KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz Quelle: Amgen 2020a</p>									

Für den Endpunkt „Veränderung der Anzahl oraler Aphthen zu Woche 12“ lag bei der Subgruppenanalyse für das Merkmal Anzahl oraler Aphthen zu Baseline ein signifikantes Ergebnis hinsichtlich des Interaktionstests vor ($p < 0,05$). Die Effektschätzer der Kategorien waren, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. In der Kategorie „Anzahl oraler Aphthen von ≥ 2 bis ≤ 5 “ zeigten sich, analog zur Gesamtpopulation, statistisch signifikante Unterschiede, die Ergebnisse der Kategorien „Anzahl oraler Aphthen > 5 bis ≤ 10 “ und „Anzahl oraler Aphthen > 10 “ fielen dagegen statistisch nicht signifikant aus. In diesen beiden Kategorien lagen relativ niedrige Patientenzahlen vor mit zehn bzw. 11 Patienten in der Kategorie „Anzahl oraler Aphthen > 5 bis ≤ 10 “ und weniger als zehn Patienten in der Kategorie „Anzahl oraler Aphthen > 10 “ (Tabelle 4-72). Aufgrund der sich daraus ergebenden statistischen Unschärfe sind die Ergebnisse qualitativ nicht interpretierbar. Der Subgruppeneffekt wird somit als zufällig erachtet.

4.3.1.3.2.2 Subgruppenanalysen für den Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission

Tabelle 4-73: Subgruppenanalysen für Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Apremilast		Placebo		Apremilast vs. Placebo	
	n/N (%)	Median ^a (Wochen) [95 %-KI]	n/N (%)	Median ^a (Wochen) [95 %-KI]	HR [95 %-KI] ^b p-Wert ^c	p-Wert des Interaktionstests
Alle Patienten	57/83 (68,67)	4,57 [3,14; 6,14]	43/53 (81,13)	2,29 [2,14; 4,14]	0,61 [0,41; 0,92] 0,0159	Nicht zutreffend
Alkoholkonsum						
Ja	23/28 (82,1)	3,21 [2,14; 5,14]	11/15 (73,3)	4,29 [1,86; 6,14]	1,17 [0,57; 2,42] 0,6377	0,0284
Nein	34/55 (61,8)	5,14 [3,14; 10,14]	32/38 (84,2)	2,14 [2,14; 4,00]	0,42 [0,25; 0,69] 0,0005	
a: Der Median wurde auf Basis der Kaplan-Meier-Methode ermittelt. b: Das HR mit dem dazugehörigen 95%-KI basiert auf dem stratifizierten Cox-Modell mit Geschlecht und Region als Stratifizierungsfaktoren. c: Der p-Wert basiert auf dem stratifizierten Log-rank-Test mit Geschlecht und Region als Stratifizierungsfaktoren. HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis Quelle: Amgen 2020a						

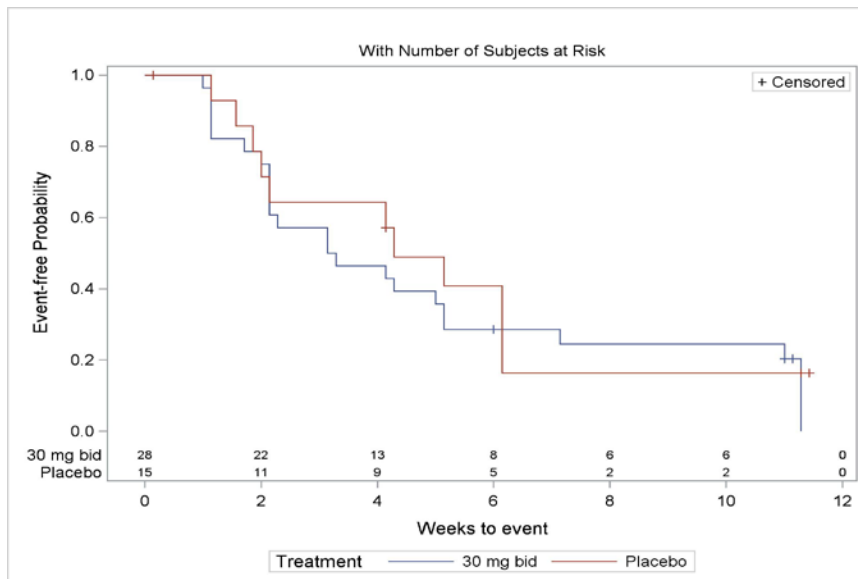


Abbildung 12: Kaplan-Meier-Kurven für den Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission, Subgruppe Alkoholkonsum: Ja, aus Studie RELIEF mit dem zu bewertenden Arzneimittel
Quelle: Amgen 2020a

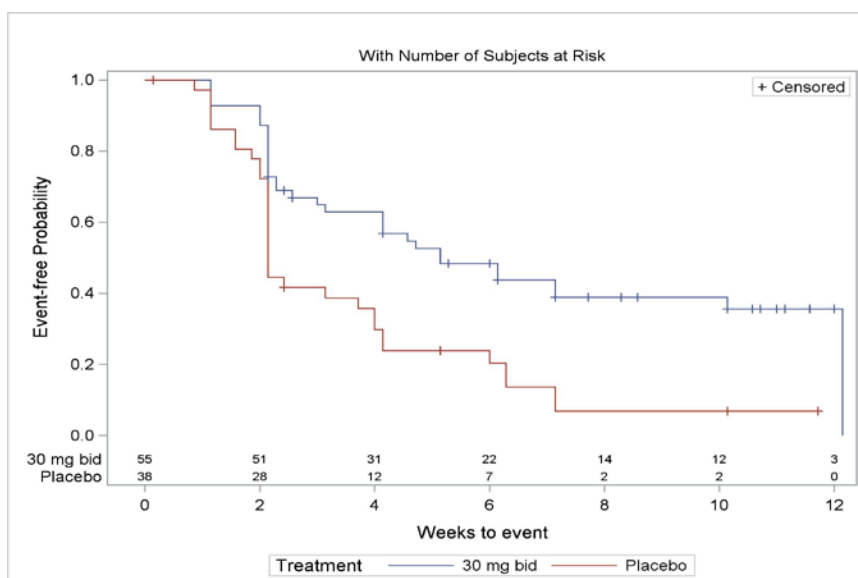


Abbildung 13: Kaplan-Meier-Kurven für den Endpunkt Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission, Subgruppe Alkoholkonsum: Nein, aus Studie RELIEF mit dem zu bewertenden Arzneimittel
Quelle: Amgen 2020a

Für den Endpunkt „Zeit bis zum Rezidiv der oralen Aphthen nach vollständiger Remission“ lag bei der Subgruppenanalyse für das Merkmal Alkoholkonsum ein signifikantes Ergebnis hinsichtlich des Interaktionstests vor ($p < 0,05$). Die Effektschätzer der Kategorie „nein“ waren,

analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. In der Kategorie „ja“ war der Effekt entgegengerichtet, statistisch aber nicht signifikant (Tabelle 4-73, Abbildung 12, Abbildung 13).

Die statistische Signifikanz bezüglich dieser Ergebnisse sollte mit Vorsicht bewertet werden. Aufgrund der post-hoc durchgeführten explorativen Analyse und der Größe der Subgruppen kann davon ausgegangen werden, dass keine Modifikation vorliegt.

4.3.1.3.2.3 Subgruppenanalysen für den Endpunkt Schmerzhaftigkeit oraler Aphthen mittels VAS

Tabelle 4-74: Subgruppenanalysen für Endpunkt Schmerzhaftigkeit oraler Aphthen mittels VAS (MMRM) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppen	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktions-tests
Alle Patienten	102	61,02 (27,63)	-39,04 (1,46)	101	61,18 (26,77)	-18,77 (1,47)	-20,27 [-23,19; -17,35] <0,0001	-1,90 [-2,24; -1,57]	Nicht zutreffend
Alkoholkonsum									
Ja	35	54,70 (29,30)	-32,89 (1,88)	30	57,80 (25,90)	-18,78 (2,07)	-14,11 [-19,23; -8,98] <0,0001	-1,33 [-1,87; -0,79]	0,0058
Nein	67	64,30 (26,30)	-41,63 (1,52)	71	62,60 (27,20)	-18,59 (1,51)	-23,04 [-26,60; -19,48] <0,0001	-2,15 [-2,57; -1,73]	
Vorherige Colchicin-Einnahme									
Ja	50	62,60 (23,60)	-36,05 (1,48)	56	60,70 (27,40)	-18,92 (1,46)	-17,13 [-21,21; -13,05] <0,0001	-1,59 [-2,03; -1,15]	0,0254
Nein	52	59,50 (31,20)	-41,51 (1,69)	45	61,80 (26,20)	-17,61 (1,79)	-23,90 [-28,07; -19,72] <0,0001	-2,27 [-2,78; -1,76]	

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

Subgruppen	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktions-tests
Vorherige Einnahme oraler Kortikosteroide									
Ja	15	67,10 (22,30)	-51,22 (2,52)	15	69,70 (22,10)	-21,14 (2,52)	-30,08 [-36,74; -23,41] <0,0001	-3,16 [-4,24; -2,09]	0,0049
Nein	87	60,00 (28,40)	-36,47 (1,26)	86	59,70 (27,30)	-18,00 (1,30)	-18,46 [-21,70; -15,22] <0,0001	-1,69 [-2,04; -1,35]	
<p>a: p-Wert des MMRM. ITT: Intention-to-treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MMRM: Gemischtes Modell mit Messwiederholungen (mixed-effects model for repeated measures); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz Quelle: Amgen 2020a</p>									

Für den Endpunkt „Schmerzhaftigkeit oraler Aphthen mittels VAS“ (MMRM) ergaben die Subgruppenanalysen für folgende Merkmale mit einem p-Wert des Interaktionstests von $< 0,05$ ein signifikantes Ergebnis:

- Alkoholkonsum
- Vorherige Colchicin-Einnahme
- Vorherige Einnahme oraler Kortikosteroide

Die Effektschätzer waren unabhängig von der Ausprägung dieser drei Subgruppenmerkmale, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast (p-Wert $< 0,0001$, unabhängig von Ausprägung und Merkmal) (Tabelle 4-74).

Die statistische Signifikanz bezüglich dieser Ergebnisse sollte mit Vorsicht bewertet werden. Aufgrund der post-hoc durchgeführten explorativen Analyse und der Größe der Subgruppen kann davon ausgegangen werden, dass keine Modifikation vorliegt.

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

4.3.1.3.2.4 Subgruppenanalysen für den Endpunkt Veränderung des BDCAF zu Woche 12

Tabelle 4-75: Subgruppenanalysen für Endpunkt Veränderung des BDCAF zu Woche 12 (LOCF) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppen	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktions-tests
Indexwert									
Alle Patienten	104	3,70 (1,58)	-0,92 (0,20)	102	3,63 (1,67)	-0,41 (0,20)	-0,52 [-0,99; -0,04] 0,0335	-0,30 [-0,57; -0,02]	Nicht zutreffend
Vorherige Colchicin-Einnahme									
Ja	52	3,60 (1,70)	-1,23 (0,23)	56	3,40 (1,60)	-0,21 (0,22)	-1,02 [-1,64; -0,40] 0,0015	-0,62 [-1,01; -0,24]	0,0289
Nein	52	3,80 (1,50)	-0,93 (0,25)	46	4,00 (1,70)	-0,97 (0,27)	0,04 [-0,70; 0,77] 0,9167	0,02 [-0,38; 0,42]	
Patientenwahrnehmung									
Alle Patienten	104	4,65 (1,39)	-1,68 (0,18)	102	4,78 (1,33)	-0,67 (0,18)	-1,01 [-1,44; -0,58] <0,0001	-0,64 [-0,92; -0,36]	Nicht zutreffend

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

Subgruppen	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktions-tests
BMI zu Baseline (kg/m²)									
< 18,5	7	3,70 (2,10)	0,55 (0,43)	3	4,00 (1,70)	-0,94 (0,66)	1,49 [-0,39; 3,37] 0,1025	1,17 [-0,28; 2,62]	0,0349
≥ 18,5 bis < 25	60	4,80 (1,30)	-2,07 (0,20)	50	4,50 (1,30)	-0,70 (0,22)	-1,37 [-1,96; -0,78] <0,0001	-0,87 [-1,27; -0,48]	
≥ 25 bis < 30	29	4,60 (1,50)	-1,36 (0,27)	29	5,10 (1,30)	-0,33 (0,27)	-1,03 [-1,80; -0,26] 0,0096	-0,70 [-1,23; -0,17]	
≥ 30	8	4,80 (1,30)	-1,66 (0,53)	20	5,10 (1,30)	-1,43 (0,34)	-0,23 [-1,53; 1,07] 0,7187	-0,15 [-0,97; 0,67]	
Erkrankungsdauer (Jahre)									
≤ 2	32	4,70 (1,40)	-1,15 (0,29)	33	4,50 (1,50)	-0,52 (0,28)	-0,63 [-1,43; 0,18] 0,1242	-0,38 [-0,87; 0,11]	0,0406
> 2 bis ≤ 6	30	4,80 (1,30)	-2,28 (0,29)	25	5,00 (1,20)	-0,83 (0,32)	-1,45 [-2,32; -0,57] 0,0016	-0,89 [-1,44; -0,33]	
> 6 bis ≤ 10	15	4,30 (1,50)	-2,19 (0,37)	16	5,00 (1,30)	-0,13 (0,36)	-2,06 [-3,13; -0,99] 0,0005	-1,38 [-2,16; -0,60]	
> 10	26	4,70 (1,50)	-1,43 (0,29)	27	4,80 (1,30)	-1,14 (0,28)	-0,29 [-1,10; 0,52] 0,4765	-0,19 [-0,73; 0,35]	

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

Subgruppen	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktions-tests
Untersucherwahrnehmung									
Alle Patienten	104	4,47 (1,21)	-1,64 (0,17)	102	4,45 (1,12)	-0,69 (0,17)	-0,95 [-1,35; -0,55] <0,0001	-0,65 [-0,93; -0,37]	Nicht zutreffend
Alter (Jahre)									
< 40	58	4,40 (1,30)	-1,83 (0,20)	44	4,40 (1,20)	-0,66 (0,23)	-1,17 [-1,78; -0,57] 0,0002	-0,76 [-1,17; -0,36]	0,0279
≥ 40 bis < 65	43	4,60 (1,00)	-1,81 (0,21)	54	4,60 (1,10)	-0,91 (0,18)	-0,90 [-1,46; -0,35] 0,0015	-0,66 [-1,07; -0,25]	
≥ 65	3	5,00 (1,00)	0,60 (0,86)	4	3,80 (1,00)	-0,95 (0,72)	1,54 [-1,91; 4,99] 0,2820	0,80 [-0,76; 2,35]	
Erkrankungsdauer (Jahre)									
≤ 2	32	4,50 (1,30)	-1,32 (0,28)	33	4,40 (1,10)	-0,66 (0,28)	-0,66 [-1,45; 0,12] 0,0967	-0,41 [-0,90; 0,08]	0,0434
> 2 bis ≤ 6	30	4,40 (1,00)	-2,26 (0,27)	25	4,60 (1,10)	-0,77 (0,29)	-1,48 [-2,28; -0,68] 0,0005	-0,99 [-1,56; -0,43]	
> 6 bis ≤ 10	15	4,50 (1,10)	-2,12 (0,32)	16	4,70 (1,40)	-0,45 (0,31)	-1,66 [-2,57; -0,75] 0,0008	-1,31 [-2,08; -0,53]	
> 10	26	4,50 (1,40)	-1,40 (0,26)	27	4,20 (1,00)	-1,28 (0,25)	-0,12 [-0,85; 0,62] 0,7499	-0,09 [-0,63; 0,45]	

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

Subgruppen	Apremilast			Placebo			Apremilast vs. Placebo		
	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	N	MW (SD) zu Baseline	LS-MW (SE) Veränderung zu Woche 12	LS-MWD [95 %-KI] p-Wert ^a	SMD [95 %-KI]	p-Wert des Interaktions-tests
<p>a: p-Wert der ANCOVA. ANCOVA: Kovarianzanalyse (analysis of covariance); BMI: Body Mass Index; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (least square-mean value difference); MW: Mittelwert; N: Anzahl der Patienten; SD: Standardabweichung (standard deviation); SE: Standardfehler (standard error); SMD: Standardisierte Mittelwertdifferenz Quelle: Amgen 2020a</p>									

Für den Endpunkt „Veränderung des BDCAF zu Woche 12“ ergaben die Subgruppenanalysen für folgende Merkmale mit einem p-Wert des Interaktionstests von $< 0,05$ ein signifikantes Ergebnis:

- Indexwert
 - Vorherige Colchicin-Einnahme
- Patientenwahrnehmung
 - BMI zu Baseline
 - Erkrankungsdauer
- Untersucherwahrnehmung
 - Alter
 - Erkrankungsdauer

Bezüglich des Subgruppenmerkmals „Indexwert – vorherige Colchicin-Einnahme“ waren die Effekte bei den Patienten mit vorheriger Colchicin-Einnahme, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. Bei den Patienten ohne vorherige Colchicin-Einnahme war kein statistischer Effekt nachweisbar (Tabelle 4-75).

Für das Subgruppenmerkmal „Patientenwahrnehmung – BMI zu Baseline“ zeigten sich in den Kategorien „ $\geq 18,5$ bis < 25 “ und „ ≥ 25 bis < 30 “ und „ ≥ 30 “, analog zur Gesamtpopulation, gleichgerichtete Effekte zugunsten von Apremilast. In den Kategorien „ $< 18,5$ “ und „ ≥ 30 “, war kein statistischer Effekt nachweisbar (Tabelle 4-75).

Das Subgruppenmerkmal „Patientenwahrnehmung – Erkrankungsdauer“ zeigte sich in den Kategorien „ > 2 bis ≤ 6 “ und „ > 6 bis ≤ 10 “, analog zur Gesamtpopulation, gleichgerichtete Effektschätzer zugunsten von Apremilast. In den Kategorien „ ≤ 2 “ und „ > 10 “ war kein statistischer Effekt nachweisbar (Tabelle 4-75).

Für das Subgruppenmerkmal „Untersucherwahrnehmung -Alter“ lagen die Effektschätzer in den Kategorien „ < 40 “ und „ ≥ 40 bis < 65 “, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. In der Kategorie „ ≥ 65 “ war kein statistischer Effekt nachweisbar (Tabelle 4-75).

Das Subgruppenmerkmal „Untersucherwahrnehmung – Erkrankungsdauer“ zeigte sich in den Kategorien „ > 2 bis ≤ 6 “ und „ > 6 bis ≤ 10 “ für alle vier Kategorien, analog zur Gesamtpopulation, gleichgerichtete Effektschätzer zugunsten von Apremilast. In den Kategorien „ ≤ 2 “ und „ > 10 “ war kein statistischer Effekt nachweisbar (Tabelle 4-75).

Die statistische Signifikanz bezüglich dieser Ergebnisse sollte mit Vorsicht bewertet werden. Aufgrund der post-hoc durchgeführten explorativen Analyse und der Größe der Subgruppen kann davon ausgegangen werden, dass keine Modifikation vorliegt.

4.3.1.3.2.5 Subgruppenanalysen für den Endpunkt Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12

Tabelle 4-76: Subgruppenanalysen für Endpunkt Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12 (NRI) aus RCT (RELIEF) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppen	Apremilast	Placebo	Apremilast vs. Placebo				p-Wert ^a	p-Wert des Interaktions-tests
	n/N (%)	n/N (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]			
Körperliche Funktionsfähigkeit								
Alle Patienten	38/104 (36,5)	26/103 (25,2)	1,42 [0,94; 2,14]	1,72 [0,94; 3,16]	0,11 [-0,02; 0,23]	0,0804	Nicht zutreffend	
BMI zu Baseline (kg/m²)								
< 18,5	1/7 (14,3)	0/3 (0,0)	1,50 [0,08; 29,15]	1,62 [0,05; 51,11]	n.b.	0,5127	0,0459	
$\geq 18,5$ bis < 25	23/60 (38,3)	9/51 (17,6)	2,17 [1,11; 4,26]	2,90 [1,19; 7,05]	0,21 [0,05; 0,37]	0,0170		
≥ 25 bis < 30	12/29 (41,4)	6/29 (20,7)	2,00 [0,87; 4,60]	2,71 [0,85; 8,66]	0,21 [-0,03; 0,44]	0,0914		
≥ 30	2/8 (25,0)	11/20 (55,0)	0,45 [0,13; 1,61]	0,27 [0,04; 1,69]	-0,30 [-0,67; 0,07]	0,1579		
Körperliche Rollenfunktion								
Alle Patienten	45/104 (43,3)	25/103 (24,3)	1,78 [1,19; 2,67]	2,40 [1,32; 4,36]	0,19 [0,06; 0,31]	0,0042	Nicht zutreffend	
Tabakkonsum								
Aktiver Raucher	9/20 (45,0)	1/12 (8,3)	5,40 [0,78; 37,50]	9,00 [0,97; 83,58]	0,37 [0,10; 0,63]	0,0330	0,0212	
Ehemaliger Raucher	4/16 (25,0)	9/22 (40,9)	0,61 [0,23; 1,64]	0,48 [0,12; 1,98]	-0,16 [-0,45; 0,14]	0,3139		
Nichtraucher	32/68 (47,1)	15/69 (21,7)	2,16 [1,29; 3,62]	3,20 [1,52; 6,74]	0,25 [0,10; 0,41]	0,0019		

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

Subgruppen	Apremilast	Placebo	Apremilast vs. Placebo				
	n/N (%)	n/N (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a	p-Wert des Interaktions-tests
Vorherige Colchicin-Einnahme							
Ja	27/52 (51,9)	11/57 (19,3)	2,69 [1,49; 4,86]	4,52 [1,92; 10,61]	0,33 [0,16; 0,50]	0,0004	0,0389
Nein	18/52 (34,6)	14/46 (30,4)	1,14 [0,64; 2,02]	1,21 [0,52; 2,83]	0,04 [-0,14; 0,23]	0,6613	
Vitalität							
Alle Patienten	63/104 (60,6)	43/103 (41,8)	1,45 [1,11; 1,90]	2,18 [1,24; 3,85]	0,18 [0,05; 0,31]	0,0069	Nicht zutreffend
Tabakkonsum							
Aktiver Raucher	13/20 (65,0)	1/12 (8,3)	7,80 [1,16; 52,35]	20,43 [2,17; 192,64]	0,57 [0,31; 0,83]	0,0021	0,0209
Ehemaliger Raucher	9/16 (56,2)	12/22 (54,5)	1,03 [0,58; 1,84]	1,07 [0,29; 3,92]	0,02 [-0,30; 0,34]	0,9180	
Nichtraucher	41/68 (60,3)	30/69 (43,5)	1,39 [1,00; 1,93]	1,97 [1,00; 3,90]	0,17 [0,00; 0,33]	0,0497	
Vorherige Colchicin-Einnahme							
Ja	32/52 (61,5)	17/57 (29,8)	2,06 [1,31; 3,24]	3,76 [1,70; 8,35]	0,32 [0,14; 0,49]	0,0009	0,0176
Nein	31/52 (59,6)	26/46 (56,5)	1,05 [0,75; 1,48]	1,14 [0,51; 2,54]	0,03 [-0,16; 0,23]	0,7579	
PCS							
Alle Patienten	50/104 (48,1)	36/103 (35,0)	1,38 [0,99; 1,91]	1,73 [0,98; 3,04]	0,12 [-0,01; 0,26]	0,0606	Nicht zutreffend
Alkoholkonsum							
Ja	12/35 (34,3)	13/31 (41,9)	0,82 [0,44; 1,52]	0,72 [0,27; 1,96]	-0,08 [-0,31; 0,16]	0,5257	0,0473
Nein	38/69 (55,1)	23/72 (31,9)	1,72 [1,16; 2,57]	2,61 [1,32; 5,19]	0,23 [0,07; 0,39]	0,0058	

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

Subgruppen	Apremilast	Placebo	Apremilast vs. Placebo				
	n/N (%)	n/N (%)	RR [95 %-KI]	OR [95 %-KI]	RD [95 %-KI]	p-Wert ^a	p-Wert des Interaktions-tests
MCS							
Alle Patienten	57/104 (54,8)	34/103 (33,0)	1,72 [1,26; 2,36]	2,52 [1,42; 4,47]	0,21 [0,07; 0,34]	0,0016	Nicht zutreffend
Vorherige Colchicin-Einnahme							
Ja	31/52 (59,6)	12/57 (21,1)	2,83 [1,63; 4,91]	5,54 [2,38; 12,87]	0,39 [0,22; 0,56]	<0,0001	0,0031
Nein	26/52 (50,0)	22/46 (47,8)	1,05 [0,70; 1,57]	1,09 [0,49; 2,41]	0,02 [-0,18; 0,22]	0,8307	
<p>a: p-Wert des Cochran-Mantel-Haenszel-Tests. BMI: Body Mass Index; ITT: Intention-to-treat; KI: Konfidenzintervall; MCS: Psychische Summenskala (mental component summary); N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; NRI: Non-Responder-Imputation; OR: Odds Ratio; PCS: Körperliche Summenskala (physical component summary); RD: Risikodifferenz; RR: Relatives Risiko Quelle: Amgen 2020a</p>							

Für den Endpunkt „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ ergaben die Subgruppenanalysen für folgende Merkmale mit einem p-Wert des Interaktionstests von $< 0,05$ ein signifikantes Ergebnis:

- Körperliche Funktionsfähigkeit
 - BMI zu Baseline
- Körperliche Rollenfunktion
 - Tabakkonsum
 - Vorherige Colchicin-Einnahme
- Vitalität
 - Tabakkonsum
 - Vorherige Colchicin-Einnahme
- PCS
 - Alkoholkonsum
- MCS
 - Vorherige Colchicin-Einnahme

Das Subgruppenmerkmal „Körperliche Funktionsfähigkeit – BMI zu Baseline“ zeigte für die Kategorien „ $< 18,5$ “, „ $\geq 18,5$ bis < 25 “ und „ ≥ 25 bis < 30 “, analog zur Gesamtpopulation, gleichgerichtete Effektschätzer zugunsten von Apremilast. In der Kategorie „ ≥ 30 “ lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Für das Subgruppenmerkmal „Körperliche Rollenfunktion – Tabakkonsum“ waren die Effektschätzer für die Kategorien „aktiver Raucher“ und „Nichtraucher“, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. In der Kategorie „ehemaliger Raucher“ lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Das Subgruppenmerkmal „Körperliche Rollenfunktion – vorherige Colchicin-Einnahme“ ließ für die Kategorie „ja“, analog zur Gesamtpopulation, gleichgerichtete Effektschätzer zugunsten von Apremilast erkennen. In der Kategorie „nein“ lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Für das Subgruppenmerkmal „Vitalität – Tabakkonsum“ lagen für die Kategorien „aktiver Raucher“ und „Nichtraucher“, analog zur Gesamtpopulation, gleichgerichtete Effektschätzer zugunsten von Apremilast vor. In der Kategorie „ehemaliger Raucher“ lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Das Subgruppenmerkmal „Vitalität – vorherige Colchicin-Einnahme“ zeigte in der Kategorie „ja“, analog zur Gesamtpopulation, gleichgerichtete Effekte zugunsten von Apremilast. In der Kategorie „nein“ lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Für das Subgruppenmerkmal „PCS – Alkoholkonsum“ waren die Effektschätzer bei den Patienten ohne Alkoholkonsum, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. Bei den Patienten mit Alkoholkonsum lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Bezüglich des Subgruppenmerkmals „MCS – vorherige Colchicin-Einnahme“ war der Effektschätzer bei den Patienten mit vorheriger Colchicin-Einnahme, analog zur Gesamtpopulation, gleichgerichtet zugunsten von Apremilast. Bei den Patienten ohne vorherige Colchicin-Einnahme lag kein statistisch signifikanter Effekt vor (Tabelle 4-76).

Die statistische Signifikanz bezüglich dieser Ergebnisse sollte mit Vorsicht bewertet werden. Aufgrund der post-hoc durchgeführten explorativen Analyse und der Größe der Subgruppen kann davon ausgegangen werden, dass keine Modifikation vorliegt.

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 4-77: Übersicht der eingeschlossenen Studie RELIEF

Studienbericht (ja/nein [Zitat])	Registereintrag (ja/nein [Zitat])	Publikation (ja/nein [Zitat])	Post-hoc-Analysen (ja/nein [Zitat])
ja (Amgen 2019)	ja clinicaltrials.gov (ClinicalTrials.gov 2014a) EU Clinical Trials Register (EU-CTR 2014)	ja (Hatemi 2019)	ja (Amgen 2020a)

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie

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

4.3.2.1.1.1 Studien des pharmazeutischen Unternehmers

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

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*
- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

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

Tabelle 4-78: Liste der Studien des pharmazeutischen Unternehmers – RCT für den adjustierten indirekten Vergleich

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
RELIEF	ja	ja	abgeschlossen	Placebo-kontrollierte Doppelblindphase: 12 Wochen Aktive Behandlungsphase: 52 Wochen	Apremilast 30mg bid, Placebo
BCT-001	ja	ja	abgeschlossen	Placebo-kontrollierte Doppelblindphase: 12 Wochen Aktive Behandlungsphase: 12 Wochen	Apremilast 30mg bid, Placebo
bid: zweimal täglich (bis in die)					

Geben Sie an, welchen Stand die Information in Tabelle 4-78 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Informationen in Tabelle 4-78 entsprechen dem Stand vom 30. März 2020.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-78 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-79: Studien des pharmazeutischen Unternehmers, die nicht für den indirekten Vergleich herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
BCT-001 ^a	Ausschlussgrund 1 ("Patientenpopulationen, die nicht dem zugelassenen Anwendungsgebiet entsprechen")
a: Medikamentöse Vorbehandlung des Behçet-Syndroms nicht verpflichtend. Damit kann keine Aussage darüber getroffen werden, ob Patienten für eine systemische Therapie infrage kommen.	

4.3.2.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem

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

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

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

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

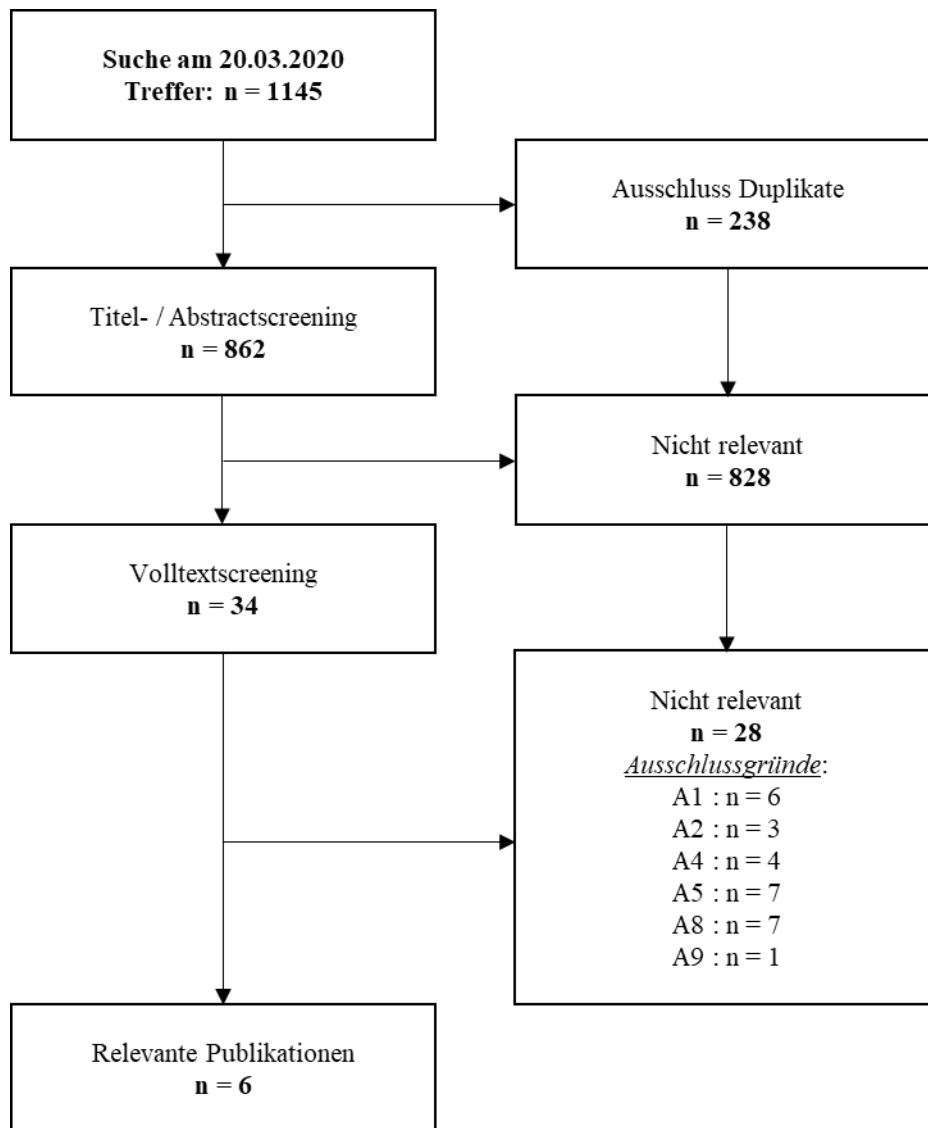


Abbildung 14: Flussdiagramm der bibliografischen Literaturrecherche - Suche nach randomisierten, kontrollierten Studien für die zVT

Zur Identifizierung relevanter Primärpublikationen für die zVT wurde am 20. März 2020 eine systematische bibliografische Literaturrecherche durchgeführt. Das vollständige Suchprofil befindet sich in Anhang 4-A.

Die bibliografische Literaturrecherche zu RCT für die zVT Apremilast im vorliegenden Anwendungsgebiet ergab 1.145 Treffer, darunter 283 Duplikate. Von den verbliebenen 862 Treffern wurden 828 nach Titel- / Abstractscreening und 28 nach anschließendem Volltextscreening begründet ausgeschlossen. Somit konnten sechs für die Nutzenbewertung relevante Publikationen identifiziert werden.

Bei vier der sechs in der bibliografischen systematischen Literaturrecherche identifizierten Studien handelt es sich um systematische Übersichtsarbeiten oder Zusammenfassungen von

mehreren Studien im AWG (Alpsoy 2002; Saenz 2000; Taylor 2014; Yazici 1990). Bei den beiden identifizierten Studien handelt es sich um RCT zu den zVT Etanercept bzw. Thalidomid über den Brückenkomparator Placebo (Hamuryudan 1998; Melikoğlu 2005).

4.3.2.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-78) 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 4-80: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich

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)
Suche nach RCT mit dem zu bewertenden Arzneimittel				
RELIEF BCT-002 NCT02307513 2014-002108-25	clinicaltrials.gov (ClinicalTrials.gov 2014a) EU Clinical Trials Register (EU-CTR 2014)	ja	ja	abgeschlossen
BCT-001 NCT00866359 2008-002722-11 ^b	clinicaltrials.gov (ClinicalTrials.gov 2014b) WHO ICTRP (ICTRP 2014)	ja	ja	abgeschlossen
Suche nach RCT der zVT für den adjustierten indirekten Vergleich				
Nicht zutreffend	-	-	-	-

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)
<p>a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>b: Ausschluss der Studie da Medikamentöse Vorbehandlung des Behçet-Syndroms nicht verpflichtend.</p> <p>EU-CTR: European Union Clinical Trials Register; ICTRP: International Clinical Trials Registry Platform; WHO: World Health Organization</p>				

Geben Sie an, welchen Stand die Information in Tabelle 4-80 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.

Stand der Informationen in Tabelle 4-80 ist der 15. April 2020.

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

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-78) 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 4-81: Relevante Studien aus der Suche auf der Internetseite des G BA – RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich

Studie	Relevante Quellen ^a	Studie in Liste der Studien des pharmazeutischen Unternehmens enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Studie durch Suche in Studienregistern / Studienergebnisdatenbanken identifiziert (ja/nein)
Suche nach RCT mit dem zu bewertenden Arzneimittel				
Nicht zutreffend	-	-	-	-
Suche nach RCT der zVT für den adjustierten indirekten Vergleich				
Nicht zutreffend	-	-	-	-
a: Quellen aus der Suche auf der Internetseite des G-BA.				

Geben Sie an, welchen Stand die Information in Tabelle 4-81 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.

Stand der Informationen in Tabelle 4-81 ist der 15. April 2020.

4.3.2.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.2.1.1.1, 4.3.2.1.1.2, 4.3.2.1.1.3 und 4.3.2.1.1.4) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich aller verfügbaren Quellen in Abschnitt 4.3.2.1.1.5 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

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

Tabelle 4-82: Studienpool – RCT mit dem zu bewertenden Arzneimittel und zVT für einen adjustierten indirekten Vergleich

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponsernte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Suche nach RCT mit dem zu bewertenden Arzneimittel						
RELIEF	ja	ja	nein	ja	ja clinicaltrials.gov (ClinicalTrials.gov 2014a) EU Clinical Trials Register (EU-CTR 2014)	ja (Hatemi 2019)
Suche nach RCT der zVT für den adjustierten indirekten Vergleich						
Melikoglu et al.	nein	nein	ja	nein	nein	ja (Melikoğlu 2005)
Hamuryuda n et al.	nein	nein	ja	nein	nein	ja (Hamuryudan 1998)
Leccesse et al.	nein	nein	ja	nein	nein	ja (Leccese 2019)
Taylor et al.	nein	nein	ja	nein	nein	ja (Taylor 2014)
Hatemi et al.	nein	nein	ja	nein	nein	ja (Hatemi 2009)
Saenz et al.	nein	nein	ja	nein	nein	ja (Saenz 2000)
<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 4.3.2.1.1.5 (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>EU-CTR: European Union Clinical Trials Register; ICTRP: International Clinical Trials Registry Platform; WHO: World Health Organization</p>						

Geben Sie an, welchen Stand die Information in Tabelle 4-82 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.

Stand der Informationen in Tabelle 4-82 ist der 15. April 2020.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

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

4.3.2.1.2.1 Studiendesign und Studienpopulationen

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

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

Die eingeschlossenen Studien für den indirekten Vergleich zwischen Apremilast und den Arzneimitteln der zVT werden in Tabelle 4-83 charakterisiert. Die Interventionen der jeweiligen Studien werden in Tabelle 4-84 charakterisiert.

Bei vier der sechs in der bibliografischen systematischen Literaturrecherche identifizierten Studien handelt es sich um systematische Übersichtsarbeiten oder Zusammenfassungen von mehreren Studien im AWG (Alpsoy 2002; Saenz 2000; Taylor 2014; Yazici 1990). Auf die weitere Darstellung dieser Übersichtsarbeiten wird verzichtet. Eine Relevanz der Übersichtsarbeiten ergäbe sich nur bei Vergleichbarkeit und somit Durchführbarkeit des adjustierten indirekten Vergleichs.

Dargestellt werden die Studien von Melikoglu et al. und Hamuryudan et al. Dementsprechend besteht der resultierende Studienpool für den indirekten Vergleich zwischen Apremilast und der zVT Etanercept bzw. Thalidomid über den Brückenkomparator Placebo (Hamuryudan 1998; Melikoğlu 2005).

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

Tabelle 4-83: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
Zu bewertendes Arzneimittel						
RELIEF	RCT, doppelblind, multizentrisch, Placebo-kontrolliert, parallel	Erwachsene Patienten (≥ 18 Jahre) mit diagnostiziertem Behçet-Syndrom anhand der ISG-Kriterien, mindestens eine nicht-biologische Vorbehandlung des Behçet-Syndroms	Gruppe 1: Apremilast (n = 104) Gruppe 2: Placebo (n = 103)	Screening-Phase: bis zu sechs Wochen Placebo-kontrollierte Doppelblindphase: 12 Wochen Aktive Behandlungsphase: 52 Wochen (beide Gruppen Apremilast) Nachbeobachtung: vier Wochen Datenschnitte nach Komplettierung der Woche 12 sowie nach Woche 28 (Wirksamkeit) und nach Behandlung von 100 Patienten über ein Jahr (Sicherheit)	53 Studienzentren in zehn Ländern: Asien (Japan, Südkorea) Europa (Frankreich, Deutschland, Griechenland, Italien) USA Rest der Welt (Libanon, Israel, Türkei) 12/2014 – 9/2018	<u>Primärer Endpunkt:</u> AUC für die Anzahl oraler Aphthen <u>Sekundäre Endpunkte:</u> Anzahl, Ansprechrate und Schmerzhaftigkeit der oralen Aphthen, Zeit bis zu vollständiger Remission, Zeit bis zum Rezidiv, Krankheitsaktivität über PRO (mittels BSAS und BDCAF) und gesundheitsbezogener Lebensqualität (mittels BD-QoL und SF-36), Sicherheit (UE)

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
Zweckmäßige Vergleichstherapie						
Melikoglu et al. 2005	RCT, doppelblind, Einzelzentrum, Placebo-kontrolliert, parallel	Erwachsene Patienten (≥ 18 Jahre) mit diagnostiziertem Behçet-Syndrom anhand der ISG-Kriterien	Gruppe 1: Etanercept (n = 20) Gruppe 2: Placebo (n = 20)	Screening: Initialer Studienbesuch Placebo-kontrollierte Doppelblindphase: 4 Wochen Nachbeobachtungsphase: 12 Wochen	Ein Studienzentrum in der Türkei (Istanbul)	<u>Primärer Endpunkt:</u> Das Ausmaß der Unterdrückung der pathogenen Response und des MSU Tests <u>Sekundäre Endpunkte:</u> Anzahl der mukokutanösen Läsionen, Anzahl der geschwollenen Gelenke, völlige Freiheit von Aphthen (oral, genital), völlige Freiheit von Läsionen (papulopustular, nodular); völlige Freiheit von Arthritis
Hamuryudan et al. 1998	RCT, doppelblind, Einzelzentrum, Placebo-kontrolliert, parallel	Erwachsene, männliche Patienten (18 – 35 Jahre) mit diagnostiziertem Behçet-Syndrom anhand der ISG-Kriterien, mindestens zwei Episoden oraler oder genitaler Aphthen innerhalb der letzten 3 Monate vor Studienbeginn	Gruppe 1: Thalidomid 100 mg (n = 32) Gruppe 2: Thalidomid 300 mg (n=31) Gruppe 3: Placebo (n = 32)	Screening: Initialer Studienbesuch Placebo-kontrollierte Doppelblindphase: 24 Wochen Nachbeobachtungsphase: 4 Wochen	Ein Studienzentrum in der Türkei (Istanbul)	<u>Primärer Endpunkt:</u> Freiheit von oralen und genitalen Aphthen während der 24-wöchigen Behandlungsphase <u>Sekundäre Endpunkte:</u> Anzahl der mukokutanösen Läsionen, Ansprechen der Augenerkrankung auf die Behandlung
<p>AUC: Fläche unter der Kurve (area under the curve); BDCAF: Behçet’s Disease Current Activity Form; BD-QoL: Behçet’s disease Quality of Life; BSAS: Behçet’s Syndrome Activity Score; ISG: International Study Group for Behçet’s Disease; MSU: Monosodium urate; PRO: Patientenberichtete Endpunkte (patient-reported outcome); RCT: randomisierte kontrollierte Studie (randomized controlled trial); SF-36: 36-Item Short Form Health Survey, Version 2; UE: Unerwünschte/s Ereignis/se; USA: Vereinigte Staaten von Amerika (United States of America)</p> <p>Quellen: Amgen 2019; Hamuryudan 1998; Melikoglu 2005</p>						

Tabelle 4-84: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Interventionsgruppe	Vergleichsgruppe	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
RELIEF	Apremilast 30 mg p.o. bid im Abstand von ca. 12 Stunden für einen Zeitraum von 64 Wochen	Placebo p.o. bid im Abstand von ca. 12 Stunden für einen Zeitraum von 12 Wochen im Anschluss daran: Apremilast 30 mg p.o. bid im Abstand von ca. 12 Stunden für einen Zeitraum von 52 Wochen	Die Patienten waren mit mindestens einer nicht-biologischen Therapie des Behçet-Syndroms, z. B. mit topischen Kortikosteroiden oder systemischen Wirkstoffen, vorbehandelt und für eine (weitere) systemische Therapie zur Behandlung oraler Aphthen geeignet. Eine Dosisänderung von Apremilast im Studienverlauf war nicht vorgesehen.
Melikoglu et al. 2005	Etanercept 25mg, zweimal pro Woche als subkutane Injektion	Placebo, zweimal pro Woche als subkutane Injektion	Eine Vorbehandlung war keine Voraussetzung für den Studieneinschluss. In der Etanercept-Gruppe wurden 11 von 20 Patienten mit Colchicin vorbehandelt. In der Placebo-Gruppe wurden 7 von 20 Patienten mit Colchicin und 5 von 20 Patienten mit Azathioprin vorbehandelt. Zwei Patienten in der Etanercept-Gruppe und ein Patient in der Placebo-Gruppe führten die Behandlung während der Studie fort.
Hamuryudan et al. 1998	Thalidomid 100mg; p.o. bis zu dreimal täglich Gruppe 1: eine Placebo-Tablette morgens und eine Placebo-Tablette und eine Thalidomid-Tablette abends Gruppe 2: eine Thalidomid-Tablette morgens und zwei Thalidomid-Tabletten abends	Placebo p.o. dreimal täglich Gruppe 3: eine Placebo-Tablette morgens und zwei Placebo-Tabletten abends	Keine Angaben zur Vortherapie der Patienten verfügbar
bid: zweimal täglich (bis in die); p.o.: oral (per os) Quellen: Amgen 2019; Hamuryudan 1998; Melikoğlu 2005			

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.

Um eine hinreichende Homogenität zwischen der Studie RELIEF und den Studien zur zVT von Melikoglu et al. und Hamuryudan et al. zu prüfen, wird im nächsten Schritt die Vergleichbarkeit der Studien anhand folgender Merkmale für die Durchführbarkeit eines adjustierten indirekten Vergleichs bewertet:

- Vergleichbarkeit der Patientenpopulation hinsichtlich:
 - o Ein- und Ausschlusskriterien
 - o Baseline-Charakteristika
- Vergleichbarkeit der Intervention (Brückenkomparator):
 - o Effekte in den Placebo-Gruppen
- Vergleichbarkeit der Operationalisierung der patientenrelevanten Endpunkte

Nur dann, wenn diese Voraussetzungen erfüllt sind, darf ein adjustierter indirekter Vergleich durchgeführt werden. Fehlende Informationen führten zum Ausschluss der betreffenden Studie.

Vergleichbarkeit der Patientenpopulation hinsichtlich Ein- und Ausschlusskriterien

In Tabelle 4-85 werden die Ein- und Ausschlusskriterien, die in den Volltext Publikationen zu den beiden relevanten Studien, berichtet wurden, gegenübergestellt.

Tabelle 4-85: Ein- und Ausschlusskriterien der RCT für den indirekten Vergleich

Ein- / Ausschlusskriterium	RELIEF	Melikoglu 2005	Hamuryudan 1998
Einschlusskriterien			
Geschlecht	Männliche und weibliche Studienteilnehmer	Ausschließlich männliche Studienteilnehmer	Ausschließlich männliche Studienteilnehmer
Alter	Ab 18 Jahre	18 bis 45 Jahre	18 bis 35 Jahre

Ein- / Ausschlusskriterium	RELIEF	Melikoglu 2005	Hamuryudan 1998
Krankheitsaktivität	Orale Aphthen, die in den vergangenen 12 Monaten mindestens dreimal aufgetreten sind, einschließlich oraler Aphthen zum Zeitpunkt des Screenings	Mindestens eine der folgenden Krankheitsmerkmale innerhalb von 3 Monaten vor Studienbeginn: orale Aphthen, genitale Aphthen, noduläre Läsionen, geschwollene Gelenke	Mindestens zwei Episoden von Aphthen innerhalb von 3 Monaten vor Studienbeginn: orale Aphthen oder genitale Aphthen
Pathergietest	Keine Angabe	Positiver Pathergietest	Keine Angaben
MSU Test	Keine Angabe	Positiver MSU-Test	Keine Angaben
Ausschlusskriterien			
Schwerwiegende Organbeteiligung	Im Rahmen des Behçet-Syndroms auftretende aktive Beteiligung wichtiger Organe: pulmonale (z.B. Pulmonalarterienaneurysma), vaskuläre (z. B. Thrombophlebitis), gastrointestinale (z. B. Ulzera im Bereich des Gastrointestinaltrakts), zentralnervöse (z. B. Meningoenzephalitis) Manifestationen, sowie okuläre Läsionen (z. B. Uveitis), die eine immunsuppressive Therapie erforderten	Im Rahmen des Behçet-Syndroms auftretende aktive Beteiligung von Augen, zentralem Nervensystem oder großen Arterien	Im Rahmen des Behçet-Syndroms auftretende aktive moderate oder schwere Beteiligung der Augen, jegliche Organbeteiligung, die eine immunsuppressive Therapie erfordert oder erfordert hat. Das Vorliegen von Neuropathie.
Infektionen	Keine Angaben	Vorliegen lokaler oder systemischer Infektionen	Keine Angaben
Laborwerte	Hämoglobin < 9 g/dl Leukozytenzahl $\leq 3.000/\mu\text{l}$ ($\leq 3,0 \times 10^9/\text{l}$) und $\geq 14.000/\mu\text{l}$ ($\geq 14 \times 10^9/\text{l}$) Thrombozytenzahl $\leq 100.000/\mu\text{l}$ ($\leq 100 \times 10^9/\text{l}$) Serumkreatinin $\geq 1,5 \text{ mg/dl}$ ($\geq 132,6 \mu\text{mol/l}$) Gesamtbilirubin $\geq 2,0 \text{ mg/dl}$ AST und ALT > 1,5 x ULN	4 Wochen vor Studienbeginn: ALT und AST zweifach über der oberen Normgrenze Hämatokrit < 28 % WBC < 4.000/mm ³ Plättchen Anzahl < 140.000/mm ³ Serumkreatinin > 1,9 mg/dl	Keine Angaben
<p>ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; MSU: Monosodium urate; ULN: obere Normgrenze (upper limit of normal), WBC: Weiße Blutkörperchen (white blood cell)</p> <p>Quellen: Amgen 2019; Hamuryudan 1998; Melikoglu 2005</p>			

Adjustierter indirekter Vergleich zwischen Apremilast und Etanercept

Eine ausreichende Vergleichbarkeit der Zulassungsstudie RELIEF und der Studie Melikoglu 2005 kann nicht mit hinreichender Sicherheit angenommen werden, sodass ein indirekter Vergleich zwischen beiden Studien als hoch verzerrt eingestuft werden müsste (Amgen 2019; Melikoğlu 2005).

Im Wesentlichen begründet sich dies an den unterschiedlichen Altersspannen, an Unterschieden in der Geschlechterverteilung und an der Unsicherheit bezüglich des Anteils von Patienten mit oralen Aphthen bei Studienbeginn, da Patienten auch aufgrund von anderen Symptomen eingeschlossen werden konnten.

Tabelle 4-86: Relevante Unterschiede in den Baseline Charakteristika zwischen RELIEF und Melikoglu 2005

Baseline Charakteristikum	RELIEF n = 207	Melikoglu 2005^a n = 40
Geschlecht (%)	m = 80 (38,5 %) w = 127 (61,5 %)	m = 40 (100 %) w = -
Alter in Jahren MW (SD)	Apremilast: 39,4 (12,1) Placebo: 40,6 (12,7)	Etanercept: 28,5 Placebo: 30,8
Dauer des Behçet-Syndroms in Jahren MW (SD)	Apremilast: 6,7 (7,4) Placebo: 6,9 (8,0)	Etanercept: 2,8 Placebo: 4,3
Patienten mit oralen Aphthen zu Studienbeginn (%)	Apremilast: 104 (100 %) Placebo: 103 (100 %)	Etanercept: k.a. Placebo: k.a.
Mittlere Anzahl der oralen Aphthen zu Studienbeginn MW (SD)	Apremilast: 4,2 (3,7) Placebo: 3,9 (2,7)	Etanercept: 2,0 (1,3) Placebo: 1,8 (1,6)
Patienten mit genitalen Aphthen zu Studienbeginn (%)	Apremilast: 17 (16,3 %) Placebo: 17 (16,5 %)	Etanercept: k.a. Placebo: k.a.
Mittlere Anzahl der genitalen Aphthen zu Studienbeginn MW (SD)	Apremilast: 2,9 (2,9) Placebo: 2,6 (2,0)	Etanercept: 1,3 (1,5) Placebo: 0,8 (1,3)
a: Nicht für alle Kategorien wurden in den Publikationen von Melikoglu 2005 Standardabweichungen angegeben. k.a.: keine Angabe; m: männlich; MW: Mittelwert; n: Anzahl; SD: Standardabweichung (standard deviation); w: weiblich Quellen: Amgen 2019; Melikoğlu 2005		

Hinsichtlich des Alters und der Dauer der Erkrankung sind die Patienten in der Studie RELIEF im Schnitt um zehn Jahre älter und fast doppelt so lange erkrankt wie die Patienten in der Studie von Melikoglu.

Hinsichtlich der Häufigkeit der oralen und genitalen Aphthen und der Anzahl der jeweiligen Aphthen zu Studienbeginn zeigt sich ebenfalls, dass die Patienten in der Studie RELIEF

deutlich schwerer betroffen sind. Die Patienten in der Studie RELIEF wiesen im Mittel beinahe doppelt so viele orale Aphthen zu Studienbeginn und fast dreimal so viele genitale Aphthen auf.

Es ist unsicher, ob alle Patienten in der Studie von Melikoglu zu Studienbeginn orale Aphthen aufwiesen. Das Vorhandensein von oralen Aphthen war kein Einschlussgrund, die Patienten konnten auch andere Manifestationen zu Studienbeginn aufweisen, um in die Studie aufgenommen zu werden. Es wäre damit nicht möglich zu überprüfen, ob ein relevanter Anteil der Patienten den Patienten aus der Studie RELIEF entspricht.

Zusätzlich war eine Vortherapie kein Einschlussgrund in die Studie von Melikoglu et al. (Melikoğlu 2005). Aus der Beschreibung der Patienten geht hervor, dass in der Etanercept-Gruppe 11 von 20 Patienten mit Colchicin vorbehandelt wurden. In der Placebo-Gruppe waren sieben von 20 Patienten mit Colchicin und fünf von 20 Patienten mit Azathioprin vorbehandelt. Zwei Patienten in der Etanercept-Gruppe und ein Patient in der Placebo-Gruppe führten die Behandlung während der Studie fort. Des Weiteren wendeten sechs Patienten in der Etanercept-Gruppe und acht Patienten in der Placebo-Gruppe im Studienverlauf zusätzliche Arzneimittel gegen Symptome des Behçet-Syndroms an (Amgen 2019; Melikoğlu 2005).

Aufgrund der Tatsache, dass sowohl die Ausprägung von oralen Aphthen als auch eine Vortherapie keine ausschlaggebenden Selektionskriterien der Studie von Melikoglu et al. waren, lässt sich nicht mit Sicherheit sagen, ob das Patientenkollektiv der Studie dem AWG von Apremilast entspricht.

Darüber hinaus ergeben sich aus dem Studiendesign weitere Faktoren, die zu einem erhöhten Verzerrungspotenzial führen oder den Vergleich der Ergebnisse limitieren. Zum einen ist die Studie mit 20 Patienten pro Behandlungsgruppe sehr klein und damit von geringer statistischer Aussagekraft. Zum anderen wurde die Studie monozentrisch durchgeführt, sodass Zentrumseffekte nicht ausgeschlossen werden können (Amgen 2019; Melikoğlu 2005).

Aufgrund der oben beschriebenen Limitationen und Unsicherheiten wird kein indirekter Vergleich zwischen Apremilast und Etanercept, basierend auf der Zulassungsstudie RELIEF und der Studie von Melikoglu et al., durchgeführt (Melikoğlu 2005).

Adjustierter indirekter Vergleich zwischen Apremilast und Thalidomid

Eine ausreichende Vergleichbarkeit der Zulassungsstudie RELIEF und der Studie Hamuryudan 1998 kann nicht mit hinreichender Sicherheit angenommen werden, sodass ein indirekter Vergleich zwischen beiden Studien als hoch verzerrt eingestuft werden müsste (Amgen 2019; Hamuryudan 1998).

Im Wesentlichen begründet sich dies an den unterschiedlichen Altersspannen, an Unterschieden in der Geschlechterverteilung, an der Unsicherheit bezüglich des Anteils von Patienten mit oralen Aphthen bei Studienbeginn und dem Fehlen von Angaben zu Laborparametern, die Aufschluss über die Krankheitsschwere des Patienten geben.

Tabelle 4-87: Relevante Unterschiede in den Baseline Charakteristika zwischen RELIEF und Hamuryudan 1998

Baseline Charakteristikum	RELIEF n = 207	Hamuryudan 1998 ^a n = 98
Geschlecht (%)	m = 80 (38,5 %) w = 127 (61,5 %)	m = 98 (100 %) w = -
Alter in Jahren MW (SD)	Apremilast: 39,4 (12,1) Placebo: 40,6 (12,7)	Thalidomid 100 mg: 28,8 Thalidomid 300 mg: 27,6 Placebo: 26,7
Dauer des Behçet-Syndroms in Jahren MW (SD)	Apremilast: 6,7 (7,4) Placebo: 6,9 (8,0)	Thalidomid 100 mg: 3,0 Thalidomid 300 mg: 2,7 Placebo: 2,7
Patienten mit oralen Aphthen zu Studienbeginn (%)	Apremilast: 104 (100 %) Placebo: 103 (100 %)	Thalidomid 100 mg: 28 (87,5 %) Thalidomid 300 mg: 26 (84 %) Placebo: 24 (75 %)
Mittlere Anzahl der oralen Aphthen zu Studienbeginn MW (SD)	Apremilast: 4,2 (3,7) Placebo: 3,9 (2,7)	Thalidomid 100 mg: 2,3 Thalidomid 300 mg: 2,3 Placebo: 1,8
Patienten mit genitalen Aphthen zu Studienbeginn (%)	Apremilast: 17 (16,3 %) Placebo: 17 (16,5 %)	Thalidomid 100 mg: 12 (38 %) Thalidomid 300 mg: 8 (26 %) Placebo: 11 (34 %)
Mittlere Anzahl der genitalen Aphthen zu Studienbeginn MW (SD)	Apremilast: 2,9 (2,9) Placebo: 2,6 (2,0)	Thalidomid 100 mg: 0,7 Thalidomid 300 mg: 0,5 Placebo: 1,1
a: Nicht für alle Kategorien wurden in den Publikationen von Hamuryudan 1998 Standardabweichungen angegeben. k.a.: keine Angabe; m: männlich; MW: Mittelwert; n: Anzahl; SD: Standardabweichung (standard deviation); w: weiblich Quellen: Amgen 2019; Hamuryudan 1998		

Hinsichtlich des Alters und der Dauer der Erkrankung waren die Patienten in der Studie RELIEF im Schnitt um zehn Jahre älter und bereits doppelt so lange erkrankt.

Hinsichtlich der Häufigkeit der oralen und genitalen Aphthen und der Anzahl der jeweiligen Aphthen zu Studienbeginn zeigte sich weiterhin, dass die Patienten in der Studie RELIEF deutlich schwerer betroffen waren. Die Patienten in der Studie RELIEF wiesen im Mittel beinahe doppelt so viele orale Aphthen zu Studienbeginn und fast viermal so viele genitale Aphthen auf.

Nicht alle Patienten in der Studie von Hamuryudan wiesen zu Studienbeginn orale Aphthen auf. Das Vorhandensein von oralen Aphthen war kein ausgewiesener Einschlussgrund, und die Patienten konnten auch andere Manifestationen zu Studienbeginn aufweisen, um in die Studie aufgenommen zu werden. Insbesondere hervorzuheben ist die Placebo-Gruppe. Hier lag der Anteil der oralen Aphthen bei 75 %, sodass für 25 % der Patienten im Brückenkomparator nicht dieselben Bedingungen wie in der Studie RELIEF zu Studienbeginn vorliegen können. Dies stellt eine wesentliche Verzerrung dar, die im Rahmen des adjustierten indirekten Vergleichs nicht behoben werden kann.

Zusätzlich war eine Vortherapie kein Einschlussgrund in die Studie von Hamuryudan et al. Aus der Beschreibung in der Publikation können keine Angaben zu Vortherapien oder begleitenden Therapien entnommen werden (Amgen 2019; Hamuryudan 1998).

Aufgrund der Tatsache, dass sowohl die Ausprägung von oralen Aphthen und eine Vortherapie keine ausschlaggebenden Selektionskriterien der Studie von Hamuryudan et al. waren, lässt sich nicht mit Sicherheit sagen, ob das Patientenkollektiv der Studie dem AWG von Apremilast entspricht.

Darüber hinaus ergeben sich aus dem Studiendesign weitere Faktoren, die zu einem erhöhten Verzerrungspotenzial führen oder den Vergleich der Ergebnisse limitieren. Die Studie wurde in einem Studienzentrum durchgeführt, sodass Zentrumsffekte nicht ausgeschlossen werden können. Daten für orale Aphthen werden nicht getrennt aufgeführt, sodass es nicht möglich ist, bei der Auswertung zwischen oralen Aphthen und genitalen Aphthen zu unterscheiden, wodurch keine Vergleichbarkeit der Endpunkte angenommen werden kann (Amgen 2019; Hamuryudan 1998).

Aufgrund der oben beschriebenen Limitationen und Unsicherheiten wird kein indirekter Vergleich zwischen Apremilast und Thalidomid, basierend auf der Zulassungsstudie RELIEF und der Studie von Hamuryudan 1998, durchgeführt.

Vergleichbarkeit der Patientenpopulation hinsichtlich Baseline-Charakteristika

Nicht relevant, da beide Studien bereits nach der Betrachtung der Ein- und Ausschlusskriterien ausgeschlossen wurden.

Vergleichbarkeit der Intervention (Brückenkomparator)

Nicht relevant, da beide Studien bereits nach der Betrachtung der Ein- und Ausschlusskriterien ausgeschlossen wurden.

Vergleichbarkeit der Operationalisierung der patientenrelevanten Endpunkte

Nicht relevant, da beide Studien bereits nach der Betrachtung der Ein- und Ausschlusskriterien ausgeschlossen wurden.

Abschließende Beurteilung des resultierenden Studienpools für einen adjustierten indirekten Vergleich

Alle identifizierten Studien, die über eine bibliografische systematische Literaturrecherche identifiziert wurden, waren nicht geeignet, um einen indirekten Vergleich mit der Studie RELIEF durchzuführen.

Die beiden Studien mit den Wirkstoffen Etanercept und Thalidomid waren nicht geeignet, um einen adjustierten indirekten Vergleich mit der Studie RELIEF durchzuführen und wurden im ersten Schritt der Überprüfung der Ähnlichkeit ausgeschlossen, da abweichende Ein- und Ausschlusskriterien bzw. Unterschiede zwischen den Studiendesignen vorliegen. Die wesentlichen Gründe für den Ausschluss der Studien waren die fehlenden Angaben zur Vorbehandlung der Patienten und das Fehlen von Angaben zum Anteil der oralen Aphthen.

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

Tabelle 4-88: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

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 4-89: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits-bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend					

Nicht zutreffend.

4.3.2.1.3.1 Endpunkt – 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 4-90: 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>
Nicht zutreffend					

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 4-91: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
Nicht zutreffend	

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

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

Tabelle 4-92: Bewertung des Verzerrungspotenzials für <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
Nicht zutreffend						

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 4-93: 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)
Nicht zutreffend	

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 4-94: 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>
Nicht zutreffend						

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 4-95: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend					

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 4-96: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
Nicht zutreffend	

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 4-97: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
Nicht zutreffend				

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 4-98: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend					

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-99: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
Nicht zutreffend	

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.

4.3.2.3.2 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.

Apremilast (Otezla[®]) ist gemäß Fachinformation indiziert zur Behandlung von erwachsenen Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt (Amgen 2020b).

Für die Bestimmung des medizinischen Nutzens und Zusatznutzens von Apremilast im AWG liegt die klinische Zulassungsstudie RELIEF vor. Die Studie RELIEF wurde anhand des CONSORT-Statements und der Cochrane-Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse auf die Planungs-, Durchführungs- und Auswertungsqualität überprüft (Anhang 4-E, Anhang 4-F). Für die Bestimmung des Ausmaßes des medizinischen Nutzens und Zusatznutzens wurden die dazugehörigen Dokumente der Studie RELIEF (Studienprotokoll, statistischer Analyseplan, Studienbericht, Publikation) herangezogen und die Ergebnisse patientenrelevanter Endpunkte im Dossier dargestellt.

Studienqualität

Für Apremilast wurde das bisher umfangreichste Studienprogramm zur Behandlung von erwachsenen Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt, durchgeführt. Die Studienpopulation der Studie RELIEF ist aufgrund der Vorbehandlung mit den Wirkstoffen Azathioprin, Prednisolon oder Prednison, die eine Zulassung im AWG besitzen, repräsentativ für das Patientenkollektiv. Es wurde zusätzlich die Studie BCT-001 identifiziert. Da jedoch keine Vortherapie für den

Studieneinschluss erfolgen musste, kann nicht davon ausgegangen werden, dass das teilnehmende Patientenkollektiv ausreichend repräsentativ für das AWG war. Daher wurde für die Bestimmung des medizinischen Nutzens und Zusatznutzens ausschließlich die Studie RELIEF herangezogen.

Bei RELIEF handelt es sich um eine multinationale, multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte Phase-III-Studie. Es wurden erwachsene Patienten im Alter ab 18 Jahren mit gesichertem Behçet-Syndrom untersucht. Dabei mussten die Patienten mindestens eine nicht-biologische Vortherapie erhalten haben und zur Screeningvisite sowie zu Baseline mindestens zwei orale Aphthen aufweisen, deren Ausmaß eine topische Behandlung als ungeeignet oder unzureichend ausschloss.

Das Verzerrungspotenzial auf Studienebene wurde für die Studie RELIEF als niedrig bewertet (Tabelle 4-18). Die Auswertung basierte auf dem ITT-Prinzip, welches adäquat umgesetzt wurde. Studienteilnehmer, die die Studie vorzeitig abbrachen, wurden im Studienbericht transparent und nachvollziehbar dargestellt. Die Studie RELIEF entspricht gemäß Verfahrensordnung der Evidenzstufe Ib (G-BA 2019). Die Randomisierung wurde adäquat durchgeführt, wobei die Gruppenzuteilung (Apremilast oder Placebo) verdeckt über eine interaktive Response-Technologie erfolgte. Patienten und behandelnde Personen waren gegenüber der Studienmedikation verblindet. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan (SAP) der Studie RELIEF bzw. im Post-hoc-SAP definiert waren. Es lagen keine sonstigen Aspekte für ein erhöhtes Verzerrungspotenzial vor. Für die Bewertung werden patientenrelevante Endpunkte aus den Kategorien zu Mortalität, gesundheitsbezogener Lebensqualität und Sicherheit herangezogen. In der Studie RELIEF ereignete sich kein Todesfall. Daher wird auf eine Ergebnisdarstellung der Mortalität verzichtet.

Insgesamt handelt es sich somit bei RELIEF um eine Studie der höchsten Evidenzstufe (Ib). Das Verzerrungspotenzial ist sehr niedrig. Die Studie ist damit als sehr valide zur Darstellung des Zusatznutzens einzustufen.

Prüfung eines indirekten Vergleich

Für die Behandlung der Manifestationen des Behçet-Syndroms besteht insgesamt ein Bedarf an robuster Evidenz sowie ein Mangel an zugelassenen Therapieoptionen. Dies spiegelt sich auch in der zVT im AWG wider. Hierzu wurde für Apremilast im vorliegenden AWG eine „Therapie nach Maßgabe des Arztes“ festgelegt. International und auch auf europäischer Ebene gibt es außer Apremilast kein weiteres Arzneimittel, das zur Behandlung von oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt, zugelassen ist. In Deutschland ist Azathioprin für die Behandlung der schweren Manifestationen zugelassen. Prednison und Prednisolon besitzen eine Zulassung für Erkrankungen, die einer systemischen Therapie mit Glucokortikoiden bedürfen (u. a. Behçet-Syndrom) – Adalimumab und Ciclosporin dagegen für die Behandlung der Uveitis, die (auch) mit dem Behçet-Syndrom einhergehen kann. In den zugrundeliegenden Fachinformationen fanden sich keine Angaben zur Wirksamkeit und Sicherheit zur Behandlung von oralen

Aphthen, die mit dem Behçet-Syndrom assoziiert sind. Als potenzielle klinische Komparatoren sind weiterhin Dapson, Interferon-alfa, Thalidomid und weitere, über Adalimumab hinausgehende TNF- α -Inhibitoren auf Basis der Leitlinien der European League Against Rheumatism (EULAR) vom G-BA benannt worden. Für diese Arzneimittel liegt in Deutschland keine Zulassung zur Behandlung von Manifestationen des Behçet-Syndroms vor.

Auf dieser Basis wurde die Durchführung eines adjustierten indirekten Vergleichs geprüft. Mit Hilfe einer systematischen bibliografischen Literaturrecherche (Stand 20. März 2020) wurde jeweils eine Studie mit der zVT Etanercept und eine Studie mit der zVT Thalidomid identifiziert. Bei beiden Studien wäre ein Vergleich über den Brückenkomparator Placebo infrage gekommen (Hamuryudan 1998; Melikoğlu 2005). Für keine der beiden Studien kann hinsichtlich der untersuchten Patientenpopulationen eine ausreichende Übereinstimmung mit den Kriterien der Studie RELIEF angenommen werden. Zudem konnte keine ausreichend große Übereinstimmung mit der Zielpopulation im AWG identifiziert werden. Ein indirekter Vergleich zwischen den Studien wäre somit weder zielführend noch valide (siehe Abschnitt 4.3.2.1.2.1).

Auf Basis dessen ergibt sich, dass sich die Studie RELIEF als bestmögliche Evidenz im AWG zur Darstellung des medizinischen Nutzens und Zusatznutzens eignet. Die Beratungen mit den Zulassungsbehörden im Rahmen des klinischen Studienprogramms von Apremilast bestätigten dies. Die Zulassungsbehörde Health Products Regulatory Authority (HPRA) stimmte zu, dass aufgrund der Tatsache, dass im AWG keine international, einheitlich zugelassenen Therapien verfügbar sind, ein Placebovergleich zielführend und valide ist, um den medizinischen Nutzen von Apremilast im AWG darzustellen (Amgen 2018b).

Validität der Endpunkte

Das Verzerrungspotenzial der Morbiditätsendpunkte „vollständige Remission der oralen Aphthen (komplettes Ansprechen)“, „Schmerzhaftigkeit oraler Aphthen mittels VAS“, „Veränderung des BSAS zu Woche 12“ und „Veränderung des BDCAF zu Woche 12“ wird jeweils als niedrig eingestuft, da sowohl Patienten als auch behandelnde oder weiterbehandelnde Personen während der Randomisierungsphase verblindet waren. Die Analysen erfolgten für die ITT-Population, die alle Patienten umfasste, die wenigstens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war (BSAS und BDCAF). Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da die Endpunkte anhand adäquater Methoden erhoben wurden.

Das Verzerrungspotenzial der Endpunkte zur gesundheitsbezogenen Lebensqualität „Veränderung des BD-QoL zu Woche 12“ und „Verbesserung des SF-36 um $\geq 2,5$ Punkte zu Woche 12“ wird für beide Endpunkte als niedrig eingestuft, da sowohl Patienten als auch behandelnde oder weiterbehandelnde Personen während der Randomisierungsphase verblindet waren. Die Analysen erfolgten für die ITT-Population, die alle Patienten umfasste, die wenigstens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine

ergebnisgesteuerte Berichterstattung hindeuten, da die Endpunkte anhand adäquater Methoden erhoben wurden.

Das Verzerrungspotenzial des Endpunkts „Sicherheit“ wird als niedrig eingestuft, da sowohl Patienten als auch behandelnde oder weiterbehandelnde Personen während der Randomisierungsphase verblindet waren. Die Analyse erfolgte für die Sicherheitspopulation, die alle mit Studienmedikation behandelten Patienten (as-treated) einschloss. Es fanden sich keine Aspekte in den Studienunterlagen, die auf eine ergebnisgesteuerte Berichterstattung hindeuten, da der Endpunkt anhand objektiver und adäquater Methoden erhoben wurde.

Das sowohl auf Studien- als auch auf Endpunktebene als niedrig bewertete Verzerrungspotenzial belegt die hohe Studienqualität sowie die Validität der Studie und der erhobenen Endpunkte.

Zusammenfassung

Die Studie RELIEF ist die bisher größte Studie zur Behandlung oraler Aphthen bei Patienten mit Behçet-Syndrom, für die eine systemische Therapie infrage kommt. Im AWG liegt keine weitere Studie mit qualitativ gleichwertigem Design vor. Bei der Studie RELIEF handelt es sich um eine randomisierte Placebo-kontrollierte Zulassungsstudie, für die keine weiteren verzerrenden Aspekte vorliegen und deren Ergebnissicherheit als hoch bewertet werden kann.

Für die Aussagen zum Zusatznutzen von Apremilast zur Behandlung erwachsener Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt, wird bezüglich der **Aussagekraft und Ergebnissicherheit** der Status eines **Hinweises** beansprucht.

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.

Für die Bestimmung des medizinischen Zusatznutzens von Apremilast wurde die zweiarmige randomisierte, Placebo-kontrollierte Studie RELIEF herangezogen (Evidenzstufe Ib).

Nachfolgend werden in Tabelle 4-100 alle Ergebnisse zum medizinischen Nutzen und Zusatznutzen unter Angabe des Ausmaßes des endpunktspezifischen Zusatznutzens für die im Dossier berücksichtigten Endpunkte dargestellt.

Tabelle 4-100: Ausmaß des Zusatznutzens auf Endpunktebene

Endpunkte	Studie RELIEF (Apremilast vs. Placebo)	Ausmaß des Zusatznutzens
Morbidität		
Vollständige Remission der oralen Aphthen	Vollständiges Ansprechen Ergebnisse: 56,7 % vs. 25,2 % RR [95 %-KI]: 2,25 [1,55; 3,27] p-Wert < 0,0001	beträchtlich
Zeit bis zur vollständigen Remission der oralen Aphthen (in Wochen)	Ergebnis: 2,14 vs. 8,14 HR [95 %-KI]: 2,40 [1,69; 3,40] p-Wert < 0,0001	
Schmerzhaftigkeit oraler Aphthen mittels VAS	Ergebnis (LS-MW): -39,04 vs. -18,77 SMD [95 %-KI]: -1,90 [-2,24; -1,57]	
Veränderung BSAS	Ergebnis (LS-MW): -17,35 vs. -5,41 SMD [95 %-KI]: -0,76 [-1,05; -0,48]	
Veränderung BDCAF	Patientenwahrnehmung: Ergebnis (LS-MW): -1,68 vs. -0,67 SMD [95 %-KI]: -0,64 [-0,92; -0,36]	
	Untersucherwahrnehmung: Ergebnis (LS-MW): -1,64 vs. -0,69 SMD [95 %-KI]: -0,65 [-0,93; -0,37]	

Endpunkte	Studie RELIEF (Apremilast vs. Placebo)	Ausmaß des Zusatznutzens
Gesundheitsbezogene Lebensqualität		
Veränderung BD-QoL	Ergebnis (LS-MW): -3,47 vs. -0,52 SMD [95 %-KI]: -0,51 [-0,79; -0,24]	beträchtlich
Verbesserung der gesundheitsbezogenen Lebensqualität anhand SF-36	Körperliche Rollenfunktion: Ergebnisse: 43,3 % vs. 24,3 % RR [95 %-KI]: 1,78 [1,19; 2,67] p-Wert = 0,0042	
	Körperliche Schmerzen Ergebnisse: 58,6 % vs. 42,7 % RR [95 %-KI]: 1,37 [1,04; 1,80] p-Wert = 0,0226	
	Allgemeine Gesundheitswahrnehmung: Ergebnisse: 47,1 % vs. 30,1 % RR [95 %-KI]: 1,57 [1,11; 2,24] p-Wert = 0,0322	
	Vitalität: Ergebnisse: 60,6 % vs. 41,8 % RR [95 %-KI]: 1,45 [1,11; 1,90] p-Wert = 0,0069	
	Soziale Funktionsfähigkeit: Ergebnisse: 51,9 % vs. 26,2 % RR [95 %-KI]: 1,99 [1,37; 2,88] p-Wert = 0,0002	
	Emotionale Rollenfunktion: Ergebnisse: 46,1 % vs. 32,0 % RR [95 %-KI]: 1,43 [1,01; 2,02] p-Wert = 0,0405	
	Psychisches Wohlbefinden: Ergebnisse: 54,8 % vs. 39,8 % RR [95 %-KI]: 1,42 [1,06; 1,89] p-Wert = 0,0322	
	Psychische Summenskala: Ergebnisse: 54,8 % vs. 33,0 % RR [95 %-KI]: 1,72 [1,26; 2,36] p-Wert = 0,0016	

Endpunkte	Studie RELIEF (Apremilast vs. Placebo)	Ausmaß des Zusatznutzens
Sicherheit		
UE gesamt	Ergebnisse: 78,8 % vs. 71,8 % RR [95 %-KI]: 1,10 [0,94; 1,28] p-Wert = 0,2436	kein geringerer oder höherer Schaden
SUE	Ergebnisse: 2,9 % vs. 3,9 % RR [95 %-KI]: 0,74 [0,17; 3,24] p-Wert = 0,6917	
Therapieabbruch aufgrund von UE	Ergebnisse: 2,9 % vs. 4,8 % RR [95 %-KI]: 0,59 [0,15; 2,42] p-Wert = 0,4633	
Schwere UE (CTCAE \geq 3)	Ergebnisse: 5,8 % vs. 5,8 % RR [95 %-KI]: 0,99 [0,33; 2,97] p-Wert = 0,9863	
UE von besonderem Interesse: Diarrhoe	Gesamtrate: Ergebnisse: 31,7 % vs. 17,5 % RR [95 %-KI]: 1,82 [1,10; 3,01] p-Wert = 0,0176	
UE bei mindestens 10 % der Patienten in einem Studienarm nach SOC und PT zu Woche 12	SOC: Erkrankungen des Gastrointestinaltrakts PT: Diarrhoe Ergebnisse: 41,3 % vs. 20,4 % RR [95 %-KI]: 2,03 [1,30; 3,16] p-Wert = 0,0011	
BDCAF: Behçet's Disease Current Activity Form; BD-QoL: Behçet's disease quality of life; BSAS: Behçet's Syndrome Activity Score; CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; LOCF: Letzte verfügbare Beobachtung (last observation carried forward); LS-MW: Kleinste-Quadrate-Mittelwert (least squares-mean); n.b.: nicht berechenbar; PT: Bevorzugter Begriff (preferred term); RR: Relatives Risiko; SF-36v2: 36-Item Short Form Health Survey, Version 2; SMD: Standardisierte Mittelwertdifferenz; SOC: Systemorganklasse (System Organ Class); SUE: Schwerwiegende unerwünschte/s Ereignis/se UE: Unerwünschte/s Ereignis/se; VAS: Visuelle Analogskala		

Mortalität

In der Studie RELIEF traten keine Todesfälle auf.

Morbidität

Vollständige Remission oraler Aphthen

Unter Apremilast erreichten mehr als doppelt so viele Patienten eine vollständige Remission der oralen Aphthen wie unter Placebo (56,7 % vs. 25,2 %). Das Relative Risiko für eine vollständige Remission der oralen Aphthen zwischen Apremilast und Placebo betrug 2,25 (95 %-KI: [1,55; 3,27]; $p < 0,0001$). Die Zeit bis zum Erreichen der vollständigen Remission wurde im Apremilast-Arm im Median nach etwa zwei Wochen erreicht. Bei Patienten, die Placebo erhielten, trat dies erst nach etwa acht Wochen ein. Der Unterschied entspricht einem

HR von 2,40 (95 %-KI: [1,69; 3,40]). Bei über 30 % der Patienten bildeten sich zudem nach vollständiger Remission bis Woche 12 keine neuen oralen Aphthen.

In einer zusätzlichen 52-wöchigen Extensionsphase wurden diese positiven Effekte von Apremilast bestätigt. Die zum Ende der Placebo-kontrollierten Phase (Woche 12) erzielte Remission der oralen Aphthen blieb für Patienten, die weiterhin Apremilast erhielten, bis zum Behandlungsende (Woche 64) erhalten. Patienten, die in der kontrollierten Phase Placebo erhielten, erhielten stattdessen unter Beibehaltung der Verblindung in der Extensionsphase ebenfalls Apremilast. Für dieses Patientenkollektiv wurde bis Woche 64 eine mit dem Apremilast/Apremilast-Arm vergleichbare Remission der oralen Aphthen berichtet. Dies verdeutlicht den raschen Wirkeintritt und die langanhaltende Wirksamkeit von Apremilast. Nach dem Absetzen der Behandlung mit Apremilast zu Woche 64 trat bei allen Patienten, unabhängig davon, ob diese in der Placebo-kontrollierten Phase mit Apremilast oder Placebo behandelt wurden, ein Wiederauftreten von oralen Aphthen ein, wodurch die Kausalität mit dem Behandlungserfolg unter Apremilast unterstrichen wird.

Schmerzhaftigkeit der oralen Aphthen

Die patientenberichtete Schmerzhaftigkeit der oralen Aphthen mittels VAS ging unter Apremilast im Vergleich zu Placebo um das 1,9-fache zurück (SMD = -1,90; 95 %-KI: [-2,24; -1,57]). In der Extensionsphase hielt dieser Effekt bis zum Behandlungsende in Woche 64 an. Patienten aus der Placebo-Gruppe, die zum Ende der Placebo-kontrollierten Phase Apremilast erhielten, zeigten eine vergleichbare Reduktion der Schmerzhaftigkeit der oralen Aphthen, wodurch die Kausalität mit dem Behandlungserfolg unter Apremilast unterstrichen wird.

Behçet's Syndrome Activity Score (BSAS) und Behçet's Disease Current Activity Score (BDCAF)

Die Ergebnisse der Symptom- und Schmerzerfassung werden durch die Ergebnisse der Fragebögen BSAS und BDCAF gestützt. Beide Fragebögen sind valide und krankheitsspezifische Instrumente zur Beurteilung der manifestationsübergreifenden, klinischen Aktivität des Behçet-Syndroms, die weitverbreitet eingesetzt werden. Die Ergebnisse bestätigen die vom Patienten empfundene, rasch einsetzende und starke Verbesserung der Krankheitssymptomatik durch Apremilast im gesamten Behandlungsverlauf. Der validierte patientenberichtete BSAS zeigte eine Verringerung der Krankheitssymptomatik unter Apremilast (SMD = -0,76; 95 %-KI: [-1,05; -0,48]). Der validierte Bogen BDCAF bestätigte einheitlich in allen Skalen eine Verringerung der Krankheitsaktivität (Index: SMD = -0,30; 95 %-KI: [-0,57; -0,02]; Patientenwahrnehmung: SMD = -0,64; 95 %-KI: [-0,92; -0,36]; Untersucherwahrnehmung: SMD = -0,65; 95 %-KI: [-0,93; -0,37]).

Die zum Ende der Placebo-kontrollierten Phase beobachtete Verbesserung in den drei Einzelkomponenten der Krankheitsaktivität anhand des BDCAF-Fragebogens in der Apremilast-Gruppe hielt bis zum Ende der Extensionsphase an. In der Gruppe der Patienten, die in der Placebo-kontrollierten Phase Placebo und in der Extensionsphase Apremilast

erhielten, verbesserte sich die Krankheitsaktivität zu Woche 64 ebenfalls in allen drei Einzelkomponenten und lag jeweils auf vergleichbarem Niveau zur Apremilast-Gruppe.

Gesamtschau Morbidität

Patienten, die am Behçet-Syndrom leiden, sind besonders von den Ausprägungen der oralen Aphthen und den damit verbundenen Schmerzen betroffen. Die signifikante Remission der oralen Aphthen sowie der ebenfalls signifikante Rückgang des mit den oralen Aphthen verbundene Schmerzempfinden zugunsten von Apremilast zur Remission der oralen Aphthen verdeutlichen die hohe therapeutische Wirksamkeit von Apremilast im AWG und den medizinischen Zusatznutzen. Die generelle Krankheitsaktivität durch das Behçet-Syndrom hat einen maßgeblichen Einfluss auf das Wohlbefinden der Patienten. Die validierten Fragebögen BSAS und BDCAF wurden durch Patienten und Untersucher unabhängig erhoben und zeigten einen signifikanten positiven Effekt von Apremilast auf die generelle Krankheitsaktivität des Behçet-Syndroms.

Gemäß AM-NutzenV § 5 Absatz 7 liegt somit ein **beträchtlicher** Zusatznutzen hinsichtlich der Morbidität vor.

Gesundheitsbezogene Lebensqualität

Behçet's disease quality of life (BD-QoL)

Die gesundheitsbezogene Lebensqualität anhand des krankheitsspezifischen Fragebogens BD-QoL verbesserte sich deutlich unter Apremilast (SMD = -0,51; 95 %-KI: [-0,79; -0,24]). Wie das Ansprechen hinsichtlich der Symptome des Behçet-Syndroms bleibt auch das erzielte positive Ergebnis am Ende der Placebo-kontrollierten Phase bei Patienten, die Apremilast erhielten, bis zum Behandlungsende in Woche 64 auf einem vergleichbaren Niveau. Patienten, die zu Woche 12 von Placebo auf Apremilast wechselten, erreichten in der Extensionsphase ein vergleichbares Ansprechen wie die Patienten in der Apremilast-Gruppe.

36-Item Short Form Health Survey, Version 2 (SF-36)

Bei den Ergebnissen des generischen Fragebogen SF-36 erreichten Patienten unter Apremilast in der Mehrheit der Subskalen eine patientenrelevante Verbesserung um $\geq 2,5$ Punkte. Dieser Wert wurde je nach Subskala bei 36,5 % (körperliche Funktionsfähigkeit) bis zu 60,6 % der Patienten (Vitalität) erreicht. Bei sieben von acht einzelnen Funktionsskalen sowie der Summenskala der psychischen Komponente wurden signifikante Vorteile zugunsten von Apremilast festgestellt. Besonders auffällig ist das sehr hohe Ansprechen im Bereich der sozialen Aspekte des Fragebogens, d. h., der Unterschied zwischen Apremilast und Placebo ist besonders ausgeprägt bei der sozialen Funktionsfähigkeit (RR: 1,99 [95%-KI: 1,37; 2,88] und der psychischen Summenskala (RR: 1,72 [95%-KI: 1,26; 2,36]).

Gesamtschau gesundheitsbezogene Lebensqualität

Die Symptomatik und Krankheitsaktivität des Behçet-Syndroms beeinflussen unmittelbar die gesundheitsbezogene Lebensqualität der betroffenen Patienten. Neben einer Linderung der Symptomatik und den Ausprägungen der Manifestationen ist somit die Verbesserung der Lebensqualität der Patienten ein wichtiges Behandlungsziel. Die anhand des BD-QoL

Fragebogens erfasste krankheitsspezifische gesundheitsbezogene Lebensqualität verbesserte sich bis zum Ende der kontrollierten Studienphase wesentlich. Diese Ergebnisse werden durch den generischen Fragebogen SF-36 gestützt. Bei sieben der acht Funktionsskalen sowie der MCS wurde eine statistisch signifikante und patientenrelevante Verbesserung der gesundheitsbezogenen Lebensqualität unter Apremilast gegenüber Placebo berichtet.

Die patientenrelevanten Effekte in der Kategorie gesundheitsbezogene Lebensqualität wurden in einer zusätzlichen 52-wöchigen Extensionsphase bestätigt: Sowohl bei Patienten, die statt Placebo eine Apremilast-Behandlung erhielten als auch bei zuvor mit Apremilast behandelten Patienten.

Gemäß AM-NutzenV § 5 Absatz 7 liegt somit ein **beträchtlicher** Zusatznutzen hinsichtlich der gesundheitsbezogenen Lebensqualität vor.

Sicherheit

Das am häufigsten berichtete UE in der Placebo-kontrollierten Phase der Studie war Diarrhoe (41,3 % Apremilast vs. 20,7 % Placebo). In keiner der Fälle wurde die Diarrhoe als schwer (CTCAE ≥ 3) oder schwerwiegend berichtet. In allen weiteren einzelnen UE nach SOC, PT und von besonderem Interesse ergaben sich keine Unterschiede zwischen den beiden Studienarmen. Zusammengefasste UE der Kategorien gesamt, schwer und schwerwiegend sowie jene, die zum Therapieabbruch führten, wiesen vergleichbare Werte auf. Insbesondere da es sich um einen Placebovergleich handelt, liegt ein sehr ausgeglichenes Sicherheitsprofil von Apremilast vor.

Gesamtschau Sicherheit

Insgesamt ist damit im Hinblick auf die Nutzendimension Sicherheit **kein geringerer oder höherer Schaden** belegt.

Weitere Aspekte

Subgruppenanalysen

In keiner der Subgruppenanalysen konnten Belege für eine Modifikation des Behandlungseffektes bzgl. der Mortalität, Morbidität, Sicherheit und gesundheitsbezogene Lebensqualität festgestellt werden (siehe Abschnitt 4.3.1.3.2). Für die Ableitung des Zusatznutzens werden daher die Ergebnisse der Gesamtpopulation herangezogen.

Übertragbarkeit der Ergebnisse

Die Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext ist gegeben (siehe Abschnitt 4.3.1.2.1).

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Insgesamt ergibt sich auf Basis der Kriterien in § 5 Absatz 7 der AM-NutzenV unter Berücksichtigung des Schweregrades der Erkrankung für Apremilast im Vergleich zur zVT „Therapie nach Maßgabe des Arztes“ ein **Hinweis für einen beträchtlichen Zusatznutzen**.

International und auch auf europäischer Ebene gibt es außer Apremilast kein weiteres Arzneimittel, das zur Behandlung von oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind, zugelassen ist. Bei Patienten, für die eine systemische Therapie infrage kommt, zeigte die Studie RELIEF, dass Apremilast im Vergleich zu Placebo:

- signifikant häufiger und früher zu einer vollständigen Remission der oralen Aphthen führt,
- die mit der Krankheit und ihrer Ausprägung verbundenen Schmerzen und Krankheitsaktivität signifikant verringert,
- die gesundheitsbezogene Lebensqualität verbessert und
- ein sehr gutes Sicherheitsprofil aufweist.

Diese positiven Effekte konnten zudem bei einem Wechsel von Placebo zu Apremilast oder bei Fortführung der Therapie über weitere 52 Wochen bestätigt und beibehalten werden.

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 4-101: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit oralen Aphthen, die mit dem Behçet-Syndrom assoziiert sind und für die eine systemische Therapie infrage kommt	Hinweis auf einen beträchtlichen 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.

Es wurde kein adjustierter indirekter Vergleich durchgeführt.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht 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

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.

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.

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. Alpsy E. 2016. *Behçet's disease: A comprehensive review with a focus on epidemiology, etiology and clinical features, and management of mucocutaneous lesions*. The Journal of Dermatology 43 (6), S. 620–632.
2. Alpsy E., Donmez L., Onder M. et al. 2007. *Clinical features and natural course of Behçet's disease in 661 cases: a multicentre study*. The British Journal of Dermatology 157 (5), S. 901–906.

²⁰ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

3. Alpsy E., Durusoy C., Yilmaz E. et al. 2002. *Interferon alfa-2a in the treatment of Behçet's disease: A randomized placebo-controlled and double-blind study*. Archives of Dermatology 138 (4), S. 467–471.
4. Altenburg A., Mahr A., Maldini C. et al. 2012. *Epidemiologie und Klinik des Morbus Adamantiades-Beçet in Deutschland. Aktuelle Daten*. Der Ophthalmologe : Zeitschrift der Deutschen Ophthalmologischen Gesellschaft 109 (6), S. 531–541.
5. Amgen Europe B. V. (Amgen) 2014a. *BDQOL - Language: German* Mapi: Screen Shots Certificate. Data on file.
6. Amgen Europe B. V. (Amgen) 2014b. *BDQOL - Language: Turkish* Mapi: Screen Shots Certificate. Data on file.
7. Amgen Europe B. V. (Amgen) 2014c. *BSAS - Language: German* Mapi: Screen Shots Certificate. Data on file.
8. Amgen Europe B. V. (Amgen) 2017. *Apremilast (CC-10004) Clinical Development Program - Data Monitoring Committee Charter Version 5.0*. Data on file.
9. Amgen Europe B. V. (Amgen) 2018a. *Clinical Study Report CC-10004-BCT-002 (Apremilast): A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Group Study Followed By An Active Treatment Phase To Evaluate The Efficacy And Safety of Apremilast (CC-10004) In The Treatment Of Subjects With Active Behçet's Disease* CC-10004-BCT-002_Table 14. Data on file.
10. Amgen Europe B. V. (Amgen) 2018b. *Otezla New Indication Pre-submission Meeting - 7 November 2018* HPRA Meeting minutes. Data on file.
11. Amgen Europe B. V. (Amgen) 2019. *Clinical Study Report CC-10004-BCT-002 (Apremilast): A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Group Study Followed By An Active Treatment Phase To Evaluate The Efficacy And Safety of Apremilast (CC-10004) In The Treatment Of Subjects With Active Behçet's Disease* CC-10004-BCT-002. Data on file.
12. Amgen Europe B. V. (Amgen) 2020a. *Apremilast Modul 4 - Post-hoc-Analysen*. Data on file.
13. Amgen Europe B. V. (Amgen) 2020b. *Fachinformation Otezla® Filmtabletten: Stand: April 2020*. Verfügbar unter: www.fachinfo.de, abgerufen am: 29.04.2020.
14. Bellach B.-M., Ellert U. und Radoschewski M. 2000. *Epidemiologie des Schmerzes – Ergebnisse des Bundes-Gesundheitssurveys 1998*. Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz 43 (6), S. 424–431.
15. Bhakta B. B., Brennan P., James T. E. et al. 1999. *Behçet's disease: evaluation of a new instrument to measure clinical activity*. Rheumatology (Oxford, England) 38 (8), S. 728–733.
16. Bodur H., Borman P., Ozdemir Y. et al. 2006. *Quality of life and life satisfaction in patients with Behçet's disease: relationship with disease activity*. Clinical rheumatology 25 (3), S. 329–333.

17. Bucher H. C., Guyatt G. H., Griffith L. E. et al. 1997. *The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials*. Journal of clinical epidemiology 50 (6), S. 683–691.
18. Bullinger M., Kirchberger I. und Ware J. 1995. *Der deutsche SF-36 Health Survey. Übersetzung und psychometrische Testung eines krankheitsübergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualität*. Z. f. Gesundheitswiss. 3 (1), S. 21–36.
19. Bundesministerium der Justiz und für Verbraucherschutz (BMJV) 2019. *Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV)*. Verfügbar unter: <https://www.gesetze-im-internet.de/am-nutzenv/BJNR232400010.html>, abgerufen am: 09.12.2019.
20. Canpolat O. und Yurtsever S. 2011. *The Quality of Life in Patients With Behçet's Disease*. Asian nursing research 5 (4), S. 229–235.
21. Choi H. J., Seo M. R., Ryu H. J. et al. 2015. *Cross-cultural adaptation and validation of the Behçet's Disease Current Activity Form in Korea*. The Korean journal of internal medicine 30 (5), S. 714–718.
22. Choi H. J., Seo M. R., Ryu H. J. et al. 2016. *Validation and reliability of a Behçet's Syndrome Activity Scale in Korea*. The Korean journal of internal medicine 31 (1), S. 170–175.
23. ClinicalTrials.gov 2014a. *A Phase 3 Randomized, Double-blind Study to Evaluate the Efficacy and Safety of Apremilast (CC-10004) in Subjects With Active Behçet's Disease*. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/study/NCT02307513?term=apremilast&cond=Beh%C3%A7et&rank=1>, abgerufen am: 20.03.2020.
24. ClinicalTrials.gov 2014b. *A Study to Evaluate the Efficacy and Safety of Apremilast (CC-10004) in the Treatment of Behçet Disease*. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT00866359>, abgerufen am: 20.03.2020.
25. Davatchi F. 2014. *Behçet's disease*. International journal of rheumatic diseases 17 (4), S. 355–357.
26. Davatchi F., Shahram F., Chams-Davatchi C. et al. 2010. *Behçet's disease: from East to West*. Clinical rheumatology 29 (8), S. 823–833.
27. Dworkin R. H., Turk D. C., Wyrwich K. W. et al. 2008. *Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations*. The journal of pain: official journal of the American Pain Society 9 (2), S. 105–121.
28. Ellert U. und Kurth B. M. 2004. *Methodische Betrachtungen zu den Summenscores des SF-36 anhand der erwachsenen bundesdeutschen Bevölkerung*. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 47, S. 1027–1032.

29. Ellert U. und Kurth B. M. 2013. *Gesundheitsbezogene Lebensqualität bei Erwachsenen in Deutschland*. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 56, S. 643–649.
30. Ertam I., Kitapcioglu G., Aksu K. et al. 2009. *Quality of life and its relation with disease severity in Behçet's disease*. Clinical and experimental rheumatology 27 (2 Suppl 53), S. S18-22.
31. European Union Clinical Trials Register (EU-CTR) 2014. *A PHASE 3, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP STUDY, FOLLOWED BY AN ACTIVE-TREATMENT PHASE TO EVALUATE THE EFFICACY AND SAFETY OF APREMILAST (CC-10004) IN THE TREATMENT OF SUBJECTS WITH ACTIVE BEHÇET'S DISEASE*. Verfügbar unter: <https://www.clinicaltrialsregister.eu/ctr-search/search?query=%28%22Beh%20C3%A7et%22+OR+%E2%80%9CBehcet%E2%80%9D%29+AND+%28%E2%80%9CApremilast%E2%80%9D+OR+%E2%80%9COtezla%E2%80%9D+OR+%E2%80%9C608141-41-9%E2%80%9D+OR+%E2%80%9CCC10004%E2%80%9D+OR+%E2%80%9CCC-10004%E2%80%9D%29>, abgerufen am: 20.03.2020.
32. Fabiani C., Vitale A., Orlando I. et al. 2017. *Quality of life impairment in Behçet's disease and relationship with disease activity: a prospective study*. Internal and emergency medicine 12 (7), S. 947–955.
33. Gemeinsamer Bundesausschuss (G-BA) 2018. *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 – Guselkumab*. Verfügbar unter: https://www.g-ba.de/downloads/40-268-4987/2018-05-17_AM-RL-XII_Guselkumab_D-330_TrG.pdf, abgerufen am: 18.09.2019.
34. Gemeinsamer Bundesausschuss (G-BA) 2019. *Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV Beratungsanforderung 2019-B-136* Korrespondenz. Data on file.
35. Gemeinsamer Bundesausschuss (G-BA) 2020. *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 und zuletzt geändert am 19. September 2019, veröffentlicht im Bundesanzeiger BAnz AT 13.01.2020 B5, in Kraft getreten am 14. Januar 2020*. Verfügbar unter: https://www.g-ba.de/downloads/62-492-2021/VerfO_2019-09-19_iK-2020-01-14.pdf, abgerufen am: 14.04.2020.
36. Gilworth G., Chamberlain M. A., Bhakta B. B. et al. 2004. *Development of the BD-QoL: a quality of life measure specific to Behçet's disease*. The Journal of Rheumatology 31 (5), S. 931–937.
37. Greco A., Virgilio A. de, Ralli M. et al. 2018. *Behçet's disease: New insights into pathophysiology, clinical features and treatment options*. Autoimmunity Reviews 17 (6), S. 567–575.

38. Hamuryudan V., Fresko I., Direskeneli H. et al. 1999. *Evaluation of the Turkish translation of a disease activity form for Behçet's syndrome*. Rheumatology (Oxford, England) 38 (8), S. 734–736.
39. Hamuryudan V., Mat C., Saip S. et al. 1998. *Thalidomide in the treatment of the mucocutaneous lesions of the Behçet syndrome. A randomized, double-blind, placebo-controlled trial*. Annals of Internal Medicine 128 (6), S. 443–450.
40. Harzallah O., Meksi Y., Hamzaoui A. et al. 2013. *Evaluation of a new Tunisian version of behçet's disease current activity form**. Open Journal of Internal Medicine 03 (04), S. 135–140.
41. Hatemi G., Christensen R., Bang D. et al. 2018. *2018 update of the EULAR recommendations for the management of Behçet's syndrome*. Annals of the Rheumatic Diseases 77 (6), S. 808–818.
42. Hatemi G., Mahr A., Ishigatsubo Y. et al. 2019. *Trial of Apremilast for Oral Ulcers in Behçet's Syndrome*. The New England Journal of Medicine 381 (20), S. 1918–1928.
43. Hatemi G., Silman A., Bang D. et al. 2009. *Management of Behçet disease: a systematic literature review for the European League Against Rheumatism evidence-based recommendations for the management of Behçet disease*. Annals of the Rheumatic Diseases 68 (10), S. 1528–1534.
44. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2017. *Allgemeine Methoden: Version 5.0 vom 10.07.2017*. Verfügbar unter: <https://www.iqwig.de/de/methoden/methodenpapier.3020.html>, abgerufen am: 22.05.2019.
45. International Clinical Trials Registry Platform (ICTRP) 2014. *A Study to Evaluate the Efficacy and Safety of Apremilast (CC-10004) in the Treatment of Behçet Disease*. Verfügbar unter: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00866359>, abgerufen am: 20.03.2020.
46. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2011. *Introductory Guide: MedDRA Version 14.0*. Verfügbar unter: https://www.who.int/medical_devices/innovation/MedDRAintroguide_version14_0_March2011.pdf, abgerufen am: 18.09.2019.
47. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2017. *MedDRA Distribution: Format Document - Version 20.0*. Verfügbar unter: https://www.meddra.org/sites/default/files/guidance/file/dist_file_format_20_0_english.pdf, abgerufen am: 19.09.2019.
48. Lawton G., Bhakta B. B., Chamberlain M. A. et al. 2004. *The Behçet's disease activity index*. Rheumatology (Oxford, England) 43 (1), S. 73–78.
49. Leccese P., Ozguler Y., Christensen R. et al. 2019. *Management of skin, mucosa and joint involvement of Behçet's syndrome: A systematic review for update of the EULAR recommendations for the management of Behçet's syndrome // Management of skin,*

mucosa and joint involvement of Behçet's syndrome: A systematic review for update of the EULAR recommendations for the management of Behçet's syndrome. Seminars in arthritis and rheumatism 4 // 48 (4), S. 752–762.

50. Main D. M. und Chamberlain M. A. 1992. *Clinical differentiation of oral ulceration in Behcet's disease.* British Journal of Rheumatology 31 (11), S. 767–770.
51. Melikoğlu M., Fresko I., Mat C. et al. 2005. *Short-term trial of etanercept in Behçet's disease: A double blind, placebo controlled study.* The Journal of Rheumatology 32 (1), S. 98–105.
52. Melikoğlu M. und Melikoğlu M. A. 2014. *What affects the quality of life in patients with Behcet's disease?* Acta reumatologica portuguesa 39 (1), S. 46–53.
53. Morfeld M. und Bullinger M. 2008. *Der SF36 Health Survey zur Erhebung und Dokumentation gesundheitsbezogener Lebensqualität.* Physikalische Medizin Rehabilitationsmedizin Kurortmedizin - PHYSIK MED REHABIL KURORTMEDI 18, S. 250–255.
54. Morfeld M., Bullinger M., Nantke J. et al. 2005. *Die Version 2.0 des SF-36 Health Survey - Ergebnisse einer bevölkerungsrepräsentativen Studie.* Sozial- und Präventivmedizin SPM 50, S. 292–300.
55. Moses Alder N., Fisher M. und Yazici Y. 2008. *Behçet's syndrome patients have high levels of functional disability, fatigue and pain as measured by a Multi-dimensional Health Assessment Questionnaire (MDHAQ).* Clinical and experimental rheumatology 26 (4 Suppl 50), S. S110-3.
56. Mumcu G., Inanc N., Ergun T. et al. 2006. *Oral health related quality of life is affected by disease activity in Behçet's disease.* Oral diseases 12 (2), S. 145–151.
57. Mumcu G., Niazi S., Stewart J. et al. 2009. *Oral health and related quality of life status in patients from UK and Turkey: a comparative study in Behcet's disease.* Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology 38 (5), S. 406–409.
58. Neves F. d. S., Moraes J. C. B. de, Kowalski S. C. et al. 2007. *Cross-cultural adaptation of the Behcet's Disease Current Activity Form (BDCAF) to Brazilian Portuguese language.* Clinical rheumatology 26 (8), S. 1263–1267.
59. Saenz A., Ausejo M., Shea B. et al. 2000. *Pharmacotherapy for Behcet's syndrome.* The Cochrane database of systematic reviews k. A. (2), S. CD001084.
60. Schomacher J. 2008. *Gütekriterien der visuellen Analogskala zur Schmerzbewertung.* Physioscience 4 (-), S. 125–133.
61. Sibley C., Yazici Y., Tascilar K. et al. 2014. *Behcet syndrome manifestations and activity in the United States versus Turkey -- a cross-sectional cohort comparison.* The Journal of Rheumatology 41 (7), S. 1379–1384.
62. Taylor J., Glenny A.-M., Walsh T. et al. 2014. *Interventions for the management of oral ulcers in Behçet's disease: Review.* Verfügbar unter:

<https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011018.pub2/epdf/full>,
abgerufen am: 07.05.2019.

63. Touma Z., Ghandour L., Sibai A. et al. 2011. *Cross-cultural adaptation and validation of Behcet's disease quality of life questionnaire*. BMC medical research methodology 11, S. 52.
64. Ware J. E., JR 2000. *SF-36 health survey update*. Spine 25 (24), S. 3130–3139.
65. Ware J. E., JR und Sherbourne C. D. 1992. *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection*. Medical care 30 (6), S. 473–483.
66. Wong S. S.-L., Wilczynski N. L. und Haynes R. B. 2006. *Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE*. Journal of the Medical Library Association : JMLA 94 (4), S. 451–455.
67. World Health Organization (WHO) 2019. *Body Mass Index (BMI)*. Verfügbar unter: <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>, abgerufen am: 09.12.2019.
68. Yazici H., Pazarli H., Barnes C. G. et al. 1990. *A controlled trial of azathioprine in Behçet's syndrome*. The New England Journal of Medicine 322 (5), S. 281–285.
69. Yazici H., Tüzün Y., Pazarli H. et al. 1984. *Influence of age of onset and patient's sex on the prevalence and severity of manifestations of Behçet's syndrome*. Annals of the Rheumatic Diseases 43 (6), S. 783–789.
70. Yi S. W., Kim J.-H., Lim K.-Y. et al. 2008. *The Behcet's Disease Quality of Life: reliability and validity of the Korean version*. Yonsei medical journal 49 (5), S. 698–704.
71. Yilmaz S., Simşek I., Cinar M. et al. 2013. *Patient-driven assessment of disease activity in Behçet's syndrome: cross-cultural adaptation, reliability and validity of the Turkish version of the Behçet's Syndrome Activity Score*. Clinical and experimental rheumatology 31 (3 Suppl 77), S. 77–83.
72. Zouboulis C. C., Kötter I., Djawari D. et al. 1997. *Epidemiological features of Adamantiades-Behçet's disease in Germany and in Europe*. Yonsei medical journal 38 (6), S. 411–422.

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	March 20, 2020	
Zeitsegment	1974 to 2020 March 19	
Suchfilter	Filter for randomized controlled trials according to Wong 2006 - Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Behcet disease/	15050
2	(Behcet* disease or Behcet*-disease).af.	17360
3	(Behcet* syndrom* or Behcet*-syndrom*).af.	2477
4	(Gilbert* Behcet* syndrom* or Gilbert*-Behcet* syndrom* or Gilbert* Behcet* disease or Gilbert*-Behcet* disease).af.	0
5	Morbus Behcet.af.	120
6	(Adamantiades* adj1 Behcet* disease).af.	193
7	(Adamantiades* adj1 Behcet* syndrom*).af.	13
8	Behcet* ulcer*.af.	7
9	Malignant aphthosis.af.	3
10	Silk road disease.af.	23
11	exp mouth ulcer/ or mouth ulcer*.af.	15990
12	oral ulcer*.af.	4071
13	exp aphthous stomatitis/ or aphthous stomatitis.af.	4820
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	32339

Intervention		
15	exp Apremilast/	1862
16	Apremilast.af.	1917
17	Otezla.af.	106
18	608141-41-9.af.	1582
19	CC10004.af.	1
20	CC-10004.af.	52
21	15 or 16 or 17 or 18 or 19 or 20	1938
Studiendesign		
22	random*.tw.	1513396
23	placebo*.mp.	449451
24	double-blind*.tw.	207191
25	22 or 23 or 24	1766166
Zusammenfassung und Limitationen		
26	14 and 21 and 25	47
27	remove duplicates from 26	46

Datenbankname	MEDLINE(R)	
Suchoberfläche	Ovid	
Datum der Suche	March 20, 2020	
Zeitsegment	1946 to March 19, 2020	
Suchfilter	Filter for randomized controlled trials according to Wong 2006 - Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Behcet Syndrome/	9009
2	(Behcet* disease or Behcet*-disease).af.	8721
3	(Behcet* syndrom* or Behcet*-syndrom*).af.	9314
4	(Gilbert* Behcet* syndrom* or Gilbert*-Behcet* syndrom* or Gilbert* Behcet* disease or Gilbert*-Behcet* disease).af.	3
5	Morbus Behcet.af.	92
6	(Adamantiades* adj1 Behcet* disease).af.	145
7	(Adamantiades* adj1 Behcet* syndrom*).af.	11
8	Behcet* ulcer*.af.	5
9	Malignant aphthosis.af.	4
10	Silk road disease.af.	11
11	mouth ulcer*.af.	470
12	exp Oral Ulcer/ or oral ulcer*.af.	3946
13	exp Stomatitis, Aphthous/ or aphthous stomatitis.af.	3850
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	17790
Intervention		
15	Apremilast.af.	587
16	Otezla.af.	30
17	608141-41-9.af.	1
18	CC10004.af.	1
19	CC-10004.af.	7
20	15 or 16 or 17 or 18 or 19	590

Studiendesign		
21	randomized controlled trial.pt.	502283
22	randomized.mp.	830932
23	placebo.mp.	212981
24	21 or 22 or 23	893777
Zusammenfassung und Limitationen		
25	14 and 20 and 24	8
26	remove duplicates from 25	8

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	March 20, 2020	
Zeitsegment	February 2020	
Suchfilter	-	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Behcet Syndrome/	125
2	(Behcet* disease or Behcet*-disease).af.	289
3	(Behcet* syndrom* or Behcet*-syndrom*).af.	166
4	(Gilbert* Behcet* syndrom* or Gilbert*-Behcet* syndrom* or Gilbert* Behcet* disease or Gilbert*-Behcet* disease).af.	0
5	Morbus Behcet.af.	0
6	(Adamantiades* adj1 Behcet* disease).af.	1
7	(Adamantiades* adj1 Behcet* syndrom*).af.	0
8	Behcet* ulcer*.af.	0
9	Malignant aphthosis.af.	0
10	Silk road disease.af.	0
11	mouth ulcer*.af.	461
12	exp Oral Ulcer/ or oral ulcer*.af.	317
13	exp Stomatitis, Aphthous/ or aphthous stomatitis.af.	538
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	1421
Intervention		
15	Apremilast.af.	432
16	Otezla.af.	13
17	608141-41-9.af.	20
18	CC10004.af.	0
19	CC-10004.af.	52
20	15 or 16 or 17 or 18 or 19	434
Zusammenfassung und Limitationen		
21	14 and 20	28
22	remove duplicates from 21	25

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Datenbankname	Embase	
Suchoberfläche	Ovid	
Datum der Suche	March 20, 2020	
Zeitsegment	1974 to 2020 March 19	
Suchfilter	Filter for randomized controlled trials according to Wong 2006 - Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Behcet disease/	15050
2	(Behcet* disease or Behcet*-disease).af.	17360
3	(Behcet* syndrom* or Behcet*-syndrom*).af.	2477
4	(Gilbert* Behcet* syndrom* or Gilbert*-Behcet* syndrom* or Gilbert* Behcet* disease or Gilbert*-Behcet* disease).af.	0
5	Morbus Behcet.af.	120
6	(Adamantiades* adj1 Behcet* disease).af.	193
7	(Adamantiades* adj1 Behcet* syndrom*).af.	13
8	Behcet* ulcer*.af.	7
9	Malignant aphthosis.af.	3
10	Silk road disease.af.	23
11	exp mouth ulcer/ or mouth ulcer*.af.	15990
12	oral ulcer*.af.	4071
13	exp aphthous stomatitis/ or aphthous stomatitis.af.	4820
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	32339

Intervention		
15	exp azathioprine/ or (Azathioprin* or BW 57-322 or 446-86-6).af.	94057
16	exp prednisolone/ or (Prednisolon* or Metacortandralon or Deltahydrocortison or 50-24-8).af.	135170
17	exp prednisone/ or (Prednison* or Metacortandracin or 53-03-2).af.	173065
18	exp thalidomide/ or (Thalidomid* or 50-35-1).af.	28477
19	exp alpha interferon/ or (Interferon alfa* or Interferon alpha* or Interferon-alpha* or Interferon-alfa* or alpha*Interferon or alfa*Interferon or IFN alpha* or IFN alfa* or IFN-alpha* or IFN-alfa*).af.	84188
20	exp infliximab/ or (Infliximab or TA-650 or 170277-31-3).af. or exp etanercept/ or (Etanercept or TNFR-Fc* or 185243-69-0).af. or exp adalimumab/ or (Adalimumab or 331731-18-1).af. or exp certolizumab pegol/ or (Certolizumab* or CDP870 or 428863-50-7).af. or exp golimumab/ or (Golimumab or CNTO148 or 476181-74-5).af. or exp tumor necrosis factor inhibitor/ or exp tumor necrosis factor antibody/ or (Tumor Necrosis Factor* inhibit* or TNF* inhibit*).af.	97931
21	exp cyclosporine/ or (Cyclosporin* or Ciclosporin* or 59865-13-3).af.	158683
22	exp dapsone/ or (Dapson* or Diaphenylsulfon* or 80-08-0).af.	19110
23	15 or 16 or 17 or 18 or 19 or 20 or 21 or 22	613874
Studiendesign		
24	random*.tw.	1513396
25	placebo*.mp.	449451
26	double-blind*.tw.	207191
27	24 or 25 or 26	1766166
Zusammenfassung und Limitationen		
28	14 and 23 and 27	752
29	remove duplicates from 28	743

Datenbankname	MEDLINE(R)	
Suchoberfläche	Ovid	
Datum der Suche	March 20, 2020	
1946 to March 19, 2020	1946 to March 19, 2020	
Suchfilter	Filter for randomized controlled trials according to Wong 2006 - Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Behcet Syndrome/	9009
2	(Behcet* disease or Behcet*-disease).af.	8721
3	(Behcet* syndrom* or Behcet*-syndrom*).af.	9314
4	(Gilbert* Behcet* syndrom* or Gilbert*-Behcet* syndrom* or Gilbert* Behcet* disease or Gilbert*-Behcet* disease).af.	3
5	Morbus Behcet.af.	92
6	(Adamantiades* adj1 Behcet* disease).af.	145
7	(Adamantiades* adj1 Behcet* syndrom*).af.	11
8	Behcet* ulcer*.af.	5
9	Malignant aphthosis.af.	4
10	Silk road disease.af.	11
11	mouth ulcer*.af.	472
12	exp Oral Ulcer/ or oral ulcer*.af.	3948
13	exp Stomatitis, Aphthous/ or aphthous stomatitis.af.	3850
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	17792
Intervention		
15	exp Azathioprine/ or (Azathioprin* or BW 57-322 or 446-86-6).af.	22890
16	exp Prednisolone/ or (Prednisolon* or Metacortandralon or Deltahydrocortison or 50-24-8).af.	62899
17	exp Prednisone/ or (Prednison* or Metacortandracin or 53-03-2).af.	52919
18	exp Thalidomide/ or (Thalidomid* or 50-35-1).af.	11367
19	exp Interferon-alpha/ or (Interferon alfa* or Interferon alpha* or Interferon-alpha* or Interferon-alfa* or alpha*Interferon or	45132

	alfa*Interferon or IFN alpha* or IFN alfa* or IFN-alpha* or IFN-alfa*).af.	
20	exp Infliximab/ or (Infliximab or TA-650 or 170277-31-3).af. or exp Etanercept/ or (Etanercept or TNFR-Fc* or 185243-69-0).af. or exp Adalimumab/ or (Adalimumab or 331731-18-1).af. or exp Certolizumab Pegol/ or (Certolizumab* or CDP870 or 428863-50-7).af. or (Golimumab or CNTO148 or 476181-74-5).af. or exp Tumor Necrosis Factor-alpha/ or (Tumor Necrosis Factor* inhibit* or TNF* inhibit*).af.	139572
21	exp Cyclosporine/ or (Cyclosporin* or Ciclosporin* or 59865-13-3).af.	58638
22	exp Dapsone/ or (Dapson* or Diaphenylsulfon* or 80-08-0).af.	6549
23	15 or 16 or 17 or 18 or 19 or 20 or 21 or 22	364150
Studiendesign		
24	randomized controlled trial.pt.	502288
25	randomized.mp.	830970
26	placebo.mp.	212994
27	24 or 25 or 26	893825
Zusammenfassung und Limitationen		
28	14 and 23 and 27	139
29	remove duplicates from 28	138

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	March 20, 2020	
Zeitsegment	February 2020	
Suchfilter	-	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Behcet Syndrome/	125
2	(Behcet* disease or Behcet*-disease).af.	289
3	(Behcet* syndrom* or Behcet*-syndrom*).af.	166
4	(Gilbert* Behcet* syndrom* or Gilbert*-Behcet* syndrom* or Gilbert* Behcet* disease or Gilbert*-Behcet* disease).af.	0
5	Morbus Behcet.af.	0
6	(Adamantiades* adj1 Behcet* disease).af.	1
7	(Adamantiades* adj1 Behcet* syndrom*).af.	0
8	Behcet* ulcer*.af.	0
9	Malignant aphthosis.af.	0
10	Silk road disease.af.	0
11	mouth ulcer*.af.	461
12	exp Oral Ulcer/ or oral ulcer*.af.	317
13	exp Stomatitis, Aphthous/ or aphthous stomatitis.af.	538
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	1421
Intervention		
15	exp Azathioprine/ or (Azathioprin* or BW 57-322 or 446-86-6).af.	3388
16	exp Prednisolone/ or (Prednisolon* or Metacortandralon or Deltahydrocortison or 50-24-8).af.	8911
17	exp Prednisone/ or (Prednison* or Metacortandracin or 53-03-2).af.	9772
18	exp Thalidomide/ or (Thalidomid* or 50-35-1).af.	1920
19	exp Interferon-alpha/ or (Interferon alfa* or Interferon alpha* or Interferon-alpha* or Interferon-alfa* or alpha*Interferon or alfa*Interferon or IFN alpha* or IFN alfa* or IFN-alpha* or IFN-alfa*).af.	6694

20	(Infliximab or TA-650 or 170277-31-3 or (Etanercept or TNFR-Fc* or 185243-69-0) or (Adalimumab or 331731-18-1) or (Certolizumab* or CDP870 or 428863-50-7) or (Golimumab or CNTO148 or 476181-74-5)).af. or exp Tumor Necrosis Factor-alpha/ or (Tumor Necrosis Factor* inhibit* or TNF* inhibit*).af.	10759
21	exp Cyclosporine/ or (Cyclosporin* or Ciclosporin* or 59865-13-3).af.	7651
22	exp Dapsone/ or (Dapson* or Diaphenylsulfon* or 80-08-0).af.	524
23	15 or 16 or 17 or 18 or 19 or 20 or 21 or 22	43107
Zusammenfassung und Limitationen		
24	14 and 23	272
25	remove duplicates from 24	264

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	clinicaltrials.gov
Internetadresse	https://clinicaltrials.gov/
Datum der Suche	20.03.2020
Eingabeoberfläche	Standard Search
Suchstrategie	(Behcet OR Behçet) AND (Apremilast OR Otezla OR 608141-41-9 OR CC10004 OR CC-10004)
Treffer	2

Studienregister	ICTRP (WHO)
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	20.03.2020
Eingabeoberfläche	Standard Search
Suchstrategie	Behcet AND Apremilast OR Behcet AND Otezla OR Behcet AND 608141-41-9 OR Behcet AND CC10004 OR Behcet AND CC-10004 OR Behçet AND Apremilast OR Behçet AND Otezla OR Behçet AND 608141-41-9 OR Behçet AND CC10004 OR Behçet AND CC-10004
Treffer	1

Studienregister	EU Clinical Trials Register
Internetadresse	www.clinicaltrialsregister.eu
Datum der Suche	20.03.2020
Eingabeoberfläche	Standard Search
Suchstrategie	("Behcet" OR "Behçet") AND ("Apremilast" OR "Otezla" OR "608141-41-9" OR "CC10004" OR "CC-10004")
Treffer	1

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Studienregister	clinicaltrials.gov
Internetadresse	https://clinicaltrials.gov/
Datum der Suche	20.03.2020
Eingabeoberfläche	Standard Search
Suchstrategie	(Behcet OR Behçet) AND (Azathioprine OR BW 57-322 OR 446-86-6) (4 Treffer)
	(Behcet OR Behçet) AND (Prednisolone OR Metacortandralon OR Deltahydrocortison OR 50-24-8) (5 Treffer)
	(Behcet OR Behçet) AND (Prednisone OR Metacortandracin OR 53-03-2) (6 Treffer)
	(Behcet OR Behçet) AND (Thalidomid OR Thalidomide OR 50-35-1) (3 Treffer)
	(Behcet OR Behçet) AND (Interferon alfa* OR Interferon alpha* OR Interferon-alpha* OR Interferon-alfa* OR alpha*Interferon OR alfa*Interferon OR IFN alpha* OR IFN alfa* OR IFN-alpha* OR IFN-alfa*) (3 Treffer)
	(Behcet OR Behçet) AND (Infliximab OR 170277-31-3 OR Etanercept OR TNFR-Fc* OR 185243-69-0 OR Adalimumab OR 331731-18-1 OR Certolizumab* OR CDP870 OR 428863-50-7 OR Golimumab OR CNTO148 OR 476181-74-5) (12 Treffer)
	(Behcet OR Behçet) AND (Tumor Necrosis Factor* α * OR Tumor Necrosis Factor*alpha* OR Tumor Necrosis Factor*alfa* OR TNF* α * or TNF*alpha* OR TNF*alfa* OR anti*TNF) (6 Treffer)
	(Behcet OR Behçet) AND (Cyclosporine OR Ciclosporin OR 59865-13-3) (3 Treffer)
	(Behcet OR Behçet) AND (Dapsone OR Diaphenylsulfone OR 80-08-0) (0 Treffer)
Treffer	42

Studienregister	ICTRP (WHO)
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	15.04.2020
Eingabeoberfläche	Standard Search
Suchstrategie	Behcet AND Azathioprine OR Behcet AND BW*57-322 OR Behcet AND 446-86-6 OR Behçet AND Azathioprine OR Behçet AND BW*57-322 OR Behçet AND 446-86-6 (1 Treffer)
	Behcet AND Prednisolone OR Behcet AND Metacortandralon OR Behcet AND Deltahydrocortison OR Behcet AND 50-24-8 OR Behçet AND Prednisolone OR Behçet AND Metacortandralon OR Behçet AND Deltahydrocortison OR Behçet AND 50-24-8 (3 Treffer)
	Behcet AND Prednisone OR Behcet AND Metacortandracin OR Behcet AND 53-03-2 OR Behçet AND Prednisone OR Behçet AND Metacortandracin OR Behçet AND 53-03-2 (3 Treffer)
	Behcet AND Thalidomide OR Behcet AND 50-35-1 OR Behçet AND Thalidomide OR Behçet AND 50-35-1 (1 Treffer)
	Behcet AND Interferon alfa* OR Behcet AND Interferon alpha* OR Behcet AND Interferon-alpha* OR Behcet AND Interferon-alfa* OR Behcet AND alpha*Interferon OR Behcet AND alfa*Interferon OR Behcet AND IFN alpha* OR Behcet AND IFN alfa* OR Behcet AND IFN-alpha* OR Behcet AND IFN-alfa* OR Behçet AND Interferon alfa* OR Behçet AND Interferon alpha* OR Behçet AND Interferon-alpha* OR Behçet AND Interferon-alfa* OR Behçet AND alpha*Interferon OR Behçet AND alfa*Interferon OR Behçet AND IFN alpha* OR Behçet AND IFN alfa* OR Behçet AND IFN-alpha* OR Behçet t AND IFN-alfa* (4 Treffer)
	Behcet AND Infliximab OR Behcet AND 170277-31-3 OR Behçet AND Infliximab OR Behçet AND 170277-31-3 (8 Treffer)
	Behcet AND Etanercept OR Behcet AND TNFR-Fc* OR Behcet AND 185243-69-0 OR Behçet AND Etanercept OR Behçet AND TNFR-Fc* OR Behçet AND 185243-69-0 (1 Treffer)
	Behcet AND Adalimumab OR Behcet AND 331731-18-1 OR Behçet AND Adalimumab OR Behçet t AND 331731-18-1 (2 Treffer)
	Behcet AND Certolizumab* OR Behcet AND CDP870 OR Behcet AND 428863-50-7 OR Behçet AND Certolizumab* OR Behçet AND CDP870 OR Behçet AND 428863-50-7 (0 Treffer)
	Behcet AND Golimumab OR Behcet AND CNTO148 OR 476181-74-5 OR Behçet AND Golimumab OR Behçet AND CNTO148 OR 476181-74-5 (34 Treffer)

	Behcet AND Tumor Necrosis Factor* α * OR Behcet AND Tumor Necrosis Factor*alpha* OR Behcet AND Tumor Necrosis Factor*alfa* OR Behcet AND TNF* α * OR Behcet AND TNF*alpha* OR Behcet AND TNF*alfa* OR Behcet AND anti*TNF OR Behçet AND Tumor Necrosis Factor* α * OR Behçet AND Tumor Necrosis Factor*alpha* OR Behçet AND Tumor Necrosis Factor*alfa* OR Behçet AND TNF* α * OR Behçet AND TNF*alpha* OR Behçet AND TNF*alfa* OR Behçet AND anti*TNF (1 Treffer)
	Behcet AND Cyclosporine OR Behcet AND Ciclosporin OR Behcet AND 59865-13-3 OR Behçet AND Cyclosporine OR Behçet AND Ciclosporin OR Behçet AND 59865-13-3 (3 Treffer)
	Behcet AND Dapsone OR Behcet AND Diaphenylsulfone OR Behcet AND 80-08-0 OR Behçet AND Dapsone OR Behçet AND Diaphenylsulfone OR Behçet AND 80-08-0 (2 Treffer)
Treffer	63

Studienregister	EU Clinical Trials Register
Internetadresse	www.clinicaltrialsregister.eu
Datum der Suche	23.03.2020
Eingabeoberfläche	Standard Search
Suchstrategie	("Behcet" OR "Behçet") AND ("Azathioprine" OR "BW 57-322" OR "446-86-6") (3 Treffer)
	("Behcet" OR "Behçet") AND ("Prednisolone" OR "Metacortandralon" OR "Deltahydrocortison" OR "50-24-8") (1 Treffer)
	("Behcet" OR "Behçet") AND ("Prednisone" OR "Metacortandracin" OR "53-03-2") (4 Treffer)
	("Behcet" OR "Behçet") AND ("Thalidomide" OR "50-35-1") (0 Treffer)
	("Behcet" OR "Behçet") AND ("Interferon alfa*" OR "Interferon alpha*" OR "Interferon-alpha*" OR "Interferon-alfa*" OR "alpha*Interferon" OR "alfa*Interferon" OR "IFN alpha*" OR "IFN alfa*" OR "IFN-alpha*" OR "IFN-alfa*") (4 Treffer)
	("Behcet" OR "Behçet") AND ("Infliximab" OR "170277-31-3" OR "Etanercept" OR "TNFR-Fc*" OR "185243-69-0" OR "Adalimumab" OR "331731-18-1" OR "Certolizumab*" OR "CDP870" OR "428863-50-7" OR "Golimumab" OR "CNT0148" OR "476181-74-5") (2 Treffer)
	("Behcet" OR "Behçet") AND ("Tumor Necrosis Factor* α " OR "Tumor Necrosis Factor*alpha*" OR "Tumor Necrosis Factor*alfa*" OR "TNF* α " OR "TNF*alpha*" OR "TNF*alfa*" OR "anti*TNF") (0 Treffer)
	("Behcet" OR "Behçet") AND ("Cyclosporine" OR "Ciclosporin" OR "59865-13-3") (3 Treffer)
("Behcet" OR "Behçet") AND ("Dapsone" OR "Diaphenylsulfone" OR "80-08-0") (0 Treffer)	
Treffer	17

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der /den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Titel (Autor) der ausgeschlossenen Dokumente	Ausschlussgrund
1	Takeo M., Tanaka Y., et al. 2018. Apremilast for behcet's syndrome: results from a phase III, randomized, double-blind, placebo-controlled study in a Japanese subgroup. <i>Arthritis and rheumatology</i> 70: 3197-3198.	Publikationstyp
2	Leccese P., Ozguler Y., et al. 2016. A systematic literature review on the treatment of skin, mucosa and joint involvement of Behcet's syndrome informing the Eular recommendations for the management of Behcet's syndrome. <i>Arthritis & rheumatology</i> 68: 3971-3972.	Studiendesign
3	Deeks E.D., 2020. Apremilast: A Review in Oral Ulcers of Behcet's Disease. <i>Drugs</i> 80 (2): 181-188.	Publikationstyp
4	Yazici Y., 2020. Management of Behcet syndrome. <i>Current Opinion in Rheumatology</i> 32 (1): 35-40.	Publikationstyp
5	Leccese P., Ozguler Y., et al. 2019. Management of skin, mucosa and joint involvement of Behcet's syndrome: A systematic review for update of the EULAR recommendations for the management of Behcet's syndrome. <i>Seminars in Arthritis and Rheumatism</i> 48 (4): 752-762.	Publikationstyp
6	Hatemi G., Seyahi E., et al. 2015. Behcet's syndrome: a critical digest of the 2014-2015 literature. <i>Clinical and experimental rheumatology</i> 33 (6 Supplement 94): S3-S14.	Publikationstyp

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nr	Titel (Autor) der ausgeschlossenen Dokumente	Ausschlussgrund
1	Alpsoy E., Durusoy C., et al. 2002. Interferon alfa-2a in the treatment of Behcet disease: A randomized placebo-controlled and double-blind study. Archives of Dermatology 138 (4): 467-471.	Studienpopulation
2	[No authors listed] 1998. Thalidomide: new preparation. For well-defined indications. Prescrire International 7 (37): 131-4.	Publikationstyp
3	Arayssi T., Hamdan A., 2005. Update on the therapy of Behcet's disease. Current Medicinal Chemistry: Anti-Inflammatory and Anti-Allergy Agents 4 (3): 339-348.	Publikationstyp
4	Arida A., Fragiadaki K., et al. 2011. Anti-TNF Agents for Behcet's Disease: Analysis of Published Data on 369 Patients. Seminars in Arthritis and Rheumatism 41 (1): 61-70.	Patientenrelevante Endpunkte
5	Atzeni F., Sarzi-Puttini P., et al. 2005. Potential off-label use of infliximab in autoimmune and non-autoimmune diseases: a review. Autoimmunity Reviews 4 (3): 144-52.	Studiendesign
6	Atzeni F., Doria A., et al. 2007. Potential target of infliximab in autoimmune and inflammatory diseases. Autoimmunity Reviews 6 (8): 529-536.	Studiendesign
7	BenEzra D., Cohen E., et al. 1988. Evaluation of conventional therapy versus cyclosporine A in Behcet's syndrome. Transplantation Proceedings 20 (3 SUPPL. 4): 136- 143.	Patientenrelevante Endpunkte
8	Dalvi S.R., Yildirim R., et al. 2012. Behcets Syndrome. Drugs 72 (17): 2223-2241.	Publikationstyp
9	Davatchi F., 2004. New and innovative therapies for Behcet's disease. APLAR Journal of Rheumatology 7 (2): 141-145.	Studiendesign
10	Eggleston D.J., Nally F.F., 1972. Treatment of aphthous ulceration with topical azathioprine. A double blind trial. British Journal of Oral Surgery 9 (3): 233-6.	Studienpopulation
11	EUCTR2014-005390-36-GB, 2015. BIO BEHCET'S TRIAL. http://www.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-005390-36-GB .	Studienstatus
12	Gisoni P., Girolomoni G., 2013. Impact of TNF-alpha antagonists on the quality of life in selected skin diseases. Giornale Italiano di Dermatologia e Venereologia 148 (3): 243-248.	Studienpopulation
13	Gul A., 2007. Standard and novel therapeutic approaches to Behcet's disease. Drugs 67 (14): 2013-2022.	Studiendesign

14	Hamuryudan V., Ozyazgan Y., et al. 1997. Azathioprine in Behcet's syndrome. <i>Arthritis and Rheumatism</i> 40 (4): 769-774.	Publikationstyp
15	Hamuryudan V., Yurdakul S., et al. 1991. Inefficacy of topical alpha interferon in the treatment of oral ulcers of Behcet's syndrome: A randomized, double blind trial. <i>British Journal of Rheumatology</i> 30 (5): 395-396.	Intervention
16	Hatemi G., Seyahi E., et al. 2015. Behcet's syndrome: a critical digest of the 2014-2015 literature. <i>Clinical & Experimental Rheumatology</i> 33 (6 Suppl 94): S3-14.	Publikationstyp
17	Hello M., Barbarot S., et al. 2010. Use of thalidomide for severe recurrent aphthous stomatitis: a multicenter cohort analysis. <i>Medicine</i> 89 (3):176-82.	Studiendesign
18	Hirohata S., 2006. Is the long-term use of systemic corticosteroids beneficial in the management of Behcet's syndrome? <i>Nature clinical practice. Rheumatology</i> 2 (7): 358-359.	Intervention
19	Jenkins J.S., Powell R.J., et al.1984. Thalidomide in severe orogenital ulceration. <i>Lancet</i> 2 (8417-8418): 1424-1426.	Studienpopulation
20	Kotter I., 2015. The challenging treatment of Behcets disease. <i>Expert Opinion on Orphan Drugs</i> 3 (10): 1101-1110.	Studiendesign
21	Masuda K., Nakajima A., et al. 1989. Double-masked trial of cyclosporin versus colchicine and long-term open study of cyclosporin in Behcet's disease. <i>Lancet</i> 1 (8647): 1093-1096.	Studienpopulation
22	Mat C., Yurdakul S., et al. 2006. A double-blind trial of depot corticosteroids in Behcet's syndrome. <i>Rheumatology</i> 45(3): 348-352.	Intervention
23	Nair J.R., Moots R.J., 2015. Behcet's disease: Review of management. <i>Indian Journal of Rheumatology</i> 10: S84-S94.	Patientenrelevante Endpunkte
24	Nowatzky J., Yazici Y., 2011. Current management of Behcet's syndrome. <i>Drug Development Research</i> 72 (7): 647-656.	Publikationstyp
25	Pipitone N., Olivieri I., et al. 2006. New approaches in the treatment of Adamantiades-Behcet's disease. <i>Current Opinion in Rheumatology</i> 18 (1): 3-9.	Publikationstyp
26	Saeterdal I., Mork N.J., et al. 2007. TNF- α Inhibitors and Efalizumab for the Treatment of Skin Diseases. Knowledge Centre for the Health Services at The Norwegian Institute of Public Health (NIPH). Report from Norwegian Knowledge Centre for the Health Services (NOKC) No. 15-2007.	Patientenrelevante Endpunkte
27	Saleh Z., Arayssi T., 2014. Update on the therapy of Behcet disease. <i>Therapeutic Advances in Chronic Disease</i> 5 (3): 112-134.	Studiendesign

28	Sharquie K.E., Najim R.A., et al. 2002. Dapsone in Behcet's disease: A double-blind, placebo-controlled, cross-over study. <i>Journal of Dermatology</i> 29 (5): 267-279.	Studienpopulation
----	---	-------------------

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

Nicht zutreffend.

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
Clinical Trials.GOV		
1	Tehran University of Medical Sciences. 2011 Mar 2. ClinicalTrials.gov: The Efficacy of Methylprednisolone in the Treatment of Patients With Ocular Involvement in Behcet's Disease. https://clinicaltrials.gov/ct2/show/NCT01306955	Studienpopulation
2	National Institutes of Health Clinical Center (CC) National Eye Institute (NEI). 2008 Mar 4. ClinicalTrials.gov: HAT in Eye Complications of Behcet's Disease. https://clinicaltrials.gov/ct2/show/NCT00001865	Studienpopulation
3	Tehran University of Medical Sciences Hoffmann-La Roche. 2008 Apr 30. ClinicalTrials.gov: Rituximab for the Treatment of Severe Ocular Manifestations of Behcet's Disease. https://clinicaltrials.gov/ct2/show/NCT00664599	Studienpopulation
4	Simon Taylor, Royal Surrey County Hospital NHS Foundation Trust Royal Surrey County Hospital NHS Foundation Trust. 2012 Jun 7. ClinicalTrials.gov: Causes of Visual Loss in Retinal Disease. https://clinicaltrials.gov/ct2/show/NCT01613963	Studienpopulation
5	Meifen Zhang, Peking Union Medical College Hospital Peking Union Medical College Hospital. 2018 Jun 21. ClinicalTrials.gov: Interferon α 2a Versus Cyclosporine for	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
	Refractory Behçet`s Disease Uveitis. https://clinicaltrials.gov/ct2/show/NCT03209219	
6	University Hospital Tuebingen. 2018 May 17. ClinicalTrials.gov: Interferon-alpha2a Versus Cyclosporin A for Severe Ocular Behçet`s Disease (INCYTOB). https://clinicaltrials.gov/ct2/show/NCT00167583	Studienpopulation
7	Nobel Pharmaceuticals Amarillo Biosciences, Inc.. 2009 Jun 16. ClinicalTrials.gov: Low Dose Interferon Alpha Treatment for Oral Ulcers of Behçet's Disease. https://clinicaltrials.gov/ct2/show/NCT00483184	Intervention
8	Tehran University of Medical Sciences Wyeth, LLC. 2010 Nov 22. ClinicalTrials.gov: Etanercept: Single Blind Control Study in Ocular Manifestations of Behçet's Disease. https://clinicaltrials.gov/ct2/show/NCT00931957	Studienpopulation
9	Assistance Publique - Hôpitaux de Paris. 2018 Feb 27. ClinicalTrials.gov: Induction Therapy With Anti-TNFa vs Cyclophosphamide in Severe Behçet Disease. https://ClinicalTrials.gov/show/NCT03371095	Studienpopulation
10	NYU Langone Health Genentech, Inc.. 2019 Feb 15. ClinicalTrials.gov: Tocilizumab for the Treatment of Behçet's Syndrome. https://clinicaltrials.gov/ct2/show/NCT01693653	Intervention
11	Assistance Publique - Hôpitaux de Paris. 2017 Dec 8. ClinicalTrials.gov: Randomized Trial Comparing Efficacy of Adalimumab, Anakinra and Tocilizumab in Non-infectious Refractory Uveitis (RUBI). https://clinicaltrials.gov/ct2/show/NCT02929251	Studienpopulation
12	Novartis Pharmaceuticals. 2015 Aug 31. ClinicalTrials.gov: Phase III Study in Refractory Behçet's Disease. https://ClinicalTrials.gov/show/NCT00995709	Studienpopulation
13	Wenjie Zheng Peking Union Medical College Hospital. 2020 Mar 10. ClinicalTrials.gov: Golimumab for the Treatment of Refractory Behçet's Uveitis. https://clinicaltrials.gov/ct2/show/NCT04218565	Studienpopulation
14	Novartis Pharmaceuticals. 2017 Jul 17. ClinicalTrials.gov: Safety and Efficacy of AIN457 in Noninfectious Uveitis. https://clinicaltrials.gov/ct2/show/NCT00685399	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
15	IRCCS Burlo Garofolo Centro di Riferimento Oncologico – Aviano Azienda Ospedaliera Spedali Civili di Brescia. 2020 Feb 12. ClinicalTrials.gov: Thalidomide Versus Infliximab in New Onset Crohn's Disease With Poor Prognostic Factors. https://clinicaltrials.gov/ct2/show/NCT03221166	Studienpopulation
16	Celgene. 2020 Mar 2. ClinicalTrials.gov: A Phase 3 Randomized, Double-blind Study to Evaluate the Efficacy and Safety of Apremilast (CC-10004) in Subjects With Active Behçet's Disease. https://clinicaltrials.gov/ct2/show/NCT02307513	Intervention
17	Celgene. 2019 Nov 15. ClinicalTrials.gov: A Study to Evaluate the Efficacy and Safety of Apremilast (CC-10004) in the Treatment of Behçet Disease. https://clinicaltrials.gov/ct2/show/NCT00866359	Intervention
18	Tamer Ahmed Macky Cairo University. 2015 Dec 3. ClinicalTrials.gov: Intravitreal Infliximab in Refractory Uveitis in Behçet's Disease: A Safety and Efficacy Clinical Study. https://clinicaltrials.gov/ct2/show/NCT02620618	Studienpopulation
19	Dr. Yolanda Braun Rambam Health Care Campus Abbott. 2016 Oct 26. ClinicalTrials.gov: Efficacy of Humira in Behçet Patients With Arthritis. https://clinicaltrials.gov/ct2/show/NCT01497717	Studienpopulation
20	Dr. Rona Smith Cambridge University Hospitals NHS Foundation Trust. 2018 May 11. ClinicalTrials.gov: Anti IL-18 (GSK1070806) in Behçet's Disease. https://clinicaltrials.gov/ct2/show/NCT03522662	Studiendesign
21	Janssen Korea Ltd. 2019 Oct 17. ClinicalTrials.gov: A Study to Evaluate Efficacy and Safety of Infliximab in Participant With Moderate-to-Severe Refractory Intestinal Behçet's Disease (BEGIN). https://clinicaltrials.gov/ct2/show/NCT02505568	Studienpopulation
22	AbbVie Eisai Co. Ltd. 2014 Jul 25. ClinicalTrials.gov: A Study of Adalimumab in Japanese Subjects With Intestinal Behçet's Disease. https://clinicaltrials.gov/ct2/show/NCT01243671	Studiendesign
23	AbbVie. 2020 Feb 13. ClinicalTrials.gov: A Prospective, Mono-Country and Multi-center Study to Observe Safety and Effectiveness of Adalimumab in KoREan Intestinal Behçet's Disease(BD) Patients (AMORE). https://clinicaltrials.gov/ct2/show/NCT02687828	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
24	AbbVie. 2019 Mar 22. ClinicalTrials.gov: Special Investigation in Patients With Intestinal Behcet's Disease (All Case Investigation). https://clinicaltrials.gov/ct2/show/NCT01960790	Studienpopulation
EU Clinical Trials Register		
25	The University of Liverpool. 2016 Feb 17. ClinicalTrialsRegister.eu: Optimal utilisation of biologic drugs in Behçet's Disease: a randomised controlled trial of infliximab (IFX) versus alpha interferon (αIFN), with genotyping and metabolomic profiling, towards a stratified medicines approach to treatment. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005390-36	Studienstatus
26	University Hospital of Tübingen. 2011 Sep 2. ClinicalTrialsRegister.eu: Canakinumab for Behçet's Disease Resistant to Standard Treatment (CanBeDisT). https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024152-29	Intervention
27	ErasmusMC. 2007 Mar 20. ClinicalTrialsRegister.eu: Mycophenolate sodium (Myfortic®) in the Treatment of Uveitis: a Pilot Study. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2006-004709-24	Studienpopulation
28	Department of Ophthalmology at St. Franziskus Hospital. 2007 Aug 6. ClinicalTrialsRegister.eu: Single center study on safety and efficacy of everolimus in patients with endogenous intermediate and posterior uveitis. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2006-004876-10	Studienpopulation
29	Assistance Publique – Hôpitaux de Paris (AP-HP). N.d. ClinicalTrialsRegister.eu: Etude randomisée multicentrique évaluant l'efficacité et la tolérance de l'infliximab comparativement au Cyclophosphamide dans les formes sévères de maladie de Behçet " Multicenter, randomized, prospective trial comparing the Efficacy and Safety of Infliximab to that of Cyclophosphamide in severe Behçet's disease". https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-002264-41	Studienpopulation
International Clinical Trials Registry Platform WHO		
30	Zanjan University of Medical Sciences. 2018 May 21. WHO ICTRP: Efficacy of Infliximab comparing with	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
	Cyclophosphamide in refractory uveitis in Behcet's diseases. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=IRCT20180110038297N1	
31	Yokohama City University Hospital. 2019 Jul 22. WHO ICTRP: Clinical studies on discontinuation of infliximab and making the shift to cyclosporine for refractory uveitis of Behcet's disease. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000025451	Studienpopulation
32	Moorfields Eye Hospital NHS Foundation Trust (UK). 2015 Jan 13. WHO ICTRP: Does addition of subcutaneous interferon alpha 2b for 6 months only to the treatment of Behcet's disease significantly reduce the number of patients requiring immunosuppressive agents to control their disease over the following 3 years? http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ISRCTN36354474	Intervention
33	Novartis Pharmaceuticals. 2019 Mar 04. WHO ICTRP: 1 Year Canakinumab Treatment in Behcet's Disease Patients With Neurologic or Vascular Involvement. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02756650	Intervention
34	Castle Study Group. 2019 Oct 07. WHO ICTRP: Comparison of Adalimumab and Steroid. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000012469	Studienpopulation
35	The First Affiliated Hospital of Heilongjiang Chinese Medicine Colledge. 2017 Nov 11. WHO ICTRP: Clinical Study on Treatment of Behcet Syndrome with Xiaoyao San in Traditional Chinese Medicine. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ChiCTR-INR-17013467	Patientenrelevante Endpunkte
36	Kyorin University. 2019 Apr 02. WHO ICTRP: Evaluation of uveoretinitis associated with Behcet's disease after discontinuation of successful infliximab therapy. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000025670	Studienpopulation
37	Department of Ophthalmology. 2019 Apr 02. WHO ICTRP: Botulinum toxin-induced uveitis in a patient with Behcet's	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
	disease under infliximab treatment: Case report. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000023887	
38	Sanofi-aventis Recherche & Développement. 2012 Mar 19. WHO ICTRP: A randomized, double-blind, parallel-group, placebo- and active calibrator-controlled. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-021020-94-DE	Studienpopulation
39	FRANCISCO J. BLANCO GARCÍA. 2013 Jul 22. WHO ICTRP: Evaluation of a protocol for the reduction of doses in patients with rheumatoid arthritis (RA) in clinical remission in treatment with biological therapies. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-004482-40-ES	Studienpopulation
40	Janssen Biologics B.V. 2014 Feb 24. WHO ICTRP: A Golimumab Phase 3b, Multicenter, Switch Assessment of Subcutaneous and Intravenous Efficacy in Rheumatoid Arthritis Patients Who Have Inadequate Disease Control Despite Treatment with Etanercept (ENBREL(r)) or Adalimumab (HUMIRA(r)). http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-010582-23-DE	Studienpopulation
41	MSD Belgium BVBA/SPRL. 2014 Nov 24. WHO ICTRP: Study to evaluate delivery preference of Simponi (golimumab) by Ulcerative Colitis patients: delivery using a prefilled syringe or the Smartject(tm) device. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-000656-29-BE	Studienpopulation
42	Schering-Plough Research Institute, a Division of Schering Corporation. 2014 Dec 08. WHO ICTRP: GO-AHEAD. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-000311-34-DE	Studienpopulation
43	Universitätsklinikum Erlangen. 2015 Mar 23. WHO ICTRP: A PHASE 3, MULTICENTER, RANDOMIZED, OPEN, PROSPECTIVE, CONTROLLED, PARALLEL-GROUP STUDY OF REDUCTION OF THERAPY IN PATIENTS WITH RHEUMATOID ARTHRITIS IN ONGOING REMISSION. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-015740-42-DE	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
44	MedImmune Ltd. 2015 Jul 10. WHO ICTRP: An exploratory clinical study to investigate mavrilimumab, an antibody being developed for the treatment of moderate to severe rheumatoid arthritis, an inflammatory condition that affects the joints versus a different antibody whose mechanism works by inhibiting tumor necrosis factor. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-005649-10-DE	Studienpopulation
45	CLINICAL PHARMACOLOGY SPANISH SOCIETY. 2015 Aug 04. WHO ICTRP: Evaluation of clinical value of a standardized protocol for dose reduction in patients with axial Spondyloarthritis and clinical remission with anti-TNF therapy. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-005871-18-ES	Studienpopulation
46	Janssen Biologics B.V. 2015 Oct 06. WHO ICTRP: A Study to Assess the Safety and Pharmacokinetics (Serum Levels) of Golimumab in Children With Moderately to Severely Active Ulcerative Colitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-004366-18-AT	Studienpopulation
47	The Leeds Teaching Hospitals NHS Trust. 2016 Jun 27. WHO ICTRP: GOLMePsA. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-004122-28-GB	Studienpopulation
48	Helse More og Romsdal Hospital Trust. 2016 Dec 12. WHO ICTRP: A study to evaluate the safety of stopping versus continuing anti TNF therapy in ulcerative colitis patients in remission, and to evaluate the safety and efficacy of restarting anti TNF therapy in patients with disease recurrence. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2016-001409-18-NO	Studienpopulation
49	MSD. 2017 Apr 03. WHO ICTRP: Go-colitis: Golimumab: A Phase 4 UK Study on its Utilisation and Impact in Ulcerative Colitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-004583-56-GB	Studienpopulation
50	Janssen Research & Development, LLC. 2017 Dec 16. WHO ICTRP: A Study to Evaluate the Accuracy of a Subset of the Length-109 Probe Set Panel (a Genetic Test) in Predicting Response to Golimumab in Participants With Moderately to	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
	Severely Active Ulcerative Colitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01988961	
51	Janssen Research & Development, LLC. 2017 Dec 16. WHO ICTRP: A Study of Golimumab in Participants With Active Ankylosing Spondylitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02186873	Studienpopulation
52	Janssen Research & Development, LLC. 2018 Jan 08. WHO ICTRP: A Study of Golimumab in Participants With Active Psoriatic Arthritis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02181673	Studienpopulation
53	UCD. 2018 Jul 23. WHO ICTRP: Golimumab dose variation to achieve response in Colitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-004724-62-IE	Studienpopulation
54	Istituto Clinico Humanitas. 2019 Jan 07. WHO ICTRP: Pilot study on the efficacy of golimumab in normalizing the symptoms and lesions of the colon-rectum in patients with ulcerative colitis who only respond to therapy with cortisone. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-002579-27-IT	Studienpopulation
55	Janssen Research & Development, LLC. 2019 Sep 09. WHO ICTRP: A Study to Evaluate the Efficacy and Safety of Ustekinumab in the Treatment of Anti-TNF α Naive Participants With Active Radiographic Axial Spondyloarthritis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02437162	Studienpopulation
56	Janssen Research & Development, LLC. 2019 Sep 09. WHO ICTRP: A Study to Evaluate the Efficacy and Safety of Ustekinumab in the Treatment of Anti-TNF(Alpha) Refractory Participants With Active Radiographic Axial Spondyloarthritis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02438787	Studienpopulation
57	Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc. 2019 Feb 28. WHO ICTRP: Evaluate the Efficacy and Safety of Golimumab in Subjects With Non-Radiographic Axial Spondyloarthritis.	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-004020-65-DE	
58	CHU SAINT-Etienne. 2019 Feb 28. WHO ICTRP: Rotation or Change of Biotherapy After TNF blocker treatment failure. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-004700-22-FR	Studienpopulation
59	Fundación de Investigación Biomédica del Hospital Universitario de La Princesa. 2019 Feb 28. WHO ICTRP: prediction of response to therapy in inflammatory bowel disease. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-002925-47-ES	Studienpopulation
60	Med Uni Graz, Univ. Klinik Innere Medizin, Abt. Rheumatologie und Immunologie. 2019 Apr 30. WHO ICTRP: Could ultrasound help to identify the patients with rheumatoid arthritis, in those the treatment with biological DMARDs could be stopped? http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-005204-15-AT	Studienpopulation
61	MERCK SHARP & DOHME B.V. 2019 Jun 30. WHO ICTRP: Open-label, Phase 4 Study, investigation of the frequency of Extra-Articular exposure in patients with Ankylosing Spondylitis treated with Golimumab. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002458-21-NL	Studienpopulation
62	University Medical Center Utrecht (UMCU). 2019 Jun 30. WHO ICTRP: REMission INDuction in very early Rheumatoid Arthritis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-004858-17-NL	Studienpopulation
63	Ghent University Hospital. 2019 Jun 30. WHO ICTRP: Exploration of TNF-alpha Blockade with golimumab in the Induction of Clinical Remission in Patients with Early Peripheral Spondyloarthritis (SpA) according to ASAS-criteria. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-003678-97-BE	Studienpopulation
64	Instituto de Medicina Molecular. 2019 Dec 10. WHO ICTRP: Golimumab efficacy in psoriatic arthritis patients with active dactylitis.	Studienpopulation

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-000643-40-PT	
65	Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc. 2019 Dec 10. WHO ICTRP: Study to correlate a new biomarker (ST2) with clinical activity in Ulcerative Colitis patients under golimumab. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-003262-25-PT	Studienpopulation
66	Ghent University. 2019 Dec 10. WHO ICTRP: Efficacy of golimumab in early axial spondyloarthritis (axSpA) in relation to gut inflammation, an early remission induction study (GO GUT). http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-001728-23-BE	Studienpopulation
67	Janssen Biologics BV. 2019 Dec 16. WHO ICTRP: A Phase 1 Study to Evaluate Multiple Doses of golimumab in Participants At-Risk for Developing Type-1 Diabetes. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-000225-12-SE	Studienpopulation
68	Sint Maartenskliniek. 2020 Feb 01. WHO ICTRP: Research into injecting golimumab less frequently by using increased doses. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2019-004101-27-NL	Studienpopulation
69	Janssen Research & Development, LLC. 2020 Mar 30. WHO ICTRP: A Study to Assess the Efficacy and Safety of Golimumab in Pediatric Participants With Moderately to Severely Active Ulcerative Colitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03596645	Studienpopulation
70	Janssen Research & Development, LLC. 2020 Mar 30. WHO ICTRP: A Study of Efficacy and Safety of Combination Therapy With Guselkumab and Golimumab in Participants With Moderately to Severely Active Ulcerative Colitis. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03662542	Studienpopulation
<p>a: Wird eine Studie in mehreren Registern identifiziert, wird diese hier nur einmal gelistet.</p> <p>Bedeutung des jeweils aufgelisteten Datums:</p>		

Nr	Referenz der ausgeschlossenen Dokumente ^a	Ausschlussgrund
Clinical Trials.GOV: Last updated EU Clinical Trials Register: Start Date		

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 4-102 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 4-102 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-102 (Anhang): Studiendesign und -methodik für Studie RELIEF

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2b	Genaue Ziele, Fragestellung und Hypothesen	<p><u>Fragestellung:</u> Bewertung der Wirksamkeit und Sicherheit von Apremilast bei der Behandlung von Patienten mit aktivem Behçet-Syndrom.</p> <p><u>Primäres Ziel:</u> Das primäre Studienziel war die Bewertung der Wirksamkeit von Apremilast bei der Behandlung von oralen Aphthen bei Patienten mit aktivem Behçet-Syndrom.</p> <p><u>Sekundäre Ziele:</u> Die sekundären Studienziele waren die Bewertung der Wirksamkeit von Apremilast bei Patienten mit aktivem Behçet-Syndrom und die Bewertung des Effektes von Apremilast auf die patientenberichteten Endpunkte (Patient Reported Outcomes, PRO) bei Patienten mit aktivem Behçet-Syndrom.</p> <p><u>Sicherheitsziel:</u> Das Sicherheitsziel war die Bewertung der Sicherheit und Verträglichkeit von Apremilast bei Patienten mit aktivem Behçet-Syndrom.</p> <p><u>Exploratives Ziel:</u> Das explorative Ziel war die Bewertung des Effektes von Apremilast auf die Krankheitsaktivität bei Patienten mit aktivem Behçet-Syndrom.</p>
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Bei der vorliegenden Studie handelt es sich um eine multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte, parallele Phase-III-Studie mit einer</p> <ul style="list-style-type: none"> • Screeningphase von bis zu sechs Wochen • Placebo-kontrollierten Doppelblindphase von 12 Wochen • aktiven Behandlungsphase von 52 Wochen • Nachbeobachtungsphase von vier Wochen

Item ^a	Charakteristikum	Studieninformation
		Das Randomisierungsverhältnis der Patienten betrug 1:1 für Apremilast vs. Placebo.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Nach Studienbeginn gab es ein globales und drei länderspezifische Amendments zum Originalprotokoll vom 10.09.2014:</p> <p><u>Protokoll-Amendment 0.1 JA, Datum 27.10.2014 (Japan)</u></p> <ul style="list-style-type: none"> • Änderung des Mindestalters der Patienten zum Zeitpunkt der Einwilligungserklärung auf ≥ 20 Jahre zur Berücksichtigung der länderspezifischen Bedingungen • Ergänzung der Informationen über die Voraussetzungen zur Bereitstellung der Studienmedikation in Japan • Ergänzung der Informationen zum Schadensersatz bei studienbedingten Gesundheitsschäden <p><u>Protokoll-Amendment 1, Datum: 02.02.2016 (global)</u></p> <ul style="list-style-type: none"> • Erweiterung der möglichen Studienteilnahme auf Patienten, die zuvor eine biologische Therapie erhalten hatten • Änderungen in Bezug auf die frühere und aktuelle Medikation: <ul style="list-style-type: none"> a. Colchicin: Es war gestattet, die Einnahme von Colchicin bis zu sieben Tage vor der Randomisierung fortzuführen b. Orale und topische Kortikosteroide: Das Ausschleichen und anschließende Absetzen oraler und topischer Kortikosteroide war bis kurz vor dem Tag der Randomisierung gestattet • Klarstellung bzgl. der erforderlichen Mindestanzahl an vorhandenen oralen Aphthen zum Studieneinschluss: Alle Patienten mussten bei Visite 1 (Screening Visite) mindestens zwei orale Aphthen aufweisen. Waren mindestens drei Aphthen vorhanden, war ein Patient unabhängig von der Zeitdauer seit der Screening Visite für die Randomisierung geeignet, sofern er alle weiteren Einschlusskriterien erfüllte. Waren bei Visite 2 nur zwei orale Aphthen vorhanden, mussten vor der Randomisierung des Patienten mindestens 14 Tage seit Visite 1 vergangen sein. • Klarstellung und Überarbeitung des Static Physician's Global Assessment (PGA) of Skin Lesions: Die Bewertung der Anzahl der Knötchen/Läsionen sollte NICHT auf <u>eine</u> Körperstelle beschränkt werden, sondern am gesamten Körpers erfolgen. <p><u>Protokoll-Amendment 1.1 JA, Datum: 02.02.2016 (Japan)</u></p> <ul style="list-style-type: none"> • Die Notwendigkeit von Hepatitis-Tests wurde aktualisiert: Laboruntersuchungen und klinische Interpretationen von Tests auf Hepatitis B- und C-Antigene oder -Antikörper für Studieneinschluss/-ausschlusszwecke wurden dem Ermessen des Prüfarztes überlassen <p><u>Protokoll-Amendment 1.1 DE, Datum: 05.07.2017 (Deutschland)</u></p> <ul style="list-style-type: none"> • Eine optionale offene Verlängerungsphase wurde hinzugefügt. Die Patienten hatten die Möglichkeit, nach der 52-wöchigen aktiven

Item ^a	Charakteristikum	Studieninformation
		Behandlungsphase in eine optionale offene Verlängerungsphase einzutreten, um ihnen eine Weiterbehandlung bis zur Marktausbietung von Apremilast zu ermöglichen.
4	Probanden / Patienten	
4a	Ein- /Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Die Patienten mussten alle nachfolgenden Kriterien erfüllen, um für die Teilnahme an der Studie geeignet zu sein:</p> <p><u>Allgemeine Kriterien:</u></p> <ul style="list-style-type: none"> • Verständnis und freiwillige Unterzeichnung der Einwilligungserklärung vor der Durchführung studienbezogener Untersuchungen • Männer und Frauen ab 18 Jahren zum Zeitpunkt der Unterzeichnung der Einwilligungserklärung • Befähigung zur Einhaltung des Visitenplans der Studie und anderer Protokollanforderungen <p><u>Krankheitsspezifische Parameter:</u></p> <ul style="list-style-type: none"> • Diagnostiziertes Behçet-Syndrom anhand der Kriterien der International Study Group (ISG) for Behçet's Disease • Mindestens dreimaliges Auftreten von oralen Aphthen, in den vergangenen 12 Monaten, unter Berücksichtigung der aufgetretenen oralen Aphthen zum Zeitpunkt des Screenings • Mindestens zwei orale Aphthen zur Visite 1 (Screening Visite) und <ul style="list-style-type: none"> a. weiterhin mindestens zwei orale Aphthen zur Visite 2 (Tag der Randomisierung), wenn die Visite 2 mindestens 14 Tage nach der Visite 1 erfolgte ODER <ul style="list-style-type: none"> b. mindestens drei orale Aphthen zur Visite 2 (Tag der Randomisierung), wenn die Visite 2 zwischen einem und 42 Tagen nach der Visite 1 erfolgte • Vorbehandlung mit mindestens einer nicht-biologischen Therapie des Behçet-Syndroms, z. B. topische Kortikosteroide oder systemische Behandlungen • Geeignet für eine systemische Therapie zur Behandlung oraler Aphthen <ul style="list-style-type: none"> a. Die Patienten wurden vom Prüfarzt als geeignet für eine systemische Therapie beurteilt, wenn ihre Schleimhautulzera aufgrund der Schwere der Erkrankung und des Ausmaßes des betroffenen Bereichs als ungeeignet für eine topische Therapie erachtet wurden oder ihre oralen Aphthen mit einer topischen Therapie nicht ausreichend kontrollierbar waren • Erfüllung der folgenden Laborparameter: <ul style="list-style-type: none"> ○ Hämoglobin > 9 g/dl ○ Leukozytenzahl $\geq 3.000/\mu\text{l}$ ($\geq 3,0 \times 10^9/\text{l}$) und $\leq 14.000/\mu\text{l}$ ($\leq 14 \times 10^9/\text{l}$) ○ Thrombozytenzahl $\geq 100.000/\mu\text{l}$ ($\geq 100 \times 10^9/\text{l}$) ○ Serumkreatinin $\leq 1,5 \text{ mg/dl}$ ($\leq 132,6 \mu\text{mol/l}$)

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Gesamtbilirubin $\leq 2,0$ mg/dl ○ Aspartat-Aminotransferase (AST) und Alanin-Aminotransferase (ALT) $< 1,5$ x obere Normgrenze (upper limit of normal, ULN). Bei Patienten, die im Screening einen ALT- und / oder AST-Wert von $\geq 1,5$ x ULN aufwiesen, konnten die Tests innerhalb der Screening-Phase einmalig wiederholt werden. Im Falle einer Wiederholung sollten die Ergebnisse innerhalb des Referenzbereichs liegen, um für den Studieneinschluss zulässig zu sein. Laboruntersuchungen konnten dabei, in Absprache mit dem medizinischen Monitor, einmal wiederholt werden, wenn nach Einschätzung des Prüfarztes Grund zur Annahme bestand, dass die Ergebnisse der Wiederholungstests innerhalb der zulässigen Grenzbereiche lagen. • Anforderungen an die Kontrazeption: <ul style="list-style-type: none"> ○ Alle gebärfähigen Patientinnen²² mussten während der Einnahme von Apremilast bis mindestens 28 Tage nach Anwendung der letzten Dosis von Apremilast eine der u.g. beschriebenen zugelassenen empfängnisverhütenden Maßnahmen anwenden. ○ Wenn sich zu Beginn oder zu einem beliebigen Zeitpunkt während der Studie die Verhütungsmaßnahme oder die Gebärfähigkeit der Frauen veränderten, sollte der Prüfarzt die Patientinnen hinsichtlich der Möglichkeiten zur Empfängnisverhütung und des korrekten und konsequenten Einsatzes von wirksamen Verhütungsmitteln aufklären, um eine Schwangerschaft zu verhindern. ○ Alle gebärfähigen Frauen mussten bei Visite 1 und 2 einen negativen Schwangerschaftstest vorweisen. Alle sexuell aktiven, gebärfähigen Patientinnen mussten eine der folgenden zugelassenen Verhütungsmethoden anwenden: <ol style="list-style-type: none"> 1. Eine der folgenden hochwirksamen Methoden: hormonelle Verhütung (Oral, Injektion, Implantat, transdermales Pflaster, Vaginalring); Intrauterinpressar, Tubenligatur oder Vasektomie des Partners ODER 2. Kondom für Männer oder Frauen (Latexkondom oder Nicht-Latexkondom, das nicht aus einer natürlichen (tierischen) Membran bestand (z.B. Polyurethan)) sowie eine zusätzliche Barrieremethode: <ol style="list-style-type: none"> (a) Diaphragma mit Spermizid (b) Portiokappe mit Spermizid (c) Verhütungsschwamm mit Spermizid ○ Männliche Patienten (einschließlich vasktomierter), die sexuell aktiv waren, mussten eine Barrieremethode (Latexkondom oder Nicht-Latexkondom, das nicht aus einer natürlichen (tierischen) Membran bestand (z.B. Polyurethan)) zur Verhütung anwenden.

²² Geschlechtsreife Frau, die: 1) keiner Hysterektomie (operative Entfernung der Gebärmutter) oder einer beidseitigen Ovariectomie (operative Entfernung beider Eierstöcke) unterzogen wurde; oder 2) für mindestens 24 aufeinanderfolgende Monate nicht postmenopausal war (d. h. die Monatsblutung zu jeder Zeit während der vorangegangenen 24 aufeinanderfolgenden Monate hatte).

Item ^a	Charakteristikum	Studieninformation
		<p>Diese Anforderung bestand für den Zeitraum ab Einnahme der Studienmedikation und musste bis mindestens 28 Tage nach der letzten Verabreichung der Studienmedikation eingehalten werden</p> <p>Ausschlusskriterien: Das Vorliegen von mindestens einem der folgenden Kriterien führte zum Ausschluss eines Patienten von der Studienteilnahme:</p> <p><u>Krankheitsspezifische Ausschlussgründe:</u></p> <ul style="list-style-type: none"> • Im Rahmen des Behçet-Syndroms auftretende aktive Beteiligung wichtiger Organe: pulmonale (z. B. Pulmonalarterienaneurysma), vaskuläre (z. B. Thrombophlebitis), gastrointestinale (z. B. Ulzera im Bereich des Gastrointestinaltrakts), zentralnervöse (z. B. Meningoenzephalitis) Manifestationen, sowie okuläre Läsionen (z. B. Uveitis), die eine immunsuppressive Therapie erforderten <p>Ausnahmen:</p> <ol style="list-style-type: none"> a. Eine vorherige Beteiligung wichtiger Organe war zulässig, wenn sie mindestens ein Jahr vor der Visite 1 (Screening Visite) auftrat und zum Zeitpunkt des Studienbeginns nicht aktiv war b. Patienten mit im Rahmen des Behçet-Syndroms auftretenden leichten Formen okulärer Läsionen, für die keine systemische immunsuppressive Therapie erforderlich war c. Patienten mit im Rahmen des Behçet-Syndroms auftretender Arthritis und Hautmanifestationen <p><u>Vorangegangene und aktuelle Medikation:</u></p> <ul style="list-style-type: none"> • Vorangegangene Anwendung biologischer Therapien für die Behandlung oraler Aphthen, die mit dem Behçet-Syndrom assoziiert waren. <ul style="list-style-type: none"> ○ Eine vorherige biologische Anwendung war für andere Indikationen zulässig, einschließlich anderer Manifestationen dem Behçet-Syndrom • Vorangegangene Einnahme von Apremilast • Anwendung von Prüfpräparaten innerhalb der letzten vier Wochen vor der Visite 2 oder fünf pharmakokinetischen / pharmakodynamischen Halbwertszeiten (unter Berücksichtigung der längeren Option) • Aktuelle Einnahme starker Cytochrom P450-Induktoren (z. B. Rifampicin, Phenobarbital, Carbamazepin, Phenytoin) • Erhalt einer begleitenden immunmodulierenden Therapie (außer oralen oder topischen Kortikosteroiden) innerhalb folgender Zeitabschnitte: <ul style="list-style-type: none"> ○ Sieben Tage vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Colchicin ○ Zehn Tage vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Azathioprin und Mycophenolat-Mofetil ○ Vier Wochen (28 Tage) vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Ciclosporin, Methotrexat, Cyclophosphamid, Thalidomid und Dapson

Item ^a	Charakteristikum	Studieninformation
		<p>Anmerkung: Orale und topische Kortikosteroide mussten angemessen ausgeschlichen und abgesetzt worden sein vor dem Tag der Visite 2 (Tag der Randomisierung)</p> <ul style="list-style-type: none"> ○ Mindestens fünf terminale Halbwertszeiten für alle Biologika, einschließlich (aber nicht darauf beschränkt) der nachfolgend genannten innerhalb von: <ul style="list-style-type: none"> - Vier Wochen vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Etanercept - Acht Wochen vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Infiximab - Zehn Wochen vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Adalimumab, Golimumab, Certolizumab, Abatacept und Tocilizumab - Sechs Monate vor Visite 2 (Baseline Visite; Tag der Randomisierung) für Secukinumab ● Erhalt intraartikulärer oder parenteraler Kortikosteroide innerhalb von sechs Wochen (42 Tage) vor Visite 2 (Baseline Visite; Tag der Randomisierung) <p><u>Allgemeiner Gesundheitszustand:</u></p> <ul style="list-style-type: none"> ● Jegliche schwerwiegende Erkrankung, auffällige Laborwerte oder psychische Erkrankung, die den Patienten daran gehindert hätten, an der Studie teilzunehmen ● Jede Gegebenheit, einschließlich des Vorliegens abweichender Laborwerte, die den Patienten im Falle einer Studienteilnahme einem nicht akzeptablen Risiko ausgesetzt hätten ● Unvermögen eine freiwillige Einverständniserklärung abzugeben ● Schwangere oder stillende Frauen ● Systemische oder opportunistische Pilzinfektionen ● Bekannte aktive aktuelle oder in der Vorgeschichte wiederkehrende bakterielle, virale, mykotische, mykobakterielle oder andere Infektionen (einschließlich, aber nicht beschränkt auf Tuberkulose und atypische mykobakterielle Erkrankungen, Hepatitis B und C und Herpes Zoster, Histoplasmose, Coccidiomykose, jedoch unter Ausschluss der Onychomykose) oder jede längere Episode einer Infektion, die eine Hospitalisierung erforderlich machte oder eine Behandlung mit intravenösen oder oralen Antibiotika innerhalb von vier Wochen während der Screening-Phase ● Klinisch signifikante Abnormalität in der Röntgenaufnahme der Brust ● Klinisch signifikante Abnormalität im 12-Kanal-Elektrokardiogramm (EKG) ● Vorgeschichte eines positiven Tests oder eines klinischen Verdachts auf das Humane Immundefizienz-Virus (HIV) oder angeborene oder erworbene Immunschwäche (z. B. variables Immundefektsyndrom) ● Vorliegen eines bösartiger Tumors oder eine vorangegangene Erkrankung an einem bösartigen Tumor, mit der Ausnahme von: <ul style="list-style-type: none"> a. Kurativ behandeltes kutanes Basalzell- oder Plattenepithelzellkarzinom in situ

Item ^a	Charakteristikum	Studieninformation
		<p>b. Kurativ behandelte zervikale intraepitheliale Neoplasie oder Gebärmutterhalskarzinom in situ ohne Anzeichen eines Rezidivs innerhalb der letzten fünf Jahre vor Visite 1</p> <ul style="list-style-type: none"> • Jede Gegebenheit, die die Interpretation der Studienergebnisse hätte vereiteln können. • Geplante Operation oder andere Eingriffe, die die Teilnahme des Patienten an der Studie unterbrochen hätten. • Vorgeschichte eines Suizidversuchs zu einem beliebigen Zeitpunkt im Leben des Patienten bis einschließlich zum Zeitpunkt zur Visite 2 (Baseline Visite; Tag der Randomisierung) oder schwere psychische Erkrankungen, die in den letzten drei Jahren vor Visite 2 (Baseline Visite; Tag der Randomisierung) eine Hospitalisierung erforderlich machten.
4b	Studienorganisation und Ort der Studiendurchführung	Sponsor der Studie war die Celgene Corporation, Summit, New Jersey, USA. Die Studie wurde an 53 Zentren in zehn Ländern durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Die Patienten wurden im Verhältnis 1:1 einer der beiden Gruppen zugeteilt:</p> <p><u>Placebo-kontrollierte Behandlungsphase (bis einschließlich Woche 12):</u></p> <p>Gruppe 1 Apremilast 30 mg oral (p.o.) zweimal täglich (bid) im Abstand von etwa 12 Stunden für einen Behandlungszeitraum von 12 Wochen (ausgenommen Dosistitrationphase in Woche 1: 10 mg an Tag 1, 20 mg an Tag 2, 30 mg an Tag 3, 40 mg an Tag 4, 50 mg an Tag 5, 60 mg ab Tag 6)</p> <p>Gruppe 2 Placebo p.o. bid im Abstand von etwa 12 Stunden für einen Zeitraum von 12 Wochen</p> <p><u>Aktive Behandlungsphase (Woche 13-64):</u></p> <p>Gruppe 1 Apremilast 30 mg p.o. bid im Abstand von etwa 12 Stunden für einen Behandlungszeitraum von 52 Wochen</p> <p>Gruppe 2 Apremilast 30 mg p.o. bid (nach Dosistitration in der ersten Woche, wie vorstehend für Gruppe 1 beschrieben) im Abstand von etwa 12 Stunden für einen Behandlungszeitraum von 52 Wochen</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden	<p>Primärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Fläche unter der Kurve (area under the curve, AUC) für die Anzahl oraler Aphthen von Baseline bis einschließlich Woche 12 (AUC_{W0-12}) <p>Sekundäre Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Komplettes Ansprechen (vollständige Remission) der oralen Aphthen zu Woche 12, definiert als Anteil der Patienten, der frei von oralen Aphthen war

Item ^a	Charakteristikum	Studieninformation
	(z. B. Mehrfachbeobachtungen, Training der Prüfer und ggf. Angaben zur Validierung von Erhebungsinstrumenten)	<ul style="list-style-type: none"> • Veränderung der Schmerzhaftigkeit der oralen Aphthen mittels visueller Analogskala (VAS) zu Woche 12 gegenüber Baseline • Komplettes Ansprechen (vollständige Remission) der genitalen Aphthen zu Woche 12 für Patienten mit genitalen Aphthen zu Baseline Komplettes Ansprechen war definiert als Anteil der Patienten, der frei von genitalen Aphthen war. • Veränderung der Schmerzhaftigkeit der genitalen Aphthen mittels VAS zu Woche 12 gegenüber Baseline bei Patienten mit genitalen Aphthen zu Baseline • Veränderung der Krankheitsaktivität mittels Behçet's Disease Activity Form (BDCAF) zu Woche 12 gegenüber Baseline • Veränderung des Behçet's Disease Quality of Life (BD-QoL) Score zu Woche 12 gegenüber Baseline • Veränderung des Behçet's Syndrome Activity Score (BSAS) zu Woche 12 gegenüber Baseline • Zeit bis zur Auflösung der oralen Aphthen (komplettes Ansprechen) • Anteil der Patienten ohne orale Aphthen nach vollständigem Ansprechen • Anzahl der oralen Aphthen nach Verlust des kompletten Ansprechens • Zeit bis zum Rezidiv der oralen Aphthen nach Verlust des kompletten Ansprechens • Veränderung der Gesamtbewertung des statischen PGA von Hautläsionen (Akne-ähnliche Läsionen, Follikulitis und Erythema nodosum) des Behçet-Syndroms zu Woche 12 gegenüber Baseline bei Patienten mit Hautläsionen im Rahmen des Behçet-Syndroms zu Baseline • Anteil an Patienten, die ein komplettes Ansprechen der oralen Aphthen bis spätestens Woche 6 nach Behandlungsbeginn erreichten und die für mindestens sechs zusätzliche Wochen während der 12-wöchigen Placebo-kontrollierten Behandlungsphase frei von oralen Aphthen blieben <p>Sicherheitsbezogene Endpunkte:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse (UE) unter Berücksichtigung der Art, Häufigkeit, des Schweregrads und des Zusammenhangs zur Behandlung mit Apremilast • Anteil der Patienten, die aufgrund von UE die Einnahme der Prüfmedikation abgebrochen hatten • Veränderungen der Vitalparameter und / oder Laborbefunde <p>Explorative Endpunkte:</p> <ul style="list-style-type: none"> - zur Wirksamkeit - zur Pharmakokinetik / Pharmakodynamik - zur Pharmakogenetik
6b	Änderungen der Zielkriterien nach	Nicht relevant.

Item ^a	Charakteristikum	Studieninformation
	Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die geplante Stichprobengröße basierte auf den Ergebnissen der Phase-II-Studie CC-10004-BCT-001. Herangezogen wurde ein zweiseitiger t-Test auf einem Signifikanzniveau von 0,05, der eine Power von 90 % aufweisen sollte. Bezogen auf den primären Wirksamkeitsendpunkt, die AUC der Anzahl oraler Aphthen von Tag 1 bis Woche 12, sollte sich dabei eine Behandlungsdifferenz von 66 oralen Aphthen ergeben ($AUC_{W0-12} \text{ Placebo} - \text{Wirkstoff} = 66$). Es ergab sich eine Stichprobengröße von 102 für beide Vergleichsgruppen. Angenommen wurde eine gemeinsame Standardabweichung von 144.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es waren keine Zwischenanalysen vor der Analyse des primären Wirksamkeitsendpunkts geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die zufällige Zuteilung der Patienten zu einer der beiden Behandlungsgruppen (APR 30 bid oder Placebo) erfolgte unter Verwendung einer interaktiven Response Technologie (IRT).
8b	Einzelheiten (z. B. Block-randomisierung, Stratifizierung)	Die Patienten, die die Studieneinschlusskriterien erfüllten, wurden mittels permutierter Blockrandomisierung im Verhältnis 1:1 einer der beiden Gruppen zugeteilt. Geeignete Patienten wurden stratifiziert nach Geschlecht, Vorgeschichte einer Uveitis und Region (Japan oder Nicht-Japan).
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Das Studienpersonal an den Prüfzentren erhielt passwortgeschützte, verschlüsselte Identifikationsnummern zur Randomisierung der Patienten mithilfe der IRT. Während der Studienvisiten verteilten der Krankenhausapotheker oder autorisiertes Studienpersonal am Prüfzentrum kodierte Kits der Studienmedikation entsprechend der vom IRT vergebenen Randomisierungsnummer.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten	Patienten, die die Voraussetzungen zur Studienteilnahme erfüllten, wurden zur Visite 2 (Baseline-Visite) vom Prüfarzt entsprechend der vom IRT vergebenen Randomisierungsnummer einer der beiden Behandlungsgruppen zugeteilt.

Item ^a	Charakteristikum	Studieninformation
	in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>a) verblindet b) verblindet c) verblindet</p> <p>Die Verblindung für die Behandlungszuweisung wurde an allen Prüfzentren bis zum Datenbankschluss in Woche 68 aufrechterhalten. Eine Aufhebung der Verblindung war nur im Notfall mit Hilfe der IRT erlaubt, wenn die Kenntnis des Prüfpräparats zur Entscheidung über die Weiterbehandlung des Patienten notwendig war. Der betreffende Patient musste aus der Studie ausgeschlossen werden, sobald sein Behandlungscode entblindet war.</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Placebo-Tabletten glichen dem Aussehen nach den Apremilast-Tabletten (in allen Dosierungsstufen, 10 mg-, 20 mg- und 30 mg-Tabletten).
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Analyse der Wirksamkeitsendpunkte:</p> <p>Die Wirksamkeitsanalysen wurden anhand der Intention-to-Treat (ITT)-Population beurteilt. Ergebnisse zur Wirksamkeit wurden nach Berücksichtigung der Multiplizitätskontrolle als statistisch signifikant betrachtet. Alle statistischen Tests wurden zweiseitig mit einem Signifikanzniveau von 0,05 durchgeführt, p-Werte und oberes sowie unteres Konfidenzintervalle (KI) wurden berichtet.</p> <p>Die primären, sekundären und exploratoriven Endpunkte wurden für die Placebo-kontrollierte Behandlungsphase analysiert und gemäß den randomisierten Behandlungsgruppen dargestellt: Placebo und Apremilast. Für die während der aktiven Behandlungsphase bewerteten Endpunkte wurden die Daten gemäß der Behandlungsfolge dargestellt, d. h. Placebo/Apremilast oder Apremilast/Apremilast.</p> <p>Die Multiplizitätskontrolle erfolgte über sequentielles Testen. Dabei erfolgten so lange sequentielle Testungen, bis eine Analyse das auf 0,05 gesetzte Signifikanzniveau nicht erreichte. Die Hierarchie der Hypothesentests wurde a priori im SAP definiert.</p> <p>Primärer Wirksamkeitsendpunkt</p> <p>Der primäre Wirksamkeitsendpunkt war die AUC für die Anzahl der oralen Aphthen ab Tag 1 (Baseline) bis Woche 12/Tag 85 (AUC_{W0-12}). Die AUC_{W0-12} für die Anzahl oraler Aphthen wurde zwischen der Placebo- und der Apremilast-Gruppe verglichen unter Verwendung einer zweiseitigen parametrischen Kovarianzanalyse (ANCOVA) mit einem Signifikanzniveau von 0,05. Das Modell beinhaltete die AUC_{W0-12} für die Anzahl oraler Aphthen als Zielvariable; Behandlung, Geschlecht und</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Region als Faktoren; die Anzahl der oralen Aphthen zu Studienbeginn bildeten die Kovariablen. Least-Squares (LS)-Mittelwerte der jeweiligen Gruppe, die zugehörigen Standardfehler (standard errors, SE), die Behandlungsdifferenzen in LS-Mittelwerten und die mit diesen assoziierten SE, zweiseitige 95 %-KI sowie zweiseitige p-Werte wurden aus dem Modell abgeleitet.</p> <p>Die primäre Wirksamkeitsanalyse wurde auf Basis der ITT- und Per-Protocol (PP)-Populationen durchgeführt.</p> <p>Die ITT-Population schloss alle randomisierten Patienten ein, die mindestens eine Dosis der Studienmedikation erhalten hatten.</p> <p>Die PP-Population schloss alle randomisierten, mit mindestens einer Dosis der Studienmedikation behandelten, Patienten ein, für die eine Baseline- und mindestens eine Post-Baseline-Erfassung der oralen Aphthen vorlag und die keine bedeutsame Protokollverletzung während der 12-wöchigen Placebo-kontrollierten Behandlungsphase aufwiesen.</p> <p>Sekundäre Wirksamkeitsendpunkte und explorativen Endpunkte</p> <p>Sekundäre Wirksamkeitsendpunkte und explorative Endpunkte wurden jeweils einzeln zusammengefasst und in einer ähnlichen Weise analysiert wie für den primären Wirksamkeitsendpunkt beschrieben. Für kontinuierliche Endpunkte wurden deskriptive Statistiken (N, Mittelwert, Median, Standardabweichung, Quartile, Minimum und Maximum) von der Behandlungsgruppe bei spezifizierten Visiten pro Studienphase ermittelt. Endpunkte, die zum Ende der Placebo-kontrollierten Phase ausgewertet wurden, wurden unter Verwendung eines ANCOVA-Modells verglichen (ähnlich dem primären Endpunkt).</p> <p>Die Responder-Analyse auf Basis der ITT-Population zu Woche 12 wurde mit dem Cochran-Mantel-Haenszel (CMH)-Test auf Signifikanz getestet ($p < 0,05$). Der Test erfolgte unter Berücksichtigung der a priori im SAP spezifizierten Stratifizierungsfaktoren. Patienten, die vor Woche 12 aus der Studie ausschieden oder für die zu Woche 12 keine Daten vorlagen, wurden als Non-Responder zu Woche 12 eingestuft.</p> <p>Sicherheitsanalysen</p> <p>Die Sicherheitsanalysen wurden anhand der Safety Population durchgeführt. Die Sicherheit wurde durch klinische Prüfung aller relevanten Parameter bewertet, einschließlich UE, UE von besonderem Interesse (d. h. Diarrhoe), Labortests und Vitalparameter; es wurde kein Interferenztest für statistische Signifikanz durchgeführt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Analysen des primären Wirksamkeitsendpunkts (AUC_{W0-12} für die Anzahl der oralen Aphthen) basierend auf der ITT-Population wurden für die folgenden Subgruppen durchgeführt:</p> <ul style="list-style-type: none"> • Geschlecht (männlich, weiblich) • Ethnie (Weiße, Nicht-Weiße) • Alter (< 40, ≥ 40 bis < 65, ≥ 65 Jahre) • Region (Japan, Nicht-Japan; Europa, Nordamerika, Asien, Rest der Welt) • Baseline-Gewichtskategorie (< 50, ≥ 50 bis < 60, ≥ 60 bis < 70, ≥ 70 bis < 85, ≥ 85 kg) • Baseline-BMI-Kategorie ($< 18,5$, $\geq 18,5$ bis < 25, ≥ 25 bis < 30, ≥ 30 kg/m²)

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Alkoholkonsum (ja, nein) • Tabakkonsum (aktueller Raucher, ehemaliger Raucher, Nichtraucher) • Krankheitsdauer (≤ 2, > 2 bis ≤ 6, > 6 bis ≤ 10, > 10 Jahre) • Baseline-Kategorie oraler Aphthen (≥ 2 bis ≤ 5, > 5 bis ≤ 10, > 10) • Vorangegangene Colchicin-Einnahme (ja, nein) • Vorangegangene Einnahme oraler Kortikosteroide (ja, nein)
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Eine Darstellung des Patientenflusses findet sich in Abbildung 15 im Anschluss an diese 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	<p><u>Placebo-kontrollierte Behandlungsphase (bis einschließlich Woche 12):</u></p> <p>a) Anzahl aller randomisierten Patienten: n = 207</p> <ul style="list-style-type: none"> • Apremilast: n = 104 • Placebo: n = 103 <p>b) Anzahl der Patienten, die tatsächlich die geplante Behandlung erhalten haben: n = 207</p> <ul style="list-style-type: none"> • Apremilast: n = 104 • Placebo: n=103 <p>c) Anzahl der Patienten, die in der Analyse des primären Zielkriteriums berücksichtigt wurden: n = 207</p> <ul style="list-style-type: none"> • Apremilast: n = 104 • Placebo: n = 103 <p><u>Aktive Behandlungsphase (Woche 13 – 64) begonnen:</u> <u>n = 178</u></p> <ul style="list-style-type: none"> • Apremilast/Apremilast: n = 95 • Placebo/Apremilast: n = 83 <p><u>Aktive Behandlungsphase (Woche 13 – 64) komplettiert:</u> n = 143</p> <ul style="list-style-type: none"> • Apremilast/Apremilast: n = 75 • Placebo/Apremilast: n = 68
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p><u>Placebo-kontrollierte Behandlungsphase (bis einschließlich Woche 12):</u></p> <p>Apremilast: n = 8 (7,7 %)</p> <ul style="list-style-type: none"> • UE: n = 3 (2,9 %) • Patientenwunsch: n = 4 (3,8 %) • Protokollverletzung: n = 1 (1,0 %) <p>Placebo: n = 20 (19,4 %)</p>

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • UE: n = 4 (3,9 %) • Fehlende Wirksamkeit: n = 8 (7,8 %) • Noncompliance: n = 1 (1,0 %) • Patientenwunsch: n = 5 (4,9 %) • Lost-to-follow-up: n = 1 (1,0 %) • Protokollverletzung: n = 1 (1,0 %) <p><u>Aktive Behandlungsphase (Woche 13 – 64):</u></p> <p>Apremilast/Apremilast: n = 20 (21,1 %)</p> <ul style="list-style-type: none"> • UE: n = 9 (9,5 %) • Fehlende Wirksamkeit: n = 2 (2,1 %) • Patientenwunsch: n = 7 (7,4 %) • Lost-to-follow-up: n = 1 (1,1 %) • Andere: n = 1 (1,1 %) <p>Placebo/Apremilast: n = 15 (18,1 %)</p> <ul style="list-style-type: none"> • UE: n = 4 (4,8 %) • Fehlende Wirksamkeit: n = 2 (2,4 %) • Patientenwunsch: n = 7 (8,4 %) • Lost to follow-up: n = 1 (1,2 %) • Schwangerschaft: n = 1 (1,2 %)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Screening-Phase erstreckte sich über bis zu sechs Wochen. Nach der 12-wöchigen Placebo-kontrollierten Behandlungsphase und der anschließenden 52-wöchigen aktiven Behandlungsphase erfolgte eine Nachbeobachtungszeit über vier Wochen.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde planmäßig beendet.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

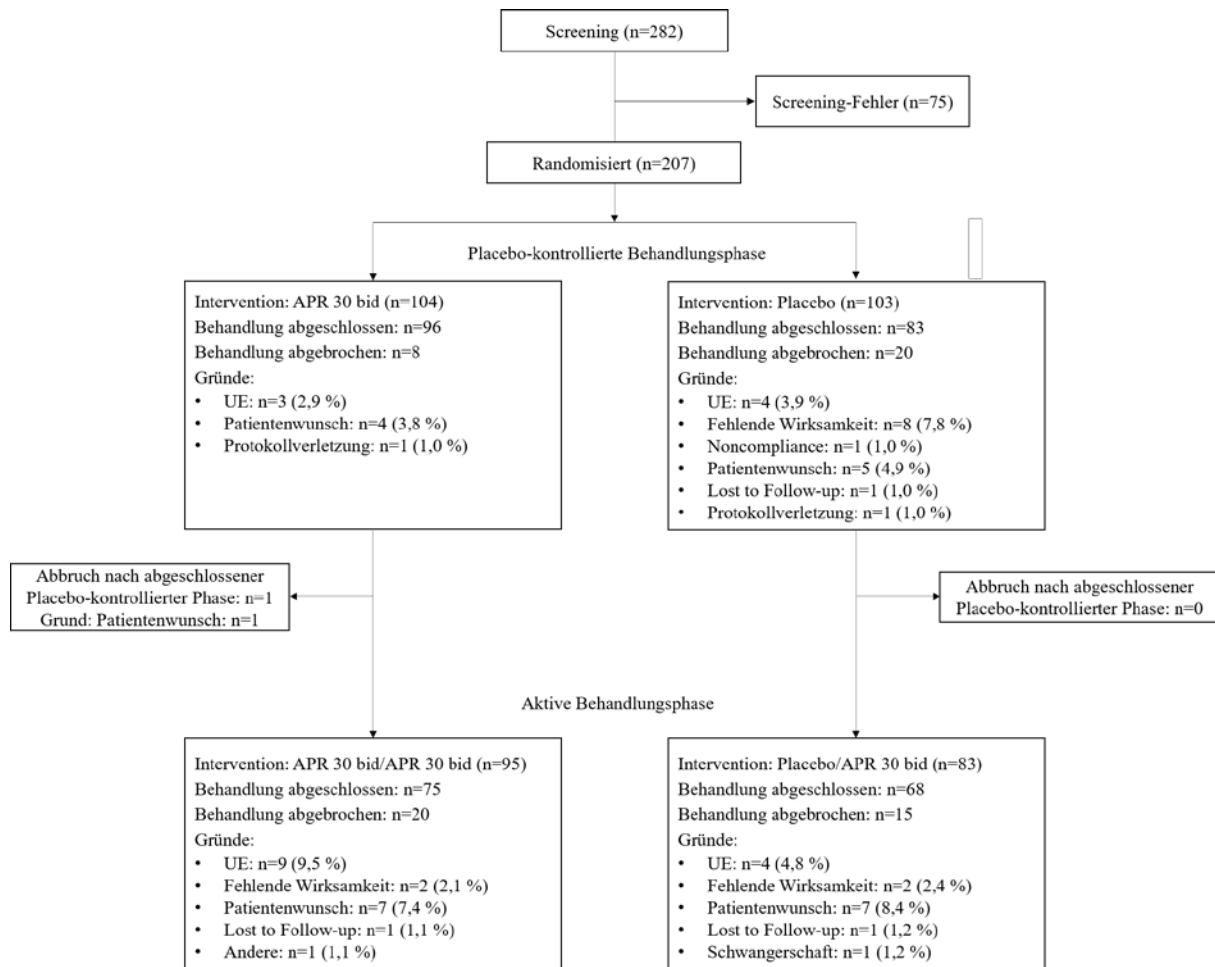


Abbildung 15: Flow-Chart zur Studie REFLIEF

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 4-103 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie CC-10004-BCT-002

Studie: RELIEF

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Clinical Study Report, RELIEF (CC-10004-BCT-002); A Phase 3 Randomized, Double-blind, Placebo-controlled, Parallel Group Study Followed by an Active Treatment Phase to Evaluate the Efficacy and Safety of Apremilast (CC-10004) in the Treatment of Subjects With Active Behçet's Disease Final Report: 12.04.2019	RELIEF

A Verzerrungsaspekte auf Studienebene:

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Studie wurde randomisiert durchgeführt.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Zuteilung der Patienten erfolgte zufällig unter Verwendung einer interaktiven Response Technologie (IRT).

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 neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die zufällige Zuteilung der Patienten zu einer der beiden Behandlungsgruppen erfolgte unter Verwendung einer IRT.

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:

Es handelt sich um eine doppelblinde Studie.

behandelnde bzw. weiterbehandelnde Personen: ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

4. **Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte** ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung.

5. **Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können** ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

In den Studienunterlagen finden sich keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Die Randomisierung wurde adäquat durchgeführt, und die Gruppenzuteilung erfolgte rein zufällig mittels einer IRT. In den Studienunterlagen finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es liegen auch keine weiteren, das Ergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial als niedrig eingestuft wird. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten und mit mindestens einer Dosis der Studienmedikation behandelten Patienten einschließt.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

**Endpunkt: Vollständige Remission der oralen Aphthen (komplettes Ansprechen)
(Der Endpunkt umfasst den Anteil der Patienten mit vollständiger Remission der oralen Aphthen, die Zeit bis zur vollständigen Remission der oralen Aphthen, die Anzahl der oralen Aphthen und die Dauer der vollständigen Remission der oralen Aphthen)**

1. Verblindung der Endpunkterheber

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten und mit mindestens einer Dosis der Studienmedikation behandelten Patienten umfasste. Alle Patienten wurden in die Auswertung einbezogen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Vollständige Remission der oralen Aphthen (komplettes Ansprechen).

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Vollständige Remission der oralen Aphthen (komplettes Ansprechen) als niedrig eingestuft wird.

Endpunkt: Schmerzhaftigkeit der oralen Aphthen mittels VAS

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war. In beiden Behandlungsgruppen wurden jeweils 99 % der Patienten in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei weniger als 5 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Schmerzhaftigkeit der oralen Aphthen mittels VAS.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war. In beiden Behandlungsgruppen wurden jeweils 99 % der Patienten in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei weniger als 5 %. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Schmerzhaftigkeit der oralen Aphthen mittels VAS als niedrig eingestuft wird.

Endpunkt: Veränderung des Behçet's Syndrome Activity Score (BSAS) zu Woche 12**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten und mit mindestens einer Dosis der Studienmedikation behandelten Patienten umfasste und für die eine Erhebung zu Baseline verfügbar war. Alle Patienten wurden in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei 0 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Veränderung des BSAS zu Woche 12.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Veränderung des BSAS zu Woche 12 als niedrig eingestuft wird.

Endpunkt: Veränderung der Behçet's Disease Current Activity Form (BDCAF) zu Woche 12

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war. In der Apremilast-Gruppe wurden alle Patienten, in der Placebo-Gruppe 99 % der Patienten in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei weniger als 5 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Veränderung des BDCAF zu Woche 12.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die eine Erhebung zu Baseline verfügbar war. In der Apremilast-Gruppe wurden alle Patienten, in der Placebo-Gruppe 99 % der Patienten in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei weniger als 5 %. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Veränderung des BDCAF zu Woche 12 als niedrig eingestuft wird.

Endpunkt: Veränderung des Behçet's disease Quality of Life (BD-QoL) zu Woche 12**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten und mit mindestens einer Dosis der Studienmedikation behandelten Patienten umfasste und für die eine Erhebung zu Baseline verfügbar war. Alle Patienten wurden in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei 0 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Veränderung des BD-QoL zu Woche 12.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die ITT-Population, die alle randomisierten

Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Veränderung des BD-QoL zu Woche 12 als niedrig eingestuft wird.

Endpunkt: Verbesserung des 36-Item Short Form Health Survey, Version 2 (SF-36) Fragebogens um $\geq 2,5$ Punkte zu Woche 12

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die statistische Analyse basierte auf der ITT-Population, die alle randomisierten und mit mindestens einer Dosis der Studienmedikation behandelten Patienten umfasste und für die eine Erhebung zu Baseline verfügbar war. Alle Patienten wurden in die Auswertung einbezogen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Veränderung des SF-36 zu Woche 12.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten und für die ein Baseline-Wert verfügbar war. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Veränderung des SF-36 zu Woche 12 als niedrig eingestuft wird.

Endpunkt: Sicherheit

(Der Endpunkt umfasst UE gesamt, SUE, schwere UE, Therapieabbruch aufgrund von UE und UE von besonderem Interesse)

1. Verblindung der Endpunkterheber

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse erfolgte für die Sicherheitspopulation, die alle mit Studienmedikation behandelten Patienten (as-treated) einschloss. Die Sicherheitspopulation stimmte mit der ITT-Population vollständig überein, da alle behandelten Patienten die ihnen zugewiesene Studienmedikation erhielten. Alle Patienten wurden in die Auswertung einbezogen. Der Nichtberücksichtigungsanteil lag somit bei 0 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts Sicherheit.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial wird als niedrig eingestuft, da eine Beeinflussung des Ergebnisses durch die doppelte Verblindung weder durch die Patienten, die behandelnden bzw. weiterbehandelnden Personen, noch durch die Endpunkterheber zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte für die Sicherheitspopulation, die alle mit Studienmedikation behandelten Patienten (as-treated) einschloss. Es liegen auch keine sonstigen, das Endpunktergebnis möglicherweise verzerrenden Aspekte vor, sodass das Verzerrungspotenzial des Endpunkts Sicherheit als niedrig eingestuft wird.
