

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Seladelpar (Seladelpar Gilead®)

Gilead Sciences GmbH

Modul 4A

*Zur Behandlung der primär biliären Cholangitis (PBC)
in Kombination mit UDCA bei Erwachsenen, die nicht
ausreichend auf UDCA alleine ansprechen, oder als
Monotherapie bei Patienten, die UDCA nicht vertragen*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Inhaltsverzeichnis.....	1
Tabellenverzeichnis.....	4
Abbildungsverzeichnis.....	8
Abkürzungsverzeichnis.....	10
4 Modul 4 – allgemeine Informationen.....	13
4.1 Zusammenfassung der Inhalte von Modul 4.....	14
4.2 Methodik.....	24
4.2.1 Fragestellung.....	24
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	26
4.2.3 Informationsbeschaffung.....	28
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	29
4.2.3.2 Bibliografische Literaturrecherche.....	29
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken.....	30
4.2.3.4 Suche auf der Internetseite des G-BA.....	32
4.2.3.5 Selektion relevanter Studien.....	33
4.2.4 Bewertung der Aussagekraft der Nachweise.....	33
4.2.5 Informationssynthese und -analyse.....	35
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	35
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	36
4.2.5.3 Meta-Analysen.....	48
4.2.5.4 Sensitivitätsanalysen.....	50
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	51
4.2.5.6 Indirekte Vergleiche.....	53
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	55
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	55
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	55
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	55
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	57
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken.....	59
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA.....	60
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	61
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	62
4.3.1.2.1 Studiendesign und Studienpopulationen.....	62
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	73
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	74
4.3.1.3.1 Endpunkte – RCT.....	74
4.3.1.3.1.1 Mortalität – RCT.....	77
4.3.1.3.1.2 Biochemisches Ansprechen – RCT.....	78

4.3.1.3.1.3	Pruritus NRS – RCT	88
4.3.1.3.1.4	5-D Pruritus – RCT.....	92
4.3.1.3.1.5	PGI-C & PGI-S – RCT	101
4.3.1.3.1.6	Fibrose – RCT	106
4.3.1.3.1.7	Hospitalisierung – RCT	109
4.3.1.3.1.8	Gesundheitsbezogene Lebensqualität PBC-40 – RCT	111
4.3.1.3.1.9	Unerwünschte Ereignisse – RCT.....	119
4.3.1.3.2	Subgruppenanalysen – RCT.....	129
4.3.1.3.2.1	Biochemisches Ansprechen – RCT	149
4.3.1.3.2.2	Pruritus NRS – RCT	153
4.3.1.3.2.3	5-D Pruritus – RCT.....	155
4.3.1.3.2.4	PGI-C & PGI-S – RCT	166
4.3.1.3.2.5	Fibrose – RCT	169
4.3.1.3.2.6	Gesundheitsbezogene Lebensqualität PBC-40 – RCT	172
4.3.1.3.2.7	Unerwünschte Ereignisse – RCT.....	176
4.3.1.3.2.8	Fazit Subgruppenanalysen	176
4.3.1.4	Liste der eingeschlossenen Studien - RCT	177
4.3.2	Weitere Unterlagen.....	179
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	179
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	179
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	179
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	180
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	180
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	182
4.3.2.1.4	Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT ...	183
4.3.2.2	Nicht randomisierte vergleichende Studien.....	183
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	183
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	183
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	184
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	184
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	186
4.3.2.2.4	Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien	186
4.3.2.3	Weitere Untersuchungen.....	186
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	186
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	187
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	187
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	187
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	188
4.3.2.3.4	Liste der eingeschlossenen Studien – weitere Untersuchungen.....	188
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	188
4.4.1	Beurteilung der Aussagekraft der Nachweise	188
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	189

4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	207
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	207
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	207
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	208
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	208
4.5.4	Verwendung von Surrogatendpunkten.....	208
4.6	Referenzliste.....	213
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		219
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken.....		223
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		225
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken).....		226
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		234
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		253
Anhang 4-G : Ergänzende Unterlagen		272

Tabellenverzeichnis

	Seite
Tabelle 4-1: Ausmaß des Zusatznutzen von SEL gegenüber Placebo auf Endpunktebene (Studie RESPONSE)	16
Tabelle 4-2: Übersicht der Kriterien für den Studieneinschluss für den Vergleich von SEL bei erwachsenen Patienten	26
Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zbAM.....	56
Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zbAM	57
Tabelle 4-5: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zbAM.....	60
Tabelle 4-6: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zbAM.....	61
Tabelle 4-7: Studienpool – RCT mit dem zbAM.....	62
Tabelle 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zbAM	63
Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zbAM	64
Tabelle 4-10: Charakterisierung der Studienpopulationen der Studie RESPONSE (SAS) – demografische und krankheitsbedingte Charakteristika aus RCT mit dem zbAM.....	64
Tabelle 4-11: Charakterisierung der Studienpopulation der Studie RESPONSE – Patientendisposition – RCT mit dem zbAM	68
Tabelle 4-12: Begleitmedikationen ($\geq 10\%$) der Studie RESPONSE (SAS) – RCT mit dem zbAM.....	69
Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zbAM	73
Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zbAM	74
Tabelle 4-15: Operationalisierung von Mortalität.....	77
Tabelle 4-16: Bewertung des Verzerrungspotenzials für Mortalität in RCT mit dem zbAM..	78
Tabelle 4-17: Operationalisierung von biochemischem Ansprechen	79
Tabelle 4-18: Bewertung des Verzerrungspotenzials für biochemisches Ansprechen in RCT mit dem zbAM.....	79
Tabelle 4-19: Ergebnisse für biochemisches Ansprechen aus RCT mit dem zbAM (Studie RESPONSE, FAS)	80
Tabelle 4-20: Ergebnisse für biochemische Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)	82
Tabelle 4-21: Operationalisierung von Pruritus NRS	88
Tabelle 4-22: Bewertung des Verzerrungspotenzials für Pruritus NRS in RCT mit dem zbAM.....	89
Tabelle 4-23: Rücklaufquoten für Pruritus NRS aus RCT mit zbAM (Studie RESPONSE, FAS)	89

Tabelle 4-24: Ergebnisse für Pruritus NRS aus RCT mit dem zbAM (Studie RESPONSE, FAS)	90
Tabelle 4-25: Ergebnisse für Pruritus NRS Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)	90
Tabelle 4-26: Operationalisierung von 5-D Pruritus.....	92
Tabelle 4-27: Bewertung des Verzerrungspotenzials für 5-D Pruritus in RCT mit dem zbAM.....	93
Tabelle 4-28: Rücklaufquoten für 5-D Pruritus aus RCT mit zbAM (Studie RESPONSE, FAS)	93
Tabelle 4-29: Ergebnisse für 5-D Pruritus Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)	95
Tabelle 4-30: Operationalisierung der Bewertungsskalen PGI-C & PGI-S.....	102
Tabelle 4-31: Bewertung des Verzerrungspotenzials für PGI-C & PGI-S in RCT mit dem zbAM.....	102
Tabelle 4-32: Rücklaufquoten für PGI aus RCT mit zbAM (Studie RESPONSE, FAS).....	103
Tabelle 4-33: Ergebnisse für PGI – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)	103
Tabelle 4-34: Operationalisierung von Fibrose.....	106
Tabelle 4-35: Bewertung des Verzerrungspotenzials für Fibrose in RCT mit dem zbAM....	106
Tabelle 4-36: Ergebnisse für Fibrose – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)	107
Tabelle 4-37: Operationalisierung von Hospitalisierung	109
Tabelle 4-38: Bewertung des Verzerrungspotenzials für Hospitalisierung in RCT mit dem zbAM.....	110
Tabelle 4-39: Ergebnisse für Hospitalisierung aus RCT mit dem zbAM (Studie RESPONSE, FAS)	110
Tabelle 4-40: Operationalisierung von PBC-40.....	111
Tabelle 4-41: Bewertung des Verzerrungspotenzials für PBC-40 in RCT mit dem zbAM... 112	
Tabelle 4-42: Rücklaufquoten für PBC-40 aus RCT mit zbAM (Studie RESPONSE, FAS) 112	
Tabelle 4-43: Ergebnisse für PBC-40 Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	114
Tabelle 4-44: Operationalisierung von UE	120
Tabelle 4-45: Bewertung des Verzerrungspotenzials für UE in RCT mit dem zbAM	121
Tabelle 4-46: Ergebnisse für UE (Gesamtraten) aus RCT mit dem zbAM (Studie RESPONSE, SAS)	121
Tabelle 4-47: Ergebnisse für UE ohne erkrankungsbezogene Ereignisse aus RCT mit dem zbAM (Studie RESPONSE, SAS)	122
Tabelle 4-48: Ergebnisse für UESI aus RCT mit dem zbAM (Studie RESPONSE, SAS)....	123

Tabelle 4-49: Ergebnisse für UESI (SUE) aus RCT mit dem zbAM (Studie RESPONSE, SAS)	124
Tabelle 4-50: Ergebnisse für UESI (Schwere UE [CTCAE-Grad $\geq$ 3]) aus RCT mit dem zbAM (Studie RESPONSE, SAS)	125
Tabelle 4-51: Ergebnisse für Jegliche UE nach SOC und PT aus RCT mit dem zbAM (Studie RESPONSE, SAS).....	126
Tabelle 4-52: Ergebnisse für UE, die zum Therapieabbruch führten (ergänzend, deskriptiv) nach SOC und PT aus RCT mit dem zbAM (Studie RESPONSE, SAS).....	128
Tabelle 4-53 Matrix der durchgeführten Subgruppenanalysen.....	131
Tabelle 4-54: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt (Studie RESPONSE).....	133
Tabelle 4-55: Präspezifizierte und für die Nutzenbewertung herangezogenen Subgruppenvariablen.....	146
Tabelle 4-56: Ergebnisse für Subgruppenanalysen zum Endpunkt Gesamt-Bilirubin – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	149
Tabelle 4-57: Ergebnisse für Subgruppenanalysen zum Endpunkt ALP – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)	151
Tabelle 4-58: Ergebnisse für Subgruppenanalysen zum Endpunkt Pruritus NRS – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)	153
Tabelle 4-59: Ergebnisse für Subgruppenanalysen zum Endpunkt 5-D Pruritus – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12.....	155
Tabelle 4-60: Ergebnisse für Subgruppenanalysen zum Endpunkt PGI-S – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12.....	166
Tabelle 4-61: Ergebnisse für Subgruppenanalysen zum Endpunkt Lebersteifigkeit (FibroScan [®]) aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12.....	169
Tabelle 4-62: Ergebnisse Subgruppenanalysen zum Endpunkt für PBC-40 – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12.....	172
Tabelle 4-63: Ergebnisse Subgruppenanalysen zum Endpunkt für Kardiovaskuläre Toxizität (UESI) aus RCT mit dem zbAM (Studie RESPONSE, SAS).....	176
Tabelle 4-64: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	180
Tabelle 4-65: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	180
Tabelle 4-66: Operationalisierung von <Endpunkt xxx>.....	181
Tabelle 4-67: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	181
Tabelle 4-68: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	182
Tabelle 4-69: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	184
Tabelle 4-70: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien	184

Tabelle 4-71: Operationalisierung von <Endpunkt xxx>.....	185
Tabelle 4-72: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	185
Tabelle 4-73: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen	187
Tabelle 4-74: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	187
Tabelle 4-75: Ausmaß des Zusatznutzen von SEL gegenüber Placebo auf Endpunktebene (Studie RESPONSE).....	193
Tabelle 4-76: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	207
Tabelle 4-77 (Anhang): Ovid: Suchstrategie EMBASE Classic+EMBASE 1947 bis heute: 14.01.2025	220
Tabelle 4-78 (Anhang): Ovid: Suchstrategie: MEDLINE(R) and Epub Ahead of Print, In- Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions 1946 bis heute: 14.01.2025	221
Tabelle 4-79 (Anhang): Ovid Suchstrategie: EBM Reviews – Cochrane Central Register of Controlled Trials bis Dezember 2024	222
Tabelle 4-80 (Anhang): Übersicht der ausgeschlossenen Studien aus der Studienregistersuche zum zbAM	226
Tabelle 4-81 (Anhang): Studiendesign und -methodik für Studie RESPONSE	235
Tabelle 4-82 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie RESPONSE.....	254

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zbAM	58
Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach RCT mit dem zbAM SEL	59
Abbildung 3: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der ALP aus RCT mit dem zbAM (Studie RESPONSE, FAS)	84
Abbildung 4: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des Gesamt-Bilirubins aus RCT mit dem zbAM (Studie RESPONSE, FAS)	84
Abbildung 5: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des direkten Bilirubins aus RCT mit dem zbAM (Studie RESPONSE, FAS)	85
Abbildung 6: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des indirekten Bilirubins aus RCT mit dem zbAM (Studie RESPONSE, FAS)	85
Abbildung 7: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der ALT aus RCT mit dem zbAM (Studie RESPONSE, FAS)	86
Abbildung 8: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der AST aus RCT mit dem zbAM (Studie RESPONSE, FAS)	86
Abbildung 9: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der GGT aus RCT mit dem zbAM (Studie RESPONSE, FAS)	87
Abbildung 10: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der 5'-Nucleotidase aus RCT mit dem zbAM (Studie RESPONSE, FAS)	87
Abbildung 11: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Pruritus NRS aus RCT mit dem zbAM (Studie RESPONSE, FAS)	91
Abbildung 12: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des Gesamt-Scores des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	98
Abbildung 13: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Grad des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	98
Abbildung 14: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Dauer des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	99
Abbildung 15: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Beeinträchtigung des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	99
Abbildung 16: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Verlauf des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	100
Abbildung 17: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Lokalisation des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	100
Abbildung 18: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des modifizierten Gesamt-Scores des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)	101

Abbildung 19: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des PGI-S aus RCT mit dem zbAM (Studie RESPONSE, FAS)	105
Abbildung 20: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des PGI-C aus RCT mit dem zbAM (Studie RESPONSE, FAS)	105
Abbildung 21: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Lebersteifigkeit (FibroScan®) aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	108
Abbildung 22: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des erweiterten Leberfibrosetests (ELF-Score) aus RCT mit dem zbAM (Studie RESPONSE, FAS)	108
Abbildung 23: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Gesamt-Scores aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	116
Abbildung 24: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Kognitive Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	116
Abbildung 25: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Emotionale Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS)	117
Abbildung 26: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Fatigue Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	117
Abbildung 27: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Pruritus Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	118
Abbildung 28: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Soziale Domäne RCT mit dem zbAM (Studie RESPONSE, FAS)	118
Abbildung 29: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Symptomatische Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS).....	119
Abbildung 30: Mittlere Pruritus NRS, Patienten mit mittlerem bis schweren Pruritus zu Baseline, Subgruppe NRS ≥ 4 Punkte zu Baseline	206
Abbildung 31 (Anhang): Flow-Chart der Studie RESPONSE.....	252

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALP	Alkalische Phosphatase
ALT	Alanin-Aminotransferase
AMA	Antimitochondriale Antikörper
AMIce	Arzneimittel-Informationssystem
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANA	Antinukleäre Antikörper
AST	Aspartat-Aminotransferase
ATC	Anatomisch-Therapeutisch-Chemisch
BMI	Body Mass Index
CD	Cluster of Differentiation
CMH	Cochran-Mantel-Haenszel
CONSORT	Consolidated Standards of Reporting Trials
COVID-19	Coronavirus-Krankheit (Coronavirus Disease) 2019
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	Datenüberwachungskomitee (Data Safety Monitoring Board)
EG	Europäische Gemeinschaft
eGFR	Geschätzte glomeruläre Filtrationsrate (Estimated Glomerular Filtration Rate)
ELF	Fortgeschrittene Leberfibrose (Enhanced Liver Fibrosis)
ELISA	Enzyme-Linked Immunosorbent Assay
EMA	Europäische Arzneimittel-Agentur (European Medicines Agency)
EMBASE	Excerpta Medica Database
EU	European Union
EU-CTR	European Union Clinical Trials Register
FAS	Full-Analysis-Set
FDA	Food and Drug Administration
FTC	Emtricitabin
G-BA	Gemeinsamer Bundesausschuss
GGT	Gamma-Glutamyltransferase
HCC	Hepatozelluläres Karzinom (Hepatocellular Carcinoma)
HDL-C	High Density Lipoprotein-Cholesterol

Abkürzung	Bedeutung
HIV	Humanes Immundefizienzvirus
ICTRP	International Clinical Trials Registry Platform
INR	International Normalized Ratio
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
IWRS	Interaktives Webdialogsystem (Interactive Web Response System)
KI (CI)	Konfidenzintervall (Confidence Interval)
LDL-C	Low Density Lipoprotein-Cholesterol
LS-MW	Kleinste-Quadrate (Least-Squares)-Mittelwert
LS-MWD	Kleinste-Quadrate (Least-Squares)-Mittelwertdifferenz
MedDRA	Medical Dictionary for Regulatory Activities
MEDLINE	Medical Literature Analysis and Retrieval System Online
MELD	Model for End-Stage Liver Disease
MID	Minimal Important Difference
MMRM	Mixed Effect Model Repeat Measurement
MSPN	Moderater bis schwerer Pruritus Numerical Rating Scale
MW	Mittelwert
NCT	National Clinical Trial
NA	Nicht auswertbar
N.D.	Nicht durchgeführt
NRS	Numerical Rating Scale
OCA	Obeticholsäure (Obeticholic Acid)
OR	Odds Ratio
PBC	Primär biliäre Cholangitis
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PPAR	Peroxisom-Proliferator-aktivierter Rezeptor
PT	Preferred Term
RCT	Randomisierte kontrollierte Studie (Randomized Controlled Trial)
RD	Risikodifferenz
RR	Relatives Risiko
SAP	Statistischer Analyseplan

Abkürzung	Bedeutung
SAS	Safety-Analysis-Set
SD	Standardabweichung (Standard Deviation)
SE	Standardfehler (Standard Error)
SEL	Seladelpar
SGB	Sozialgesetzbuch
SMD	Standardisierte Mittelwertdifferenz
SOC	System Organ Class
SUE	Schwerwiegendes unerwünschtes Ereignis
TE	Transiente Elastografie
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UDCA	Ursodeoxycholsäure (Ursodeoxycholic Acid)
UE	Unerwünschtes Ereignis
UESI	Unerwünschtes Ereignis von speziellem Interesse
UK	United Kingdom
ULN	Oberer Normwert (Upper Limit of Normal)
VerfO	Verfahrensordnung
WHO	Weltgesundheitsorganisation (World Health Organization)
zbAM	Zu bewertendes Arzneimittel
zVT	Zweckmäßige Vergleichstherapie

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

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

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Ziel der vorliegenden Untersuchung ist die Bewertung des medizinischen Nutzens und medizinischen Zusatznutzens von Seladelpar (SEL). Gemäß Fachinformation findet SEL Anwendung in der Behandlung der primär biliären Cholangitis (PBC) in Kombination mit Ursodeoxycholsäure (UDCA) bei Erwachsenen, die nicht ausreichend auf UDCA alleine ansprechen, oder als Monotherapie bei Patienten, die UDCA nicht vertragen.

Bei SEL handelt es sich um ein Arzneimittel zur Behandlung einer seltenen Erkrankung (Orphan Drug) nach der Verordnung (Europäische Gemeinschaft [EG]) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. Der Orphan Drug-Status wurde am 5. Dezember 2024 durch die Europäische Arzneimittel-Agentur (EMA) bestätigt. Gemäß § 35a Abs. 1 Satz 11 Sozialgesetzbuch (SGB) V gilt der medizinische Zusatznutzen für Orphan Drugs durch die Zulassung als belegt. Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie (zVT) müssen nach § 35a Abs. 1 Satz 12 SGB V nur nach Aufforderung durch den Gemeinsamen Bundesausschuss (G-BA) vorgelegt werden, wenn der Umsatz des Arzneimittels in den letzten zwölf Kalendermonaten einen Betrag von 30 Millionen Euro übersteigt. Es ist lediglich das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien für die Anzahl der Patienten und Patientengruppen nachzuweisen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht. Es bedarf keiner Ausführungen zu einer zVT.

Die zu bewertende Intervention SEL soll einmal täglich in einer Dosierung von 10 mg oral als Hartkapsel verabreicht werden.

Die Nutzenbewertung erfolgt, sofern verfügbar, nach den patientenrelevanten Endpunkten Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse (UE).

Datenquellen

Die Identifizierung der für die Nutzenbewertung geeigneten Studien erfolgte über firmeninterne Informationsquellen der Gilead Sciences GmbH (im Folgenden Gilead), eine systematische bibliografische Literaturrecherche in den Datenbanken Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica Database (EMBASE) und Cochrane (Cochrane Central Register of Controlled Trials) sowie eine Studienregisterrecherche in den Portalen clinicaltrials.gov, European Union Clinical Trials Register (EU-CTR) und International Clinical Trials Registry Platform (ICTRP).

Als Datenquellen dienen, sofern verfügbar, Studienberichte inklusive Appendizes und zusätzlich, wenn möglich, post-hoc durchgeführte Zusatzanalysen der relevanten Studien. Datenquellen für Studien, die ausschließlich über die bibliografische Literaturrecherche oder die systematische Studienregistersuche identifiziert werden, sind Publikationen und

Registereinträge. Für schon identifizierte Studien werden außerdem Datenquellen auf der Seite des G-BA, der EMA und des Arzneimittel-Informationssystems (AMIce) gesucht.

Ein-/Ausschlusskriterien für Studien

Die Einschlusskriterien für die einzuschließenden Studien wurden wie folgt definiert:

- Population: Erwachsene Patienten mit einer PBC, die nicht ausreichend auf UDCA alleine ansprechen oder die UDCA nicht vertragen
- Intervention: SEL gemäß Fachinformation
- Vergleichstherapie: Keine Einschränkung
- Endpunkte: Mortalität, Morbidität, gesundheitsbezogene Lebensqualität, UE
- Studientyp: Randomisierte kontrollierte Studie (RCT)
- Studiendauer: ≥ 48 Wochen
- Publikationstyp: Vollpublikation, Ergebnisse aus Studienregistern oder Studienbericht verfügbar, der den Kriterien des Consolidated Standards of Reporting Trials (CONSORT)-Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht

Ausschlusskriterium ist jeweils das Nichterfüllen von mindestens einem der definierten Einschlusskriterien.

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

Die eingeschlossenen Studien wurden mittels CONSORT-Kriterien beschrieben und das Verzerrungspotenzial auf Studien- und Endpunktebene wurde gemäß den Vorgaben aus Anhang 4-F bewertet.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Im Folgenden werden die Ergebnisse zum medizinischen Zusatznutzen von SEL für alle relevanten Endpunkte beschrieben. Der Nutzenbewertung werden die Ergebnisse der RCT RESPONSE zugrunde gelegt. Erhoben wurden Ergebnisse zu einer Vielzahl von Monaten, dargestellt werden die Ergebnisse zu Monat 6 und Monat 12 (alle weiteren sind dem Anhang 4-G zu entnehmen). Sowohl bei Monat 6 als auch bei Monat 12 handelt es sich jeweils um eine Momentaufnahme des aktuellen Erkrankungszustands der Patienten. Speziell hinsichtlich der Symptomatik und Lebensqualität ist der Verlauf bei der PBC nicht linear progredient, sondern kann Tagesform-abhängigen Schwankungen unterliegen. Durch die gleichwertige Betrachtung beider Erhebungszeitpunkte kann eine umfassendere Betrachtung der Ergebnisse, unter Berücksichtigung möglicher Veränderungen, gewährleistet werden. Daher werden beide Zeitpunkte als gleich relevant angesehen und zur Ableitung des Zusatznutzens herangezogen.

Das Ausmaß des Zusatznutzens wird anhand der folgenden patientenrelevanten Endpunkte abgeleitet:

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Mortalität		
Mortalität	Es ist kein Patient verstorben.	Kein Zusatznutzen/Schaden ableitbar
Morbidität – Biochemisches Ansprechen		
ALP-Normalisierung (ALP ≤1.0 x ULN) Monat 6 Monat 12 Kombiniertes biochemisches Ansprechen Monat 6 Monat 12 ALP – Stetige Analyse Monat 6 Monat 12 Gesamt-Bilirubin – Stetige Analyse Monat 6 Monat 12	RR: 16,40 [2,32; 116,20], 0,0051 RR: 14,80 [2,08; 105,44], 0,0071 RR: 3,60 [2,16; 6,01], <0,0001 RR: 3,09 [1,87; 5,10], <0,0001 LS-MWD: -122,56 [-142,15; -102,98], <0,0001 SMD: -1,70 [-2,05; -1,35] LS-MWD: -117,02 [-144,20; -89,84], <0,0001 SMD: -1,23 [-1,56; -0,90] LS-MWD: -0,07 [-0,15; 0,01], 0,0896 SMD: -0,26 [-0,56; -0,05] LS-MWD: -0,02 [-0,14; 0,11], 0,7777 SMD: -0,04 [-0,35; 0,26]	Ausmaß: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis
Morbidität – Symptomatik^a		
Pruritus NRS – Stetige Analyse Monat 6	LS-MWD: -0,90 [-1,35; -0,45], <0,0001	<u>Gesamtpopulation:</u> Ausmaß:

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Monat 12	SMD: -0,58 [-0,89; -0,28] LS-MWD: -0,83 [-1,35; -0,31], 0,0018 SMD: -0,47 [-0,78; -0,17]	Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 12	LS-MWD: -1,77 [-2,97; -0,57], 0,0045 SMD: -0,76 [-1,28; -0,24]	<u>Für Patienten mit Pruritus NRS ≥ 4:</u> Ausmaß:
5-D Pruritus Veränderung – Gesamt-Score		Erheblicher Zusatznutzen Wahrscheinlichkeit: Hinweis
Monat 6	LS-MWD: -1,76 [-2,72; 0,81], 0,0003 SMD: -0,53 [-0,83; -0,22]	
Monat 12	LS-MWD: -2,31 [-3,40; -1,23], <0,0001 SMD: -0,62 [-0,93; -0,31]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -3,42 [-5,32; -1,51], 0,0006 SMD: -0,91 [-1,43; -0,38]	
5-D Pruritus Veränderung – Subdomäne Grad		
Monat 6	LS-MWD: -0,35 [-0,56; -0,15], 0,0008 SMD: -0,49 [-0,80; -0,19]	
Monat 12	LS-MWD: -0,38 [-0,61; -0,15], 0,0016 SMD: -0,47 [-0,77; -0,16]	
5-D Pruritus Veränderung – Subdomäne Dauer		
Monat 6	LS-MWD: -0,18 [-0,39; 0,02], 0,0731 SMD: -0,26 [-0,56; 0,04]	
Monat 12	LS-MWD: -0,40 [-0,63; -0,16], 0,0011 SMD: -0,49 [-0,79; -0,19]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 12	LS-MWD: -0,71 [-1,22; -0,20], 0,0075 SMD: -0,70 [-1,21; -0,18]	
5-D Pruritus Veränderung – Subdomäne Beeinträchtigung		
Monat 6	LS-MWD: -0,47 [-0,73; -0,20], 0,0006 SMD: -0,50 [-0,81; -0,20]	
Monat 12	LS-MWD: -0,60 [-0,91; -0,29], 0,0002 SMD: -0,57 [-0,87; -0,26]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -0,97 [-1,55; -0,40], 0,0012 SMD: -0,85 [-1,38; -0,33]	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
5-D Pruritus Veränderung – Subdomäne Verlauf		
Monat 6	LS-MWD: -0,31 [-0,68; 0,06], 0,0981 SMD: -0,25 [-0,55; 0,06]	
Monat 12	LS-MWD: -0,47 [-0,90; -0,04], 0,0340 SMD: -0,32 [-0,62; -0,02]	
5-D Pruritus Veränderung – Subdomäne Lokalisation		
Monat 6	LS-MWD: -0,41 [-0,68; -0,14], 0,0034 SMD: -0,43 [-0,73; -0,13]	
Monat 12	LS-MWD: -0,38 [-0,67; -0,10], 0,0085 SMD: -0,39 [-0,69; -0,09]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 12	LS-MWD: -1,03 [-1,61; -0,44], 0,0009 SMD: -0,89 [-1,41; -0,36]	
5-D Pruritus Veränderung – Modifizierter Gesamt-Score		
Monat 6	LS-MWD: -1,41 [-2,14; -0,68], 0,0002 SMD: -0,55 [-0,86; -0,25]	
Monat 12	LS-MWD: -1,79 [-2,60; -0,97], <0,0001 SMD: -0,63 [-0,94; -0,33]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -2,89 [-4,43; -1,34], 0,0004 SMD: -0,94 [-1,47; -0,42]	
Monat 12	LS-MWD: -3,21 [-4,65; -1,77], <0,0001 SMD: -1,12 [-1,66; -0,59]	
PGI-S – Stetige Analyse		
Monat 6	LS-MWD: -0,42 [-0,61; -0,23], <0,0001 SMD: -0,64 [-0,95; -0,32]	
Monat 12	LS-MWD: -0,10 [-0,32; 0,11], 0,3453 SMD: -0,14 [-0,44; 0,17]	
PGI-C – Stetige Analyse		
Monat 6	LS-MWD: -0,45 [-0,88; -0,02], 0,0391 SMD: -0,31 [-0,61; -0,01]	
Monat 12	LS-MWD: -0,41 [-0,85; 0,03], 0,0673 SMD: -0,27 [-0,58; 0,03]	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Fibrose – Lebersteifigkeit (FibroScan®) Monat 6 Monat 12 Fibrose – Erweiterter Leberfibrosetest (ELF-Score) Monat 6 Monat 12	LS-MWD: -1,15 [-2,67; 0,36], 0,1354 SMD: -0,22 [-0,54; 0,09] LS-MWD: -1,37 [-3,09; 0,35], 0,1174 SMD: -0,24 [-0,55; 0,08] LS-MWD: 0,10 [-0,03; 0,24], 0,1264 SMD: 0,23 [-0,08; 0,53] LS-MWD: 0,08 [-0,11; 0,26], 0,4215 SMD: 0,12 [-0,18; 0,42]	
Hospitalisierungen aufgrund der Erkrankung Monat 12	RR: 1,06 [0,27; 4,17], 0,9384	
Gesundheitsbezogene Lebensqualität^a		
PBC-40: Veränderung – Gesamt-Score Monat 6 Monat 12 PBC-40: Veränderung – Kognitive Domäne Monat 6 Monat 12 Subgruppenanalyse Pruritus NRS ≥4 zu Monat 6 PBC-40: Veränderung – Emotionale Domäne Monat 6 Monat 12	LS-MWD: -0,92 [-5,30; 3,46], 0,6787 SMD: -0,06 [-0,36; 0,24] LS-MWD: 0,33 [-4,98; 5,64], 0,9019 SMD: 0,02 [-0,28; 0,32] LS-MWD: -0,43 [-1,38; 0,52], 0,3750 SMD: -0,13 [-0,44; 0,17] LS-MWD: -0,13 [-1,20; 0,94], 0,8106 SMD: -0,04 [-0,34; 0,27] LS-MWD: -1,98 [-3,59; -0,36], 0,0172 SMD: -0,64 [-1,16; -0,12] LS-MWD: 0,21 [-0,40; 0,81], 0,5013 SMD: 0,10 [-0,20; 0,40] LS-MWD: 0,25 [-0,40; 0,90], 0,4488 SMD: 0,11 [-0,19; 0,41]	Kein Zusatznutzen/Schaden ableitbar

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
PBC-40: Veränderung – Fatigue Domäne		
Monat 6	LS-MWD: -1,19 [-3,07; 0,69], 0,2131 SMD: -0,18 [-0,49; 0,12]	
Monat 12	LS-MWD: -0,47 [-2,66; 1,71], 0,6692 SMD: -0,06 [-0,37; 0,24]	
PBC-40: Veränderung – Pruritus Domäne		
Monat 6	LS-MWD: -0,93 [-1,68; -0,18], 0,0159 SMD: -0,35 [-0,66; -0,05]	
Monat 12	LS-MWD: -0,83 [-1,68; 0,02], 0,0544 SMD: -0,28 [-0,59; 0,02]	
PBC-40: Veränderung – Soziale Domäne		
Monat 6	LS-MWD: 0,95 [-0,58; 2,47], 0,2216 SMD: 0,18 [-0,12; 0,49]	
Monat 12	LS-MWD: 1,05 [-0,47; 2,56], 0,1735 SMD: 0,20 [-0,10; 0,50]	
PBC-40: Veränderung – Symptomatische Domäne		
Monat 6	LS-MWD: 0,17 [-0,64; 0,97], 0,6827 SMD: 0,06 [-0,24; 0,36]	
Monat 12	LS-MWD: 0,09 [-0,87; 1,06], 0,8475 SMD: 0,03 [-0,27; 0,33]	
Unerwünschte Ereignisse		
UE	RR: 1,02 [0,91; 1,16], 0,6976	Ausmaß: Geringer Zusatznutzen Wahrscheinlichkeit: Hinweis
SUE	RR: 1,14 [0,37; 3,57], 0,8186	
Schwere UE (CTCAE-Grad ≥3)	RR: 1,42 [0,54; 3,78], 0,4799	
Therapieabbrüche aufgrund von UE	RR: 0,68 [0,16; 2,94], 0,6023	
UE, die zum Tod führten	Es führte kein UE zum Tod.	
UE, SUE und schwere UE von speziellem Interesse	UE von speziellem Interesse Pruritus-assoziert: RR: 0,36 [0,14; 0,89], 0,0273 SUE von speziellem Interesse Es traten keine statistisch signifikanten Ergebnisse auf ^c Schwere UE von speziellem Interesse Es traten keine statistisch signifikanten Ergebnisse auf ^d	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
UE, SUE und schwere UE ohne erkrankungsbezogene Ereignisse	Jegliche UE ohne erkrankungsbezogene Ereignisse RR: 1,04 [0,92; 1,19], 0,5144 SUE ohne erkrankungsbezogene Ereignisse RR: 1,14 [0,37; 3,57], 0,8186 Schwere UE ohne erkrankungsbezogene Ereignisse RR: 1,78 [0,61; 5,18], 0,2923	
Jegliche UE nach SOC und PT ^b	SOC Erkrankungen der Haut und des Unterhautgewebes RR: 0,68 [0,37; 1,23], 0,2021 PT Pruritus RR: 0,30 [0,12; 0,80], 0,0160	
SUE nach SOC und PT ^b	Keine SUE nach SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen	
Schwere UE nach SOC und PT ^b	Keine schweren UE nach SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen	
a: Schwerwiegende bzw. schwere Endpunkte b: Aus Gründen der Übersichtlichkeit werden an dieser Stelle nur fazitrelevante Ergebnisse (signifikanter und relevanter Unterschied in Gesamtpopulation oder Subgruppe) dargestellt. c: Kardiovaskuläre Toxizität: 1 (0,8%) Patient im SEL-Arm; Ischämische Herzkrankheit: 1 (0,8%) Patient im SEL-Arm; Lebertoxizität: 1 (0,8%) Patient im SEL-Arm d: Kardiovaskuläre Toxizität: 1 (0,8%) Patient im SEL-Arm; Ischämische Herzkrankheit: 1 (0,8%) Patient im SEL-Arm; Lebertoxizität: 1 (0,8%) Patient im SEL-Arm ALP=Alkalische Phosphatase, CTCAE=Common Terminology Criteria for Adverse Events, ELF=Fortgeschrittene Leberfibrose, KI=Konfidenzintervall, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, NRS=Numerical Rating Scale, PBC=Primär biliäre Cholangitis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, PT=Preferred Term, RR=Relatives Risiko, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, SOC=System Organ Class, SUE=Schwerwiegendes unerwünschtes Ereignis, UE=Unerwünschtes Ereignis, ULN=Oberer Normwert		

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Die PBC ist eine chronisch-entzündliche, autoimmun vermittelte, cholestatische Lebererkrankung, die unbehandelt zu Fibrose und einer biliären Zirrhose im Endstadium fortschreitet. Das Vorhandensein einer Leberzirrhose stellt auch bei PBC-Patienten den entscheidenden Risikofaktor für die Entwicklung eines hepatozellulären Karzinoms (HCC) dar.

Die Lebensqualität von Patienten mit PBC wird sowohl durch die Erkrankung selbst als auch durch sozioökonomische Umstände beeinträchtigt. Neben den bekannten Symptomen Fatigue und Pruritus sind auch soziale (Folge-)Probleme, emotionale und kognitive (Folge-)Symptome, Depression, autonome Dysfunktionen und Tagesmüdigkeit in univariater Analyse mit schlechterer Lebensqualität assoziiert. Verglichen mit anderen chronischen Erkrankungen wie

Diabetes mellitus empfanden Patienten mit PBC ihre Erkrankung als größere Belastung und wiesen somit eine reduzierte Lebensqualität auf.

Da es sich bei der PBC um eine seltene Erkrankung handelt, stellt die geringe Anzahl Betroffener und somit potenzieller Studienteilnehmer naturgemäß eine besondere Herausforderung bei der Durchführung aussagekräftiger Studien dar. Mit der Studie RESPONSE konnte trotzdem eine RCT durchgeführt und somit eine Studie von hoher Evidenzgüte und Aussagekraft für die Nutzenbewertung vorgelegt werden. Die Vorteile der Behandlung mit SEL spiegeln sich in deutlichen Effekten mit statistischer Signifikanz und teilweise klinischer Relevanz wider, die trotz des vergleichsweise kleinen Patientenkollektivs gezeigt werden konnten. Die Ergebnisse der Studie RESPONSE belegen die Wirksamkeit von SEL im Hinblick auf das biochemische Ansprechen, die Symptomatik, Lebensqualität sowie das Eintreten unerwünschter Ereignisse. Die beobachtete Verringerung des Leberwertes alkalische Phosphatase (ALP) sowie der respektiven Normalisierung ebendieser unter SEL zeigt, dass ein Fortschreiten der Lebererkrankung verlangsamt und eventuell sogar verhindert und somit der Zustand und die Funktion dieses lebensnotwendigen Organs erhalten werden kann. Damit verbessert sich auch die Prognose für die Patienten im weiteren Krankheitsverlauf, welcher potenziell zu Leberfibrose, Zirrhose und letztendlich zu einer Lebererkrankung im Endstadium bis hin zum Tod führt.

Patienten mit PBC nehmen die einschränkenden Auswirkungen des Pruritus auf ihr tägliches Leben unmittelbar wahr, was mit einer erheblichen Beeinträchtigung der Lebensqualität einhergeht. Die Intensität des Pruritus ist mit Beeinträchtigungen in vielen Lebensbereichen, wie reduzierter Schlafqualität und Arbeitsproduktivität, sowie Depressionen und Suizidalität verbunden. Pruritus wird aufgrund der direkten Spürbarkeit durch den Patienten verbunden mit starken Einschränkungen des alltäglichen Lebens des Patienten als patientenrelevanter Endpunkt betrachtet. Hinzu kommt, dass Pruritus meist nicht zufriedenstellend therapiert werden kann. Auftreten und Intensität von Pruritus korrelieren nicht mit dem Schweregrad der Lebererkrankung, können patientenindividuell sehr unterschiedlich sein und treten unabhängig vom Ausmaß der Cholestase – oft schubweise – auf. Unter SEL konnte in der Studie RESPONSE eine deutliche Verbesserung des sehr belastenden Pruritus gezeigt werden. Sowohl für die Pruritus Numerical Rating Scale (NRS) als auch für den 5-D Pruritus zeigten sich ausschließlich statistisch signifikante Vorteile für die Behandlung mit SEL. Aufgrund der deutlichen Verringerung der Pruritus-Inzidenz über den gesamten Behandlungszeitraum wird gerade auch im Vergleich zu Placebo davon ausgegangen, dass SEL Pruritus nicht nur lindert, sondern auch einen antipruritogenen Effekt aufweist, das heißt auch dessen de novo-Entstehung hemmt. Das bedeutet, dass schlussendlich alle Patienten, also Patienten mit, aber auch ohne Pruritus von einer Therapie mit SEL im Hinblick auf bestehenden sowie neu auftretenden Pruritus beträchtlich profitieren.

Zu berücksichtigen ist zudem, dass sich bei Patienten mit einer bereits erhöhten Belastung durch Pruritus zu Baseline (Subgruppe Pruritus NRS ≥ 4 zu Baseline, das entspricht Patienten mit moderat bis schwerwiegendem Pruritus) eine noch deutlichere und nachhaltigere Reduktion unter der Behandlung mit SEL zeigte (siehe auch Abbildung 30). Neben dem gezeigten Vorteil

in der Gesamtpopulation wird deutlich, dass speziell Patienten mit einer höheren Pruritus-Belastung nochmal stärker, und damit erheblich, von der Behandlung mit SEL profitieren.

Des Weiteren ist bemerkenswert, dass unter der Behandlung mit SEL statistisch signifikante Verbesserungen in den unerwünschten Ereignissen (UE) zu beobachten waren, was SEL zu einem äußerst verträglichen Arzneimittel macht. Es ist zu betonen, dass dieser Vorteil gegenüber Placebo, also keiner Therapie, gezeigt werden konnte.

Diese positiven Effekte zeigen, dass SEL einen bislang ungedeckten therapeutischen Bedarf für Patienten mit einer PBC im vorliegenden Anwendungsgebiet deckt und dabei den Gesamtzustand der Patienten zudem noch deutlich verbessert. SEL trägt maßgeblich zur Reduktion der klinischen Manifestationen der PBC bei und führt so zu einem guten Gesundheitszustand und potenziell zu einer erhöhten Lebenserwartung. Vor dem Hintergrund einer schweren, potenziell lebensverkürzenden chronischen Erkrankung mit einer progredienten Akkumulation der Symptomlast im weiteren Verlauf ist für SEL ein **beträchtlicher Zusatznutzen in der Gesamtpopulation** festzustellen. Da sich in der Population der PBC-Patienten mit moderat-schwerem Pruritus zu Baseline (Pruritus NRS ≥ 4 Punkte) der Vorteil von SEL besonders deutlich zeigt, wird für diese Teilpopulation ein **erheblicher Zusatznutzen** festgestellt.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

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

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

Ziel der vorliegenden Untersuchung ist die Bewertung des medizinischen Nutzens und medizinischen Zusatznutzens von SEL. Gemäß Fachinformation findet SEL Anwendung in der Behandlung der PBC in Kombination mit UDCA bei Erwachsenen, die nicht ausreichend auf UDCA alleine ansprechen, oder als Monotherapie bei Erwachsenen, die UDCA nicht vertragen [1].

Bei SEL handelt es sich um ein Arzneimittel zur Behandlung einer seltenen Erkrankung (Orphan Drug) nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden [2]. Der Orphan Drug-Status wurde am 5. Dezember 2024 durch die EMA bestätigt. Gemäß § 35a Abs. 1 Satz 11 SGB V gilt der medizinische Zusatznutzen für Orphan Drugs durch die Zulassung als belegt [3]. Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im Verhältnis zur zVT müssen nach § 35a Abs. 1 Satz 12 SGB V nur nach Aufforderung durch den G-BA vorgelegt werden, wenn der Umsatz des Arzneimittels in den letzten zwölf Kalendermonaten einen Betrag von 30 Millionen Euro übersteigt [3]. Es ist das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung

begründenden Studien für die Anzahl der Patienten und Patientengruppen nachzuweisen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht.

Patientenpopulation

Die gemäß Indikation zugelassene Patientenpopulation umfasst Erwachsene mit PBC, die nicht ausreichend auf UDCA alleine ansprechen oder die UDCA nicht vertragen.

Intervention

Die zu bewertende Intervention soll einmal täglich in einer Dosierung von 10 mg oral (Hartkapsel) als Monotherapie oder in Kombination mit UDCA verabreicht werden.

Vergleichstherapie

Da es sich bei SEL um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, ist keine zVT zu bestimmen. Der medizinische Nutzen und Zusatznutzen von SEL gelten bereits durch die Zulassung als belegt und das Ausmaß des Zusatznutzens ist auf Grundlage der Zulassung und der die Zulassung begründenden Studien zu bestimmen.

Endpunkte

Der medizinische Zusatznutzen für Orphan Drugs gilt durch die Zulassung als belegt. Zur Bewertung des darüber hinausgehenden medizinischen Nutzens und medizinischen Zusatznutzens von SEL werden Endpunkte herangezogen, die anerkannte patientenrelevante Therapieziele im Anwendungsgebiet abbilden. Diese Endpunkte umfassen:

- Mortalität
- Morbidität
 - Biochemisches Ansprechen
 - Symptomatik
- Gesundheitsbezogene Lebensqualität anhand patientenrelevanter Fragebögen
- UE

Eine umfassende Auflistung der im Rahmen des Dossiers eingeschlossenen patientenrelevanten Endpunkte zusammen mit der Begründung ihrer Patientenrelevanz, respektive der Begründung als valide Surrogate findet sich in Abschnitt 4.2.5.2 und 4.5.4.

Studientypen

Da RCT mit der geringsten Ergebnisunsicherheit behaftet sind, liefern sie, sofern methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt, die für die Nutzenbewertung einer medizinischen Intervention zuverlässigsten Ergebnisse. Entsprechend werden der Zusatznutzenbewertung RCT zugrunde gelegt.

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.

Gemäß den im vorherigen Abschnitt 4.2.1 definierten Komponenten der Fragestellung zur Bewertung des medizinischen Zusatznutzens von SEL wird nach Studien gesucht, die die folgenden Kriterien erfüllen (siehe Tabelle 4-2).

Tabelle 4-2: Übersicht der Kriterien für den Studieneinschluss für den Vergleich von SEL bei erwachsenen Patienten

	Einschlusskriterien	E	Ausschlusskriterien	A
Population	Erwachsene Patienten mit PBC, die nicht ausreichend auf UDCA alleine ansprechen oder die UDCA nicht vertragen	1	• Personen ohne PBC • Patienten mit PBC, die auf UDCA alleine ansprechen und die UDCA vertragen • Kinder und Jugendliche <18 Jahre • Dekompensierte Lebererkrankung	1
Intervention	SEL gemäß Fachinformation	2	Andere Intervention oder abweichende Dosierungen	2
Vergleichstherapie	-	3	-	3
Endpunkte	Mindestens einer der folgenden Endpunkte wird berichtet: • Mortalität • Morbidität • Gesundheitsbezogene Lebensqualität • UE	4	Keiner der genannten patientenrelevanten Endpunkte wird berichtet	4
Studientyp	RCT	5	Andere Studientypen	5
Studiendauer	≥48 Wochen	6	<48 Wochen	6

	Einschlusskriterien	E	Ausschlusskriterien	A
Publikationstyp	Vollpublikation, Ergebnisse aus Studienregistern oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht	7	• Review-Artikel, Meta-Analysen • Case Reports • Keine Vollpublikation (z. B. Notes, News, Short Surveys, Conference Abstracts) • Studienregistereintrag, in dem keine Ergebnisse verfügbar sind • Jeglicher Publikationstyp, in dem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden	7
CONSORT=Consolidated Standards of Reporting Trials, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, UDCA=Ursodeoxycholsäure, UE=Unerwünschtes Ereignis				

Begründung der Wahl der Einschlusskriterien für Studien

Population

Die Patientenpopulation ergibt sich aus dem Wortlaut der Zulassung für das zu bewertende Arzneimittel (zbAM) SEL. Die für dieses Dossier relevante Population sind erwachsene Patienten mit PBC, die nicht ausreichend auf UDCA alleine ansprechen oder die UDCA nicht vertragen.

Intervention

Das Kriterium der Intervention ergibt sich aus dem zbAM SEL, das laut Fachinformation einmal täglich in einer Dosierung von 10 mg oral (Hartkapsel) als Monotherapie oder in Kombination mit UDCA verabreicht werden soll.

Vergleichstherapie

Da es sich bei SEL um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, ist keine zVT zu bestimmen. Der medizinische Nutzen und medizinische Zusatznutzen von SEL gelten bereits durch die Zulassung als belegt und das Ausmaß des Zusatznutzens ist auf Grundlage der Zulassung und der die Zulassung begründenden Studien zu bestimmen.

Endpunkte

Das Kriterium der Endpunkte ergibt sich aus den in der Indikation PBC etablierten und validierten patientenrelevanten Endpunkten, wie in Abschnitt 4.2.1 definiert.

Studientyp

RCT sind, sofern der jeweiligen Fragestellung entsprechend und methodisch adäquat durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet. Es wird somit ausschließlich nach RCT gesucht.

Studiendauer

Das Kriterium der Mindeststudiendauer ergibt sich aus der Art der Erkrankung, der Art der Behandlung, dem Wirkmechanismus von SEL, der Leitlinienempfehlung sowie dem bisherigen Vorgehen des G-BA bei vergleichbaren Erkrankungen.

Bei der PBC handelt es sich um eine autoimmun-vermittelte, chronisch entzündliche Lebererkrankung [4], für die keine ursächliche Therapie existiert, deren Fortschreiten aber medikamentös verlangsamt und teilweise verhindert werden kann. Ziel der medikamentösen Therapie ist es, die Symptome zu lindern und vor allem das Fortschreiten der Krankheit zu verlangsamen. Entsprechend handelt es sich bei der Therapie mit SEL um die kontinuierliche Therapie einer chronischen Erkrankung, die zum aktuellen Zeitpunkt noch nicht heilbar ist.

Vergleichbare Indikationen hinsichtlich der Art der Erkrankung (chronische, nicht-heilbare Erkrankung) sowie der Art der Therapie (kontinuierliche, lebenslange Therapie) sind die Infektionen mit dem Hepatitis B-Virus sowie dem humanen Immundefizienzvirus (HIV) – in beiden Indikationen hat der G-BA in den bisherigen Verfahren eine Mindeststudiendauer von 48 Wochen als sachgerecht bewertet und Studien mit kürzerer Behandlungsdauer als nicht nutzenbewertungsrelevant erachtet, so beispielsweise im Verfahren zu Emtricitabin (FTC)/Rilpivirin/Tenofoviridisoproxil (Handelsname Eviplera[®], antiretroviral vorbehandelte Patienten): „Weiterhin ist die in der Studie realisierte Mindeststudiendauer der randomisierten kontrollierten Studienphase von 24 Wochen grundsätzlich als zu kurz für die Bewertung des Zusatznutzens einzustufen, da die HIV-Infektion bisher nicht heilbar ist und so eine lebenslange Therapie erfordert. Um Endpunkte in dieser Indikation valide beurteilen zu können, ist ein längerer Studienverlauf erforderlich.“ [5]. In weiteren Verfahren wurde bei Vorliegen mehrerer Datenschnitte ausschließlich der spätere Datenschnitt berücksichtigt, so im Verfahren zu Bictegravir/FTC/Tenofovirafenamid (Handelsname Biktarvy[®]): „Bei der Dossiereinreichung lagen für beide Studien Daten zu Woche 48 vor, im Verlauf des Stellungnahmeverfahrens reichte der pharmazeutische Unternehmer einen weiteren Datenschnitt zu Woche 96 für beide Studien ein. Für die Nutzenbewertung ist der Datenschnitt zu Woche 96 maßgeblich.“ [6]. Die EMA geht im Rahmen der Entwicklung von Substanzen zur Behandlung der PBC unter Berücksichtigung von Wirksamkeit und Sicherheit, aber auch abhängig vom Ausmaß der zu erwartenden Wirkung, von einer Studiendauer von etwa zwei Jahren aus [7]. Auch bei Typ-2-Diabetes mellitus wurde vom G-BA eine Ausweitung der Studiendauer auf 12 Monate – entspricht 48 Wochen – als wünschenswert angesehen, um Wirksamkeit und Sicherheit über einen längeren Zeitraum zu überprüfen [8], ebenso bei anderen chronischen Erkrankungen, die eine Langzeitbehandlung erfordern wie in der Indikation Hypercholesterinämie [9, 10]. Darauf basierend und aufgrund der Dauer der Zulassungsstudie zu SEL über 48 Wochen wird von einer Mindeststudiendauer von ≥ 48 Wochen im vorliegenden Anwendungsgebiet ausgegangen.

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

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

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

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

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

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

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von

den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die bibliografische Literaturrecherche wird in den Datenbanken MEDLINE, EMBASE und Cochrane (Cochrane Central Register of Controlled Trials) durchgeführt, wobei für jede einzelne Datenbank eine sequentielle, individuell adaptierte Suchstrategie entwickelt und separat angewandt wird. Bei allen Recherchen werden nur Studien am Menschen berücksichtigt. Studien an gesunden Teilnehmern werden nicht berücksichtigt. Mithilfe von validierten Filtern wird eine Einschränkung auf RCT vorgenommen. Für die Suche wird die Plattform OVID verwendet.

Die Ergebnisse der systematischen bibliografischen Literaturrecherche nach RCT mit dem zbAM SEL sind in Abschnitt 4.3.1.1.2 dargestellt.

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

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

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

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

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

Die Suche soll mindestens in den Studienregistern/ Studienergebnisdatenbanken clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (<https://clinicaldata.ema.europa.eu>) sowie dem Arzneimittel-Informationssystem (AMIS, <https://www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B.

krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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

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

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

Die Studienregistersuche wird in den Portalen [clinicaltrials.gov](http://www.clinicaltrials.gov) (www.clinicaltrials.gov), EU-CTR (www.clinicaltrialsregister.eu), sowie ICTRP (Suchportal der Weltgesundheitsorganisation [WHO]: <https://trialsearch.who.int/>) durchgeführt. Es wird nach abgeschlossenen, laufenden und abgebrochenen Studien gesucht.

Die Ergebnisse der Studienregistersuche nach RCT sind in Abschnitt 4.3.1.1.3 beschrieben.

Die Details zur Suchstrategie und das Datum der Suche sind in Anhang 4-B beschrieben. Die ausgeschlossenen Studien sind im Anhang 4-D aufgelistet.

Zusätzlich wird in dem Clinical Data Suchportal der EMA (<https://clinicaldata.ema.europa.eu>) sowie dem AMIS (<https://www.bfarm.de/DE/Arzneimittel/Arzneimittelinformationen/>) nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (Liste der Studien des pharmazeutischen Unternehmers, bibliografische Literaturrecherche oder Studienregistersuche) identifiziert wurden, gesucht. Für diese Suchen erfolgt keine Dokumentation der Suchstrategie.

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.

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

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

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Die Suche auf der Internetseite des G-BA erfolgt nach Einträgen zu Studien mit dem zbAM SEL, die bereits anderweitig identifiziert wurden (Liste der Studien des pharmazeutischen Unternehmers, bibliografische Literaturrecherche oder Studienregistersuche). Für diese Suche erfolgt keine Dokumentation der Suchstrategie.

Die Ergebnisse der Suche nach RCT für das zbAM auf der Internetseite des G-BA sind in Abschnitt 4.3.1.1.4 dargestellt.

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

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.

Das Screening der Ergebnisse aus der in Abschnitt 4.2.3.2 beschriebenen bibliografischen Literaturrecherche, der in Abschnitt 4.2.3.3 beschriebenen Studienregistersuche und der in Abschnitt 4.2.3.4 beschriebenen Suche auf der Internetseite des G-BA wird von zwei Personen unabhängig voneinander durchgeführt. Dabei werden die Studien anhand der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien auf ihre Relevanz geprüft. Im Falle von Diskrepanzen hinsichtlich der Bewertung der Relevanz findet eine Diskussion zur Klärung des Ein- oder Ausschlusses statt.

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.

Für die Fragestellung relevante Studien werden hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung des Verzerrungspotenzials erfolgt entsprechend den Vorgaben der Dossiervorlage. Basis der Bewertung sind Informationen aus dem klinischen Studienbericht (falls vorhanden), Registereinträgen oder aus Volltextpublikationen der jeweiligen klinischen Studien.

Bewertung des Verzerrungspotenzials

Sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte werden untersucht. Entsprechend der Verfahrensordnung (VerfO) des G-BA zur Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V werden dabei folgende Aspekte bewertet [11]:

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie der behandelnden Personen
- Ergebnisunabhängige Berichterstattung
- Sonstige Aspekte (z. B. intransparenter Patientenfluss)

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber

- Umsetzung des Intention-to-Treat (ITT)-Prinzips
- Ergebnisunabhängige Berichterstattung
- Sonstige Aspekte (z. B. Behandlung fehlender Werte und Drop-outs)

Die Bewertungen des Verzerrungspotenzials der Einzelstudien werden tabellarisch zusammengefasst. Gemäß Verfo wird das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. 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.

Abschließend wird unter Berücksichtigung der oben genannten Aspekte das Verzerrungspotenzial auf Ebene der Endpunkte als „hoch“ oder „niedrig“ eingestuft. Von einem „niedrigen“ Verzerrungspotenzial wird ausgegangen, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führt nicht zum Ausschluss aus der Bewertung. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

Die Vorgehensweise bei der Bewertung des Verzerrungspotenzials für jede Studie und jeden Endpunkt richtet sich nach den Hinweisen der Bewertungsbögen in Anhang 4-F.

Eine Bewertung des Verzerrungspotenzials gemäß Anhang 4-F wird nur für vergleichende Studien vorgenommen. Bei nicht-vergleichenden Studien ist aufgrund der fehlenden Referenzbehandlung grundsätzlich von einem erhöhten Verzerrungspotenzial auszugehen.

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-

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.

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.

In die Bewertung einfließende Studien werden anhand von Design-Charakteristika (Studiendesign, Studiendauer, Zahl der randomisierten Patienten, Ort und Zeitraum der Studiendurchführung, Interventionen, primäre Zielkriterien und relevante sekundäre Zielkriterien) beschrieben. Die Beschreibung der Design-Charakteristika deckt die im CONSORT-Statement angegebenen Items 2b bis 14 ab. Darüber hinaus wird der Patientenfluss mithilfe der Informationen aus dem CONSORT-Flow-Chart dargestellt. Nicht-randomisierte Studien werden anhand des Transparent Reporting of Evaluations with Non-Randomized Design (TREND)-Statements beschrieben.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

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

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind

⁴ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche 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.

die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

Patientencharakteristika

Für die in die Nutzenbewertung eingeschlossenen Studien werden die untersuchten Studienpopulationen charakterisiert. Hierfür werden neben demografischen und prognostischen Charakteristika auch die den Krankheitsstatus beschreibenden Faktoren der Studienpopulationen betrachtet.

Relevante demografische bzw. prognostische Charakteristika zu Baseline umfassen:

- Alter
- Geschlecht
- Ethnische Zugehörigkeit
- Region
- Body Mass Index (BMI)
- Zirrhose-Status
- Antimitochondriale Antikörper (AMA)-Status
- Model for End-Stage Liver Disease (MELD)-Score
- Fibrose-Score abgeleitet aus der Lebersteifigkeit
- GLOBE Risk Score
- Rotterdam Stage Disease
- ALP
- Gesamt-Bilirubin
- Pruritus NRS
- Fatigue
- High Density Lipoprotein-Cholesterol (HDL-C)
- Low Density Lipoprotein-Cholesterol (LDL-C)
- Alanin-Aminotransferase (ALT)
- Aspartat-Aminotransferase (AST)
- Gamma-Glutamyltransferase (GGT)
- Vorherige Einnahme von Obeticholsäure (OCA) und/oder Fibraten
- UDCA-Intoleranz
- Alter bei PBC-Diagnose

- United Kingdom (UK)-PBC Risk Score

Darüber hinaus wird die Disposition der Patienten mit Angabe der Gründe für einen vorzeitigen Abbruch der Studie und Studienmedikation beschrieben. Weiterhin wird die Dauer der Behandlung sowie die Beobachtungsdauer der unten aufgeführten Endpunkte betrachtet. Der Patientenfluss der Studien wird in einem Flow-Chart dargestellt.

Patientenrelevante Endpunkte

Die Ergebnisse der Einzelstudien werden mittels patientenrelevanter Endpunkte dargestellt und gegenübergestellt. Die folgenden Endpunkte werden als patientenrelevant und entsprechend als nutzenbewertungsrelevant erachtet:

- Mortalität
- Morbidität
 - Biochemisches Ansprechen
 - Symptomatik
- Gesundheitsbezogene Lebensqualität anhand patientenrelevanter Fragebögen
- UE

Daneben werden ergänzend die nicht-schweren UE und erkrankungsbezogene UE sowie Therapieabbrüche aufgrund von UE getrennt nach System Organ Class (SOC) und Preferred Term (PT) berichtet, bei der Ableitung des Zusatznutzens jedoch nicht herangezogen.

Statistische Auswertungen

Als Effektschätzer für die Einzelstudien werden für binäre Variablen das Relative Risiko (RR), das Odds Ratio (OR) sowie die Risikodifferenz (RD) berechnet. Zur Interpretation der Ergebnisse wird primär das RR herangezogen. Inkonsistenzen zwischen den Effektschätzern (RR und OR) werden diskutiert.

Für kontinuierliche Variablen wird die Kleinste-Quadrate-Mittelwertdifferenz (LS-MWD) berechnet.

Für Responderanalysen wird entsprechend der Methodik des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) eine Verbesserung um mindestens 15% der Skalenspannweite zugrunde gelegt [12]. Bei einem statistisch signifikanten Unterschied zwischen den Behandlungen wird – sofern keine Responderanalysen durchgeführt werden können – zur Abschätzung der klinischen Relevanz des Unterschieds zusätzlich die standardisierte Mittelwertdifferenz (SMD) mittels Hedges' g berechnet [12]. Dazu wird die Lage des Konfidenzintervalls (KI) in Bezug auf die Irrelevanzschwelle von 0,2 betrachtet [12].

Effektschätzer, KI, p-Werte und Heterogenitätsmaße werden, wenn nicht gesondert gekennzeichnet, mittels der im Cochrane Handbuch (<https://training.cochrane.org/handbook>) beschriebenen Methodik mit der Software SAS oder R unter Verwendung des Pakets meta zusätzlich berechnet.

In Fällen, in denen auf Basis der verfügbaren Evidenz die Berechnung der zuvor beschriebenen Effektschätzer, KI, p-Werte und Heterogenitätsmaße nicht möglich ist, werden, sofern möglich, abweichende Maße zur Quantifizierung des Behandlungsunterschieds berichtet und dies entsprechend gekennzeichnet.

Patientenrelevanz und Operationalisierung der Endpunkte

Im Folgenden wird die Patientenrelevanz und Validität der in die Nutzenbewertung eingeschlossenen Endpunkte begründet und deren Operationalisierung beschrieben.

Mortalität

Die Mortalität wird im Rahmen der UE erhoben. Es wird der Anteil der Patienten mit einem UE, das zum Tod führt, betrachtet. Die Operationalisierung als Tod jeglicher Ursache wird gewählt, um einen Selektionsbias aufgrund des Ausschlusses bestimmter Todesfälle auf Basis von vermuteter Kausalität zu vermeiden.

Die Mortalität zählt aufgrund ihrer eindeutigen Erhebung (nur Tod zählt als Ereignis) zu den harten, patientenrelevanten Endpunkten. Dies bestätigt auch der G-BA: Nach § 3 Abs. 1 des 5. Kapitels der Verfo des G-BA ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“ [11]. Entsprechend ist die Gesamtmortalität als patientenrelevanter Endpunkt im Sinne der Verfo anzusehen.

Morbidität

Biochemisches Ansprechen

Im vorliegenden Anwendungsgebiet basieren derzeit alle Ansprechkriterien auf eine Therapie (vorrangig mit UDCA) auf Grenzwerten für den Grad der Erhöhung der ALP und des Bilirubins. Diese Parameter sind auf Populationsebene prädiktiv für das Risiko einer Progression der Erkrankung bis zur Transplantation oder zum Tod [13].

Da sich Leberkomplikationen wie die Notwendigkeit einer Lebertransplantation und leberbedingte Mortalität erst nach längerer Zeit entwickeln, würde der direkte Nachweis der Auswirkungen der Therapie auf klinische Endpunkte eine längere Nachbeobachtungszeit erfordern. Das Fortschreiten der Erkrankung wird durch die Therapie so stark verlangsamt, dass eine unverhältnismäßig lange Beobachtungszeit in klinischen Studien nötig wäre, um klinische Endpunkte wie Morbidität und Mortalität adäquat zu erfassen, und die Ereignishäufigkeiten bei normaler Studiendauer von 48 und 96 Wochen sind nicht ausreichend. Klinische Endpunkt-Studien sind vor diesem Hintergrund kaum mehr möglich: Zum einen ist es ethisch nicht vertretbar, einen Patienten bis zu einem klinischen Ereignis auf einem Regime zu belassen, wenn dies gegebenenfalls durch die Initiierung einer Therapie verhindert werden könnte, zum anderen stellen sowohl die erforderliche Studiendauer als auch die notwendige Patientenzahl praktische Hindernisse dar. In der Indikation HIV ist eine klare Korrelation zwischen den Surrogatparametern Viruslast und Cluster of Differentiation (CD)4-Zellzahl und den entsprechenden Morbiditäts- und Mortalitäts-Endpunkten in klinischen Studien vielfach

validiert worden. Entsprechend werden Zulassungsstudien im HIV-Bereich seit 1997 im Grunde ausnahmslos unter Verwendung dieser validierten Surrogate durchgeführt. Auch im Anwendungsgebiet der PBC ist es ethisch nicht vertretbar, einen Patienten bis zu einem klinischen Ereignis zu beobachten, obwohl im Verumarm der Studie eine wirksame Therapie verfügbar wäre, um dieses zu verhindern, dies insbesondere vor dem Hintergrund, dass im Vergleichsarm keine wirksame Therapie zur Verfügung steht.

Im Rahmen des biochemischen Ansprechens werden die ALP-Normalisierung (definiert als $ALP \leq 1 \times$ oberer Normwert [ULN]), das kombinierte biochemische Ansprechen (definiert als $ALP < 1,67 \times ULN$, Gesamt-Bilirubin $\leq ULN$ und ALP-Reduktion von $\geq 15\%$) sowie die stetigen Analysen zu ALP und Gesamt-Bilirubin als Surrogate für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitschwere, respektive der Endpunkte Lebertransplantation oder Tod herangezogen. Ergänzend werden die Veränderung der Parameter ALT, AST, direktes und indirektes Bilirubin sowie GGT und 5'-Nucleotidase dargestellt.

ALP-Normalisierung

ALP ist für viele Stoffwechselprozesse im Körper essentiell und findet sich unter anderem in Knochenzellen und im Lebergewebe. Bei Verdacht auf Lebererkrankungen wird die Konzentration der ALP im Blut bestimmt. Die PBC ist eine nicht-eitrige Entzündung der kleinen intrahepatischen Gallengänge, wodurch es zu einem Stau von Gallenflüssigkeit in der Leber (Cholestase) und damit einer vermehrten Freisetzung von ALP aus den betroffenen Gallengangs-Epithelzellen kommt [14]. Ein erhöhter ALP-Wert wird typischerweise im gesamten Verlauf der PBC-Erkrankung beobachtet [14, 15].

Studiendaten belegen, dass unabhängig von Risikofaktoren wie Alter und Krankheitsstadium die transplantationsfreie Überlebenszeit steigt, je niedriger der ALP-Wert ist [15]. Dabei ist die Normalisierung der ALP, definiert als $ALP \leq 1 \times ULN$, mit dem niedrigsten Risiko für eine Lebertransplantation oder Tod verbunden: In einer Patientenkohorte der GLOBAL PBC Study Group mit UDCA-behandelten und -unbehandelten Patienten zeigten sich bei Patienten, die – unabhängig von der UDCA-Therapie – nach einem Jahr einen normalen ALP-Wert ($\leq 1 \times ULN$) aufwiesen, 10-Jahres-Überlebensraten von 93,2%. Patienten mit einem ALP-Wert zwischen $1,0$ - $1,67 \times ULN$ nach einem Jahr hatten im weiteren Verlauf 10-Jahres-Überlebensraten von 86,1% [16]. Weitere Daten zeigen ebenfalls, dass bei Patienten mit PBC das Erreichen einer ALP-Normalisierung nach einem Jahr (unter UDCA-Therapie) mit einem deutlich längeren komplikationsfreien Überleben assoziiert ist [17]. Eine weitere Studie belegt, dass jede anhaltende Erhöhung der ALP bei PBC unter einer UDCA-Therapie mit einem gewissen Grad an persistierender Krankheitsaktivität, gemessen auf Basis des Entzündungs- und Stoffwechselproteoms im Serum, verbunden ist und belegt damit die Bedeutung biochemischer Marker wie der ALP bei PBC, bei denen eine lineare Beziehung zum Krankheitsverlauf besteht. Diese anhaltende Krankheitsaktivität birgt das Potenzial für ein stilles Fortschreiten der Krankheit und stellt damit eine mechanistische Erklärung für die Ergebnisse der GLOBAL PBC Study Group dar [13]. Daher wird der ALP-Wert unter PBC-Behandlung als entscheidender Prädiktor für das Langzeitüberleben und das Risiko einer Krankheitsprogression angesehen.

Entsprechend sollte das Behandlungsziel die Normalisierung der ALP, respektive der Leberwerte sein, um eine vollständige Kontrolle der Krankheitsaktivität bei der PBC zu erreichen [13]. Es ist mittlerweile allgemein anerkannt, dass jede Form von Abnormalität bei den Leberwerten mit einer anhaltenden Leberentzündung, einem erhöhten Risiko für Fibrose und schlechteren klinischen Ergebnissen in Verbindung steht [13].

Zusammengefasst ist der nachweislich ausschlaggebende Faktor für die Krankheitsprogression bei der PBC die entzündliche Krankheitsaktivität in der Leber, die sich in den erhöhten ALP-Werten widerspiegelt. Folglich kann die ALP-Normalisierung als geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC herangezogen werden [15]. Die ALP-Normalisierung wird in der vorliegenden Nutzenbewertung als nicht-schwerer Endpunkt eingestuft.

Kombiniertes biochemisches Ansprechen

Das kombinierte biochemische Ansprechen ist definiert als ALP $<1,67 \times \text{ULN}$, Gesamt-Bilirubin $\leq \text{ULN}$ und ALP-Reduktion von $\geq 15\%$. Eine Überschreitung dieser Grenzwerte für ALP und Bilirubin wird aktuell als klinische Indikation für eine Zweitlinientherapie herangezogen [18].

Wie bereits beschrieben, handelt es sich bei der ALP-Normalisierung aufgrund der Korrelation mit dem Grad der Leberschädigung um ein geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC. Ergänzend zur Bedeutung einer ALP-Normalisierung ist auch eine Senkung des ALP-Spiegels mit einem verbesserten transplantationsfreien Überleben oder Tod assoziiert: Studiendaten belegen, dass unabhängig von Risikofaktoren wie Alter und Krankheitsstadium die transplantationsfreie Überlebenszeit steigt, je niedriger der ALP-Wert ist. In einer Kohorte mit PBC-Patienten zeigte sich die Assoziation zwischen ALP-Werten und dem Risiko einer Lebertransplantation oder Tod sowohl für Baseline-ALP-Werte, als auch im Rahmen der Nachbeobachtungszeit von bis zu 5 Jahren und unabhängig von der UDCA-Therapie. Die transplantationsfreien Überlebensraten waren nach 5, 10 und 15 Jahren bei Patienten mit ALP $\leq 2,0 \times \text{ULN}$ deutlich höher als bei Patienten mit höheren ALP-Werten ($> 2,0 \times \text{ULN}$) [15]. Folglich kann der ALP-Wert unter PBC-Behandlung (beispielsweise mit UDCA) als entscheidender Prädiktor für das Langzeitüberleben und das Risiko einer Krankheitsprogression angesehen werden [15]. Im Rahmen klinischer Studien wird bezüglich des kombinierten biochemischen Ansprechens ein Cut-off von $1,67 \times \text{ULN}$ als geeignet angesehen [15]. Eine Senkung des ALP-Wertes $< 1,67 \times \text{ULN}$ und damit unter den genannten Schwellenwert von $2,0 \times \text{ULN}$ stellt daher analog zur ALP-Normalisierung ein valides Surrogat dar.

Neben der Senkung des ALP-Wertes wird eine Veränderung des ALP-Wertes von über 10% im Krankheitsverlauf und damit oberhalb möglicher natürlicher Schwankungen als relevant im Rahmen klinischer Studien erachtet [19]. Die hier betrachtete ALP-Reduktion um $\geq 15\%$ kann

daher als adäquates Maß zur Erfassung einer klinisch relevanten Verbesserung angesehen werden.

Neben der ALP gilt auch das Gesamt-Bilirubin als etablierter prognostischer Faktor der PBC [14, 16, 19]. Bei der PBC ist besonders das Gesamt-Bilirubin von klinischer Relevanz. Es setzt sich zusammen aus dem direkten (wasserlösliches) und dem indirekten (unkonjugiertes, wasserunlösliches) Bilirubin. Ein Anstieg des Gesamt-Bilirubins erfolgt bei der PBC primär dadurch, dass es zu einer gestörten Sekretion des direkten Bilirubins über die Gallengänge kommt [20]. Primär dient der Gesamt-Bilirubinwert als wichtiges Surrogat für das Progressionsrisiko, respektive der klinischen Endpunkte Lebertransplantation oder Tod [14, 15, 19] und wird sowohl bei Erstdiagnose als auch im Krankheitsverlauf zur Ermittlung des Krankheitsstadiums der PBC herangezogen. Die mit fortschreitender PBC auftretende erhöhte Bilirubin-Konzentration ist mit einer sinkenden Überlebenswahrscheinlichkeit assoziiert [14, 15, 19]. Das Risiko für eine Lebertransplantation oder Tod steigt bei Gesamt-Bilirubinwerten jenseits des Schwellenwertes von $0,6 \times \text{ULN}$ an. Eine Senkung des Bilirubinwertes $<0,6 \times \text{ULN}$ nach einem Jahr unter UDCA-Therapie führte zu einer Verbesserung der 10-Jahres-Überlebensrate um 11%. Folglich ist das Erreichen von Gesamt-Bilirubinwerten $\leq 0,6 \times \text{ULN}$ bei Patienten mit PBC mit dem geringsten Risiko für eine Lebertransplantation oder Tod verbunden [16]. Untersuchungen in einem kanadischen Register (mit einer zusätzlichen Validierungskohorte) zum Entzündungs- und Stoffwechselproteom im Serum von PBC-Patienten unter UDCA-Therapie zeigen ebenfalls, dass für das Therapieziel einer vollständigen Kontrolle der Krankheitsaktivität bei PBC sowohl eine ALP-Normalisierung als auch eine Normalisierung des Bilirubins erforderlich ist [13]. Basierend auf den Daten der GLOBE-Kohorte zeigt sich weiter, dass erhöhte Gesamt-Bilirubinwerte (sowohl zu Baseline als auch nach 5 Jahren) mit einer schlechteren Prognose assoziiert sind. Die transplantationsfreien Überlebensraten nach 5, 10 und 15 Jahren waren bei Patienten mit normalen Gesamt-Bilirubinwerten ($\leq 1 \times \text{ULN}$) signifikant höher als bei Patienten mit erhöhten Gesamt-Bilirubinwerten [15]. Obschon die Bilirubinwerte meist relativ spät im Krankheitsverlauf ansteigen, ist der prädiktive Wert nicht nur auf das späte Krankheitsstadium beschränkt und selbst bei Patienten mit normalen Werten verbessert sich die Prognose, wenn die Werte sinken [15]. Bei anderen Lebererkrankungen hat sich der Gesamt-Bilirubinwert ebenfalls als prädiktiv für den klinischen Verlauf bestätigt und ist in gängigen Prognosemodellen in Kombination mit weiteren Leberwerten enthalten [15, 19]. Als Schwellenwert für das Gesamt-Bilirubin wird in der Literatur der im Rahmen des kombinierten biochemischen Ansprechens verwendete Cut-off von $\leq 1,0 \times \text{ULN}$ als adäquat angesehen [15]. Der Gesamt-Bilirubinwert kann daher als wichtiger Prädiktor für einen ungünstigen Verlauf der PBC und das individuelle Risiko einer Krankheitsprogression angesehen werden [14].

Untersuchungen aus dem kanadischen Register (mit einer zusätzlichen Validierungskohorte) zeigen unterstützend dazu, dass der Verlust des biochemischen Ansprechens zu einem beliebigen Zeitpunkt mit einem erhöhten Risiko für Lebertransplantation oder Tod korreliert – umgekehrt ist das (Wieder-)Erreichen eines biochemischen Ansprechens im Verlauf der Krankheit zu jedem Zeitpunkt mit besseren klinischen Ergebnissen verbunden. Ein weiterer

Beleg dafür, dass das Erreichen eines nachhaltigen biochemischen Ansprechens auf die Therapie ein wichtiges Ziel während des gesamten Krankheitsverlaufs darstellt [18].

Folglich werden die Einzelkomponenten des kombinierten biochemischen Ansprechens (ALP und Gesamt-Bilirubin) ebenfalls als geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC herangezogen. Es wird davon ausgegangen, dass sowohl eine Abnahme der ALP- als auch der Gesamt-Bilirubinwerte mit einer besseren Prognose für den Patienten assoziiert sind, da niedrigere Werte eindeutig mit einem besseren transplantationsfreien Überleben verbunden sind. Für die vorliegende Nutzenbewertung wird das kombinierte biochemische Ansprechen sowie die stetigen Analysen von ALP und Gesamt-Bilirubin als nicht-schwerer Endpunkt eingestuft.

Ergänzend werden die Veränderungen der ALT, AST, des direkten und indirekten Bilirubins sowie der GGT und der 5'-Nucleotidase dargestellt. Diese Laborparameter korrelieren ebenfalls mit dem Schweregrad der PBC-Erkrankung und können in Kombination mit ALP und/oder Bilirubin zur Beurteilung der Krankheitsaktivität herangezogen werden. Beispielsweise wird in der klinischen Praxis der hepatische Ursprung eines erhöhten ALP-Wertes meist durch eine gleichzeitige Erhöhung der GGT (oder der 5'-Nukleotidase) sowie des direkten Bilirubins bestätigt [4, 14, 15, 21, 22].

Zusammengefasst ist die positive Auswirkung der hier relevanten Surrogatparameter (ALP-Normalisierung und kombiniertes biochemisches Ansprechen) im Hinblick auf die langfristige Vermeidung der Folgekomplikation und somit die klinischen Endpunkte Lebertransplantation und Tod qualitativ belegt. Biochemisches Ansprechen stellt somit in der vorliegenden Indikation einen patientenrelevanten Endpunkt dar.

Pruritus

Die Operationalisierung von Pruritus erfolgt anhand der Fragebögen Pruritus NRS und 5-D Pruritus.

Chronischer Pruritus ist ein schwerwiegendes Symptom der PBC, dessen Ursache bislang nicht vollständig geklärt ist [4]. Er kann bei bis zu 80% der Patienten auftreten [4, 23, 24]. Auftreten und Intensität von Pruritus korreliert nicht mit dem Schweregrad der Lebererkrankung, kann patientenindividuell sehr unterschiedlich sein [25] und tritt unabhängig vom Ausmaß der Cholestase – oft schubweise – auf.

Pruritus zeigt meist eine gesteigerte nächtliche Intensität und führt so zu teils schwerem Schlafentzug, Abgeschlagenheit sowie Depressionen und kann bei einer extremen und therapierefraktären Ausprägung sogar Suizidalität zur Folge haben. Die daraus resultierende Tagesmüdigkeit und Abgeschlagenheit können die Arbeitsfähigkeit der Betroffenen erheblich einschränken. Hinzu kommt, dass Pruritus häufig nicht zufriedenstellend therapiert werden kann [4]. Jüngste Untersuchungen haben gezeigt, dass PBC-Patienten mit schwerem Pruritus eine ähnlich niedrige Lebensqualität aufweisen wie Patienten mit schwerer Parkinson-

Krankheit [26]. Therapierefraktärer, unerträglicher Pruritus stellt bei Patienten mit PBC in seltenen Fällen eine Indikation zur Lebertransplantation dar [24, 27].

Insgesamt stellt Pruritus ein schwerwiegendes und sehr belastendes Symptom für den Patienten dar, das mit einer beträchtlichen Einschränkung der Lebensqualität einhergeht und manchmal nur mit einer Lebertransplantation gelindert werden kann. Sowohl die Linderung als auch das Verhindern des Auftretens von Pruritus werden als direkt patientenrelevant angesehen.

Pruritus NRS

Die Pruritus NRS wird anhand einer 11-Punkte-Skala durch den Patienten selbst bewertet von 0 („kein Pruritus“) bis 10 („schwerster vorstellbarer Pruritus“). Dargestellt werden zum einen die präspezifizierte Operationalisierung zur Anzahl und Anteil der Patienten mit einer Verringerung um ≥ 4 Punkte gegenüber Baseline. Hierzu werden die binären Effektmaße RR, OR und RD mit zugehörigem p-Wert angegeben. Des Weiteren werden die Ergebnisse der stetigen Analyse zur Pruritus NRS anhand der Änderungen gegenüber Baseline anhand der LS-MWD mit zugehörigem p-Wert sowie der SMD zu den Monaten 6 und 12 präsentiert. Numerische Bewertungsskalen sind in der klinischen Praxis und in Studien häufig verwendete objektive Instrumente zur Bewertung der Pruritusintensität und werden daher als adäquat angesehen. Die Pruritus NRS wurde in zwei Studien bei Patienten mit chronischem Pruritus validiert und stellt ein verlässliches und multidimensionales Messinstrument dar [28, 29].

5-D Pruritus

Der Fragebogen 5-D Pruritus umfasst 27 Items in fünf Subdomänen (Grad, Dauer, Verlauf, Beeinträchtigung und Lokalisation) als Ausprägungen der Morbidität. Die Werte der einzelnen Items reichen von 1 bis 5 und werden durch den Patienten selbst bewertet. Der Bereich Beeinträchtigung besteht aus vier Items, die die Auswirkungen des Pruritus auf das tägliche Leben bewerten: Schlaf, Freizeit/Soziales, Hausarbeit/Besorgungen und Arbeit/Schule. Für den Bereich Lokalisation wird die Anzahl der betroffenen Körperteile mit einer möglichen Summe von 0 bis 16 ermittelt. Die Punkte für jede der fünf Subdomänen werden zur Ermittlung eines 5-D Gesamt-Scores aufsummiert. Der 5-D Gesamt-Score kann von 5 (kein Pruritus) bis 25 (schwerster Pruritus) reichen. Sowohl für den Gesamt-Score als auch für die einzelnen Subdomänen werden die Änderungen gegenüber Baseline anhand der LS-MWD mit zugehörigem p-Wert sowie der SMD zu den Monaten 6 und 12 präsentiert. Ergänzend wird ein modifizierter Gesamt-Score basierend auf vier der fünf Subdomänen (Grad, Dauer, Beeinträchtigung und Lokalisation) berechnet. Diese Anpassung erfolgte aufgrund von Missverständnissen bei der Bewertung der Subdomäne Verlauf: Patienten, die keinen Pruritus hatten, gaben oft „unverändert“ an, was einem Wert von 4 entspricht, und nicht „ganz verschwunden“, was einem Wert von 1 entspricht. Auch für den modifizierten Gesamt-Score wird die Änderung gegenüber Baseline anhand der LS-MWD mit zugehörigem p-Wert sowie der SMD zu den Monaten 6 und 12 berichtet.

Die Validierung der 5-D Pruritus-Skala erfolgte im Rahmen einer Studie basierend auf 234 Patienten mit chronischem Pruritus. Darunter waren 63 Patienten mit einer

Lebererkrankung. Der 5-D wurde dabei als verlässliches und multidimensionales Messinstrument für Pruritus etabliert und validiert [30].

PGI-C & PGI-S

Die Global Improvement-Skalen gibt es in einem vom Kliniker bewerteten und in einem vom Patienten bewerteten Format. Die Patient Global Impression of Severity (PGI-S)- und Patient Global Impression of Change (PGI-C)-Skalen sind patientenberichtete, valide Instrumente, anhand derer sowohl eine Einschätzung des Schweregrades der Erkrankung eines Patienten und seines klinischen Gesundheitszustands (PGI-S-Skala) als auch der Verbesserung, respektive Veränderung des klinischen Gesundheitszustands des Patienten im zeitlichen Verlauf (PGI-C-Skala) abgefragt werden kann.

Der PGI-S dient der Bewertung der Schwere der Erkrankung und der aktuellen Symptomatik. Die Beurteilung erfolgt basierend auf einer fünfstufigen Skala von Kategorie 1 (keine Symptome) bis 5 (sehr starke Symptome) in Bezug auf die vergangene Woche. Der PGI-C dient der Bewertung der Veränderung der Krankheitslast und Symptomatik unter der Behandlung für eine spezifische Erkrankung, wobei die Patienten direkt die selbsteingeschätzte Änderung gegenüber Baseline angeben. Die Frage ist auf einer siebenstufigen Skala von Kategorie 1 (sehr große Verbesserung) bis 7 (sehr große Verschlechterung) zu beurteilen [31].

Für beide Fragebögen werden die Änderungen gegenüber Baseline anhand der LS-MWD mit zugehörigem p-Wert sowie der SMD zu den Monaten 6 und 12 berichtet.

Entsprechende Endpunkte wurden bereits in früheren Nutzenbewertungsverfahren vom G-BA zur Bewertung des Gesundheitszustands als patientenrelevant anerkannt [32, 33]. Beide Fragebögen sind standardisiert und werden in klinischen Studien zur Beurteilung der Symptomschwere bzw. der Veränderung der Krankheitslast regelmäßig eingesetzt [34-36].

Beide Fragebögen ermöglichen es, die patientenindividuelle Wahrnehmung der Schwere der Erkrankung bzw. der Veränderung der Symptombelastung aus Sicht des Patienten zu erfassen und sind somit als unmittelbar patientenrelevant zu erachten.

Fibrose

Die Fibrose wird bestimmt mittels eines Scores für fortgeschrittene Leberfibrose (ELF) und der Messung der Lebersteifigkeit auf Basis des FibroScan®. Die Änderung gegenüber Baseline wird durch die Effektmaße LS-MWD und SMD mit zugehörigem p-Wert zu den Monaten 6 und 12 angegeben.

Bei der PBC kommt es zu entzündlichen Veränderungen der kleinen Gallengänge in der Leber und im weiteren Verlauf der Erkrankung zu einer Zerstörung der Gallengänge und Cholestase. Greift die Entzündung auf das gesamte Lebergewebe über, führt dies zu Leberfibrose. Während eine Fibrose in frühen Stadien reversibel ist, schreitet sie unbehandelt zur irreversiblen Zirrhose fort. Die Umstrukturierung des normalen Lebergewebes in fibrotisches und folgend zirrhotisches Gewebe führt zu einem Verlust der normalen Leberfunktion und schwerwiegenden systemischen und lebensbedrohlichen Komplikationen, die für den Patienten

direkt spürbar sind. So kann es aufgrund der Schädigung der Leber durch die Entwicklung einer portalen Hypertonie zu Ösophagusvarizen, gastrointestinalen Blutungen sowie Aszites kommen [37]. In Folge der Leberzirrhose kann es auch zum Auftreten von schweren Infektionen kommen. Des Weiteren stellt eine Leberzirrhose den entscheidenden Risikofaktor für die Entwicklung eines HCC dar [4] und kann zum Leberversagen führen [14, 38, 39]. Die Entstehung einer Fibrose beeinflusst direkt den Krankheitsverlauf und die Prognose sowie die weiteren Therapieentscheidungen und stellt einen validen und patientenrelevanten Endpunkt dar. Die Messung des Fibrosegrades mittels FibroScan® und ELF-Score erfasst somit die Krankheitsschwere sowie die Progression der PBC im Zeitverlauf.

Messung der Lebersteifigkeit (FibroScan®)

Zur nicht-invasiven Messung der Lebersteifigkeit und der Beurteilung der Festigkeit des Gewebes, respektive des Fibrosegrades gilt die vibrationskontrollierte transiente Elastografie (TE) als am besten untersuchtes und validiertes Verfahren. Diese sollte bei erwachsenen Patienten mit einer autoimmunen Lebererkrankung zur Beurteilung eines frühen oder fortgeschrittenen Krankheitsstadiums und zur Prognoseabschätzung durchgeführt werden [4]. Werte zur Lebersteifigkeit von mehr als 9,6 kPa werden mit einem 5-fach erhöhten Risiko für Leberdekompensation, Lebertransplantation oder Tod verbunden [14]. Die TE hat sich als zuverlässiges, nicht-invasives Verfahren zur Bestimmung der Lebersteifigkeit, respektive Leberfibrose bei der PBC erwiesen. Sie wird sowohl zur Diagnostik als auch zur Überwachung des Fortschreitens der Fibrose eingesetzt und stellt ein valides Messinstrument für den patientenrelevanten Endpunkt Fibrose dar [40-42].

Messung des Fibrosegrades (ELF-Score)

Der ELF-Score ist ein nicht-invasiver Bluttest, bei dem drei direkte Fibrosemarker gemessen werden. Mittels dieses Scores können Patienten mit PBC in verschiedene Risikogruppen eingeteilt werden [14]. Anhand einer Studie an 512 Patienten mit chronischer Hepatitis C wurde die Eignung des ELF-Scores als Instrument zur Früherkennung einer Leberfibrose oder -zirrhose bei einer chronischen Hepatitis C bestätigt. Damit stellt dieser Score ebenfalls ein valides Messinstrument für den patientenrelevanten Endpunkt Fibrose dar [43].

Hospitalisierung

Dargestellt werden Anzahl und Anteil der Patienten mit Hospitalisierungen zu Monat 12, sowie das RR, OR und die RD mit zugehörigem p Wert.

Bei einer Hospitalisierung ist davon auszugehen, dass sich die krankheitsspezifische Symptomatik in einem Ausmaß verschlechtert hat, das eine stationäre Behandlung unabdingbar macht. Die Hospitalisierung ist somit eine konkrete Folge des Fortschreitens der Erkrankung, die mit einer Verschlechterung der Symptomatik einhergeht. Daher kann eine Hospitalisierung als Annäherung an den klinischen Zustand einer Symptomverschlechterung angesehen werden. Außerdem kann der Patient die zunehmenden Beeinträchtigungen seiner Funktionen und Aktivitäten durch die Erkrankung, die zur Hospitalisierung führt, unmittelbar wahrnehmen. Die mit der Hospitalisierung verbundene schwere Symptomatik beeinträchtigt den

Gesundheitszustand sowie die Lebensqualität des Patienten, sodass deren Vermeidung von höchster Relevanz ist.

Folglich sind Hospitalisierungen unabhängig von ihrer Ursache patientenrelevant und bilden eine schwerwiegende Verschlechterung der Symptomatik und des Gesundheitszustands ab.

In der Studie RESPONSE wurde die Anzahl der Hospitalisierungen aufgrund der Erkrankung erfasst.

Gesundheitsbezogene Lebensqualität

PBC-40

Die Erfassung der Lebensqualität wird per se als patientenrelevant eingestuft. Entsprechend § 2 Satz 3 Arzneimittel-Nutzenbewertungsverordnung (AM-Nutzen V) ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“

Der Fragebogen PBC-40 ist ein etablierter und validierter Fragebogen für die PBC und enthält Domänen zur Morbidität sowie der gesundheitsbezogenen Lebensqualität [44]. Der G-BA sieht darin ein „weitestgehend validiertes und krankheitsspezifisches Instrument zur Messung der Lebensqualität bei Patienten mit PBC“ [45], weshalb im Rahmen des Dossiers eine Darstellung des PBC-40 ausschließlich im Rahmen der gesundheitsbezogenen Lebensqualität erfolgt.

Der PBC-40 ist in sechs Domänen unterteilt, die als Einzeldomänen sowie als Gesamt-Score dargestellt werden. Den Antworten zu den Fragen werden sogenannte Score-Punkte zugeordnet (1=überhaupt nicht, 2=selten, 3=manchmal, 4=häufig, 5=immer), wobei der Gesamt-Score die Aufsummierung über alle Einzelfragen angibt und somit Werte zwischen 32 und 200 annehmen kann.

Die klinische Relevanz der Veränderung wird basierend auf der Änderung gegenüber Baseline durch die Effektmaße LS-MWD und SMD mit zugehörigem p-Wert zu den Monaten 6 und 12 angegeben.

UE

Dem Verträglichkeits- und Sicherheitsprofil kommt generell bei jeder Therapie eine besondere Bedeutung zu: Neben den erwünschten Wirkungen eines Arzneimittels treten in vielen Fällen auch UE auf, die den Patienten neben der eigentlichen Erkrankung physisch und psychisch belasten können. Die Bedeutung von UE ist abhängig von ihrer Schwere, ihrer Häufigkeit, ihrer Behandelbarkeit, respektive der Reversibilität ihrer Symptome.

Entsprechend werden neben Wirksamkeits- und Lebensqualitätsparametern auch UE als Endpunkte in die Nutzenbewertung eingeschlossen. Sowohl der G-BA als auch das IQWiG erkennen UE als patientenrelevanten Endpunkt an [11, 12].

Die Erhebung von UE erfolgt im Rahmen klinischer Studien standardisiert mittels Medical Dictionary for Regulatory Activities (MedDRA 26.0) und nach den Regeln der International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use Good Clinical Practice.

Es wird der Anteil Patienten mit mindestens einem Ereignis betrachtet. Relevante Ereignisse sind jegliche UE, schwerwiegende UE (SUE), schwere UE (definiert als UE mit Common Terminology Criteria for Adverse Events [CTCAE]-Grad ≥ 3), UE, die zum Therapieabbruch führten, UE, die zum Tod führten, sowie UE von speziellem Interesse (UESI). Des Weiteren werden UE nach Schweregrad nach SOC und PT gemäß folgenden Häufigkeitskriterien selektiert:

- Jegliche UE nach SOC und PT, die bei $\geq 10\%$ der Patienten in einem Studienarm oder bei ≥ 10 Patienten sowie mindestens 1% in einem Studienarm aufgetreten sind
- SUE nach SOC und PT, die bei $\geq 5\%$ der Patienten in einem Studienarm oder bei ≥ 10 Patienten sowie mindestens 1% in einem Studienarm aufgetreten sind
- Schwere UE nach SOC und PT, die bei $\geq 5\%$ der Patienten in einem Studienarm oder bei ≥ 10 Patienten sowie mindestens 1% in einem Studienarm aufgetreten sind

Nicht-schwere UE, Therapieabbrüche aufgrund von UE nach SOC und PT (deskriptiv) sowie die UE-Raten unter Ausschluss erkrankungsbezogener Ereignisse werden ergänzend dargestellt.

Dargestellt werden jeweils Anzahl und Anteil der Patienten mit Ereignissen, sowie das RR, OR und die RD mit zugehörigem p-Wert.

4.2.5.3 Meta-Analysen

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

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten nach der Knapp-

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

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁶ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand geeigneter statistische Maße auf Vorliegen von Heterogenität^{7, 5} erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0,05$), soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Nicht zutreffend.

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

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

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

4.2.5.4 Sensitivitätsanalysen

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

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

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

Auf Studienebene dienen Sensitivitätsanalysen zur Überprüfung der Robustheit gegenüber methodischer Variation. Dabei können beispielsweise die Wahl der Auswertungspopulation (ITT-Population vs. alternative Population), der Imputationsalgorithmus (Observed-Auswertung vs. Last Observation Carried Forward-Imputation), die Endpunkt-Operationalisierung (Ein- oder Ausschluss bestimmter Kriterien) oder der Einschluss von Patienten, die aus besonderen Gründen nicht in der primären Analyse berücksichtigt wurden, variiert werden. Dabei ist man auf die Verfügbarkeit solcher Analysen aus den Einzelstudien angewiesen.

ALP-Normalisierung

Um die Robustheit der Ergebnisse zur ALP-Normalisierung zu untersuchen, wurde im Rahmen dieser Nutzenbewertung eine ergänzende Sensitivitätsanalyse durchgeführt. Es wurde neben dem Cut-off-Wert $ALP \leq 1 \times ULN$ (ALP-Normalisierung) auch der präspezifizierte Cut-off-Wert hinsichtlich $ALP < 1,5 \times ULN$ betrachtet. Der ALP Cut-off von $< 1,5 \times ULN$ stellt im Rahmen der Paris-II-Kriterien einen weiteren Marker für den therapeutischen Erfolg bei der PBC dar. Bereits das Erreichen eines ALP-Wertes von $< 1,5 \times ULN$ nach 12 Monaten gilt als Prädiktor für langfristig verbesserte klinische Ergebnisse. Dazu gehören ein reduziertes Risiko für leberbedingte Komplikationen und eine erhöhte transplantationsfreie Überlebensrate [46].

Kombiniertes Biochemisches Ansprechen

In der relevanten RCT RESPONSE handelt es sich bei dem Endpunkt kombiniertes biochemisches Ansprechen um einen kombinierten Endpunkt aus $ALP < 1,67 \times ULN$, Gesamt-Bilirubin $\leq ULN$ sowie ALP-Reduktion von $\geq 15\%$. Als primäre Dossieranalyse wird dieser Endpunkt dargestellt. Die im SAP präspezifizierte Operationalisierung wird in diesem Dossier als Sensitivitätsanalyse dargestellt und lautete:

Anteil der Patienten, die nach 12 Monaten basierend auf dem folgenden zusammengesetzten Endpunkt als Responder betrachtet werden können:

- ALP <1,67 x ULN
- Gesamt-Bilirubin ≤ULN
- ALP-Reduktion von ≥15%

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.

Es werden Subgruppenanalysen durchgeführt, um zu überprüfen, ob der Therapieeffekt über die Subgruppen hinweg konstant ist oder ob es Hinweise auf eine Effektmodifikation gibt. Datenquellen für Subgruppenresultate sind Subgruppenanalysen aus Studienberichten sowie post-hoc-Analysen. Folgende Subgruppen werden getrennt betrachtet:

- Alter zu Screening
- Alter zum Zeitpunkt der PBC-Diagnose
- Geschlecht
- Ethnie
- Region
- Pruritus NRS
- UDCA-Intoleranz
- Vorherige Einnahme von OCA und/oder Fibraten
- Mono-/Kombitherapie

Sowie als Maß der Krankheitsschwere:

- Zirrhose
- ALP
- GGT
- Gesamt-Bilirubin

Im Abschnitt 4.3.1.3.2 werden die Cut-offs und die Relevanz der Subgruppenanalysen begründet.

Den Modulvorlagen entsprechend werden Subgruppenanalysen nur dann dargestellt und als relevant erachtet, wenn mindestens zehn Patienten pro Subgruppe in die Analyse eingehen und bei binären Endpunkten zudem in einer Subgruppe mindestens zehn Ereignisse beobachtet wurden. Für UE nach SOC und PT werden Subgruppenanalysen durchgeführt, sofern für die jeweilige SOC bzw. den jeweiligen PT auf Basis der Gesamtpopulation ein statistisch signifikanter Behandlungsunterschied vorliegt.

Grundsätzlich werden Subgruppenanalysen für jeden Endpunkt für die primär relevante Analyse betrachtet.

Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen werden mittels Interaktionstests bewertet. Zur Identifikation von Effektmodifikationen wird für binäre und kontinuierliche Variablen ein auf dem Q-Test basierender p-Wert herangezogen. Bedeutsame Heterogenität zwischen den Subgruppen wird ab einem p-Wert von $<0,05$ angenommen. Sofern fazitrelevant, führt diese Effektmodifikation zu getrennten Nutzaussagen.

Eine detaillierte tabellarische Darstellung von Subgruppenresultaten erfolgt nur für Subgruppenanalysen, die einen statistisch signifikanten Interaktionstest (p -Wert $< 0,05$) aufweisen. Alle weiteren Subgruppenanalysen werden lediglich ergänzend in Anhang 4-G anhand der Ausgabe der statistischen Analysesoftware dargestellt.

4.2.5.6 Indirekte Vergleiche

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

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

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

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

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

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

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

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

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

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

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

Nicht zutreffend.

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

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

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

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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

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

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

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

Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zbAM

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen/abgebrochen/laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
RESPONSE CB8025-32048	Ja	Nein ^a	Abgeschlossen	12 Monate Datenschnitt: 11.08.2023	• SEL 10 mg • Placebo
ENHANCE CB8025-31735	Ja	Nein ^a	Abgeschlossen ^b	12 Monate ^b Datenschnitt: 16.02.2020	• SEL 5-10 mg • SEL 10 mg • Placebo
IDEAL CB8025-32251	Nein	Nein ^a	Laufend	52 Wochen	• SEL 10 mg • Placebo
AFFIRM CB8025-41837	Nein	Nein ^a	Laufend	36 Monate	• SEL 5 mg (Child-Pugh-A) • SEL 10 mg (Child-Pugh-B) • Placebo
CB8025-21528	Ja	Nein ^a	Abgebrochen ^c	12 Wochen	• SEL 50 mg • SEL 200 mg • Placebo
a: Die Firma CymaBay Therapeutics Inc. wurde von Gilead Sciences Inc. übernommen und ist seit März 2024 mit dem Abschluss der Übernahme eine hundertprozentige Tochtergesellschaft von Gilead Sciences Inc. Der aktuelle Zulassungsinhaber von SEL ist entsprechend Gilead Sciences Ireland UC. Das Produkt wird in Deutschland von der Gilead Sciences GmbH vertrieben. Auf ClinicalTrials.gov wird Gilead Sciences als Sponsor angegeben, in den Studiendokumenten wird auch CymaBay Therapeutics noch genannt. b: Die Studie sollte ursprünglich die Behandlung über einen Zeitraum von 12 Monaten bewerten, sie wurde jedoch vorzeitig beendet und die Endpunktanalyse nach drei Monaten durchgeführt. c: Die Studie wurde vorzeitig beendet, nachdem sich die Notwendigkeit einer Dosisreduzierung ergeben hatte. RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, zbAM=Zu bewertendes Arzneimittel					

Geben Sie an, welchen Stand die Information in Tabelle 4-3 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 Angaben in Tabelle 4-3 haben den Stand vom 15. Januar 2025.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-3 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zbAM

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
ENHANCE CB8025-31735	Zu kurze Studiendauer (<48 Wochen)
IDEAL CB8025-32251	Keine Ergebnisse vorhanden
AFFIRM CB8025-41837	Keine Ergebnisse vorhanden
CB8025-21528	Dosierung von SEL nicht zulassungskonform
SEL=Seladelpar, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel	

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.

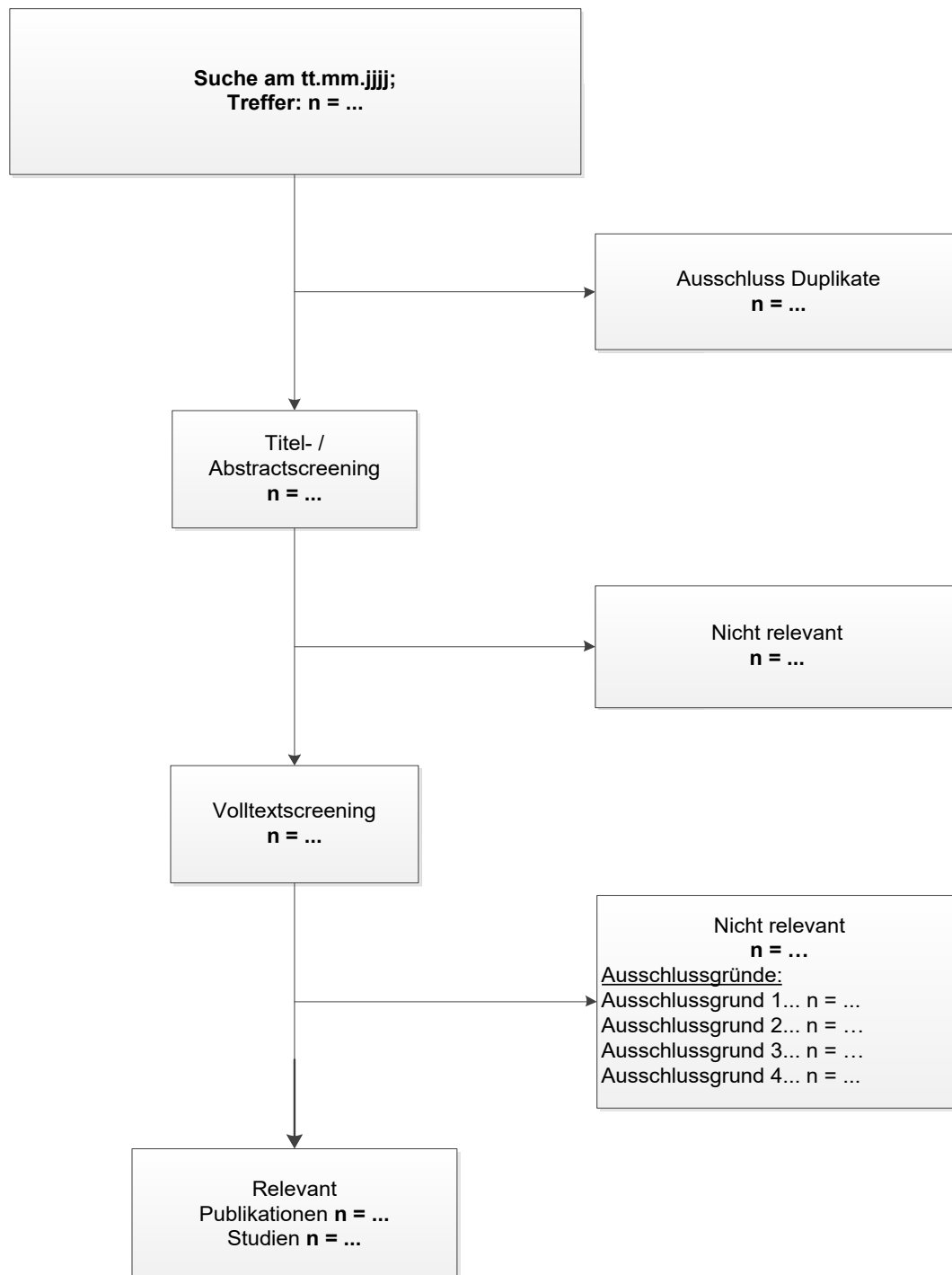


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zbAM

n=Anzahl der Patienten mit Ereignis, zbAM=Zu bewertendes Arzneimittel

Die umfassende bibliografische Literaturrecherche nach RCT mit dem zbAM SEL wurde am 15. Januar 2025 durchgeführt. Diese erzielte insgesamt 84 Treffer (ohne Duplikate), die zunächst auf Grundlage von Titel und Abstract nach den in Abschnitt 4.2.2 definierten Kriterien überprüft wurden.

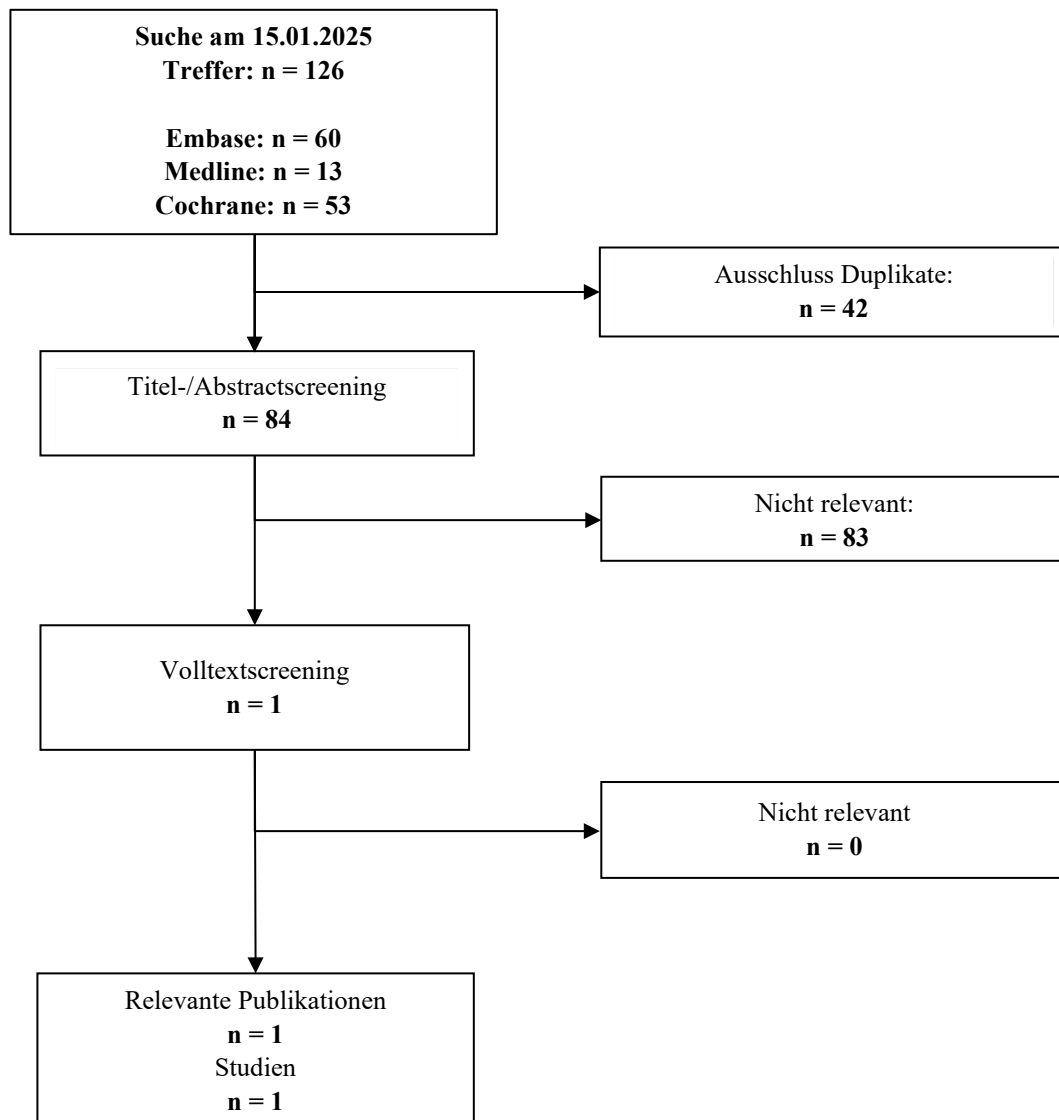


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach RCT mit dem zbAM SEL

n=Anzahl der Patienten mit Ereignis, RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, zbAM=Zu bewertendes Arzneimittel

Es konnte eine Publikation zu der RCT RESPONSE identifiziert werden, die für die Bewertung von SEL relevant ist [47].

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 4-3) 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 4-5: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zbAM

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)
RESPONSE	ClinicalTrials.gov [48] EU-CTR [49] WHO-ICTRP [50]	Ja	Ja [47]	Abgeschlossen
a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. EU-CTR=European Union Clinical Trials Register, ICTRP=International Clinical Trials Registry Platform, NCT=National Clinical Trial, RCT=Randomisierte kontrollierte Studie, WHO=Weltgesundheitsorganisation, zbAM=Zu bewertendes Arzneimittel				

Geben Sie an, welchen Stand die Information in Tabelle 4-5 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Angaben in Tabelle 4-5 haben den Stand vom 15. Januar 2025.

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 4-3) 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 4-6: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zbAM

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 G-BA=Gemeinsamer Bundesausschuss, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel				

Geben Sie an, welchen Stand die Information in Tabelle 4-6 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 Dokuments maßgeblichen Zeitpunkt liegen.

Die Suche auf der Internetseite des G-BA wurde auf Studien beschränkt, die bereits in der Liste der Studien des pharmazeutischen Unternehmers, der bibliografischen Literaturrecherche oder Studienregistersuche als relevant für die vorliegende Nutzenbewertung identifiziert wurden. Es konnten keine zusätzlichen Quellen (Daten zu Studienmethodik oder -ergebnissen) zu der Studie RESPONSE identifiziert werden, die für die Bewertung von SEL im vorliegenden Anwendungsgebiet relevant sind.

Die Angaben in Tabelle 4-6 haben den Stand vom 15. Januar 2025.

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 4-7: Studienpool – RCT mit dem zbAM

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zbAM (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienberichte (ja/nein [Zitat])	Register-einträge ^c (ja/nein [Zitat])	Publikation und sonstige Quellen ^d (ja/nein [Zitat])
placebokontrolliert						
RESPONSE	Ja	Nein ^e	Ja	Ja [51]	Ja [48-50]	Ja [47]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Nicht zutreffend						
a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden. b: Studie, für die der Unternehmer Sponsor war. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA. e: Die Firma CymaBay Therapeutics Inc. wurde von Gilead Sciences Inc. übernommen und ist seit März 2024 mit dem Abschluss der Übernahme eine hundertprozentige Tochtergesellschaft von Gilead Sciences Inc. Der aktuelle Zulassungsinhaber von SEL ist entsprechend Gilead Sciences Ireland UC. Das Produkt wird in Deutschland von der Gilead Sciences GmbH vertrieben. Auf ClinicalTrials.gov wird Gilead Sciences als Sponsor angegeben, in den Studiendokumenten wird auch CymaBay Therapeutics noch genannt. G-BA=Gemeinsamer Bundesausschuss, RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, zbAM=Zu bewertendes Arzneimittel						

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 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zbAM

Studie	Studiendesign <RCT, doppelblind/einfach, verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
RESPONSE	RCT, doppelblind, parallel, multizentrisch, Phase 3	Erwachsene mit PBC, die inadäquat auf UDCA angesprochen haben oder UDCA nicht vertragen haben	Gruppe A: SEL 10 mg (n=128) Gruppe B: Placebo (n=65)	Screening: bis zu 3 Monate Run in: 2 Wochen Behandlung: 12 Monate Follow-up: 17 Tage Datenschnitte: siehe Tabelle 4-3	90 Standorte aus 24 Ländern Randomisierung erster Patient: 21. April 2021 Letzter Patient letzte Visite; 11. August 2023	Primärer Endpunkt Biochemisches Ansprechen ^a Patientenrelevante sekundäre Endpunkte: • ALP- Normalisierung • Bilirubin- Normalisierung • Pruritus-NRS • PBC-40 • 5-D Pruritus • PGI-C & PGI-S • Fibrose • Hospitalisierung • UE
a: Anteil Patienten mit ALP <1,67 x ULN, Gesamt-Bilirubin ≤ULN und ALP-Reduktion von ≥15% ALP=Alkalische Phosphatase, n=Anzahl der Patienten, NRS=Numerical Rating Scale, PBC=Primär biliäre Cholangitis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, UDCA=Ursodeoxycholsäure, UE=Unerwünschtes Ereignis, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel						

Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zbAM

Studie	Gruppe A	Gruppe B	ggf. weitere Spalten mit Behandlungsscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in- Phase etc.
RESPONSE	10 mg SEL (oder 5 mg SEL bei Heruntertitration) oral, einmal täglich für 12 Monate	Placebo, oral, einmal täglich für 12 Monate	<u>Heruntertitration:</u> Bei Probanden, die spezifische Sicherheitsüberwachungskriterien erfüllen oder bei denen Probleme mit der Verträglichkeit auftreten, kann eine Dosisreduktion vorgenommen werden. Probanden, die ursprünglich 10 mg SEL erhalten haben, werden verblindet auf 5 mg heruntertitriert. Bei Probanden, die ursprünglich Placebo erhalten haben, erfolgt eine verblindete Heruntertitration und sie bleiben in der Placebo-Gruppe. Probanden, bei denen ein klinisch bedeutsames UE auftritt, das nach dem klinischen Urteil des Prüfarztes eine Dosisreduktion rechtfertigt, kommen ebenfalls für eine Dosisreduktion in Frage. Die Dosisreduktion muss von dem medizinischen Monitor genehmigt werden und wird verblindet durchgeführt. <u>UDCA:</u> Während der Studie wird das Studienmedikament bei Probanden, die UDCA vertragen, als Zusatz zur UDCA-Therapie verabreicht; bei Patienten mit UDCA-Unverträglichkeit wird das Studienmedikament als Monotherapie verabreicht. UDCA wird oral ein- oder mehrmals täglich in der gleichen Dosis wie vor Studienbeginn und nach dem klinischen Urteil des Prüfarztes verabreicht.
RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, UDCA=Ursodeoxycholsäure, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel			

Tabelle 4-10: Charakterisierung der Studienpopulationen der Studie RESPONSE (SAS) – demografische und krankheitsbedingte Charakteristika aus RCT mit dem zbAM

	SEL	Placebo
N	128	65
Alter (Jahre)		
Mittelwert (SD)	56,6 (9,99)	57,0 (9,17)
Median	57,0	58,0
Q1; Q3	50,0; 63,5	50,0; 63,0
Min; Max	28; 75	33; 75
Geschlecht, n (%)		
Weiblich	123 (96,1)	60 (92,3)
Männlich	5 (3,9)	5 (7,7)

	SEL	Placebo
Ethnische Zugehörigkeit, n (%)		
Hispanisch oder lateinamerikanisch	29 (22,7)	27 (41,5)
Nicht-hispanisch oder lateinamerikanisch	97 (75,8)	38 (58,5)
fehlend	2 (1,6)	0
Region, n (%)		
Nordamerika	50 (39,1)	13 (20,0)
Lateinamerika	24 (18,8)	19 (29,2)
Europa, Naher Osten und Afrika	49 (38,3)	27 (41,5)
Asien-Pazifik	5 (3,9)	6 (9,2)
BMI (kg/m²)		
Mittelwert (SD)	27,2 (5,61)	26,8 (4,81)
Median	26,0	26,2
Q1; Q3	23,0; 29,8	23,7; 29,3
Min; Max	17,5; 45,0	17,4; 40,1
Zirrhose-Status, n (%)		
Ja	18 (14,1)	9 (13,8)
Nein	110 (85,9)	56 (86,2)
AMA-Status, n (%)		
Positiv ^a	106 (82,8)	55 (84,6)
Negativ	20 (15,6)	9 (13,8)
Mehrdeutig	2 (1,6)	1 (1,5)
MELD-Score, n (%)		
<12	128 (100,0)	65 (100,0)
≥12	0	0
Fibrose-Score abgeleitet aus der Lebersteifigkeit, n (%)		
F0	44 (38,3)	26 (41,9)
F1	22 (19,1)	15 (25,2)
F2	17 (14,8)	7 (11,3)
F3	21 (18,3)	10 (16,1)
F4	11 (9,6)	4 (6,5)
GLOBE Risk Score		
n	125	65
Mittelwert (SD)	0,3 (0,66)	0,3 (0,71)
Median	0,3	0,2
Q1; Q3	-0,2; 0,8	-0,1; 0,8
Min; Max	-1,2; 1,8	-1,3; 1,8
Rotterdam Stage Disease, n (%)		
Mild	106 (82,8)	60 (92,3)
Moderat fortgeschritten	22 (17,2)	5 (7,7)

	SEL	Placebo
ALP (U/L) zu Baseline		
Mittelwert (SD)	314,6 (122,96)	313,8 (117,68)
Median	278,3	281,3
Q1; Q3	227,8; 357,2	236,3; 353,3
Min; Max	182; 786	161, 698
ALP-Level, n (%)		
<350 U/L	93 (72,7)	47 (72,3)
≥350 U/L	35 (27,3)	18 (27,7)
Gesamt-Bilirubin (mg/dL) zu Baseline		
Mittelwert (SD)	0,8 (0,31)	0,7 (0,31)
Median	0,7	0,7
Q1; Q3	0,6; 0,9	0,5; 1,0
Min; Max	0,3; 1,9	0,3; 1,9
Gesamt-Bilirubin-Level, n (%)		
≤1 x ULN	108 (84,4)	60 (92,3)
>1 und ≤2 x ULN	20 (15,6)	5 (7,7)
>2 x ULN	0	0
<0,6 x ULN	59 (46,1)	32 (49,2)
≥0,6 x ULN	69 (53,9)	33 (50,8)
Pruritus NRS		
Mittelwert (SD)	3,0 (2,81)	3,0 (2,96)
Median	2,8	1,8
Q1; Q3	0,1; 5,3	0,2; 5,6
Min; Max	0; 9	0; 9
Pruritus NRS-Level, n (%)		
<4	79 (61,7)	42 (64,6)
≥4	49 (38,3)	23 (35,4)
Fatigue, n (%)		
Ja	60 (46,9)	33 (50,8)
Nein	68 (53,1)	32 (49,2)
HDL-C (mg/dL) zu Baseline(mg/dL) zu Baseline		
Mittelwert (SD)	80,5 (23,09)	75,1 (22,34)
Median	79,7	71,7
Q1; Q3	62,7; 95,8	58,7; 91,7
Min; Max	34, 165	34, 126
LDL-C		
Mittelwert (SD)	136,7 (45,35)	137,4 (51,13)
Median	135,5	131,0
Q1; Q3	101,7; 165,3	103,0; 171,0
Min; Max	54; 268	51, 295

	SEL	Placebo
ALT (U/L) zu Baseline		
Mittelwert (SD)	47,4 (23,47)	48,2 (22,83)
Median	45,3	43,0
Q1; Q3	28,0; 63,7	31,0; 58,3
Min; Max	13; 109	9; 115
AST (U/L) zu Baseline		
Mittelwert (SD)	39,6 (16,14)	41,7 (16,03)
Median	37,7	38,0
Q1; Q3	27,0; 48,0	30,3; 49,3
Min; Max	15,7; 1402,0	16; 1088
GGT (U/L) zu Baseline		
Mittelwert (SD)	269,0 (240,04)	287,5 (249,56)
Median	195,2	216,3
Q1; Q3	114,2; 336,5	120,3; 338,0
Min; Max	12,7; 1402,0	41,5; 1088,0
Vorherige Einnahme von OCA und/oder Fibraten, n (%)		
Ja	20 (15,6)	13 (20,0)
Nein	108 (84,4)	52 (80,0)
UDCA-Intoleranz, n (%)		
Ja	120 (93,8)	61 (93,8)
Nein	8 (6,3)	4 (6,2)
Alter bei PBC-Diagnose (Jahre), n (%)		
Mittelwert (SD)	49,2 (9,94)	49,3 (10,94)
Median	50,0	50,0
Q1; Q3	42,0; 55,5	42,0; 57,0
Min; Max	25; 70	26; 68
UK-PBC Risk Score (5 Jahre)		
n	125	65
Mittelwert (SD)	0,0 (0,02)	0,0 (0,02)
Median	0,0	0,0
Q1; Q3	0,0; 0,0	0,0; 0,0
Min; Max	0,0; 0,1	0,0; 0,1
UK-PBC Risk Score (10 Jahre)		
n	125	65
Mittelwert (SD)	0,1 (0,06)	0,1 (0,06)
Median	0,1	0,1
Q1; Q3	0,0; 0,1	0,0; 0,1
Min; Max	0,0; 0,3	0,0; 0,3
UK-PBC Risk Score (15 Jahre)		
n	125	65
Mittelwert (SD)	0,1 (0,10)	0,0 (0,09)
Median	0,1	0,1

	SEL	Placebo
Q1; Q3	0,1; 0,2	0,1; 0,1
Min; Max	0,0; 0,5	0,0; 0,4
a: Positiv >1:40 in der Immunfluoreszenz oder M2 positiv im ELISA ALP=Alkalische Phosphatase, ALT=Alanin-Aminotransferase, AMA=Antimitochondriale Antikörper, AST=Aspartat-Aminotransferase, BMI=Body Mass Index, ELISA=Enzyme-Linked Immunosorbent Assay, GGT=Gamma-Glutamyltransferase, HDL-C=High Density Lipoprotein-Cholesterol, LDL-C=Low Density Lipoprotein-Cholesterol, MELD=Model of Endstage Liver Disease, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OCA=Obeticholsäure, NRS=Numerical Rating Scale, Q=Quartil, RCT=Randomisierte kontrollierte Studie, SAS=Safety-Analysis-Set, SD=Standardabweichung, SEL=Seladelpar, UDCA=Ursodeoxycholsäure, UK=United Kingdom, ULN=Oberer Normwert, PBC=Primär biliäre Cholangitis, zbAM=Zu bewertendes Arzneimittel		

Tabelle 4-11: Charakterisierung der Studienpopulation der Studie RESPONSE – Patientendisposition – RCT mit dem zbAM

	SEL	Placebo
Disposition der Patienten, n (%)		
Randomisiert	128	65
Keine Studienmedikation	0	0
SAS	128 (100)	65 (100)
FAS	128 (100)	65 (100)
Behandlung beendet	118 (92,2)	57 (87,7)
Vorzeitiges Behandlungsende	10 (7,8)	8 (12,3)
Studie beendet	117 (91,4)	57 (87,7)
Vorzeitiges Ausscheiden aus der Studie	11 (8,6)	8 (12,3)
Mediane Behandlungsdauer unter Studienmedikation, Wochen (min; max)	52,1 (0,1; 54,7)	52,0 (1,3; 55,4)
Hauptgründe für den frühzeitigen Abbruch der Studienmedikation, n (%)		
UE	4 (3,1)	4 (6,2)
Einverständnis zurückgezogen	4 (3,1)	2 (3,1)
Protokollabweichung	1 (0,8)	1 (1,5)
Lost to Follow-up	1 (0,8)	1 (1,5)
Hauptgründe für den frühzeitigen Abbruch der Studie, n (%)		
UE	3 (2,3)	4 (6,2)
Einverständnis zurückgezogen	5 (3,9)	2 (3,1)
Protokollabweichung	1 (0,8)	1 (1,5)
Lost to Follow-up	1 (0,8)	1 (1,5)
Andere	1 (0,8)	0
FAS=Full-Analysis-Set, n=Anzahl der Patienten mit Ereignis, RCT=Randomisierte kontrollierte Studie, SAS=Safety-Analysis-Set, SEL=Seladelpar, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel		

Tabelle 4-12: Begleitmedikationen ($\geq 10\%$) der Studie RESPONSE (SAS) – RCT mit dem zbAM

ATC Level 2 Term Preferred Term	SEL	Placebo
N	128	65
Patienten mit mindestens einer Begleitmedikation, n (%)	125 (97,7)	64 (98,5)
Vitamine	64 (50,0)	28 (43,1)
Cholecalciferol	25 (19,5)	15 (23,1)
Vitamin D (Not Otherwise Specified)	18 (14,1)	3 (4,6)
Substanzen für säurebedingte Störungen	62 (48,4)	25 (38,5)
Omeprazol	26 (20,3)	11 (16,9)
Analgetika	57 (44,5)	28 (43,1)
Paracetamol	27 (21,1)	22 (33,8)
Lipidmodifizierende Wirkstoffe	42 (32,8)	34 (52,3)
Atorvastatin	11 (8,6)	12 (18,5)
Colestyramin	6 (4,7)	8 (12,3)
Psycholeptika	36 (28,1)	18 (27,7)
Antibakterielle Mittel zur systemischen Anwendung	35 (27,3)	20 (30,8)
Mineralstoffprodukte	35 (27,3)	16 (24,6)
Entzündungshemmende und antirheumatische Produkte	34 (26,6)	23 (35,4)
Ibuprofen	16 (12,5)	12 (18,5)
Antihistaminika zur systemischen Anwendung	33 (25,8)	21 (32,3)
Hydroxyzin	14 (10,9)	3 (4,6)
Anti-Anämie-Präparate	31 (24,2)	8 (12,3)
Wirkstoffe, die auf das Renin-Angiotensin-System wirken	27 (21,1)	19 (29,2)
Psychoanaleptika	26 (20,3)	9 (13,8)
Schilddrüsen-Therapie	24 (18,8)	9 (13,8)
Levothyroxin	14 (10,9)	7 (10,8)
Beta-Blocker	21 (16,4)	8 (12,3)
Substanzen für funktionelle gastrointestinale Störungen	19 (14,8)	9 (13,8)
Substanzen zur Behandlung von Diabetes	19 (14,8)	7 (10,8)
Ophthalmologische Behandlung	17 (13,3)	12 (18,5)
Substanzen für obstruktive Atemwegserkrankungen	17 (13,3)	6 (9,2)
Kortikosteroide, dermatologische Präparate	14 (10,9)	8 (12,3)
Kortikosteroide zur systemischen Anwendung	13 (10,2)	11 (16,9)

ATC Level 2 Term Preferred Term	SEL	Placebo
N	128	65
Patienten mit mindestens einer Begleitmedikation, n (%)	125 (97,7)	64 (98,5)
Kalziumkanal-Blocker	12 (9,4)	7 (10,8)
Antithrombotische Mittel	11 (8,6)	7 (10,8)
Acetylsalicylsäure	5 (3,9)	7 (10,8)
Impfstoffe	10 (7,8)	8 (12,3)
ATC=Anatomisch-Therapeutisch-Chemisch, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, RCT=Randomisierte kontrollierte Studie, SAS=Safety-Analysis-Set, SEL=Seladelpar, zbAM=Zu bewertendes Arzneimittel		

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

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

Studiendesign

Die Studie RESPONSE ist eine multizentrische, doppelblinde, randomisierte Parallelgruppenstudie der Phase 3 bei Erwachsenen mit PBC, die inadäquat auf UDCA angesprochen oder UDCA nicht vertragen haben.

Im Anschluss eines bis zu 3-monatigen Screening-Zeitraums und einer 2-wöchigen Run-in-Phase wurden die Patienten durch ein elektronisches Randomisierungssystem im Verhältnis 2:1 (SEL:Placebo), stratifiziert nach ALP-Level und Vorliegen von Pruritus, randomisiert und einer der zwei Behandlungsgruppen zugeteilt:

- SEL 10 mg für 12 Monate
- Placebo für 12 Monate

Für die Nutzenbewertung werden die Behandlungsgruppen nach 6-monatiger und 12-monatiger Behandlung verglichen. Im Folgenden wird der SEL 10 mg Arm als SEL-Arm bezeichnet. UDCA konnte in beiden Armen bei UDCA-Toleranz weiter eingenommen werden.

Die Gesamtstudiendauer beträgt 16 Monate (einschließlich der Screening-, Run-in-, Behandlungs- und Nachbeobachtungsphase). Der Behandlungszeitraum beträgt 12 Monate, dem 17 Tage Nachbeobachtung folgen. Patienten hatten außerdem die Möglichkeit im Anschluss an die Behandlungsdauer in eine offene Langzeitnachbeobachtungsstudie (ASSURE) eingeschlossen zu werden. Für diese Patienten entfällt die 14 ($\pm$ 3)-tägige Nachbeobachtungsphase.

Die Studie RESPONSE ist international und multizentrisch. Die Probanden kamen aus insgesamt 90 Studienzentren in Argentinien, Australien, Österreich, Belgien, Kanada, Chile, Tschechische Republik, Frankreich, Deutschland, Griechenland, Ungarn, Israel, Italien, Republik Korea, Mexiko, Neuseeland, Polen, Rumänien, Russland, Spanien, Schweiz, Türkei, Vereinigtem Königreich und den Vereinigten Staaten von Amerika.

Die Patienten im SEL-Arm waren zu Baseline durchschnittlich 56,6 Jahre alt; 96,1% der SEL-Patienten waren weiblich. Im Placebo-Arm lag das mittlere Alter bei 57,0 Jahren und 92,3% der Patienten war weiblich. Der mittlere BMI zu Baseline lag im SEL-Arm bei 27,2 kg/m² und im Placebo-Arm bei 26,8 kg/m². Der Großteil der Patienten war nicht-lateinamerikanisch oder hispanisch (75,8% vs. 58,5%). Der mittlere ALP-Wert zu Baseline betrug im SEL-Arm 314,6 U/L und im Placebo-Arm 313,8 U/L. Im Mittel hatten die Patienten eine identische Pruritus NRS (3,0 in beiden Armen). Die Mehrheit der Patienten in beiden Behandlungsarmen (85,9% vs. 86,2%) hatten zu Baseline keine Zirrhose. Weitere Patientencharakteristika sind in Tabelle 4-10 dargestellt. Insgesamt zeigt sich damit eine Vergleichbarkeit der Studienarme.

In der Studie RESPONSE brachen 7,8% der Patienten die Behandlung mit SEL und 12,3% der Patienten die Placebo-Gabe ab. Die Hauptgründe in beiden Studienarmen waren UE (3,1% vs. 6,2%), Einverständnis zurückgezogen (3,1% vs. 3,1%), Protokollabweichung (0,8% vs. 1,5%) und Lost to Follow-up (0,8% vs. 1,5%). Die Studie RESPONSE brachen 8,6% der Patienten aus dem SEL-Arm und 12,3% aus dem Placebo-Arm ab. Die Hauptgründe in beiden Studienarmen waren UE (2,3% vs. 6,2%), Einverständnis zurückgezogen (3,9% vs. 3,1%), Protokollabweichung (0,8% vs. 1,5%), Lost to Follow-up (0,8% vs. 1,5%) und Andere (0,8% vs. 0%) (siehe Tabelle 4-11).

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse der Studie RESPONSE auf den deutschen Versorgungsalltag ist aufgrund der guten Vergleichbarkeit der Studienpopulation und der deutschen Population mit PBC gegeben.

Im Hinblick auf das Alter handelt es sich bei der PBC typischerweise um eine Erkrankung älterer Patienten – so erkrankten die meisten Patienten um das 5. Lebensjahrzehnt [4, 52] und unter 25 Jahren tritt PBC im Prinzip nicht auf [27]. Frauen sind dabei mit 90% deutlich häufiger betroffen als Männer [4]. In der hier relevanten Studie RESPONSE lag das mittlere Alter der Patienten bei 56,6 (SEL-Arm), respektive 57,0 (Placebo-Arm) Jahren und es wurden fast ausschließlich Frauen eingeschlossen (96,1, respektive 92,3%); der überwiegende Anteil der Patienten war weiß.

Im SEL-Arm lag der mittlere BMI bei 27,2 kg/m², im PLACEBO-Arm bei 26,8 kg/m². Dies ist vergleichbar mit dem mittleren BMI der deutschen Bevölkerung mit 26,0 kg/m² [53].

Der Anteil der Patienten mit PBC, die einen positiven AMA-Status vorweisen, lag in der Studie RESPONSE bei 82,8%, respektive 84,6%, und bewegt sich damit in der gleichen Größenordnung wie der in einer retrospektiven deutschen Erhebung zur PBC-Erkrankung ermittelte Anteil an Patienten mit positivem AMA-Status (86,9%) [54].

Zusammenfassend ist daher von einer Übertragbarkeit der Studiendaten der Studie RESPONSE auf den deutschen Versorgungskontext auszugehen.

Vergleichstherapie

Da es sich bei SEL um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, ist keine zVT zu bestimmen. Der medizinische Nutzen und Zusatznutzen von SEL gelten bereits durch die Zulassung als belegt und das Ausmaß des Zusatznutzens ist auf Grundlage der Zulassung und der die Zulassung begründenden Studie zu bestimmen.

Bei der Zulassungsstudie RESPONSE handelt es sich um eine Placebo-kontrollierte Studie. Neben der Studienmedikation erhielten alle Patienten eine patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität.

In Tabelle 4-12 werden alle Begleittherapien dargestellt, die in der Studie RESPONSE bei mindestens 10% Patienten in mindestens einem Arm verabreicht wurden. Zu den häufigsten Begleittherapien (bei ≥10% der Patienten in mindestens einem Arm) im Verum-, respektive Kontrollarm der Studie RESPONSE zählten Vitamine, Wirkstoffe für säurebedingte Störungen (Omeprazol), Analgetika (Paracetamol) und lipidmodifizierende Wirkstoffe (Atorvastatin und Colestyramin).

Vitamin-Präparate wurden am häufigsten eingesetzt. Bei Patienten mit cholestatischer Lebererkrankung besteht das Risiko einer metabolischen Knochenerkrankung, einer Dyslipidämie und eines Mangels an fettlöslichen Vitaminen (Vitamine A, D, E und K) [27]. Zudem ist die PBC mit einem signifikant erhöhten Risiko für die Entwicklung einer Osteopenie oder Osteoporose assoziiert. Daher soll bei Patienten mit PBC bei fehlenden Kontraindikationen (beispielsweise Nierensteine in der Vorgeschichte) eine primärprophylaktische Substitution mit Cholecalciferol erfolgen [4].

Auch der Protonenpumpeninhibitor Omeprazol wurde bei 20,3% (SEL), respektive 16,9% (Placebo) der Patienten mit PBC eingesetzt. Protonenpumpeninhibitoren reduzieren die Freisetzung von Magensäure und werden zur Behandlung von Sodbrennen, Refluxösophagitis, Magen- und Zwölffingerdarmgeschwüren sowie bei Infektionen mit *Helicobacter pylori* eingesetzt. Diese Arzneimittel werden unabhängig von der PBC in Deutschland häufig verschrieben, dies zeigt sich ebenfalls in der Studie RESPONSE.

PBC-Patienten können eine Cholestase-bedingte Hypercholesterinämie aufweisen, die zu Xanthomen und/oder Xanthelasmen führen kann. Hierfür kann eine Statintherapie verabreicht werden [4]. Zur Behandlung des cholestatischen Pruritus steht beispielsweise das Austauscherharz Colestyramin zur Verfügung [4, 55].

Es ist somit davon auszugehen, dass die Patienten in der Studie grundsätzlich eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität erhalten haben und auch die PBC adäquat behandelt wurde.

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 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zbAM

Studie	Adäquate Erzeugung der Randomisierungs- sequenz	Verdeckung der Gruppenteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
RESPONSE	Ja	Ja	Ja	Ja	Ja	Ja	Niedrig
RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel							

Begründen Sie für jede Studie die abschließende Einschätzung.

In der Studie RESPONSE wurde ein elektronisches Randomisierungssystem verwendet. Es kann somit davon ausgegangen werden, dass die Randomisierungssequenz adäquat erzeugt wurde und die Gruppenteilung verdeckt erfolgte. Sowohl Patienten als auch behandelnde Personen waren verblindet. Es gibt keine Hinweise auf eine ergebnisabhängige Berichterstattung oder auf sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten.

Das Verzerrungspotenzial wird daher für diese Studie als niedrig eingestuft.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

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

Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zbAM

Studie	Mortalität	Morbidität ^a	Lebensqualität ^b	UE
RESPONSE	Ja	Ja	Ja	Ja
a: Biochemisches Ansprechen (ALP-Normalisierung und kombiniertes biochemisches Ansprechen), 5-D Pruritus, Pruritus-NRS, PGI-C & PGI-S, Fibrose; Hospitalisierung b: PBC-40 ALP=Alkalische Phosphatase, NRS=Numerical Rating Scale, PBC=Primär biliäre Cholangitis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, RCT=Randomisierte kontrollierte Studie, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel				

4.3.1.3.1 Endpunkte – RCT

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

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

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

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

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

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

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

1. Falls in einer Studie Responderanalysen unter Verwendung einer MID präspezifiziert sind und das Responsekriterium mindestens 15 % der Skalenspannweite des verwendeten Erhebungsinstruments entspricht, sind diese Responderanalysen für die Bewertung darzustellen.

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

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

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,
2. Gesamtrate schwerwiegender UE (SUE),
3. Gesamtrate der Abbrüche wegen UE,

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

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

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

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

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

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

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

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

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

vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

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

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

4.3.1.3.1.1 Mortalität – RCT

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

Tabelle 4-15: Operationalisierung von Mortalität

Studie	Operationalisierung
RESPONSE	Die Erhebung der Mortalität erfolgt im Rahmen der UE und ist dem Abschnitt 4.3.1.3.1.9 zu entnehmen.
UE=Unerwünschtes Ereignis	

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

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

Tabelle 4-16: Bewertung des Verzerrungspotenzials für Mortalität in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Die Mortalität wird im Rahmen der UE in Abschnitt 4.3.1.3.1.9 dargestellt.					
ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Mortalität wird im Rahmen der UE in Abschnitt 4.3.1.3.1.9 dargestellt.

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.

Die Mortalität wird im Rahmen der UE in Abschnitt 4.3.1.3.1.9 dargestellt. In der Studie RESPONSE ist kein Patient verstorben.

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.2 Biochemisches 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 4-17: Operationalisierung von biochemischem Ansprechen

Studie	Operationalisierung
RESPONSE	<u>ALP-Normalisierung</u> Die ALP-Normalisierung wurde definiert als $ALP \leq 1 \times ULN$. Ergänzend wird außerdem $ALP < 1,5 \times ULN$ als Sensitivitätsanalyse dargestellt. Die Darstellung erfolgt zu Monat 6 und 12, der Vergleich erfolgt anhand RR, OR und RD mit zugehörigem p-Wert. <u>Kombiniertes biochemisches Ansprechen</u> Das kombinierte biochemische Ansprechen wurde definiert als $ALP < 1,67 \times ULN$, Gesamt-Bilirubin $\leq ULN$ und ALP-Reduktion von $\geq 15\%$. Dargestellt werden Anzahl und Anteil der Patienten mit kombiniertem Ansprechen zu Monat 6 und 12, sowie das RR, OR und RD mit zugehörigem p-Wert. Neben dem zusammengesetzten Endpunkt werden auch die Einzelkomponenten getrennt dargestellt analog der Analyse des zusammengesetzten Endpunktes: • $ALP < 1,67 \times ULN$ • ALP-Reduktion von $\geq 15\%$ • Bilirubin-Abfall: Gesamt-Bilirubin $\leq ULN$ Die Darstellung erfolgt ebenfalls zu Monat 6 und 12. <u>Änderungen gegenüber Baseline</u> Zusätzlich neben der binären Analyse werden die Ergebnisse der stetigen Analyse zu ALP und Gesamt-Bilirubin sowie ergänzend weiterer Laborparameter dargestellt. Hierfür werden die Änderungen gegenüber Baseline anhand der LS-MWD mit zugehörigem p-Wert sowie der SMD zu den Monaten 6 und 12 präsentiert.
ALP=Alkalische Phosphatase, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, OR=Odds Ratio, RD=Risikodifferenz, RR=Relatives Risiko, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert	

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 4-18: Bewertung des Verzerrungspotenzials für biochemisches Ansprechen in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

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

Tabelle 4-19: Ergebnisse für biochemisches Ansprechen aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert RD [95%-KI], p-Wert
Monat 6			
ALP-Normalisierung (ALP $\leq 1,0 \times$ ULN)	34/128 (26,6)	0/65 (0,0)	16,40 [2,32; 116,20], 0,0051 25,49 [3,34; 194,87], 0,0018 0,27 [0,19; 0,34], <0,0001
ALP <1,5 x ULN	82/128 (64,1)	8/65 (12,3)	5,21 [2,70; 10,05], <0,0001 15,94 [6,41; 39,64], <0,0001 0,52 [0,40; 0,63], <0,0001
Kombiniertes biochemisches Ansprechen	85/128 (66,4)	12/65 (18,5)	3,60 [2,16; 6,01], <0,0001 12,92 [5,56; 30,03], <0,0001 0,48 [0,36; 0,60] p<0,0001
• ALP <1,67 x ULN	89/128 (69,5)	15/65 (23,1)	3,01 [1,93; 4,69], <0,0001 11,58 [5,10; 26,31], <0,0001 0,46 [0,34; 0,58], <0,0001
• Bilirubin-Abfall: Gesamt-Bilirubin $\leq$ ULN	111/128 (86,7)	54/65 (83,1)	1,05 [0,93; 1,19], 0,4456 1,41 [0,60; 3,27], 0,4284 0,04 [-0,06; 0,15], 0,4417
• ALP-Reduktion von $\geq 15\%$	118/128 (92,2)	26/65 (40,0)	2,31 [1,71; 3,13], <0,0001 18,67 [8,11; 42,98], <0,0001 0,52 [0,40; 0,65], <0,0001
Monat 12			
ALP-Normalisierung (ALP $\leq 1,0 \times$ ULN)	32/128 (25,0)	0/65 (0,0)	14,80 [2,08; 105,44], 0,0071 21,99 [2,87; 168,78], 0,0030 0,25 [0,18; 0,33], <0,0001
ALP <1,5 x ULN	75/128 (58,6)	8/65 (12,3)	4,79 [2,45; 9,37], <0,0001 9,87 [4,28; 22,73], <0,0001 0,46 [0,35; 0,58], <0,0001

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert RD [95%-KI], p-Wert
Kombiniertes biochemisches Ansprechen	79/128 (61,7)	13/65 (20,0)	3,09 [1,87; 5,10], <0,0001 7,26 [3,42; 15,41], <0,0001 0,42 [0,29; 0,54], <0,0001
ALP <1,67 x ULN	84/128 (65,6)	17/65 (26,2)	2,51 [1,65; 3,80], <0,0001 6,78 [3,22; 14,26], <0,0001 0,39 [0,27; 0,52], <0,0001
Bilirubin-Abfall: Gesamt-Bilirubin ≤ULN	104/128 (81,3)	50/65 (76,9)	1,06 [0,91; 1,24], 0,4511 1,34 [0,64; 2,79], 0,4335 0,05 [-0,07; 0,17], 0,4441
ALP-Reduktion von ≥15%	107/128 (83,6)	21/65 (32,3)	2,61 [1,81; 3,75], <0,0001 10,50 [5,18; 21,30], <0,0001 0,52 [0,39; 0,65], <0,0001
Ein RR >1 bedeutet einen Vorteil für die Behandlung mit SEL. ALP=Alkalische Phosphatase, FAS=Full-Analysis-Set, KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SEL=Seladelpar, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel			

Es zeigten sich statistisch signifikante Ergebnisse für die ALP-Normalisierung, operationalisiert durch $ALP \leq 1 \times ULN$. Dieser statistisch signifikante Vorteil zeigte sich sowohl zu Monat 6 (RR [95%-KI]: 16,40 [2,32; 116,20], 0,0051) als auch zu Monat 12 (RR [95%-KI]: 14,80 [2,08; 105,44], 0,0071). Die Sensitivitätsanalyse $ALP < 1,5 \times ULN$ unterstützt die Ergebnisse zur ALP-Normalisierung, da auch hier ein statistisch signifikanter Vorteil zu beiden Analysezeitpunkten gezeigt werden konnte (RR [95%-KI] zu Monat 6: 5,21 [2,70; 10,05], $p < 0,0001$; Monat 12: 4,79 [2,45; 9,37], $p < 0,0001$).

Für das kombinierte biochemische Ansprechen zeigte sich ebenfalls ein statistisch signifikanter Vorteil zugunsten von SEL. Der Vorteil war sowohl zu Monat 6 (RR [95%-KI]: 3,60 [2,16; 6,01]; $p < 0,0001$) als auch zu Monat 12 (RR [95%-KI]: 3,09 [1,87; 5,10]; $p < 0,0001$) statistisch signifikant. Zu Monat 6 sprachen 66,4% der Patienten unter SEL an, unter Placebo 18,5%. Dieser Unterschied im biochemischen Ansprechen bestätigt sich zu Monat 12 mit Ansprechraten von 61,7% unter SEL und 20,0% unter Placebo.

Der statistisch signifikante Vorteil für die Behandlung mit SEL zeigte sich auch in den beiden Einzelkomponenten des kombinierten biochemischen Ansprechens ($ALP < 1,67 \times ULN$ und ALP-Reduktion von $\geq 15\%$) zu Monat 6 und 12. Für die Komponente Bilirubin-Abfall (Gesamt-Bilirubin $\leq ULN$) zeigte sich für beide Beobachtungszeiträume ein numerischer Vorteil zugunsten von SEL. In der Gesamtschau zeigt sich damit ein gleichgerichteter Effekt der präspezifizierten Einzelkomponenten. Zudem sind alle Komponenten von ähnlicher Schwere und als valide Surrogate anzusehen. Entsprechend kann die Ableitung des Zusatznutzens in Abschnitt 4.4.2 ausschließlich basierend auf dem kombinierten biochemischen Ansprechen erfolgen.

Tabelle 4-20: Ergebnisse für biochemische Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 6						
Veränderung der ALP (U/L)						
126	314,55 (122,96)	-140,59 (6,52)	63	313,83 (117,68)	-18,03 (8,66)	-122,56 [-142,15; -102,98], <0,0001 -1,70 [-2,05; -1,35]
Veränderung des Gesamt-Bilirubins (mg/dL)						
126	0,77 (0,31)	-0,07 (0,02)	63	0,74 (0,31)	0,00 (0,03)	-0,07 [-0,15; 0,01], 0,0896 -0,26 [-0,56; 0,05]
Veränderung des direkten Bilirubins (mg/dL)						
126	0,24 (0,16)	-0,00 (0,01)	63	0,21 (0,14)	0,00 (0,02)	0,00 [-0,04; 0,04], 0,9811 0,00 [-0,31; 0,30]
Veränderung des indirekten Bilirubins (mg/dL)						
126	0,53 (0,19)	-0,06 (0,01)	63	0,53 (0,20)	0,01 (0,02)	-0,07 [-0,11; -0,03], 0,0007 -0,52 [-0,82; -0,21]
Veränderung der ALT (U/L)						
126	47,45 (23,47)	-10,67 (1,69)	63	48,25 (22,83)	-2,69 (2,34)	-7,98 [-13,49; -2,47], 0,0048 -0,42 [-0,73; -0,12]
Veränderung der AST (U/L)						
126	39,62 (16,14)	-2,43 (1,34)	63	41,67 (16,03)	-1,91 (1,86)	-0,52 [-4,88; 3,83], 0,8135 -0,03 [-0,34; 0,27]
Veränderung der GGT (U/L)						
126	269,04 (240,04)	-105,62 (7,40)	63	287,51 (249,56)	-35,86 (10,22)	-69,76 [-93,70; -45,81], <0,0001 -0,84 [-1,16; -0,53]
Veränderung der 5'-Nucleotidase (U/L)						
126	15,30 (11,30)	-7,15 (0,45)	63	16,73 (10,60)	-2,39 (0,63)	-4,76 [-6,19; -3,32], <0,0001 -0,94 [-1,25; -0,62]
Monat 12						
Veränderung der ALP (U/L)						
126	314,55 (122,96)	-133,93 (8,51)	63	313,83 (117,68)	-16,91 (11,67)	-117,02 [-144,20; -89,84], <0,0001 -1,23 [-1,56; -0,90]
Veränderung des Gesamt-Bilirubins (mg/dL)						
126	0,77 (0,31)	0,00 (0,04)	63	0,74 (0,31)	0,02 (0,05)	-0,02 [-0,14; 0,11], 0,7777 -0,04 [-0,35; 0,26]
Veränderung des direkten Bilirubins (mg/dL)						
126	0,24 (0,16)	0,03 (0,02)	63	0,21 (0,14)	0,01 (0,03)	0,03 [-0,05; 0,10], 0,5139 0,10 [-0,20; 0,40]

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Veränderung des indirekten Bilirubins (mg/dL)						
126	0,53 (0,19)	-0,03 (0,02)	63	0,53 (0,20)	0,01 (0,02)	-0,04 [-0,10; 0,02], 0,1470 -0,22 [-0,53; 0,08]
Veränderung der ALT (U/L)						
126	47,45 (23,47)	-12,20 (1,62)	63	48,25 (22,83)	-3,88 (2,22)	-8,32 [-13,55; -3,09], 0,0020 -0,46 [-0,77; -0,16]
Veränderung der AST (U/L)						
126	39,62 (16,14)	-2,49 (1,31)	63	41,67 (16,03)	-1,50 (1,80)	-0,98 [-5,22; 3,25], 0,6473 -0,07 [-0,37; 0,24]
Veränderung der GGT (U/L)						
126	269,04 (240,04)	-107,97 (8,49)	63	287,51 (249,56)	-18,33 (11,79)	-89,65 [-117,47; -61,82], <0,0001 -0,94 [-1,26; -0,63]
Veränderung der 5'-Nucleotidase (U/L)						
126	15,30 (11,30)	-7,11 (0,62)	63	16,73 (10,60)	-4,04 (0,86)	-3,08 [-5,10; -1,06], 0,0031 -0,45 [-0,75; -0,14]
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit SEL. a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline. ALP=Alkalische Phosphatase, ALT=Alanin-Aminotransferase, AST=Aspartat-Aminotransferase, FAS=Full-Analysis-Set, GGT=Gamma-Glutamyltransferase, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, zbAM=Zu bewertendes Arzneimittel						

Die Behandlung mit SEL führte im Vergleich zu Placebo im Mittel zu einer größeren Verbesserung bei der ALP, dem Gesamt-Bilirubin, direkten Bilirubin, indirekten Bilirubin, der ALT, der AST, der GGT sowie 5'-Nucleotidase. Die Verbesserung zeigte sich in statistisch signifikanten Unterschieden oder mindestens numerischen Vorteilen zu Monat 6 und 12, was mit einer geringeren Krankheitsaktivität der SEL-Patienten gegenüber den Placebo-Patienten zu assoziieren ist. Die jeweiligen zeitlichen Verläufe der mittleren Veränderung sind in Abbildung 3 bis Abbildung 10 dargestellt.

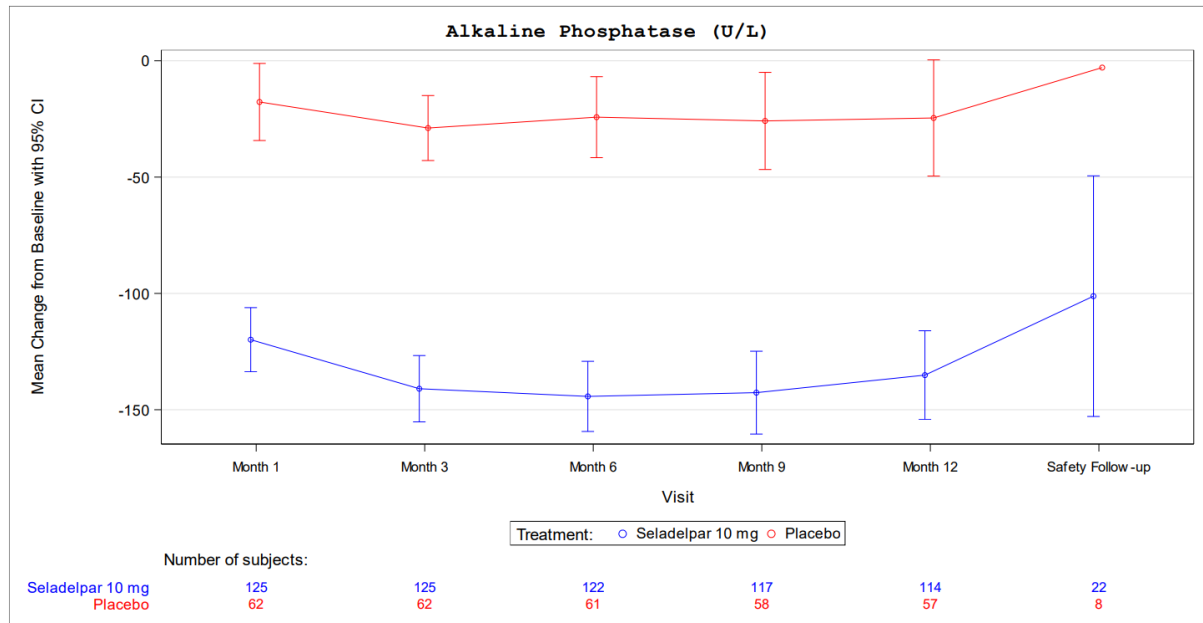


Abbildung 3: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der ALP aus RCT mit dem zbAM (Studie RESPONSE, FAS)

ALP=Alkalische Phosphatase, CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

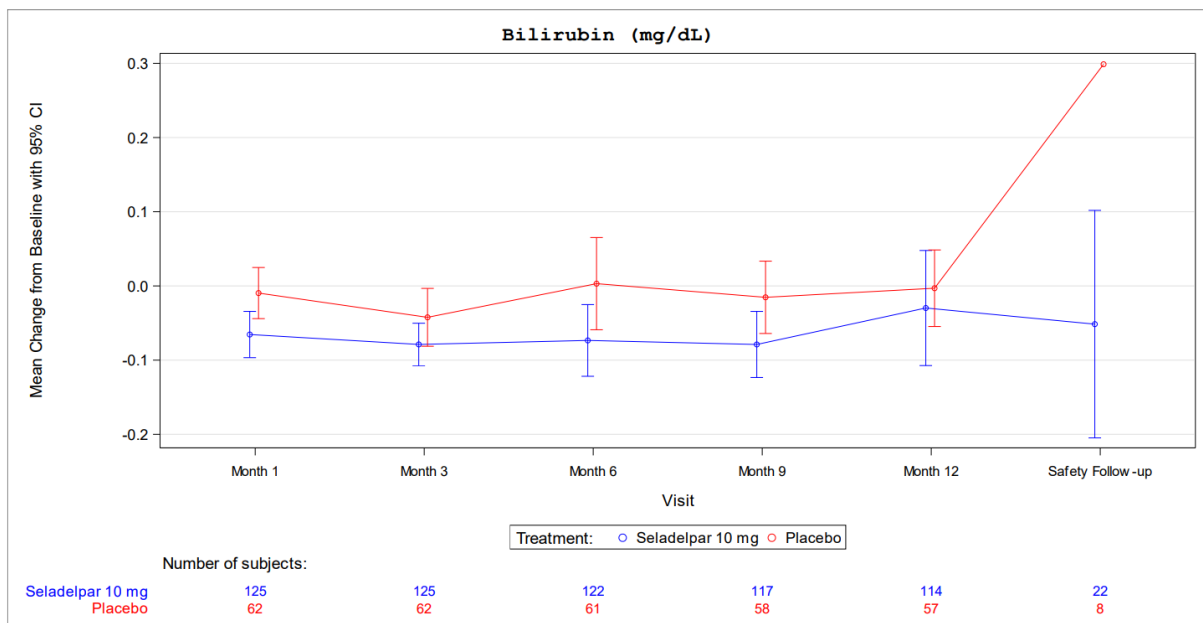


Abbildung 4: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des Gesamt-Bilirubins aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

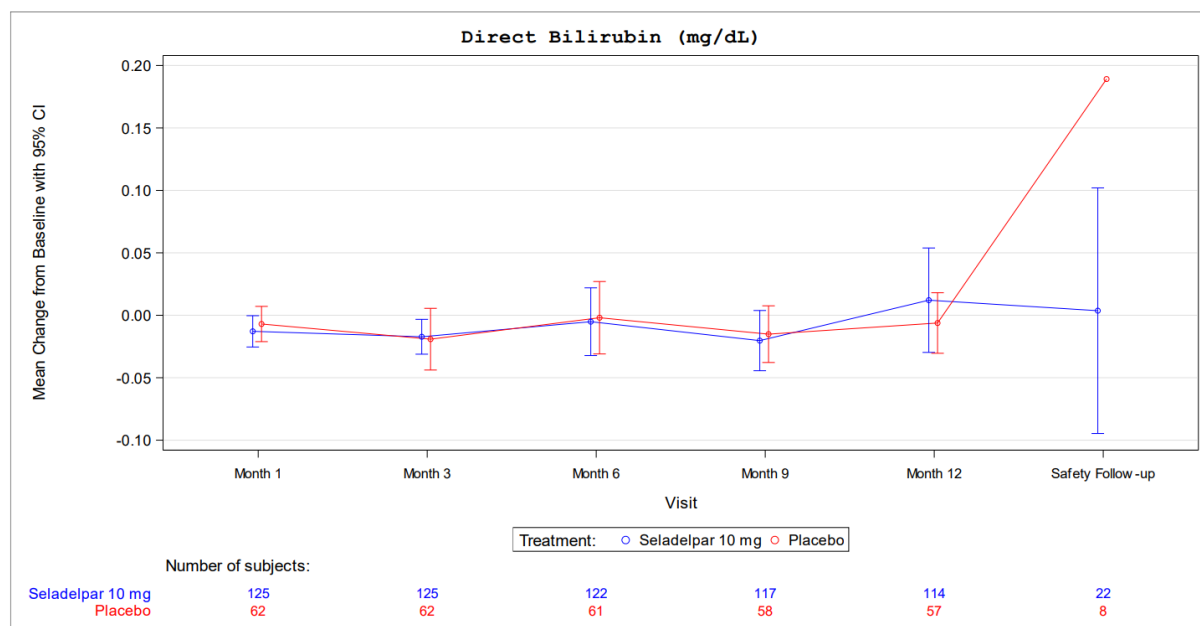


Abbildung 5: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des direkten Bilirubins aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

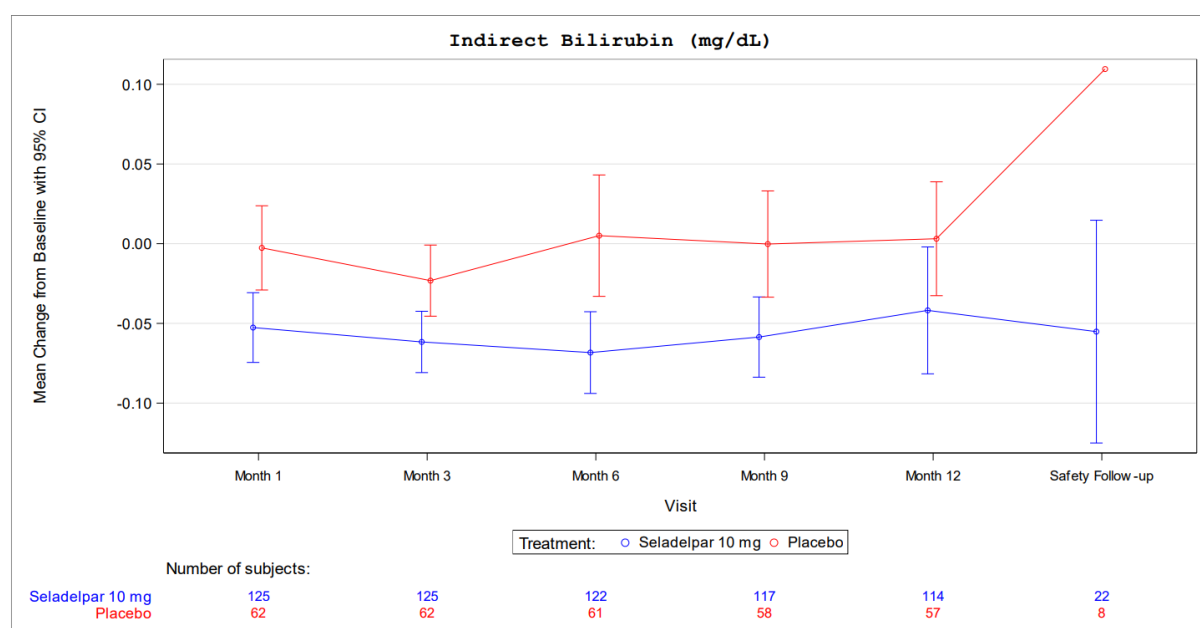


Abbildung 6: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des indirekten Bilirubins aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

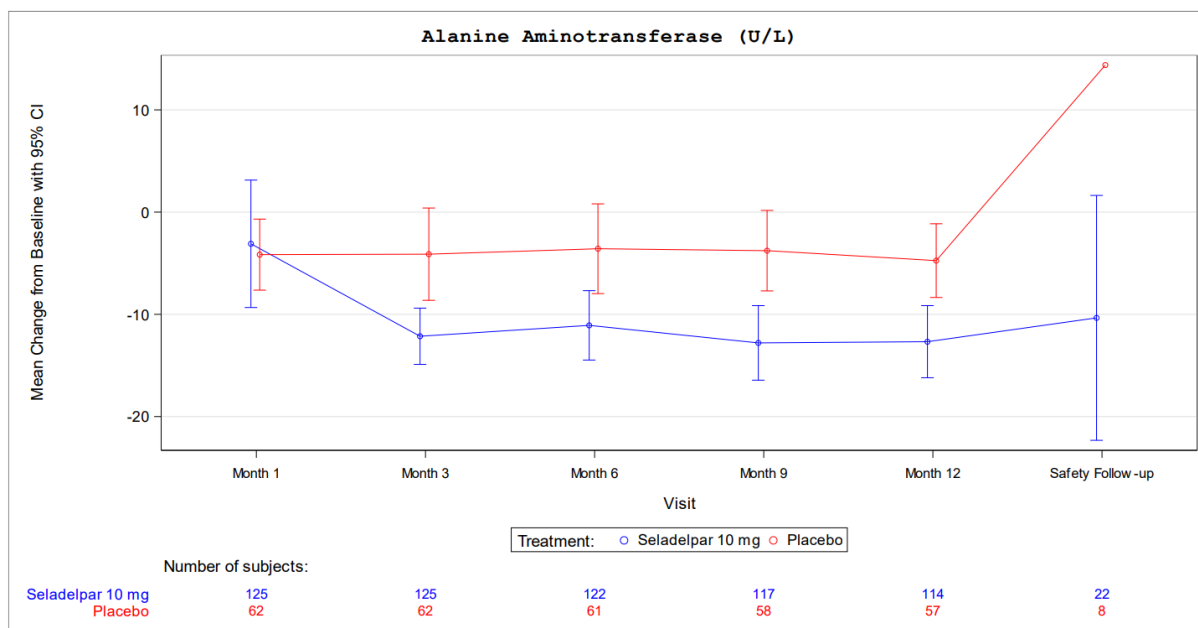


Abbildung 7: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der ALT aus RCT mit dem zbAM (Studie RESPONSE, FAS)

ALT=Alanin-Aminotransferase, CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

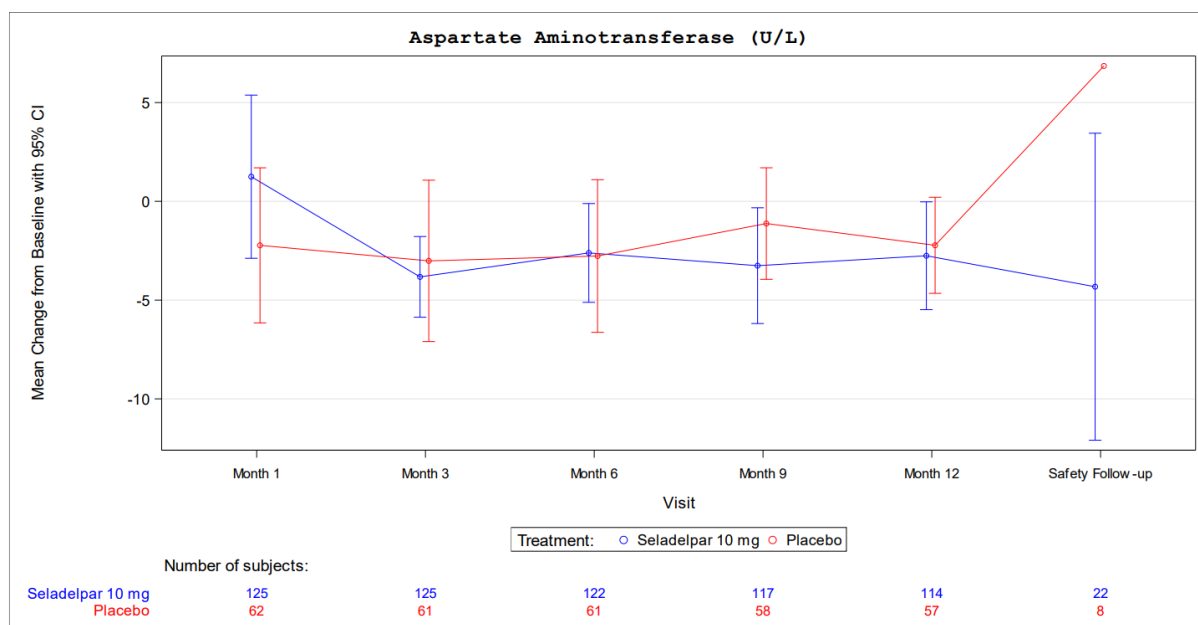


Abbildung 8: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der AST aus RCT mit dem zbAM (Studie RESPONSE, FAS)

AST=Aspartat-Aminotransferase, CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

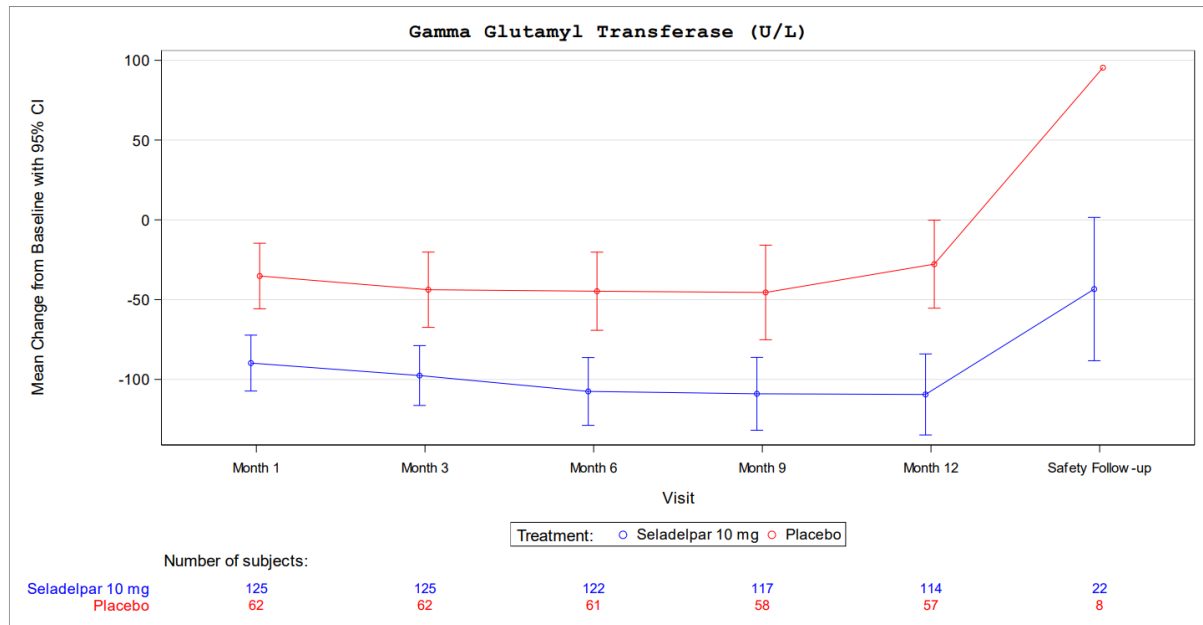


Abbildung 9: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der GGT aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, GGT= Gamma-Glutamyltransferase, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

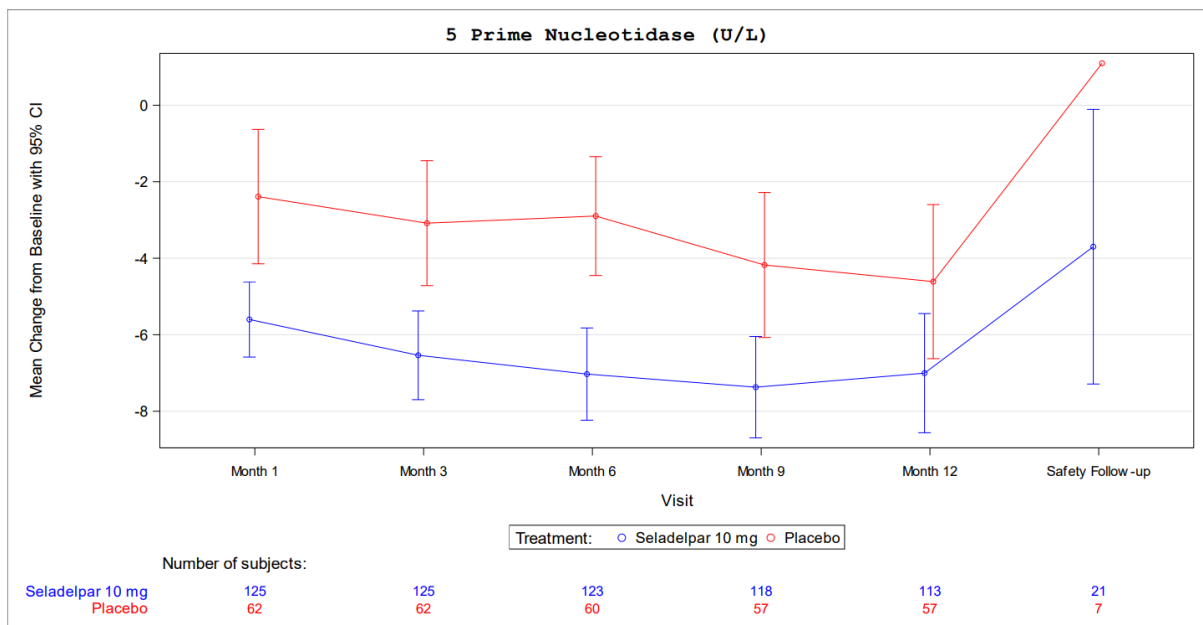


Abbildung 10: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der 5'-Nucleotidase aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

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

durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.3 Pruritus NRS – RCT

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

Tabelle 4-21: Operationalisierung von Pruritus NRS

Studie	Operationalisierung
RESPONSE	Die Pruritus NRS wird anhand einer 11-Punkte-Skala durch den Patienten selbst bewertet von 0 („kein Pruritus“) bis 10 („schwerster vorstellbarer Pruritus“). Dargestellt werden Anzahl und Anteil der Patienten mit einer Verringerung um ≥ 4 Punkte gegenüber Baseline. Hierzu werden die binären Effektmaße RR, OR und RD mit zugehörigem p-Wert angegeben. Außerdem werden die Ergebnisse der stetigen Analyse zur Pruritus NRS dargestellt. Hierfür werden die Änderungen gegenüber Baseline anhand der LS-MWD mit zugehörigem p-Wert sowie der SMD zu den Monaten 6 und 12 präsentiert.
LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, NRS=Numerical Rating Scale, OR=Odds Ratio, RD=Risikodifferenz, RR=Relatives Risiko, 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 4-22: Bewertung des Verzerrungspotenzials für Pruritus NRS in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, NRS=Numerical Rating Scale, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

Tabelle 4-23: Rücklaufquoten für Pruritus NRS aus RCT mit zbAM (Studie RESPONSE, FAS)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	124 (96,9)	65	63 (96,9)
Monat 6	128	121 (94,5)	65	61 (93,8)
Monat 9	128	108 (84,4)	65	56 (86,2)
Monat 12	128	105 (82,0)	65	48 (73,9)
FAS=Full-Analysis-Set, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, NRS=Numerical Rating Scale, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel				

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

Tabelle 4-24: Ergebnisse für Pruritus NRS aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert RD [95%-KI], p-Wert
Monat 6			
Verringerung um ≥ 4 Punkte gegenüber Baseline	14/128 (10,9)	4/65 (6,2)	1,65 [0,61; 4,47], 0,3280 1,90 [0,55; 6,59], 0,3115 0,04 [-0,03; 0,11], 0,2727
Monat 12			
Verringerung um ≥ 4 Punkte gegenüber Baseline	15/128 (11,7)	2/65 (3,1)	3,50 [0,88; 14,01], 0,0764 4,64 [0,96; 22,48], 0,0563 0,08 [0,01; 0,14], 0,0169
Ein RR >1 bedeutet einen Vorteil für die Behandlung mit SEL. FAS=Full-Analysis-Set, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, NRS=Numerical Rating Scale, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SEL=Seladelpar, zbAM=Zu bewertendes Arzneimittel			

Hinsichtlich der Pruritus NRS zeigte sich zu Monat 6 ein höheres Ansprechen im SEL-Arm gegenüber dem Placebo-Arm. Eine Verringerung um ≥ 4 Punkte gegenüber Baseline erreichten 10,9% der Patienten im SEL-Arm und 6,2% der Patienten im Placebo-Arm. Zu Monat 12 zeigte sich ebenfalls ein numerischer Vorteil zugunsten von SEL (11,7% der Patienten im SEL-Arm und 3,1% der Patienten im Placebo-Arm).

Insgesamt ist zu betonen, dass für viele Patienten sowohl im SEL- als auch im Placebo-Arm eine Verringerung um ≥ 4 Punkte gegenüber Baseline nicht möglich war, da diese im Mittel eine Pruritus NRS von 3,0 zu Baseline aufwiesen (Tabelle 4-10). Entsprechend wird die nachfolgende stetige Analyse als relevanter erachtet:

Tabelle 4-25: Ergebnisse für Pruritus NRS Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 6						
127	3,03 (2,81)	-1,33 (0,14)	64	3,02 (2,96)	-0,42 (0,19)	-0,90 [-1,35; -0,45], 0,0001 -0,58 [-0,89; -0,28]

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 12						
127	3,03 (2,81)	-1,35 (0,16)	64	3,02 (2,96)	-0,52 (0,22)	-0,83 [-1,35; -0,31], 0,0018 -0,47 [-0,78; -0,17]

Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit SEL.

a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.

FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, NRS=Numerical Rating Scale, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, zbAM=Zu bewertendes Arzneimittel

Sowohl zu Monat 6 (LS-MWD [95%-KI]: -0,90 [-1,35; -0,45]) als auch zu Monat 12 (LS-MWD [95%-KI]: -0,83 [-1,35; -0,31]) zeigte sich ein stärkerer mittlerer Abfall der Pruritus NRS im SEL-Arm als im Placebo-Arm. Der Vorteil ist sowohl zu Monat 6 (p=0,0001) als auch zu Monat 12 (p=0,0018) statistisch signifikant und zu Monat 6 zusätzlich klinisch relevant (SMD [95%-KI]: -0,58 [-0,89; -0,28]). Der zeitliche Verlauf der mittleren Veränderung der Pruritus NRS ist in Abbildung 11 dargestellt.

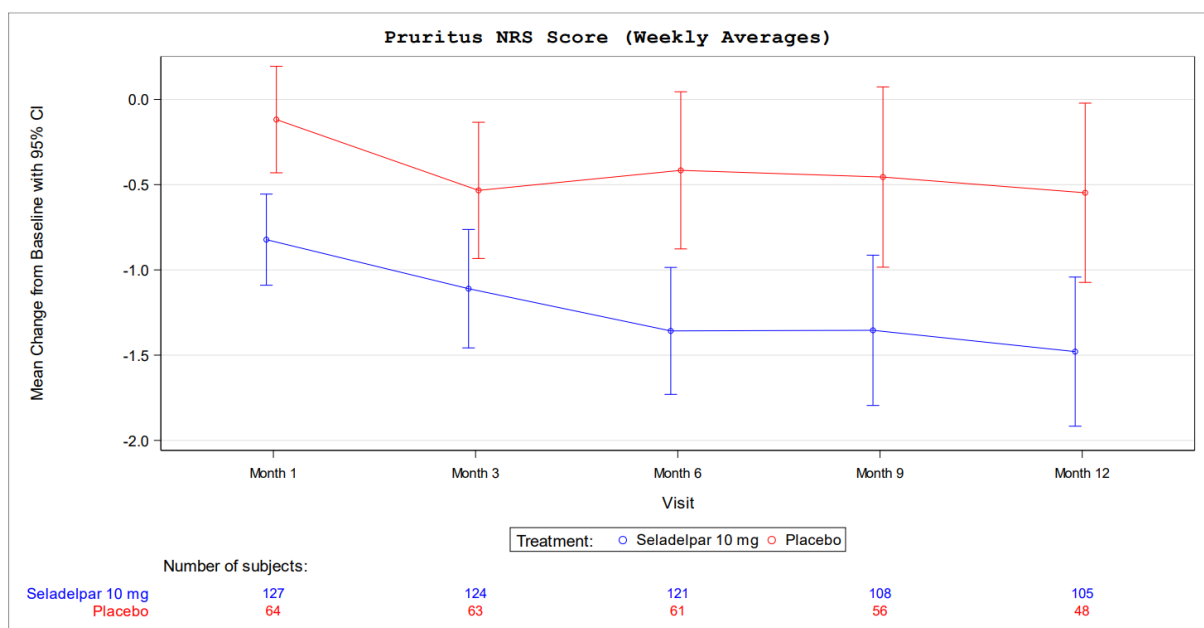


Abbildung 11: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Pruritus NRS aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, NRS=Numerical Rating Scale, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

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

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.4 5-D Pruritus – RCT

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

Tabelle 4-26: Operationalisierung von 5-D Pruritus

Studie	Operationalisierung
RESPONSE	Der 5-D Pruritus besteht aus 27 Items in fünf Subdomänen (Grad, Dauer, Beeinträchtigung, Verlauf und Lokalisation) als Ausprägungen der Morbidität. Die Werte der einzelnen Items reichen von 1 bis 5. Der Bereich Beeinträchtigung besteht aus vier Items, die die Auswirkungen des Pruritus auf das tägliche Leben bewerten: Schlaf, Freizeit/Soziales, Hausarbeit/Besorgungen und Arbeit/Schule. Für den Bereich Lokalisation wird die Anzahl der betroffenen Körperteile mit einer möglichen Summe von 0 bis 16 ermittelt. Die Punkte für jede der fünf Subdomänen werden zur Ermittlung eines 5-D Gesamt-Scores aufsummiert. Der 5-D Gesamt-Score kann von 5 (kein Pruritus) bis 25 (schwerster Pruritus) reichen. Ergänzend wurde ein modifizierter Gesamt-Score basierend auf vier Subdomänen (Grad, Dauer, Beeinträchtigung und Lokalisation) berechnet. Diese Anpassung erfolgte, da es zu Missverständnissen der Bewertung der Subdomäne Verlauf kam. Patienten, die keinen Pruritus hatten, gaben oft „unverändert“ an, was einem Wert von 4 entspricht, und nicht „ganz verschwunden“, was einem Wert von 1 entspricht. Hierzu werden für den Gesamt-Score sowie die Subdomänen und den modifizierten Gesamt-Score die Ergebnisse zur kontinuierlichen Analyse mithilfe der stetigen Effektmaße LS-MWD und SMD mit zugehörigem p-Wert zu den Monaten 6 und 12 angegeben.
LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, 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 4-27: Bewertung des Verzerrungspotenzials für 5-D Pruritus in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

Tabelle 4-28: Rücklaufquoten für 5-D Pruritus aus RCT mit zbAM (Studie RESPONSE, FAS)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
Gesamt-Score				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,9)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)
Subdomäne Grad				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,8)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
Subdomäne Dauer				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,8)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)
Subdomäne Beeinträchtigung				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,8)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)
Subdomäne Verlauf				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,8)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)
Subdomäne Lokalisation				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,9)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)
Modifizierter Gesamt-Score				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	127 (99,2)	65	64 (98,5)
Monat 3	128	126 (98,4)	65	64 (98,5)
Monat 6	128	123 (96,1)	65	61 (93,9)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
Monat 9	128	115 (89,8)	65	57 (87,7)
Monat 12	128	96 (75,0)	65	46 (70,8)
N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel				

Stellen Sie die Ergebnisse für den Endpunkt 5-D Pruritus für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-29: Ergebnisse für 5-D Pruritus Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 6						
Gesamt-Score						
128	11,63 (4,85)	-2,04 (0,30)	64	11,23 (4,65)	-0,27 (0,41)	-1,76 [-2,72; -0,81], 0,0003 -0,53 [-0,83; -0,22]
Subdomäne Grad						
128	2,26 (1,04)	-0,38 (0,06)	64	2,24 (0,99)	-0,02 (0,09)	-0,35 [-0,56; -0,15], 0,0008 -0,49 [-0,80; -0,19]
Subdomäne Dauer						
128	1,59 (1,11)	-0,28 (0,06)	64	1,56 (1,08)	-0,10 (0,09)	-0,18 [-0,39; 0,02], 0,0731 -0,26 [-0,56; 0,04]
Subdomäne Beeinträchtigung						
128	2,27 (1,37)	-0,52 (0,08)	64	2,22 (1,24)	-0,05 (0,11)	-0,47 [-0,73; -0,20], 0,0006 -0,50 [-0,81; -0,20]
Subdomäne Verlauf						
128	3,35 (1,09)	-0,62 (0,11)	64	3,13 (1,16)	-0,31 (0,15)	-0,31 [-0,68; 0,06], 0,0981 -0,25 [-0,55; 0,06]
Subdomäne Lokalisation						
128	2,16 (1,29)	-0,18 (0,08)	64	2,08 (1,20)	0,23 (0,12)	-0,41 [-0,68; -0,14], 0,0034 -0,43 [-0,73; -0,13]
Modifizierter Gesamt-Score						
128	8,28 (4,24)	-1,40 (0,23)	64	8,10 (3,96)	0,01 (0,31)	-1,41 [-2,14; -0,68], 0,0002 -0,55 [-0,86; -0,25]

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 12						
Gesamt-Score						
128	11,63 (4,85)	-2,26 (0,33)	64	11,23 (4,65)	0,05 (0,46)	-2,31 [-3,40; -1,23], <0,0001 -0,62 [-0,93; -0,31]
Subdomäne Grad						
128	2,26 (1,04)	-0,42 (0,07)	64	2,24 (0,99)	-0,05 (0,10)	-0,38 [-0,61; -0,15], 0,0016 -0,47 [-0,77; -0,16]
Subdomäne Dauer						
128	1,59 (1,11)	-0,32 (0,07)	64	1,56 (1,08)	-0,08 (0,10)	-0,40 [-0,63; -0,16], 0,0011 -0,49 [-0,79; -0,19]
Subdomäne Beeinträchtigung						
128	2,27 (1,37)	-0,55 (0,09)	64	2,22 (1,24)	0,05 (0,13)	-0,60 [-0,91; -0,29], 0,0002 -0,57 [-0,87; -0,26]
Subdomäne Verlauf						
128	3,35 (1,09)	-0,64 (0,13)	64	3,13 (1,16)	-0,17 (0,18)	-0,47 [-0,90; -0,04], 0,0340 -0,32 [-0,62; -0,02]
Subdomäne Lokalisation						
128	2,16 (1,29)	-0,25 (0,09)	64	2,08 (1,20)	0,14 (0,12)	-0,38 [-0,67; -0,10], 0,0085 -0,39 [-0,69; -0,09]
Modifizierter Gesamt-Score						
128	8,28 (4,24)	-1,59 (0,25)	64	8,10 (3,96)	0,19 (0,35)	-1,79 [-2,60; -0,97], <0,0001 -0,63 [-0,94; -0,33]
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit SEL. a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline. FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinst-Quadrate-Mittelwert, LS-MWD= Kleinst-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, zbAM=Zu bewertendes Arzneimittel						

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus im Gesamt-Score (Monat 6: LS-MWD [95%-KI]: -1,76 [-2,72; -0,81]; Monat 12: LS-MWD [95%-KI]: -2,31 [-3,40; -1,23]). Der Unterschied zwischen SEL und Placebo war zu beiden Zeitpunkten statistisch signifikant ($p=0,0003$ und $p<0,0001$) und klinisch relevant (Monat 6: SMD [95%-KI]: -0,53 [-0,83; -0,22]; Monat 12: SMD [95%-KI]: -0,62 [-0,93; -0,31]) zugunsten von SEL.

Die Vorteile einer Behandlung mit SEL gegenüber Placebo zeigten sich in allen Subdomänen sowie im modifizierten Gesamt-Score zu allen Zeitpunkten:

Zu beiden Zeitpunkten führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus in der Subdomäne Grad (Monat 6: LS-MWD [95%-KI]: -0,35 [-0,56; -0,15]; Monat 12: LS-MWD [95%-KI]: -0,38 [-0,61; -0,15]). Der Vorteil der Behandlung mit SEL gegenüber Placebo war zu beiden Zeitpunkten statistisch signifikant ($p=0,0008$ und $p=0,0016$).

Auch in der Subdomäne Dauer führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall zu beiden berichteten Zeitpunkten (Monat 6: LS-MWD [95%-KI]: -0,18 [-0,39; -0,02]; Monat 12: LS-MWD [95%-KI]: -0,40 [-0,63; -0,16]). Der Unterschied zwischen SEL und Placebo war zu Monat 12 statistisch signifikant ($p=0,0011$).

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus in der Subdomäne Beeinträchtigung (Monat 6: LS-MWD [95%-KI]: -0,47 [-0,73; -0,20]; Monat 12: LS-MWD [95%-KI]: -0,60 [-0,91; -0,29]). Der Vorteil für die Behandlung mit SEL gegenüber Placebo war zu beiden Zeitpunkten sowohl statistisch signifikant ($p=0,0006$ und $p=0,0002$) als auch klinisch relevant (Monat 6: SMD [95%-KI]: -0,50 [-0,81; -0,20]; Monat 12: SMD [95%-KI]: -0,57 [-0,87; -0,26]).

Auch in der Subdomäne Verlauf führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall zu beiden Zeitpunkten (Monat 6: LS-MWD [95%-KI]: -0,31 [-0,68; -0,06]; Monat 12: LS-MWD [95%-KI]: -0,47 [-0,90; -0,04]). Der Unterschied zwischen SEL und Placebo war zu Monat 12 statistisch signifikant ($p=0,0340$).

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus in der Subdomäne Lokalisation (Monat 6: LS-MWD [95%-KI]: -0,41 [-0,68; -0,14]; Monat 12: LS-MWD [95%-KI]: -0,38 [-0,67; -0,10]). Der Unterschied zwischen SEL und Placebo war zu beiden Zeitpunkten statistisch signifikant ($p=0,0034$ und $p=0,0085$).

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus im modifizierten Gesamt-Score (Monat 6: LS-MWD [95%-KI]: -1,41 [-2,14; -0,68]; Monat 12: LS-MWD [95%-KI]: -1,79 [-2,60; -0,97]). Der Unterschied zwischen SEL und Placebo war zu beiden Zeitpunkten statistisch signifikant ($p=0,0002$ und $p<0,0001$) und klinisch relevant (Monat 6: SMD [95%-KI]: -0,55 [-0,86; -0,25]; Monat 12: SMD [95%-KI]: -0,63 [-0,94; -0,33]) zugunsten von SEL.

Die jeweiligen zeitlichen Verläufe der mittleren Veränderung sind in Abbildung 12 bis Abbildung 18 dargestellt.

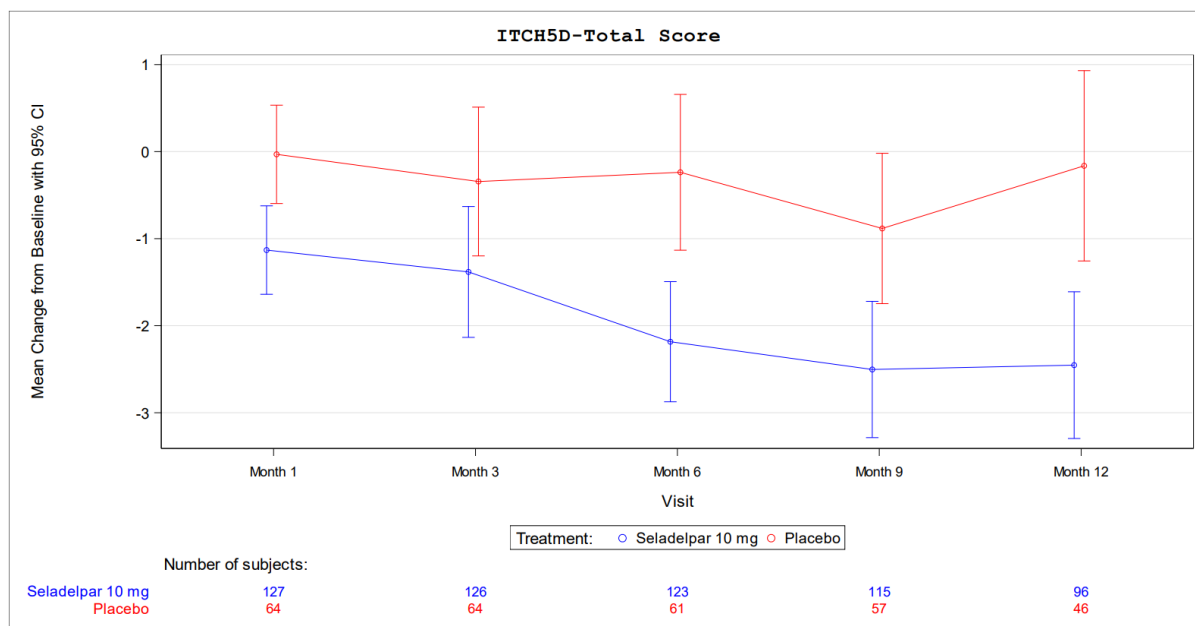


Abbildung 12: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des Gesamt-Scores des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

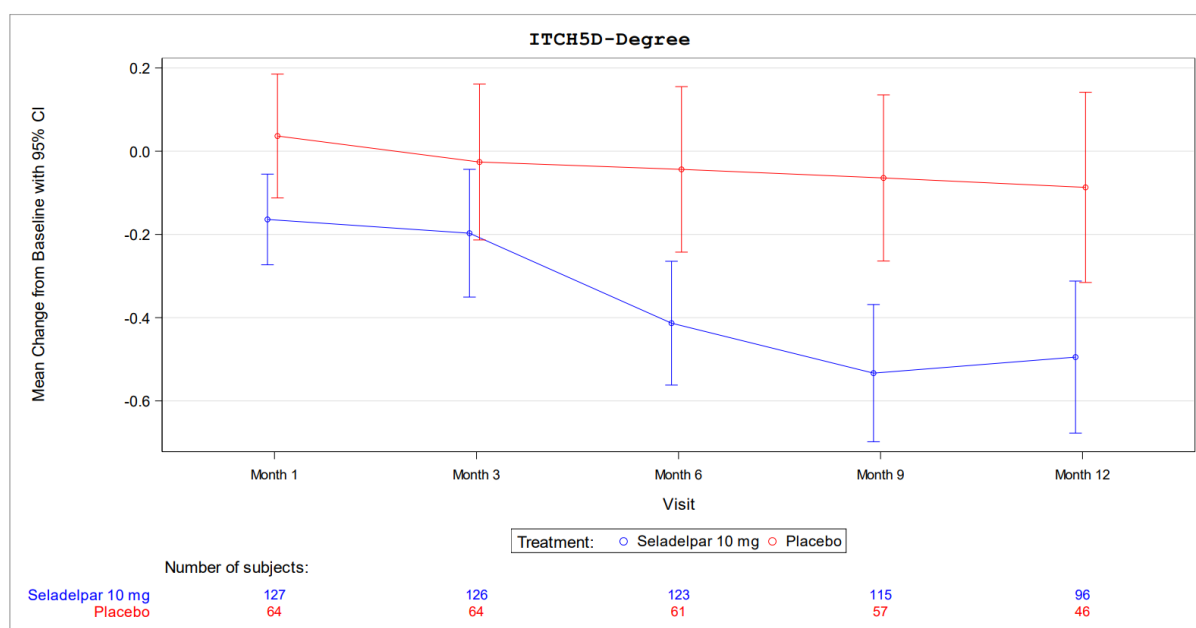


Abbildung 13: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Grad des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

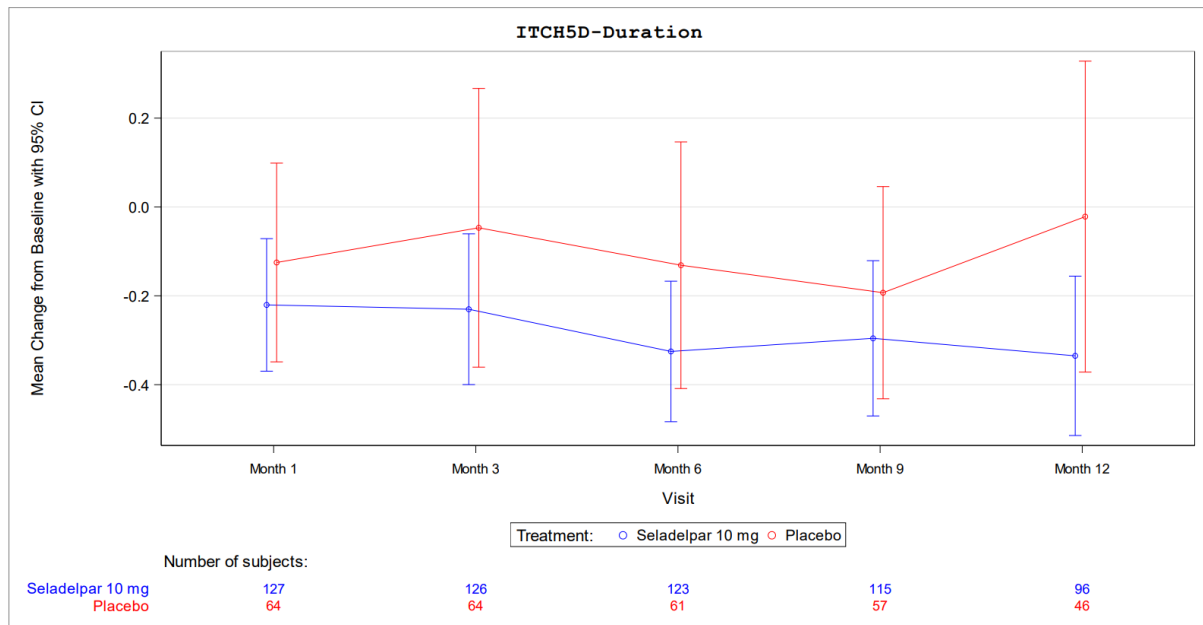


Abbildung 14: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Dauer des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

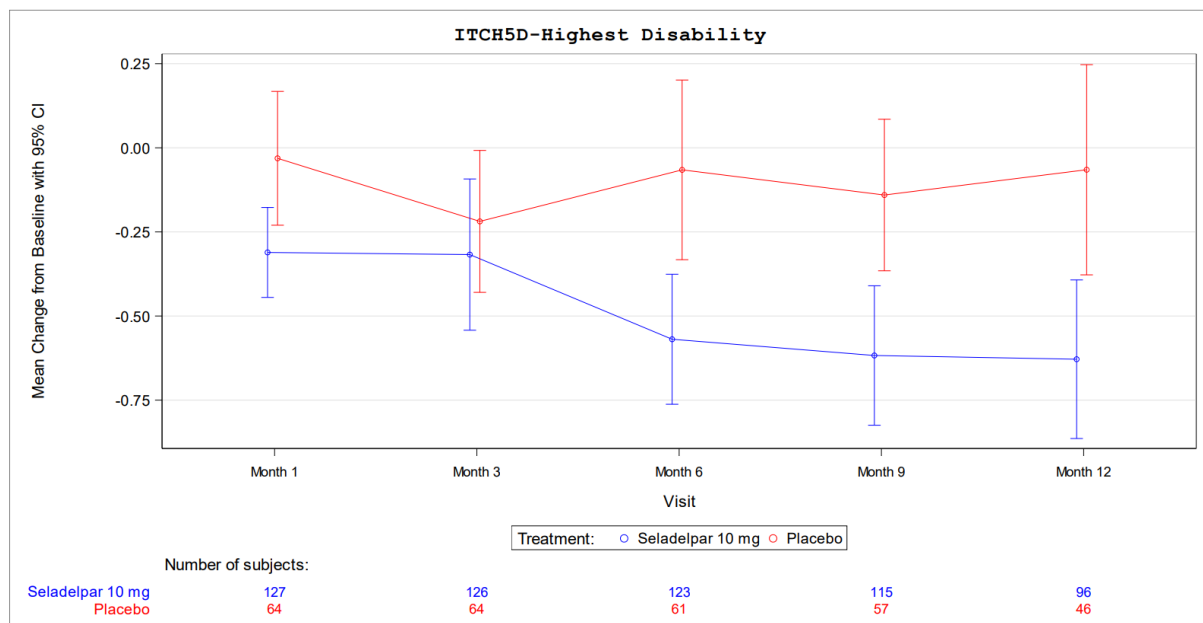


Abbildung 15: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Beeinträchtigung des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

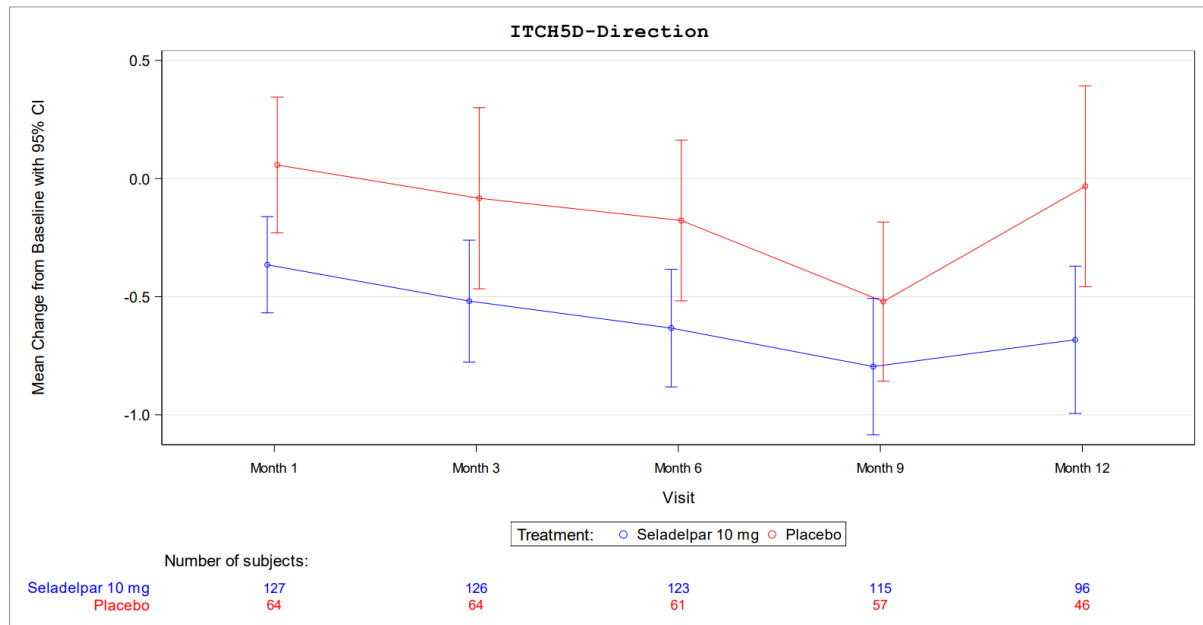


Abbildung 16: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Verlauf des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

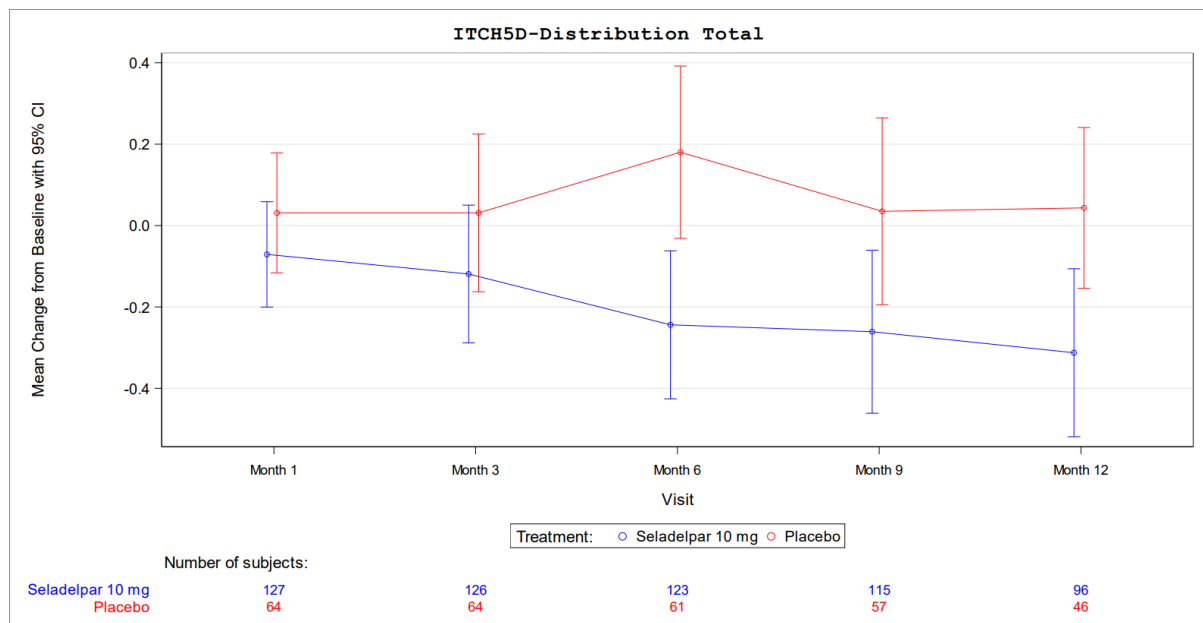


Abbildung 17: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Subdomäne Lokalisation des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

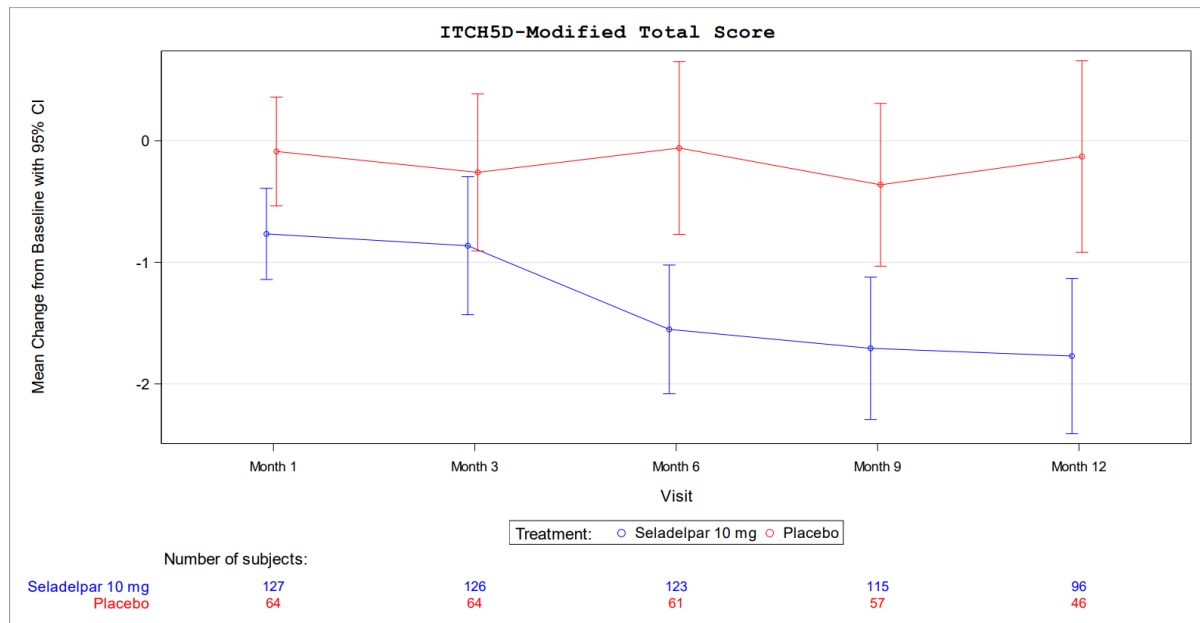


Abbildung 18: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des modifizierten Gesamt-Scores des 5-D Pruritus aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.5 PGI-C & PGI-S – RCT

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

Tabelle 4-30: Operationalisierung der Bewertungsskalen PGI-C & PGI-S

Studie	Operationalisierung
RESPONSE	Die Fragebögen PGI-S & PGI-C sind patientenberichtete Fragebögen, die beide aus nur einem zu berichteten Item bestehen. Der PGI-S dient der Bewertung der Schwere der Erkrankung und der aktuellen Symptomatik. Die Beurteilung erfolgt basierend auf einer fünfstufigen Skala von Kategorie 1 (keine Symptome) bis 5 (sehr starke Symptome) in Bezug auf die vergangene Woche. Der PGI-C dient der Bewertung der Veränderung der Krankheitslast und Symptomatik unter der Behandlung für eine spezifische Erkrankung. Die Frage ist auf einer siebenstufigen Skala von Kategorie 1 (sehr große Verbesserung) bis 7 (sehr große Verschlechterung) zu beurteilen. Für beide Fragebögen werden die Ergebnisse zur kontinuierlichen Analyse mithilfe der stetigen Effektmaße LS-MWD und SMD mit zugehörigem p-Wert zu den Monaten 6 und 12 angegeben.
LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, 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 4-31: Bewertung des Verzerrungspotenzials für PGI-C & PGI-S in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

Tabelle 4-32: Rücklaufquoten für PGI aus RCT mit zbAM (Studie RESPONSE, FAS)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
PGI-S				
Baseline	128	128 (100,0)	65	64 (98,5)
Monat 1	128	116 (90,6)	65	61 (93,9)
Monat 3	128	116 (90,6)	65	57 (87,7)
Monat 6	128	110 (85,9)	65	53 (81,5)
Monat 9	128	111 (86,7)	65	53 (81,5)
Monat 12	128	94 (73,4)	65	51 (78,5)
PGI-C				
Baseline	128	116 (90,6)	65	61 (93,9)
Monat 1	128	115 (89,8)	65	57 (87,7)
Monat 3	128	105 (82,0)	65	52 (80,0)
Monat 6	128	111 (86,7)	65	53 (81,5)
Monat 9	128	94 (73,4)	65	51 (78,5)
FAS=Full-Analysis-Set, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel				

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

Tabelle 4-33: Ergebnisse für PGI – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 6						
PGI-S						
125	2,13 (0,98)	-0,37 (0,06)	62	2,16 (0,95)	0,04 (0,08)	-0,42 [-0,61; -0,23], <0,0001 -0,64 [-0,95; -0,32]
PGI-C						
124	- ^b	2,61 (0,13)	63	- ^b	3,07 (0,18)	-0,45 [-0,88; -0,02], 0,0391 -0,31 [-0,61, -0,01]

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 12						
PGI-S						
125	2,13 (0,98)	-0,36 (0,07)	62	2,16 (0,95)	-0,26 (0,09)	-0,10 [-0,32; 0,11], 0,3453 -0,14 [-0,44; 0,17]
PGI-C						
124	- ^b	2,60 (0,14)	63	- ^b	3,01 (0,18)	-0,41 [-0,85; 0,03], 0,0673 -0,27 [-0,58; 0,03]
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit SEL. a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline. b: Es wird kein Baseline-Wert im Rahmen des PGI-C erhoben. FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, zbAM=Zu bewertendes Arzneimittel						

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des PGI-S (Monat 6: LS-MWD [95%-KI]: -0,42 [-0,61; -0,23]; Monat 12: LS-MWD [95%-KI]: -0,10 [-0,32; 0,11]). Der Unterschied zwischen SEL und Placebo war zu Monat 6 statistisch signifikant ($p < 0,0001$) und klinisch relevant SMD [95%-KI]: -0,64 [-0,95; -0,32]. Der zeitliche Verlauf der mittleren Veränderung des PGI-S ist in Abbildung 19 dargestellt.

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des PGI-S (Monat 6: LS-MWD [95%-KI]: -0,45 [-0,88; -0,02]; Monat 12: LS-MWD [95%-KI]: -0,41 [-0,85; 0,03]). Der Unterschied zwischen SEL und Placebo war zu Monat 6 statistisch signifikant ($p = 0,0391$). Der zeitliche Verlauf der mittleren Veränderung des PGI-C ist in Abbildung 20 dargestellt.

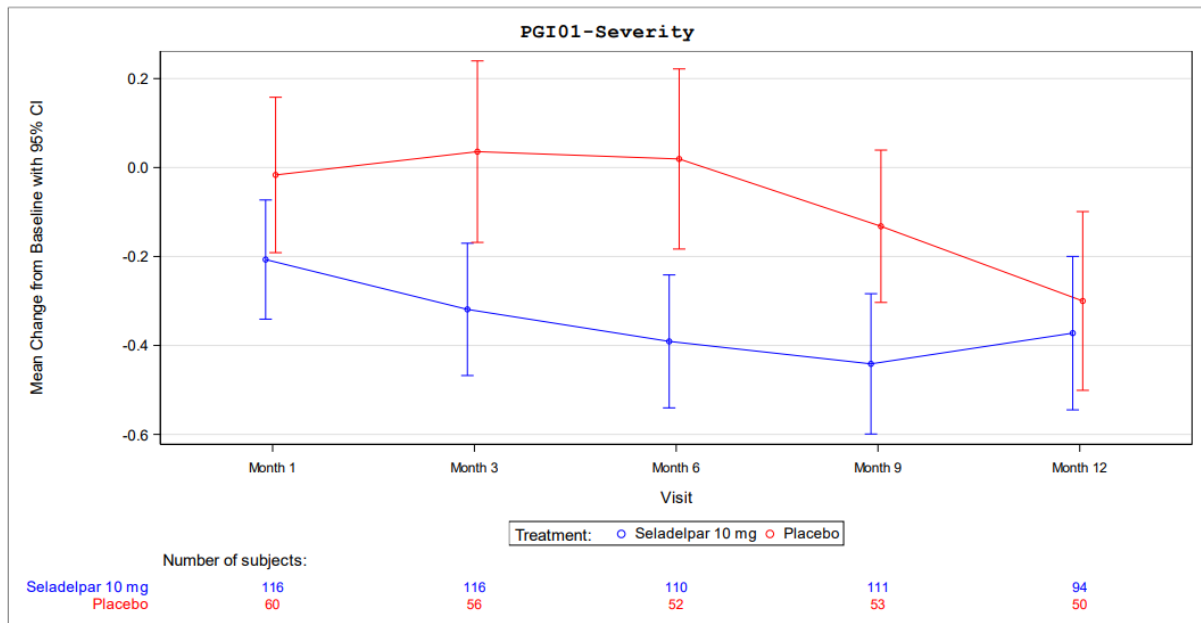


Abbildung 19: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des PGI-S aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PGI-S=Patient Global Impression of Severity, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

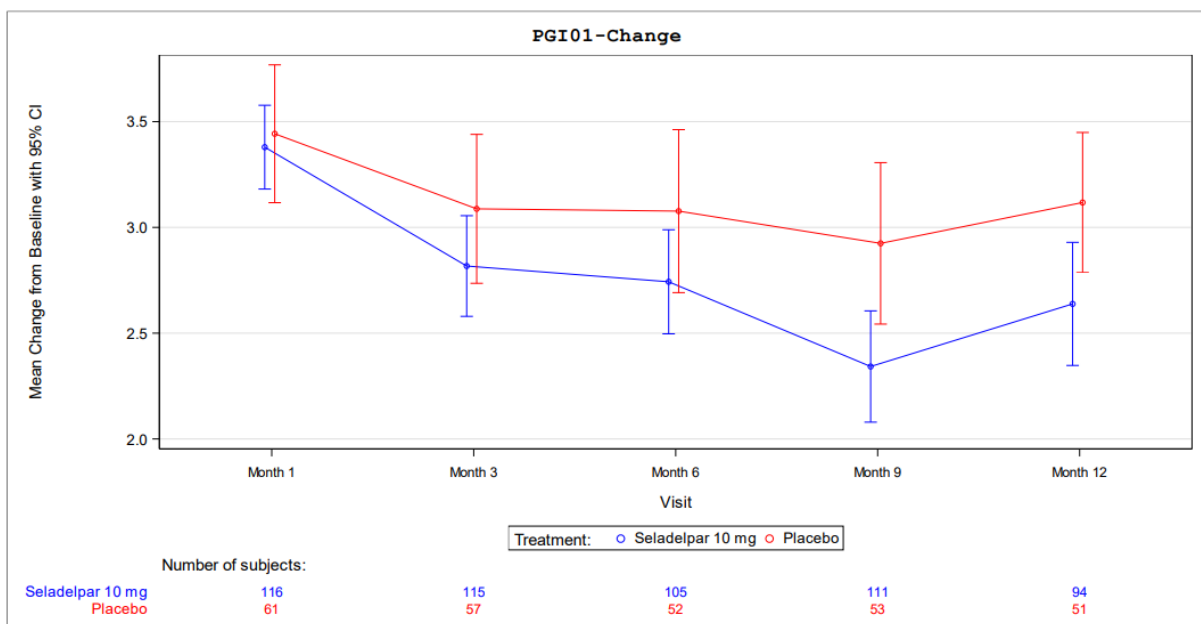


Abbildung 20: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des PGI-C aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PGI-C=Patient Global Impression of Change, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

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

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.6 Fibrose – RCT

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

Tabelle 4-34: Operationalisierung von Fibrose

Studie	Operationalisierung
RESPONSE	Die Fibrose wird bestimmt mittels ELF-Score und Lebersteifigkeit auf Basis des FibroScan®. Die Änderung gegenüber Baseline wird durch die Effektmaße LS-MWD und SMD mit zugehörigem p-Wert zu den Monaten 6 und 12 angegeben.
ELF=Enhanced Liver Fibrosis, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, 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 4-35: Bewertung des Verzerrungspotenzials für Fibrose in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

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

Tabelle 4-36: Ergebnisse für Fibrose – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 6						
<i>Veränderung Lebersteifigkeit (FibroScan®)</i>						
109	9,84 (6,16)	0,84 (0,50)	59	8,74 (4,18)	2,00 (0,65)	-1,15 [-2,67; 0,36], 0,1354 -0,22 [-0,54; 0,09]
<i>Veränderung erweiterter Leberfibrosetest (ELF-Score)</i>						
126	10,16 (1,03)	0,16 (0,04)	59	10,23 (0,85)	0,06 (0,06)	0,10 [-0,03; 0,24], 0,1264 0,23 [-0,08; 0,53]
Monat 12						
<i>Veränderung Lebersteifigkeit (FibroScan®)</i>						
109	9,84 (6,16)	0,56 (0,56)	59	8,74 (4,18)	1,92 (0,72)	-1,37 [-3,09; 0,35], 0,1174 -0,24 [-0,55; 0,08]
<i>Veränderung erweiterter Leberfibrosetest (ELF-Score)</i>						
126	10,16 (1,03)	0,10 (0,06)	63	10,23 (0,85)	0,02 (0,08)	0,08 [-0,11; 0,26], 0,4215 0,12 [-0,18; 0,42]
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit SEL.						
a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.						
ELF=Enhanced Liver Fibrosis, FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrat-Mittelwert, LS-MWD=Kleinste-Quadrat-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, zbAM=Zu bewertendes Arzneimittel						

Die Veränderung der Lebersteifigkeit (FibroScan®) ist unter SEL und Placebo vergleichbar, sowohl zu Monat 6 als auch zu Monat 12. Dies wird auch durch die Ergebnisse des erweiterten

Leberfibrosetests (ELF-Score) bestätigt. Die jeweiligen zeitlichen Verläufe sind in Abbildung 21 und Abbildung 22 dargestellt.

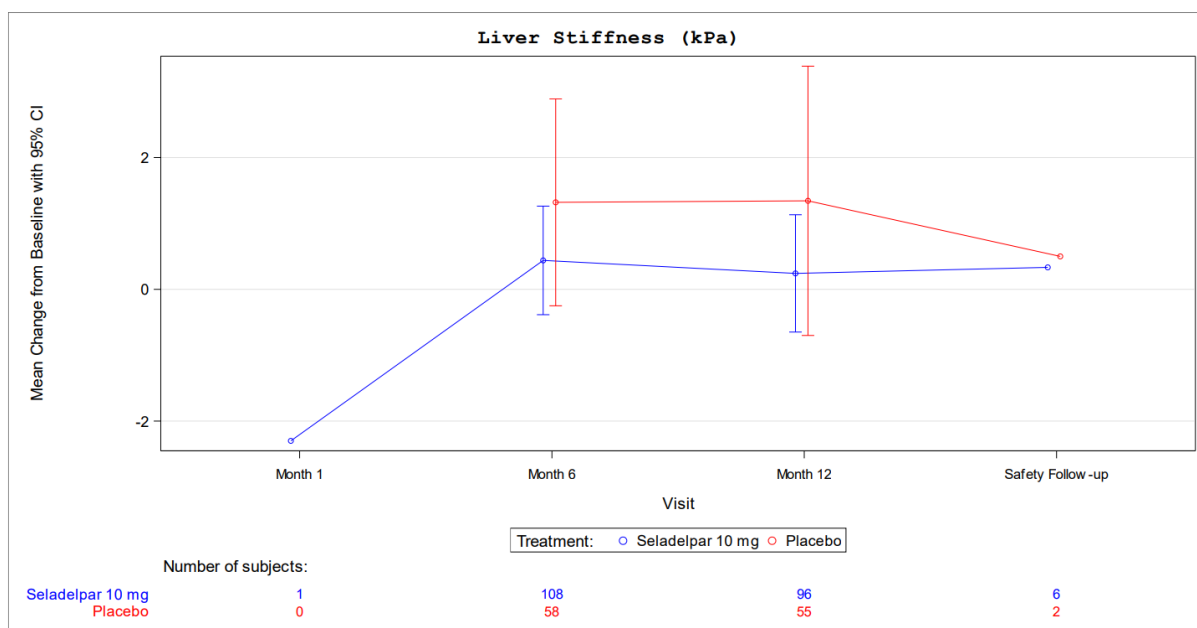


Abbildung 21: Darstellung des zeitlichen Verlaufs der mittleren Veränderung der Lebersteifigkeit (FibroScan®) aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

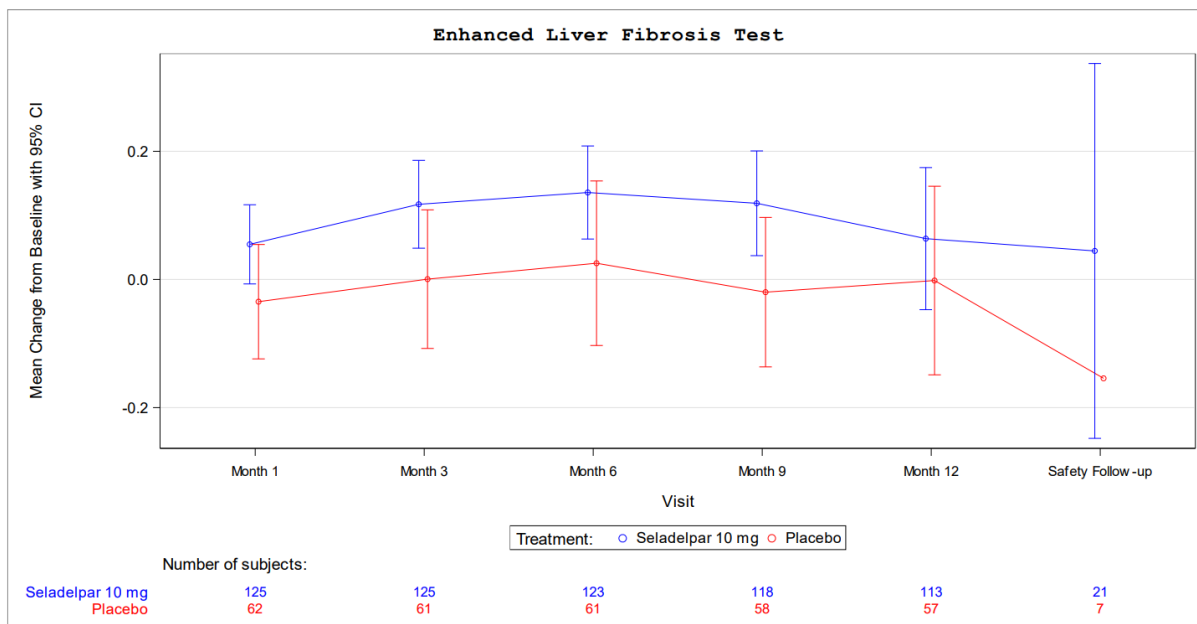


Abbildung 22: Darstellung des zeitlichen Verlaufs der mittleren Veränderung des erweiterten Leberfibrosetests (ELF-Score) aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.7 Hospitalisierung – RCT

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

Tabelle 4-37: Operationalisierung von Hospitalisierung

Studie	Operationalisierung
RESPONSE	Es wird die Anzahl der Hospitalisierungen aufgrund der Erkrankung der Patienten gemessen. Dargestellt werden Anzahl und Anteil der Patienten mit Hospitalisierungen zu Monat 12, sowie das RR, OR und RD mit zugehörigem p-Wert.
OR=Odds Ratio, RD=Risikodifferenz, RR=Relatives Risiko	

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

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

Tabelle 4-38: Bewertung des Verzerrungspotenzials für Hospitalisierung in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

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

Tabelle 4-39: Ergebnisse für Hospitalisierung aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert RD [95%-KI], p-Wert
Hospitalisierungen aufgrund der Erkrankung	6/128 (4,7)	3/65 (4,6)	1,06 [0,27; 4,17], 0,9384 1,06 [0,25; 4,47], 0,9379 0,00 [-0,06; 0,06], 0,9373
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. FAS=Full-Analysis-Set, KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, zbAM=Zu bewertendes Arzneimittel			

Der Anteil der Patienten, die aufgrund ihrer Erkrankung hospitalisiert wurden, ist mit 4,7% unter SEL und 4,6% unter Placebo in beiden Armen vergleichbar. Entsprechend zeigt sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen.

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.8 Gesundheitsbezogene Lebensqualität PBC-40 – RCT

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

Tabelle 4-40: Operationalisierung von PBC-40

Studie	Operationalisierung
RESPONSE	Der PBC-40 ist in sechs Domänen unterteilt. Die Darstellung erfolgt als Einzeldomänen sowie als Gesamt-Score. Den Antworten zu den Fragen werden sogenannte Score-Punkte zugeordnet (1=überhaupt nicht, 2=selten, 3=manchmal, 4=häufig, 5=immer), wobei der Gesamt-Score die Aufsummierung über alle Einzelfragen angibt und somit Werte zwischen 32 und 200 annehmen kann. Die Veränderung wird basierend auf der Änderung gegenüber Baseline durch die Effektmaße LS-MWD und SMD mit zugehörigem p-Wert zu den Monaten 6 und 12 angegeben.
LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, PBC=Primär biliäre Cholangitis, 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 4-41: Bewertung des Verzerrungspotenzials für PBC-40 in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

Tabelle 4-42: Rücklaufquoten für PBC-40 aus RCT mit zbAM (Studie RESPONSE, FAS)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
Gesamt-Score				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
Kognitive Domäne				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
Emotionale Domäne				
Baseline	128	128 (100,0)	65	65 (100,0)

Erhebungszeitraum	Seladelpar		Placebo	
	N	n (%)	N	n (%)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
Fatigue Domäne				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
Pruritus Domäne				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
Soziale Domäne				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
Symptomatische Domäne				
Baseline	128	128 (100,0)	65	65 (100,0)
Monat 1	128	118 (92,2)	65	62 (95,4)
Monat 3	128	121 (94,5)	65	59 (90,8)
Monat 6	128	113 (88,3)	65	53 (81,5)
Monat 9	128	112 (87,5)	65	55 (84,6)
Monat 12	128	94 (73,4)	65	51 (78,5)
FAS=Full-Analysis-Set, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel				

Stellen Sie die Ergebnisse für den Endpunkt PBC-40 für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-43: Ergebnisse für PBC-40 Veränderung zu Baseline aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Monat 6						
Gesamt-Score						
126	91,64 (29,32)	-5,57 (1,34)	63	91,44 (29,45)	-4,65 (1,88)	-0,92 [-5,30; 3,46], 0,6787 -0,06 [-0,36; 0,24]
Kognitive Domäne						
126	13,22 (5,51)	-0,42 (0,29)	63	12,84 (4,88)	0,01 (0,41)	-0,43 [-1,38; 0,52], 0,3750 -0,13 [-0,44; 0,17]
Emotionale Domäne						
126	7,79 (2,89)	-0,68 (0,18)	63	7,72 (3,15)	-0,89 (0,26)	0,21 [-0,40; 0,81], 0,5013 0,10 [-0,20; 0,40]
Fatigue Domäne						
126	27,57 (10,01)	-1,97 (0,57)	63	27,41 (10,64)	-0,78 (0,81)	-1,19 [-3,07; 0,69], 0,2131 -0,18 [-0,49; 0,12]
Pruritus Domäne						
126	5,14 (3,85)	-0,77 (0,23)	63	5,60 (3,97)	0,16 (0,33)	-0,93 [-1,68; -0,18], 0,0159 -0,35 [-0,66; -0,05]
Soziale Domäne						
126	22,85 (8,44)	-0,66 (0,46)	63	22,21 (8,32)	-1,61 (0,65)	0,95 [-0,58; 2,47], 0,2216 0,18 [-0,12; 0,49]
Symptomatische Domäne						
126	15,06 (4,56)	-0,35 (0,24)	63	15,67 (5,46)	-0,52 (0,35)	0,17 [-0,64; 0,97], 0,6827 0,06 [-0,24; 0,36]
Monat 12						
Gesamt-Score						
126	91,64 (29,32)	-5,85 (1,64)	63	91,44 (29,45)	-6,19 (2,23)	0,33 [-4,98; 5,64], 0,9019 0,02 [-0,28; 0,32]
Kognitive Domäne						
126	13,22 (5,51)	-0,55 (0,33)	63	12,84 (4,88)	-0,42 (0,44)	-0,13 [-1,20; 0,94], 0,8106 -0,04 [-0,34; 0,27]

Seladelpar			Placebo			Seladelpar vs. Placebo
N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert SMD [95%-KI]
Emotionale Domäne						
126	7,79 (2,89)	-0,70 (0,20)	63	7,72 (3,15)	-0,95 (0,27)	0,25 [-0,40; 0,90], 0,4488 0,11 [-0,19; 0,41]
Fatigue Domäne						
126	27,57 (10,01)	-1,97 (0,67)	63	27,41 (10,64)	-1,50 (0,92)	-0,47 [-2,66; 1,71], 0,6692 -0,06 [-0,37; 0,24]
Pruritus Domäne						
126	5,14 (3,85)	-1,31 (0,26)	63	5,60 (3,97)	-0,48 (0,36)	-0,83 [-1,68; 0,02], 0,0544 -0,28 [-0,59; 0,02]
Soziale Domäne						
126	22,85 (8,44)	-0,67 (0,47)	63	22,21 (8,32)	-1,72 (0,64)	1,05 [-0,47; 2,56], 0,1735 0,20 [-0,10; 0,50]
Symptomatische Domäne						
126	15,06 (4,56)	-0,10 (0,30)	65	15,67 (5,46)	-0,19 (0,41)	0,09 [-0,87; 1,06], 0,8475 0,03 [-0,27; 0,33]
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit SEL. a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline. FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, zbAM=Zu bewertendes Arzneimittel						

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des PBC-40 in der Pruritus Domäne (Monat 6: LS-MWD [95%-KI]: -0,93 [-1,68; -0,18]; Monat 12: LS-MWD [95%-KI]: -0,83 [-1,68; 0,02]). Der Unterschied zwischen SEL und Placebo war zu Monat 6 signifikant (p=0,0159) jedoch nicht klinisch relevant. Es zeigten sich keine weiteren statistisch signifikanten Unterschiede hinsichtlich des PBC-40 und den entsprechenden Domänen. Die jeweiligen zeitlichen Verläufe der mittleren Veränderung sind in Abbildung 23 bis Abbildung 29 dargestellt.

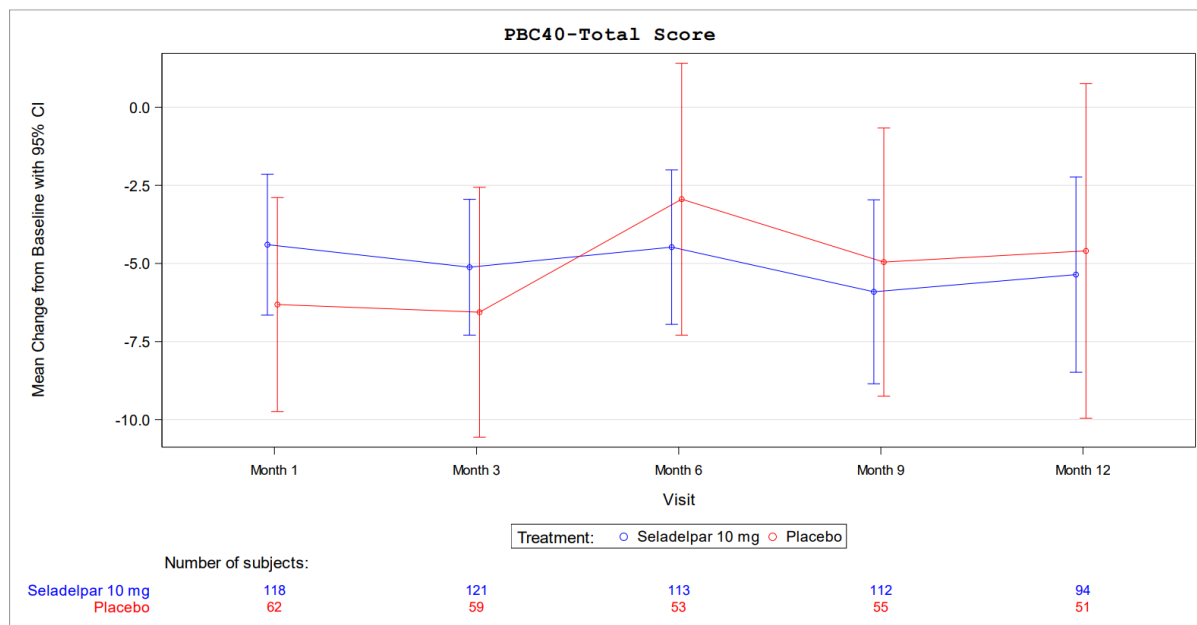


Abbildung 23: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Gesamt-Scores aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

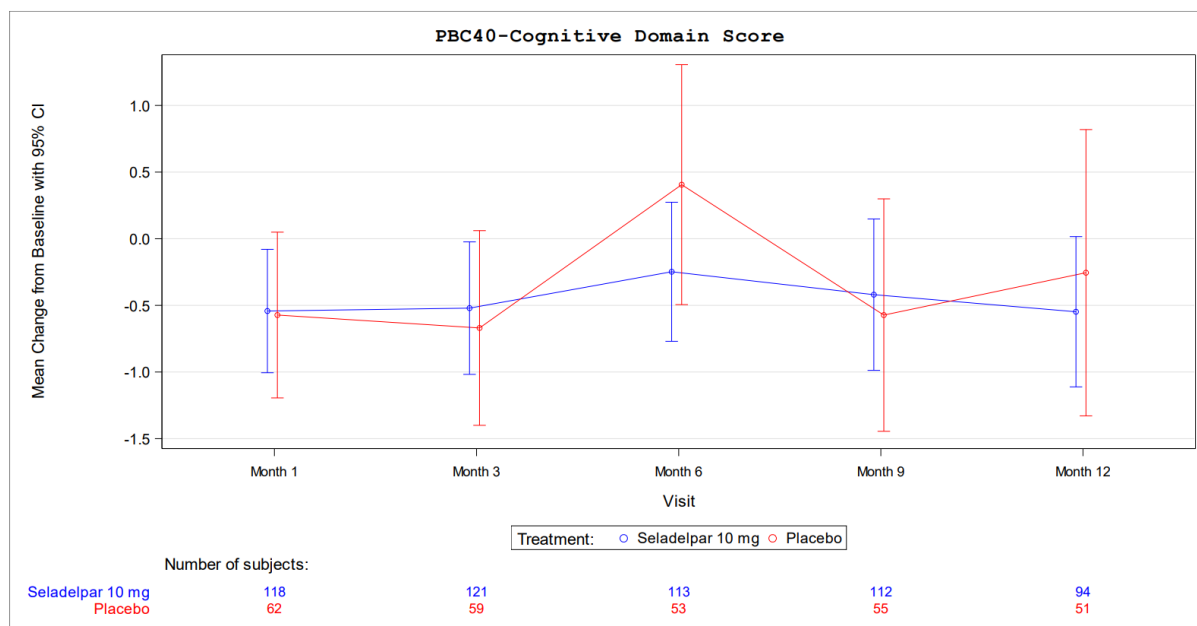


Abbildung 24: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Kognitive Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

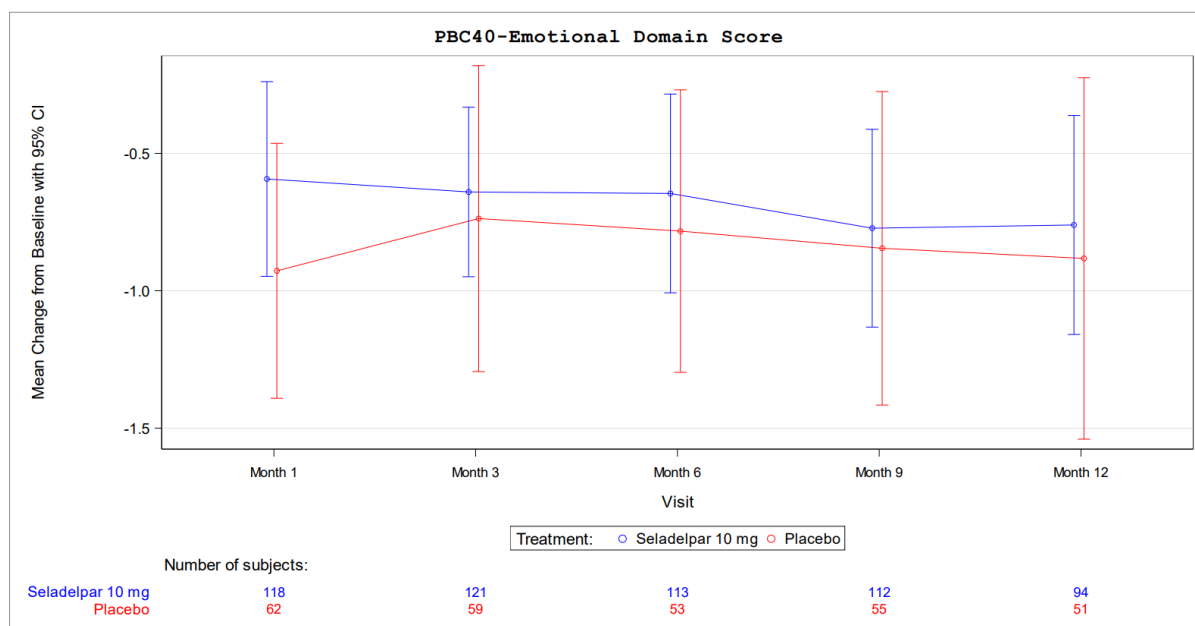


Abbildung 25: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Emotionale Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

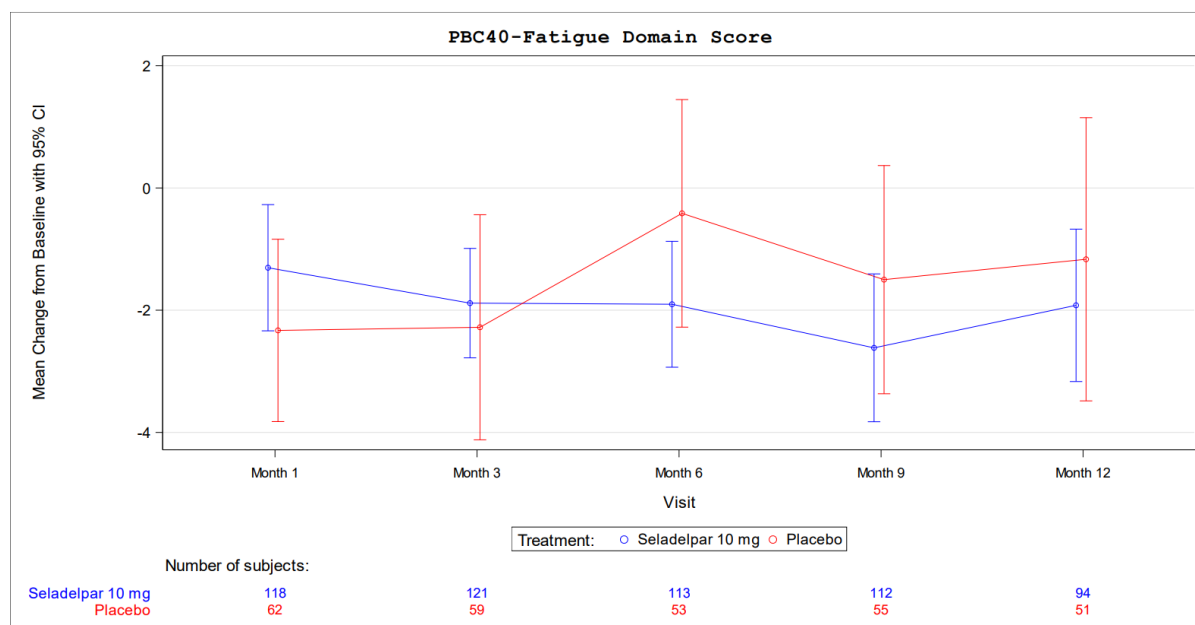


Abbildung 26: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Fatigue Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

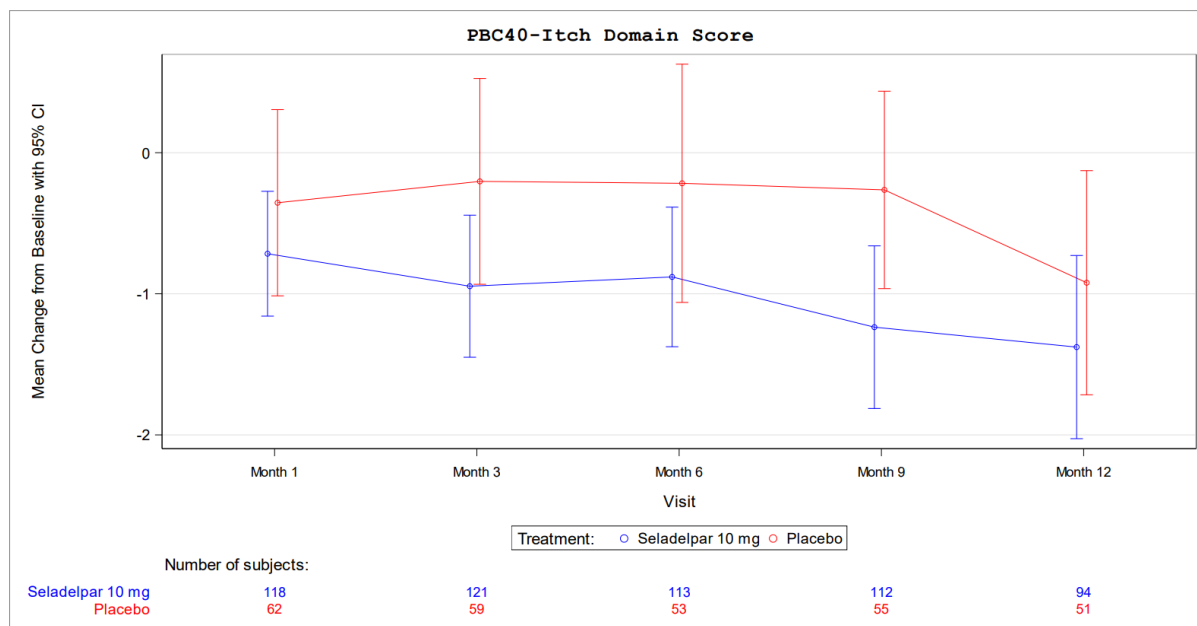


Abbildung 27: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Pruritus Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

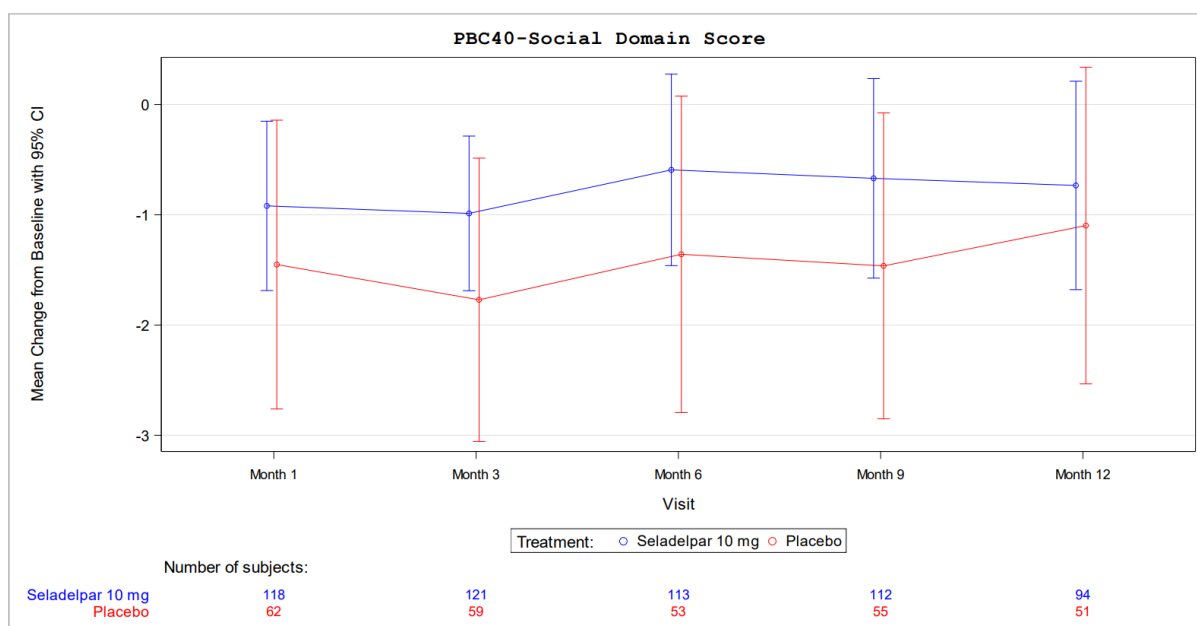


Abbildung 28: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Soziale Domäne RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

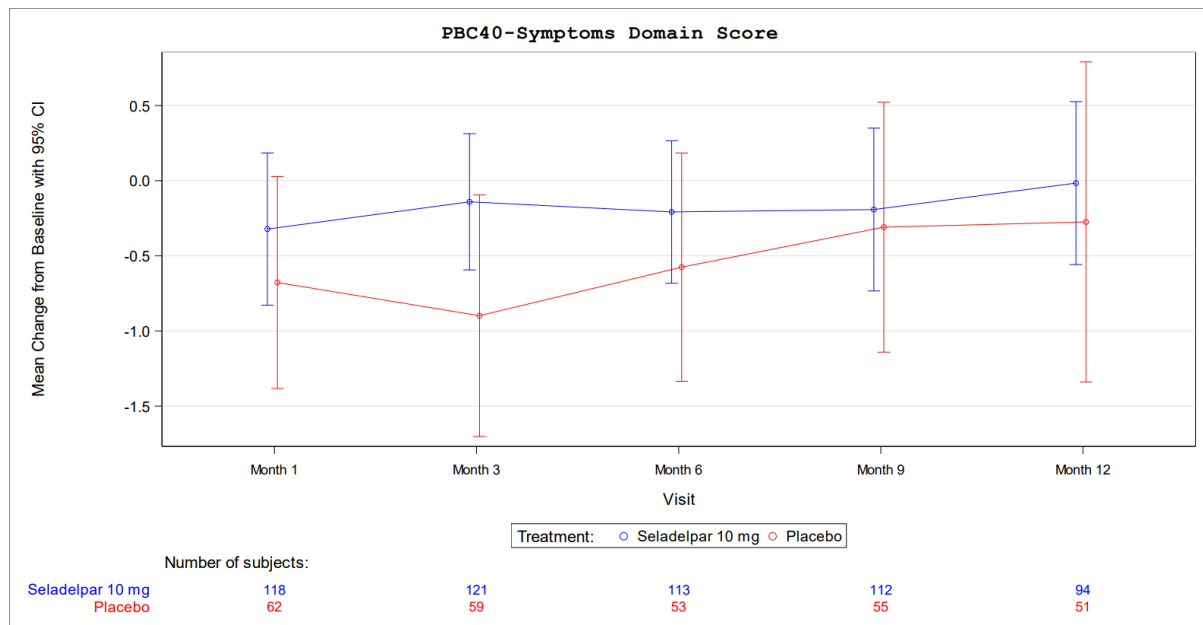


Abbildung 29: Darstellung des zeitlichen Verlaufs der mittleren Veränderung PBC-40 Symptomatische Domäne aus RCT mit dem zbAM (Studie RESPONSE, FAS)

CI=Konfidenzintervall, FAS=Full-Analysis-Set, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, zbAM=Zu bewertendes Arzneimittel

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.1.9 Unerwünschte Ereignisse – RCT

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

Tabelle 4-44: Operationalisierung von UE

Studie	Operationalisierung
RESPONSE	Folgende Ereignisse sind von Interesse: • Jegliche UE • SUE • Schwere UE (definiert als UE mit CTCAE-Grad ≥ 3) • UE, die zum Tod führten • UE, die zum Therapieabbruch führten • UE von speziellem Interesse^a ○ Pruritus-assoziiert ○ Lebertoxizität ○ Muskeltoxizität ○ Renale Toxizität ○ Pankreasbedingte Toxizität ○ Kardiovaskuläre Toxizität ▪ Herzrhythmusstörungen ▪ Herzinsuffizienz ▪ Kardiomyopathie ▪ Ischämische Herzkrankheit ○ Frakturen • UE nach SOC und PT: ○ Jegliche UE nach SOC und PT, die bei $\geq 10\%$ der Patienten in einem Studienarm oder bei ≥ 10 Patienten sowie mindestens 1% in einem Studienarm aufgetreten sind ○ SUE nach SOC und PT, die bei $\geq 5\%$ der Patienten in einem Studienarm oder bei ≥ 10 Patienten sowie mindestens 1% in einem Studienarm aufgetreten sind ○ Schwere UE nach SOC und PT, die bei $\geq 5\%$ der Patienten in einem Studienarm oder bei ≥ 10 Patienten sowie mindestens 1% in einem Studienarm aufgetreten sind Ergänzend werden außerdem folgende Ereignisse berichtet: • Nicht schwere UE (definiert als UE mit CTCAE-Grad < 3) • Therapieabbrüche aufgrund von UE nach SOC und PT (deskriptiv) • Gesamtrate UE, SUE und schwere UE ohne erkrankungsbezogene Ereignisse^b Dargestellt werden Anzahl und Anteil der Patienten mit Ereignissen, sowie das RR, OR und RD mit zugehörigem p-Wert.
a: Die Definitionen der einzelnen UE von speziellem Interesse sind dem Anhang 4-G zu entnehmen. b: Bei der Auswertung wurden post-hoc identifizierte Ereignisse (bestimmte PT) ausgeschlossen, die potenziell im Zusammenhang mit der Grunderkrankung stehen. Darstellung einschließlich einer Auflistung der ausgeschlossenen Ereignisse in Anhang 4-G. CTCAE=Common Terminology Criteria for Adverse Events, OR=Odds Ratio, PT=Preferred Term, RD=Risikodifferenz, RR=Relatives Risiko, SOC=System Organ Class, SUE=Schwerwiegendes unerwünschtes Ereignis, UE=Unerwünschtes Ereignis	

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

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

Tabelle 4-45: Bewertung des Verzerrungspotenzials für UE in RCT mit dem zbAM

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
RESPONSE	Niedrig	Ja	Ja	Ja	Ja	Niedrig
ITT=Intention-to-Treat, RCT=Randomisierte kontrollierte Studie, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wird auf Studienebene als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Auch sonstige das Verzerrungspotenzial beeinflussende Aspekte können nicht identifiziert werden.

Entsprechend wird für die Studie RESPONSE das Verzerrungspotenzial für diesen Endpunkt als niedrig bewertet.

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

Gesamtraten UE

Tabelle 4-46: Ergebnisse für UE (Gesamtraten) aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Jegliche UE	111/128 (86,7)	55/65 (84,6)	1,02 [0,91; 1,16], 0,6976 1,19 [0,51; 2,76], 0,6908 0,02 [-0,08; 0,13], 0,6962
Schwere UE (CTCAE-Grad $\geq$ 3)	14/128 (10,9)	5/65 (7,7)	1,42 [0,54; 3,78], 0,4799 1,47 [0,51; 4,29], 0,4767 0,03 [-0,05; 0,12], 0,4510
SUE	9/128 (7,0)	4/65 (6,2)	1,14 [0,37; 3,57], 0,8186 1,15 [0,34; 3,90], 0,8183 0,01 [-0,06; 0,08], 0,8145

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
UE, die zum Tod führten	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
UE, die zum Therapieabbruch führten	4/128 (3,1)	3/65 (4,6)	0,68 [0,16; 2,94], 0,6023 0,67 [0,14; 3,07], 0,6029 -0,01 [-0,07; 0,04], 0,6220
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit SEL. CTCAE=Common Terminology Criteria for Adverse Events, KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, NA=Nicht auswertbar, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, SEL=Seladelpar, SUE=Schwerwiegendes unerwünschtes Ereignis, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel			

Es zeigten sich keine statistisch signifikanten Unterschiede zwischen SEL vs. Placebo hinsichtlich jeglicher UE, SUE, schwerer UE, und UE, die zum Therapieabbruch führten. Es ist kein Patient verstorben.

Gesamtraten UE unter Ausschluss erkrankungsbezogener Ereignisse

Tabelle 4-47: Ergebnisse für UE ohne erkrankungsbezogene Ereignisse aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Jegliche UE ohne erkrankungsbezogene Ereignisse	111/128 (86,7)	54/65 (83,1)	1,04 [0,92; 1,19], 0,5144 1,33 [0,58; 3,04], 0,4981 0,04 [-0,07; 0,14], 0,5105
Schwere UE (CTCAE-Grad $\geq$ 3) ohne erkrankungsbezogene Ereignisse	14/128 (10,9)	4/65 (6,2)	1,78 [0,61; 5,18], 0,2923 1,87 [0,59; 5,94], 0,2865 0,05 [-0,03; 0,13], 0,2389
SUE ohne erkrankungsbezogene Ereignisse	9/128 (7,0)	4/65 (6,2)	1,14 [0,37; 3,57], 0,8186 1,15 [0,34; 3,90], 0,8183 0,01 [-0,06; 0,08], 0,8145
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. CTCAE=Common Terminology Criteria for Adverse Events, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, SUE=Schwerwiegendes unerwünschtes Ereignis, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel			

Es zeigten sich keine statistisch signifikanten Unterschiede zwischen SEL vs. Placebo hinsichtlich jeglicher UE ohne erkrankungsbezogene Ereignisse, SUE ohne erkrankungsbezogene Ereignisse und schwerer UE ohne erkrankungsbezogene Ereignisse.

UE von speziellem Interesse

Tabelle 4-48: Ergebnisse für UESI aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Herzrhythmusstörungen	3/128 (2,3)	2/65 (3,1)	0,76 [0,13; 4,45], 0,7624 0,76 [0,12; 4,64], 0,7626 -0,01 [-0,06; 0,04], 0,7715
Herzinsuffizienz	3/128 (2,3)	1/65 (1,5)	1,52 [0,16; 14,36], 0,7130 1,54 [0,16; 15,06], 0,7126 0,01 [-0,03; 0,05], 0,6915
Kardiovaskuläre Toxizität	13/128 (10,2)	5/65 (7,7)	1,32 [0,49; 3,54], 0,5812 1,36 [0,46; 3,98], 0,5792 0,02 [-0,06; 0,11], 0,5620
Kardiomyopathie	5/128 (3,9)	0/65 (0,0)	5,63 [0,32; 100,24], 0,2396 5,83 [0,32; 107,15], 0,2350 0,04 [0,01; 0,07], 0,0225
Ischämische Herzkrankheit	5/128 (3,9)	2/65 (3,1)	1,27 [0,25; 6,37], 0,7718 1,28 [0,24; 6,79], 0,7714 0,01 [-0,05; 0,06], 0,7623
Lebertoxizität	8/128 (6,3)	6/65 (9,2)	0,68 [0,25; 1,87], 0,4517 0,66 [0,22; 1,98], 0,4532 -0,03 [-0,11; 0,05], 0,4757
Muskelttoxizität	8/128 (6,3)	5/65 (7,7)	0,81 [0,28; 2,38], 0,7055 0,80 [0,25; 2,55], 0,7060 -0,01 [-0,09; 0,06], 0,7141
Pankreasbedingte Toxizität	2/128 (1,6)	1/65 (1,5)	1,02 [0,09; 10,99], 0,9898 1,02 [0,09; 11,42], 0,9898 0,00 [-0,04; 0,04], 0,9898
Pruritus-assoziiert	7/128 (5,5)	10/65 (15,4)	0,36 [0,14; 0,89], 0,0273 0,32 [0,12; 0,88], 0,0273 -0,10 [-0,20; 0,00], 0,0433
Renale Toxizität	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Frakturen	7/128 (5,5)	0/65 (0,0)	7,67 [0,45; 132,32], 0,1607 8,09 [0,45; 143,83], 0,1547 0,05 [0,02; 0,09], 0,0065
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, NA=Nicht auswertbar, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, UESI=Unerwünschtes Ereignis von speziellem Interesse, zbAM=Zu bewertendes Arzneimittel			

Hinsichtlich der UESI jeglichen Schweregrades zeigte sich für das UESI Pruritus-assoziierte UE ein statistisch signifikanter Vorteil zugunsten der Patienten unter SEL (RR [95%-KI]: 0,36 [0,14; 0,89], p=0,0273). Weitere statistisch signifikante Unterschiede traten nicht auf.

Tabelle 4-49: Ergebnisse für UESI (SUE) aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Herzrhythmusstörungen	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Herzinsuffizienz	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Kardiovaskuläre Toxizität	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Kardiomyopathie	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Ischämische Herzkrankheit	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Lebertoxizität	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Muskeltoxizität	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Pankreasbedingte Toxizität	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Pruritus-assoziiert	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Renale Toxizität	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Frakturen	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, NA=Nicht auswertbar, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, SUE=Schwerwiegendes unerwünschtes Ereignis, UESI=Unerwünschtes Ereignis von speziellem Interesse, zbAM=Zu bewertendes Arzneimittel			

Es zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen hinsichtlich der schwerwiegenden UESI. Es trat jeweils ein SUE in dem Endpunkt Kardiovaskuläre Toxizität, Ischämische Herzkrankheit und Lebertoxizität im SEL-Arm auf.

Tabelle 4-50: Ergebnisse für UESI (Schwere UE [CTCAE-Grad $\geq$ 3]) aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Herzrhythmusstörungen	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Herzinsuffizienz	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Kardiovaskuläre Toxizität	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Kardiomyopathie	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Ischämische Herzkrankheit	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Lebertoxizität	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Muskelttoxizität	0/128 (0,0)	1/65 (1,5)	0,17 [0,01; 4,13], 0,2767 0,17 [0,01; 4,16], 0,2757 -0,02 [-0,05; 0,01], 0,3136

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
Pankreasbedingte Toxizität	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Pruritus-assoziiert	0/128 (0,0)	1/65 (1,5)	0,17 [0,01; 4,13], 0,2767 0,17 [0,01; 4,16], 0,2757 -0,02 [-0,05; 0,01], 0,3136
Renale Toxizität	0/128 (0,0)	0/65 (0,0)	NA [NA; NA], NA NA [NA; NA], NA NA [NA; NA], NA
Frakturen	1/128 (0,8)	0/65 (0,0)	1,53 [0,06; 37,16], 0,7922 1,54 [0,06; 38,36], 0,7920 0,01 [-0,01; 0,02], 0,3154
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. CTCAE=Common Terminology Criteria for Adverse Events, KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, UE=Unerwünschtes Ereignis, UESI=Unerwünschtes Ereignis von speziellem Interesse, zbAM=Zu bewertendes Arzneimittel			

Auch für die schweren UESI zeigen sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen. Es trat jeweils ein schweres UE in dem Endpunkt Kardiovaskuläre Toxizität, Ischämische Herzkrankheit und Lebertoxizität im SEL-Arm auf.

UE nach SOC und PT

Tabelle 4-51: Ergebnisse für Jegliche UE nach SOC und PT aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
SOC Erkrankungen des Blutes und des Lymphsystems	15/128 (11,7)	3/65 (4,6)	2,54 [0,76; 8,46], 0,1290 2,74 [0,76; 9,84], 0,1216 0,07 [0,00; 0,15], 0,0653
SOC Erkrankungen des Gastrointestinaltrakts	41/128 (32,0)	24/65 (36,9)	0,87 [0,58; 1,30], 0,4924 0,81 [0,43; 1,51], 0,4971 -0,05 [-0,19; 0,09], 0,5010
SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	23/128 (18,0)	13/65 (20,0)	0,90 [0,49; 1,66], 0,7312 0,88 [0,41; 1,87], 0,7322 -0,02 [-0,14; 0,10], 0,7354
SOC Infektionen und parasitäre Erkrankungen	58/128 (45,3)	35/65 (53,8)	0,84 [0,63; 1,13], 0,2512 0,71 [0,39; 1,29], 0,2629 -0,09 [-0,23; 0,06], 0,2608

Endpunkt	Seladelpar	Placebo	Seladelpar vs. Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert
PT COVID-19	23/128 (18,0)	10/65 (15,4)	1,17 [0,59; 2,30], 0,6544 1,20 [0,54; 2,71], 0,6525 0,03 [-0,08; 0,14], 0,6454
SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	17/128 (13,3)	4/65 (6,2)	2,16 [0,76; 6,15], 0,1500 2,34 [0,75; 7,25], 0,1423 0,07 [-0,01; 0,15], 0,0919
SOC Untersuchungen	18/128 (14,1)	6/65 (9,2)	1,52 [0,64; 3,65], 0,3454 1,61 [0,61; 4,27], 0,3397 0,05 [-0,04; 0,14], 0,3066
SOC Stoffwechsel- und Ernährungsstörungen	15/128 (11,7)	10/65 (15,4)	0,76 [0,36; 1,60], 0,4724 0,73 [0,31; 1,73], 0,4747 -0,04 [-0,14; 0,07], 0,4893
SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	31/128 (24,2)	18/65 (27,7)	0,87 [0,53; 1,44], 0,5980 0,83 [0,42; 1,64], 0,6005 -0,03 [-0,17; 0,10], 0,6052
SOC Erkrankungen des Nervensystems	22/128 (17,2)	9/65 (13,8)	1,24 [0,61; 2,54], 0,5539 1,29 [0,56; 2,99], 0,5509 0,03 [-0,07; 0,14], 0,5382
PT Kopfschmerzen	10/128 (7,8)	2/65 (3,1)	2,54 [0,57; 11,25], 0,2199 2,67 [0,57; 12,56], 0,2140 0,05 [-0,02; 0,11], 0,1384
SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums	12/128 (9,4)	5/65 (7,7)	1,22 [0,45; 3,31], 0,6981 1,24 [0,42; 3,69], 0,6971 0,02 [-0,07; 0,10], 0,6880
SOC Erkrankungen der Haut und des Unterhautgewebes	20/128 (15,6)	15/65 (23,1)	0,68 [0,37; 1,23], 0,2021 0,62 [0,29; 1,31], 0,2066 -0,07 [-0,19; 0,05], 0,2243
PT Pruritus	6/128 (4,7)	10/65 (15,4)	0,30 [0,12; 0,80], 0,0160 0,27 [0,09; 0,78], 0,0157 -0,11 [-0,20; -0,01], 0,0274
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. COVID-19=Coronavirus-Krankheit 2019, KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OR=Odds Ratio, PT=Preferred Term, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, SOC=Systemorganklasse, UE=Unerwünschtes Ereignis, zbAM=Zu bewertendes Arzneimittel			

Für die UE nach SOC und PT jeglichen Schweregrades zeigt sich erneut der statistisch signifikante Vorteil im PT Pruritus für die Patienten im SEL-Arm (RR [95%-KI]: 0,30 [0,12; 0,80]; p=0,0160). Bei allen weiteren SOC und PT zeigen sich keine statistisch signifikanten Unterschiede.

SUE nach SOC und PT

In der Studie RESPONSE traten bei den SUE keine SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen auf.

Schwere UE nach SOC und PT

In der Studie RESPONSE traten bei den schweren UE keine SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen auf.

Therapieabbrüche aufgrund UE nach SOC und PT

Tabelle 4-52: Ergebnisse für UE, die zum Therapieabbruch führten (ergänzend, deskriptiv) nach SOC und PT aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Endpunkt	Seladelpar	Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
SOC Erkrankungen des Blutes und des Lymphsystems	1/128 (0,8)	0/65 (0,0)
PT Koagulopathie	1/128 (0,8)	0/65 (0,0)
SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	1/128 (0,8)	0/65 (0,0)
PT Progression einer Erkrankung	1/128 (0,8)	0/65 (0,0)
SOC Leber- und Gallenerkrankungen	0/128 (0,0)	1/65 (1,5)
PT Hyperbilirubinämie	0/128 (0,0)	1/65 (1,5)
SOC Erkrankungen des Immunsystems	0/128 (0,0)	1/65 (1,5)
PT Overlap-Syndrom	0/128 (0,0)	1/65 (1,5)
SOC Untersuchungen	1/128 (0,8)	0/65 (0,0)
PT Leberfunktionstest erhöht	1/128 (0,8)	0/65 (0,0)
SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen)	1/128 (0,8)	0/65 (0,0)
PT Papilläres Schilddrüsenkarzinom	1/128 (0,8)	0/65 (0,0)

Endpunkt	Seladelpar	Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
SOC Psychiatrische Erkrankungen	0/128 (0,0)	1/65 (1,5)
PT Suizidversuch	0/128 (0,0)	1/65 (1,5)
N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, PT=Preferred Term, SOC=Systemorganklasse		

Insgesamt gab es nur sehr wenige Therapieabbrüche aufgrund von UE in beiden Behandlungsarmen.

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

Es wurde keine Meta-Analyse durchgeführt, da nur eine relevante Studie im Anwendungsgebiet identifiziert wurde.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.2 Subgruppenanalysen – RCT

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

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

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

¹⁶ unbesetzt

- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.
- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmotifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.
- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmotifikatoren 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 4-53 Matrix der durchgeführten Subgruppenanalysen

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Mortalität														
Mortalität ^a	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Morbidität														
Biochemisches Ansprechen ^b	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Pruritus NRS	●	●	●	●	●	●	●	●	○	●	●	○	●	●
5-D Pruritus	●	●	●	●	●	●	●	●	○	●	●	○	●	●
PGI-C & PGI-S	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Fibrose	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Hospitalisierung	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Gesundheitsbezogene Lebensqualität														
PBC-40	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Unerwünschte Ereignisse														
UE ^c	●	●	●	●	●	●	●	●	○	●	●	○	●	●

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
• A priori geplante Subgruppenanalyse. ○ Post-hoc durchgeführte Subgruppenanalyse. a: Es ist kein Patient verstorben. b: Beinhaltet die Unterendpunkte ALP-Normalisierung und kombiniertes biochemisches Ansprechen. c: Es wurden die Unterendpunkte Gesamtraten, jegliche UESI, schwerwiegende UESI, schwere UESI, sowie jegliche UE, SUE und schwere UE getrennt nach SOC und PT, bei denen ein statistisch signifikanter Behandlungsunterschied vorlag, betrachtet. ALP=Alkalische Phosphatase, GGT=Gamma-Glutamyltransferase, N.D.=Nicht durchgeführt, NRS=Numerical Rating Scale, OCA=Obeticholsäure, PBC=Primär biliäre Cholangitis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, PT=Preferred Term, SOC=System Organ Class, SUE= Schwerwiegendes unerwünschtes Ereignis, UDCA=Ursodeoxycholsäure, UE=Unerwünschtes Ereignis, UESI= Unerwünschtes Ereignis von speziellem Interesse														

Stellen Sie anschließend in Tabelle 4-54 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ($p < 0,05$) Interaktionsterme.

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

Tabelle 4-54: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt (Studie RESPONSE)

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Mortalität														
Mortalität ^a	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Morbidität														
<i>Ansprechen zu Monat 6</i>														
ALP-Normalisierung	0,4983	0,7456	0,2499	N.D.	0,8244	0,2179	0,1758	N.D.	0,2196	0,6327	N.D.	0,6457	0,1138	0,9053
Kombiniertes biochemisches Ansprechen	0,7247	0,2144	0,9111	N.D.	0,7562	0,4568	0,7990	0,5048	0,6722	0,7400	0,8862	0,9425	0,9033	0,4333
ALP – stetige Analyse	0,1210	0,9948	N.D.	N.D.	0,8784	0,1769	N.D.	0,4172	N.D.	0,2246	0,0017	0,0561	0,0246	0,0529
Gesamt-Bilirubin – stetige Analyse	0,5197	0,7953	0,1824	N.D.	0,3878	0,1706	0,2756	0,0880	0,3411	0,3782	0,0737	0,8233	0,4772	0,4178

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
<i>Ansprechen zu Monat 12</i>														
ALP-Normalisierung	0,3087	0,8297	0,2624	N.D.	0,8407	0,1554	0,1862	0,2910	0,2318	0,5609	N.D.	0,7745	0,1212	0,9026
Kombiniertes biochemisches Ansprechen	0,6471	0,7806	0,9766	N.D.	0,6063	0,3869	0,8322	0,4747	0,7043	0,9635	0,5562	0,3399	0,7985	0,2807
ALP – stetige Analyse	0,6726	0,6680	N.D.	N.D.	0,8345	0,6002	N.D.	0,7901	N.D.	0,2223	0,1209	0,1201	0,3538	0,1387
Gesamt-Bilirubin – stetige Analyse	0,8116	0,6851	0,1632	N.D.	0,1141	0,6138	0,0004	0,1497	0,0645	0,3599	0,1935	0,5584	0,9829	0,6966
<i>Pruritus NRS zu Monat 6</i>														
Verbesserung um ≥ 4 Punkte	0,8630	0,1602	0,6944	N.D.	N.D.	0,3188	0,3973	0,5063	0,5645	N.D.	0,5001	0,8455	0,4117	0,1341
Stetige Analyse	0,4617	0,1905	N.D.	N.D.	0,5035	0,5467	N.D.	0,3835	0,4882	0,1519	0,7227	0,5490	0,3014	0,4101
<i>Pruritus NRS zu Monat 12</i>														
Verbesserung um ≥ 4 Punkte	0,0903	N.D.	0,9392	N.D.	N.D.	0,6416	0,7975	0,7884	0,9150	N.D.	N.D.	0,7859	0,7527	0,1593

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Stetige Analyse	0,4898	0,5763	N.D.	N.D.	0,2411	0,9704	N.D.	0,7289	0,4854	0,0256	0,2398	0,7060	0,1978	0,0120
5-D Pruritus – Stetige Analyse zu Monat 6														
Gesamt-Score	0,7628	0,8951	N.D.	N.D.	0,4791	0,0597	N.D.	0,0250	N.D.	0,0216	0,0947	0,6788	0,9534	0,0099
Subdomäne Grad	0,6778	0,5189	N.D.	N.D.	0,3653	0,1669	N.D.	0,1330	N.D.	0,0560	0,1043	0,7140	0,7122	0,0138
Subdomäne Dauer	0,8680	0,8193	N.D.	N.D.	0,6272	0,0854	N.D.	0,5498	N.D.	0,0902	0,1827	0,5607	0,9534	0,2488
Subdomäne Beeinträchtigung	0,4233	0,1778	N.D.	N.D.	0,4389	0,5624	N.D.	0,1476	N.D.	0,0147	0,4548	0,2411	0,6420	0,0493
Subdomäne Verlauf	0,7033	0,9834	N.D.	N.D.	0,4531	0,2455	N.D.	0,0217	N.D.	0,2721	0,4028	0,8940	0,9802	0,3244
Subdomäne Lokalisation	0,3845	0,1900	N.D.	N.D.	0,8215	0,0897	N.D.	0,5669	N.D.	0,0510	0,3047	0,4116	0,8684	0,0080
Modifizierter Gesamt-Score	0,5215	0,8722	N.D.	N.D.	0,5644	0,0474	N.D.	0,1414	N.D.	0,0085	0,1070	0,5144	0,9969	0,0049
5-D Pruritus – Stetige Analyse zu Monat 12														
Gesamt-Score	0,7522	0,9254	N.D.	N.D.	0,2053	0,9249	N.D.	0,0546	N.D.	0,0946	0,6282	0,3736	0,2869	0,0213

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Subdomäne Grad	0,7635	0,6583	N.D.	N.D.	0,2027	0,7134	N.D.	0,4325	N.D.	0,3169	0,7701	0,2886	0,3889	0,0382
Subdomäne Dauer	0,0413	0,2261	N.D.	N.D.	0,5316	0,9829	N.D.	0,7045	N.D.	0,0227	0,6740	0,7227	0,8925	0,2933
Subdomäne Beeinträchtigung	0,5524	0,4611	N.D.	N.D.	0,1973	0,8873	N.D.	0,1788	N.D.	0,1777	0,6646	0,6676	0,9075	0,0540
Subdomäne Verlauf	0,7839	0,8660	N.D.	N.D.	0,2674	0,6873	N.D.	0,0414	N.D.	0,9460	0,5320	0,1853	0,6868	0,3416
Subdomäne Lokalisation	0,5797	0,7404	N.D.	N.D.	0,7182	0,7862	N.D.	0,8204	N.D.	0,0025	0,2826	0,3591	0,6537	0,0730
Modifizierter Gesamt-Score	0,4500	0,9020	N.D.	N.D.	0,2835	0,9406	N.D.	0,1796	N.D.	0,0123	0,4534	0,7187	0,4230	0,0137
PGI – Stetige Analyse zu Monat 6														
PGI-S	0,4356	0,4156	N.D.	N.D.	0,6462	0,4008	N.D.	0,1951	N.D.	0,2079	0,6613	0,6215	0,5820	0,1917
PGI-C	0,2890	0,8797	N.D.	N.D.	0,1043	0,8697	N.D.	0,4761	N.D.	0,3952	0,9442	0,8029	0,3096	0,7433
PGI – Stetige Analyse zu Monat 12														
PGI-S	0,2783	0,3464	N.D.	N.D.	0,4396	0,0463	N.D.	0,6024	N.D.	0,2120	0,0221	0,3465	0,0257	0,0033

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
PGI-C	0,6261	0,3719	N.D.	N.D.	0,6642	0,8378	N.D.	0,8325	N.D.	0,5512	0,1545	0,8172	0,9885	0,8089
Fibrose – Stetige Analyse zu Monat 6														
Erweiterter Leberfibrosetest (ELF-Score)	0,9912	0,6529	0,7564	N.D.	0,4526	0,2466	N.D.	0,0635	0,6122	0,3237	0,2655	0,4824	0,8853	0,8223
Lebersteifigkeit (FibroScan [®])	0,2823	0,0512	N.D.	N.D.	0,0581	0,1335	0,3991	0,0107	0,0838	0,7203	0,4128	0,1106	0,6141	0,0109
Fibrose – Stetige Analyse zu Monat 12														
Erweiterter Leberfibrosetest (ELF-Score)	0,1207	0,7607	0,8255	N.D.	0,9592	0,3614	N.D.	0,2861	0,1906	0,3366	0,3792	0,4427	0,0726	0,1480
Lebersteifigkeit (FibroScan [®])	0,7746	0,0387	N.D.	N.D.	0,1298	0,0756	0,1036	0,0283	0,0328	0,3403	0,1584	0,3783	0,6854	0,3722
Hospitalisierung zu Monat 12														
Hospitalisierung ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Gesundheitsbezogene Lebensqualität														
<i>PBC-40 Stetige Analyse zu Monat 6</i>														
Gesamt-Score	0,7299	0,1978	N.D.	N.D.	0,5280	0,1535	N.D.	0,7448	N.D.	0,1569	0,7074	0,1796	0,1945	0,9646
Kognitive Domäne	0,1621	0,0906	N.D.	N.D.	0,7791	0,4003	N.D.	0,2241	0,0218	0,0207	0,9925	0,0314	0,1242	0,1441
Emotionale Domäne	0,4581	0,9401	N.D.	N.D.	0,1122	0,8244	N.D.	0,3315	N.D.	0,3608	0,5180	0,4233	0,6469	0,3701
Fatigue Domäne	0,7108	0,2022	N.D.	N.D.	0,3829	0,6212	N.D.	0,4743	N.D.	0,6669	0,4884	0,6176	0,3336	0,9534
Pruritus Domäne	0,3337	0,2198	N.D.	N.D.	0,9207	0,0681	N.D.	0,9641	N.D.	0,1401	0,8311	0,0924	0,1247	0,3114
Soziale Domäne	0,5382	0,5298	N.D.	N.D.	0,9996	0,8074	N.D.	0,9723	N.D.	0,0288	0,4135	0,1354	0,3816	0,8955
Symptomatische Domäne	0,0505	0,5333	N.D.	N.D.	0,5261	0,4236	N.D.	0,3100	N.D.	0,1916	0,6073	0,6401	0,7422	0,0527
<i>PBC-40 Stetige Analyse zu Monat 12</i>														
Gesamt-Score	0,7471	0,5756	N.D.	N.D.	0,0233	0,4446	N.D.	0,4703	N.D.	0,3131	0,3669	0,2796	0,2909	0,0730

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Kognitive Domäne	0,3271	0,2770	N.D.	N.D.	0,0357	0,6993	N.D.	0,7198	0,3447	0,4555	0,7549	0,2382	0,9653	0,7874
Emotionale Domäne	0,5237	0,4713	N.D.	N.D.	0,4135	0,4082	N.D.	0,5887	N.D.	0,6678	0,5293	0,1984	0,3995	0,1962
Fatigue Domäne	0,7528	0,9926	N.D.	N.D.	0,0307	0,6884	N.D.	0,8361	N.D.	0,4063	0,6281	0,8024	0,5426	0,1739
Pruritus Domäne	0,6501	0,7403	N.D.	N.D.	0,2400	0,6406	N.D.	0,1133	N.D.	0,3420	0,0704	0,3565	0,4216	0,1545
Soziale Domäne	0,4350	0,7138	N.D.	N.D.	0,3695	0,8796	N.D.	0,3420	N.D.	0,1669	0,3973	0,1518	0,5157	0,1574
Symptomatische Domäne	0,6797	0,5965	N.D.	N.D.	0,0830	0,6964	N.D.	0,5081	N.D.	0,5238	0,3835	0,8404	0,2100	0,0768
Unerwünschte Ereignisse														
Gesamtraten														
Jegliche UE	0,0987	0,6132	0,9428	N.D.	0,0993	0,8584	0,0637	0,8868	0,0623	0,4038	0,9723	0,5087	0,3999	0,2530
SUE	N.D.	N.D.	0,3726	N.D.	N.D.	0,8974	N.D.	0,6537	N.D.	N.D.	N.D.	N.D.	0,4650	N.D.
Schwere UE (CTCAE-Grad $\geq$ 3)	0,4764	0,9074	0,3138	N.D.	N.D.	0,2041	N.D.	0,5510	N.D.	0,2740	0,4121	0,8641	0,5474	0,8105

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
UE, die zum Tod führten ^b	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
UE, die zum Therapieabbruch führten ^b	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Jegliche UESI														
Herzrhythmusstörungen ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Herzinsuffizienz ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Kardiovaskuläre Toxizität	0,5171	0,1855	N.D.	N.D.	N.D.	0,8929	0,9639	0,9887	0,8470	0,7138	0,0964	0,0458	0,5328	N.D.
Kardio-myopathie ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Ischämische Herzkrankheit ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Lebertoxizität	0,8469	N.D.	0,3080	N.D.	N.D.	0,7018	0,6300	0,5353	0,2989	0,6441	N.D.	0,6576	0,4298	N.D.

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Muskeltoxizität	N.D.	N.D.	N.D.	N.D.	N.D.	0,2986	0,7058	0,1689	0,6028	N.D.	0,3070	0,7389	N.D.	N.D.
Pankreas bedingte Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Pruritus-assoziierte UE	0,8448	N.D.	0,9355	N.D.	N.D.	0,6709	N.D.	0,4764	0,6371	0,8520	0,2855	0,3371	N.D.	N.D.
Renale Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Frakturen ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Schwerwiegende UESI														
Herzrhythmusstörungen ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Herzinsuffizienz ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Kardiovaskuläre Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Kardiomyopathie ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Ischämische Herzkrankheit ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Lebertoxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Muskelttoxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Pankreas- bedingte Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Pruritus- assoziierte UE ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Renale Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Frakturen ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Schwere UESI (CTCAE-Grad ≥3)														
Herzrhythmus- störungen ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Herzinsuffizi- enz ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Kardiovaskuläre Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Kardiomyopathie ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Ischämische Herzkrankheit ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Lebertoxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Muskelttoxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Pankreasbedingte Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Pruritus-assoziierte UE ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Renale Toxizität ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Frakturen ^c	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.

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

Endpunkt	Alter zum Screening	Alter zum Zeitpunkt der PBC-Diagnose	Geschlecht	Ethnie ^b	Region	Zirrhose	UDCA-Intoleranz	Vorherige Einnahme von OCA und/oder Fibraten	Mono-/Kombitherapie	Pruritus NRS	ALP	GGT	Gesamt-Bilirubin 0,6	Gesamt-Bilirubin 1
Jegliche UE nach SOC und PT														
SOC Erkrankungen der Haut und des Unterhaut- gewebes, PT Pruritus	0,6943	N.D.	0,9877	N.D.	N.D.	0,5345	N.D.	0,3684	0,7083	0,9777	0,1954	0,4221	N.D.	N.D.
a: Es ist kein Patient verstorben. b: In mindestens einer der Subgruppenkategorien befinden sich weniger als 10 Patienten, daher wurden die Analysen nicht durchgeführt. c: In keiner der Subgruppenkategorien sind mehr als 10 Events aufgetreten, daher wurden die Analysen nicht durchgeführt. ALP=Alkalische Phosphatase, CTCAE=Common Terminology Criteria for Adverse Events, ELF= Enhanced Liver Fibrosis, GGT=Gamma-Glutamyltransferase, N.D.=Nicht durchgeführt, NRS=Numerical Rating Scale, OCA=Obeticholsäure, PBC=Primär biliäre Cholangitis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, PT=Preferred Term, SOC=Systemorganklasse, SUE=Schwerwiegendes unerwünschtes Ereignis, UDCA=Ursodeoxycholsäure, UE=Unerwünschtes Ereignis, UESI= Unerwünschtes Ereignis von speziellem Interesse														

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.

Die Subgruppenanalysen werden gemäß der in den Abschnitten 4.2.5.2 und 4.2.5.5 beschriebenen Methodik durchgeführt.

Den Modulvorlagen entsprechend werden Subgruppenanalysen nur dann dargestellt und als relevant erachtet, wenn mindestens zehn Patienten pro Subgruppe in die Analyse eingehen und bei binären Endpunkten zudem in einer Subgruppe mindestens zehn Ereignisse beobachtet wurden. Für UE nach SOC und PT werden Subgruppenanalysen durchgeführt, sofern für die jeweilige SOC bzw. den jeweiligen PT auf Basis der Gesamtpopulation ein statistisch signifikantes Ergebnis vorliegt.

Zunächst werden die p-Werte aller durchgeführten Interaktionstests tabellarisch dargestellt. Daran anschließend werden die Subgruppenanalysen, deren Interaktionstests statistisch signifikant sind (p-Wert <0,05), zusammenfassend beschrieben und die Ergebnisse detailliert tabellarisch dargestellt. Subgruppenanalysen mit nicht-signifikantem Interaktionstest (p-Wert ≥0,05) werden ergänzend in Anhang 4-G berichtet; diese Analysen haben keine Auswirkung auf das Gesamtfazit.

In Tabelle 4-55 sind sowohl die für Subgruppenanalysen präspezifizierten Subgruppenkategorisierungen als auch die für die vorliegende Nutzenbewertung herangezogenen Subgruppenaufteilungen aufgeführt.

Tabelle 4-55: Präspezifizierte und für die Nutzenbewertung herangezogenen Subgruppenvariablen

Subgruppenvariablen	Präspezifizierung gemäß SAP	Für die Nutzenbewertung herangezogene Kategorisierung
	RESPONSE	
Alter zum Screening	<65 Jahre vs. ≥65 Jahre	<65 Jahre vs. ≥65 Jahre
Alter zum Zeitpunkt der PBC-Diagnose	<50 Jahre vs. ≥50 Jahre	<50 Jahre vs. ≥50 Jahre
Geschlecht	Männlich vs. weiblich	Männlich vs. weiblich
Ethnie	Weiß vs. Schwarz vs. Asiatisch vs. Andere	Weiß vs. Schwarz vs. Asiatisch vs. Andere
Region	Nordamerika vs. Europa vs. Restliche Welt	Nordamerika vs. Europa vs. Restliche Welt
Zirrhose	Ja vs. nein	Ja vs. nein
UDCA-Intoleranz	UDCA-Toleranz vs. UDCA Intoleranz	UDCA-Toleranz vs. UDCA Intoleranz
Vorherige Einnahme von OCA und/oder Fibraten	Ja vs. nein	Ja vs. nein
Mono-/Kombitherapie	-	Monotherapie (SEL) vs. Kombinationstherapie (SEL+UDCA)
Pruritus NRS	<4 Punkte vs. ≥4 Punkte	<4 Punkte vs. ≥4 Punkte
ALP	ALP <350 U/L vs. ALP ≥350 U/L	ALP <350 U/L vs. ALP ≥350 U/L
GGT	-	≤3 x ULN vs. >3 x ULN
Gesamt-Bilirubin 0,6	<0,6 x ULN vs. ≥0,6 x ULN	<0,6 x ULN vs. ≥0,6 x ULN
Gesamt-Bilirubin 1	≤ULN vs. >ULN	≤ULN vs. >ULN
ALP=Alkalische Phosphatase, GGT=Gamma-Glutamyltransferase, NRS=Numerical Rating Scale, OCA=Obeticholsäure, PBC=Primär biliäre Cholangitis, SAP=Statistischer Analyseplan, SEL=Seladelpar, UDCA=Ursodeoxycholsäure, ULN=Oberer Normwert		

Die Relevanz der Subgruppenvariablen und der verwendeten Cut-off wird im Folgenden für die Fälle begründet, in denen die Subgruppenbildung nicht natürlich oder durch Kategorien definiert ist.

Alter zum Screening

Es wird dem im statistischen Analyseplan (SAP) präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in <65 Jahre vs. ≥65 Jahre.

Alter zum Zeitpunkt der PBC-Diagnose

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in <50 Jahre vs. ≥50 Jahre.

Geschlecht

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in männlich vs. weiblich.

Ethnie

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in weiß vs. schwarz vs. asiatisch vs. andere.

Region

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in Nordamerika vs. Europa vs. Restliche Welt.

Zirrhose

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in ja vs. nein.

UDCA-Intoleranz

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in UDCA-Einnahme vs. UDCA-Intoleranz.

Vorherige Einnahme von OCA und/oder Fibraten

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in ja vs. nein.

Mono-/Kombitherapie

Es erfolgt eine getrennte Betrachtung von Mono- und Kombinationstherapie.

Pruritus NRS

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in Pruritus NRS <4 Punkte vs. Pruritus NRS $\geq$ 4 Punkte.

ALP

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in ALP <350 U/L vs. ALP $\geq$ 350 U/L.

GGT

Als Cut-off erfolgt eine Einteilung in GGT $\leq$ 3 x ULN vs. GGT >3 x ULN. Ein erhöhter GGT-Wert weist auf eine Cholestase hin und ist somit mit einer schlechten Prognose assoziiert [14]. Als Cut-off erfolgt – basierend auf der Empfehlung der European Association for the Study of the Liver-Leitlinie zu cholestatischen Lebererkrankungen (2009) – eine Einteilung in GGT $\leq$ 3 x ULN vs. GGT >3 x ULN [56].

Gesamt-Bilirubin

Es wird dem im SAP präspezifizierten Cut-off gefolgt. Entsprechend erfolgt die Einteilung in $<0,6 \times \text{ULN}$ vs. $\geq 0,6 \times \text{ULN}$ für die Subgruppe Gesamt-Bilirubin 0,6 sowie in $\leq \text{ULN}$ vs. $> \text{ULN}$ für die Subgruppe Gesamt-Bilirubin 1.

Im Folgenden werden alle Subgruppenanalysen dargestellt, für die sich ein statistisch signifikanter Interaktionstest ergab. Alle durchgeführten Subgruppenanalysen sind Anhang 4-G zu finden.

Aufgrund sehr geringer Patientenzahlen in nur einer der Ausprägungen einer Subgruppe (UDCA-Toleranz, Therapie und vorherige Einnahme von OCA und/oder Fibraten) kommt es teils zu statistisch signifikanten Ergebnissen, deren Aussagekraft jedoch stark eingeschränkt ist.

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

4.3.1.3.2.1 Biochemisches Ansprechen – RCT

Tabelle 4-56: Ergebnisse für Subgruppenanalysen zum Endpunkt Gesamt-Bilirubin – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Monat 12									
UDCA-Intoleranz									
UDCA-Toleranz	118	0,77 (0,32)	0,00 (0,04)	60	0,73 (0,31)	0,01 (0,05)	0,00 [-0,13; 0,13], 0,9658	-0,01 [-0,32; 0,30]	0,0004
UDCA-Intoleranz	8	0,78 (0,21)	-0,09 (0,04)	3	0,85 (0,31)	0,38 (0,10)	-0,47 [-0,73; -0,20] 0,0036	-3,21 [-5,34; -1,07]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar. a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline. FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, UDCA=Ursodeoxycholsäure, zbAM=Zu bewertendes Arzneimittel									

In dem Endpunkt Gesamt-Bilirubin – stetige Analyse zeigte sich zu Monat 12 eine Effektmodifikation durch das Subgruppenmerkmal UDCA-Intoleranz ($p=0,0004$). Bei Patienten mit UDCA-Intoleranz ergab sich ein statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,47 [-0,73; -0,20]; $p=0,0036$), der auch klinisch relevant (SMD [95%-KI]: -3,21 [-5,34; -1,07]) war. Bei Patienten mit UDCA-Toleranz ergab sich kein Unterschied zwischen SEL und Placebo. Aufgrund der deutlich geringeren Patientenzahlen in der Subgruppenkategorie UDCA-Intoleranz gegenüber der Subgruppenkategorie UDCA-Toleranz (N=11 vs. N=178) ist die Aussagekraft der Analyse stark eingeschränkt und wird folglich nicht herangezogen und als nicht fazitrelevant eingestuft.

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

Tabelle 4-57: Ergebnisse für Subgruppenanalysen zum Endpunkt ALP – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Monat 6									
ALP									
<350 U/L	92	254,17 (47,01)	-113,66 (5,74)	45	253,63 (44,61)	-13,25 (8,05)	-100,41 [-119,53; -81,29], <0,0001	-1,82 [-2,24; -1,41]	0,0017
≥350 U/L	34	475,00 (118,40)	-219,72 (15,97)	18	471,04 (102,75)	-28,57 (22,17)	-191,15 [-246,47; -135,82], <0,0001	-2,02 [-2,71; -1,32]	
Gesamt-Bilirubin 1									
≤1 x ULN	106	301,65 (114,03)	-130,98 (7,24)	59	306,77 (109,72)	-17,48 (9,09)	-113,50 [-134,04; -92,96], <0,0001	-1,55 [-1,91; -1,19]	0,0246
>1 x ULN	20	384,24 (147,54)	-187,06 (14,61)	4	398,58 (185,12)	15,65 (35,48)	-202,71 [-282,83; -122,58], <0,0001	-2,96 [-4,37; -1,55]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar.									
a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.									
ALP=Alkalische Phosphatase, FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel									

In dem Endpunkt ALP – stetige Analyse zeigte sich eine Effektmodifikation für das Subgruppenmerkmal ALP zu Baseline ($p=0,0017$). Bei Patienten mit <350 U/L ergab sich ein statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -100,41 [-119,53; -81,29]; $p<0,0001$), der auch klinisch relevant (SMD [95%-KI]: -1,82 [-2,24; -1,41]) war. Bei Patienten mit ≥ 350 U/L ergab sich ebenso ein statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -191,15 [-246,47; -135,82]; $p<0,0001$), ebenfalls klinisch relevant (SMD [95%-KI]: -2,02 [-2,71; -1,32]).

In dem Endpunkt ALP – stetige Analyse zeigte sich außerdem eine Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin zu Baseline ($p=0,0246$). Bei Patienten mit ≤ 1 x ULN Gesamt-Bilirubin ergab sich ein statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -113,50 [-134,04; -92,96]; $p<0,0001$), der auch klinisch relevant (SMD [95%-KI]: -1,55 [-1,91; -1,19]) war. Bei Patienten mit >1 x ULN Gesamt-Bilirubin ergab sich ebenso ein statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -202,71 [-282,83; -122,58]; $p<0,0001$), ebenfalls klinisch relevant (SMD [95%-KI]: -2,96 [-4,37; -1,55]).

4.3.1.3.2.2 Pruritus NRS – RCT

Tabelle 4-58: Ergebnisse für Subgruppenanalysen zum Endpunkt Pruritus NRS – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS)

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Monat 12									
Pruritus NRS									
<4	79	1,11 (1,34)	-0,24 (0,15)	42	1,04 (1,12)	0,10 (0,20)	-0,33 [-0,79; 0,13], 0,1527	-0,26 [-0,63; 0,12]	0,0256
≥4	48	6,13 (1,42)	-3,27 (0,33)	22	6,63 (1,44)	-1,50 (0,50)	-1,77 [-2,97; -0,57], 0,0045	-0,76 [-1,28; -0,24]	
Gesamt-Bilirubin 0,6									
<0,6 x ULN	58	2,32 (2,48)	-0,89 (0,21)	31	2,99 (2,91)	-0,70 (0,28)	-0,19 [-0,86; 0,48], 0,5744	-0,12 [-0,55; 0,32]	0,0120
≥0,6 x ULN	69	3,64 (2,94)	-1,60 (0,22)	33	3,05 (3,05)	-0,13 (0,32)	-1,47 [-2,24; -0,71], 0,0002	-0,79 [-1,22; -0,36]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar.									
a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.									
FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, NRS=Numerical Rating Scale, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel									

In dem Endpunkt Pruritus NRS – stetige Analyse zeigte sich eine Effektmodifikation für das Subgruppenmerkmal Pruritus NRS zu Baseline ($p=0,0256$). Bei Patienten mit ≥ 4 Punkte Pruritus NRS zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter (LS-MWD [95%-KI]: -1,77 [-2,97; -0,57]; $p=0,0045$) und klinisch relevanter (SMD [95%-KI]: -0,76 [-1,28; -0,24]) Vorteil für SEL. Bei Patienten mit Verringerung um <4 Punkte zu Baseline (Pruritus NRS) war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant. Es zeigte sich demnach eine größere Verbesserung im Pruritus für Patienten mit einem stärkeren Pruritus. Die Effektmodifikation zur Pruritus NRS wird als fazitrelevant eingestuft.

In dem Endpunkt Pruritus NRS – stetige Analyse zeigte sich eine Effektmodifikation durch das Subgruppenmerkmal Gesamt-Bilirubin 0,6 ($p=0,0120$). Bei Patienten mit Gesamt-Bilirubin $\geq 0,6 \times \text{ULN}$ zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -1,47 [-2,24; -0,71]; $p=0,0002$), der auch klinisch relevant (SMD [95%-KI]: -0,79 [-1,22; -0,36]) war. Bei Patienten mit Gesamt-Bilirubin $<0,6 \times \text{ULN}$ zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant. Die unterschiedlichen Effektgrößen zwischen den Subgruppen waren zu erwarten und sind in der Art der Analyse begründet, außerdem ist ein höherer Bilirubin-Wert mit einer erhöhten Pruritus NRS assoziiert (vergleiche Abschnitt 4.3.1.3.2.8). Diese Effektmodifikation wird daher als nicht fazitrelevant eingestuft.

4.3.1.3.2.3 5-D Pruritus – RCT

Tabelle 4-59: Ergebnisse für Subgruppenanalysen zum Endpunkt 5-D Pruritus – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Gesamt-Score									
Monat 6									
Vorherige Einnahme von OCA und/oder Fibraten									
Ja	20	12,75 (5,48)	-2,03 (0,76)	12	12,77 (5,03)	-2,68 (0,96)	0,66 [-1,84; 3,15], 0,5950	-0,19 [-0,53; 0,91]	0,0250
Nein	108	11,42 (4,73)	-2,02 (0,33)	52	10,84 (4,52)	0,29 (0,45)	-2,32 [-3,35; -1,29], <0,0001	-0,69 [-1,03; -0,35]	
Pruritus NRS									
<4	79	8,76 (2,91)	-0,51 (0,34)	42	8,38 (2,47)	0,41 (0,45)	-0,92 [-1,96; 0,12], 0,0835	-0,31 [-0,68; 0,07]	0,0216
≥4	49	16,25 (3,64)	-4,73 (0,53)	22	16,41 (2,83)	-1,31 (0,79)	-3,42 [-5,32; -1,51], 0,0006	-0,91 [-1,43; -0,38]	
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	10,36 (4,00)	-1,28 (0,42)	31	11,22 (4,83)	-0,73 (0,54)	-0,55 [-1,80; 0,70], 0,3839	-0,17 [-0,61; 0,26]	0,0099

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
≥0,6 x ULN	69	12,71 (5,27)	-2,50 (0,41)	33	11,23 (4,54)	0,50 (0,60)	-3,00 [-4,41; -1,59], <0,0001	-0,87 [-1,30; -0,44]	
Monat 12									
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	10,36 (4,00)	-1,10 (0,53)	31	11,22 (4,83)	-0,07 (0,72)	-1,02 [-2,71; 0,66], 0,2292	-0,25 [-0,69; 0,18]	0,0213
≥0,6 x ULN	69	12,71 (5,27)	-3,18 (0,41)	33	11,23 (4,54)	0,38 (0,59)	-3,56 [-4,97; -2,15], <0,0001	-1,04 [-1,48; -0,60]	
Subdomäne Grad									
Monat 6									
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	2,05 (0,91)	-0,13 (0,09)	31	2,24 (1,06)	-0,03 (0,12)	-0,09 [-0,36; 0,17], 0,4851	-0,14 [-0,57; 0,30]	0,0138
≥0,6 x ULN	69	2,43 (1,12)	-0,51 (0,09)	33	2,23 (0,94)	0,09 (0,13)	-0,60 [-0,91; -0,29], 0,0002	-0,79 [-1,22; -0,36]	
Monat 12									
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	2,05 (0,91)	-0,13 (0,11)	31	2,24 (1,06)	-0,01 (0,15)	-0,12 [-0,47; 0,22], 0,4797	-0,14 [-0,58; 0,29]	0,0382

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
≥0,6 x ULN	69	2,43 (1,12)	-0,61 (0,09)	33	2,23 (0,94)	0,00 (0,13)	-0,61 [-0,93; -0,29], 0,0003	-0,79 [-1,22; -0,36]	
Subdomäne Dauer									
Monat 12									
Alter zum Screening									
<65 Jahre	99	1,62 (1,11)	-0,28 (0,08)	52	1,63 (1,16)	-0,06 (0,10)	-0,23 [-0,47; 0,02], 0,0722	-0,30 [-0,64; 0,04]	0,0413
≥65 Jahre	29	1,52 (1,13)	-0,31 (0,17)	12	1,29 (0,58)	0,58 (0,26)	-0,89 [-1,50; -0,28], 0,0055	-0,94 [-1,65; -0,24]	
Pruritus NRS									
<4	79	1,08 (0,29)	-0,01 (0,08)	42	1,10 (0,35)	0,04 (0,11)	-0,05 [-0,32; 0,22], 0,7043	-0,07 [-0,45; 0,30]	0,0227
≥4	49	2,43 (1,41)	-0,99 (0,14)	22	2,42 (1,40)	-0,28 (0,21)	-0,71 [-1,22; -0,20], 0,0075	-0,70 [-1,21; -0,18]	
Subdomäne Beeinträchtigung									
Monat 6									
Pruritus NRS									
<4	79	1,51 (0,85)	-0,18 (0,09)	42	1,51 (0,78)	0,02 (0,11)	-0,20 [-0,46; 0,07], 0,1463	-0,26 [-0,63; 0,12]	0,0147

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
≥4	49	3,50 (1,12)	-1,21 (0,16)	22	3,51 (0,80)	-0,24 (0,24)	-0,97 [-1,55; -0,40], 0,0012	-0,85 [-1,38; -0,33]	
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	2,05 (1,18)	-0,40 (0,11)	31	2,20 (1,23)	-0,21 (0,14)	-0,20 [-0,54; 0,14], 0,2472	-0,23 [-0,67; 0,20]	0,0493
≥0,6 x ULN	69	2,47 (1,48)	-0,58 (0,12)	33	2,24 (1,27)	0,14 (0,17)	-0,72 [-1,12; -0,32], 0,0006	-0,73 [-1,16; -0,31]	
Subdomäne Verlauf									
Monat 6									
Vorherige Einnahme von OCA und/oder Fibraten									
Ja	20	3,23 (1,31)	-0,26 (0,29)	12	3,32 (1,20)	-0,86 (0,36)	0,61 [-0,34; 1,55], 0,1990	0,47 [-0,26; 1,19]	0,0217
Nein	108	3,37 (1,05)	-0,69 (0,12)	52	3,08 (1,16)	-0,15 (0,17)	-0,55 [-0,94; -0,15], 0,0070	-0,44 [-0,77; -0,11]	
Monat 12									
Vorherige Einnahme von OCA und/oder Fibraten									
Ja	20	3,23 (1,31)	-0,36 (0,35)	12	3,32 (1,20)	-0,84 (0,42)	0,48 [-0,64; 1,60], 0,3849	0,31 [-0,41; 1,03]	0,0414

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Nein	108	3,37 (1,05)	-0,71 (0,13)	52	3,08 (1,16)	0,02 (0,20)	-0,73 [-1,19; -0,27], 0,0022	-0,52 [-0,85; -0,18]	
Subdomäne Lokalisation									
Monat 6									
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	1,81 (1,14)	-0,07 (0,12)	31	2,13 (1,31)	0,01 (0,16)	-0,08 [-0,44; 0,28], 0,6732	-0,08 [-0,52; 0,35]	0,0080
≥0,6 x ULN	69	2,45 (1,35)	-0,29 (0,11)	33	2,03 (1,10)	0,49 (0,16)	-0,78 [-1,16; -0,39], 0,0001	-0,83 [-1,26; -0,40]	
Monat 12									
Pruritus NRS									
<4	79	1,57 (0,93)	0,01 (0,10)	42	1,50 (0,77)	0,03 (0,14)	-0,02 [-0,34; 0,30], 0,9022	-0,02 [-0,40; 0,35]	0,0025
≥4	49	3,10 (1,23)	-0,79 (0,16)	22	3,13 (1,14)	0,24 (0,24)	-1,03 [-1,61; -0,44], 0,0009	-0,89 [-1,41; -0,36]	

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Modifizierter Gesamt-Score									
Monat 6									
Zirrhose									
Ja	18	9,72 (5,33)	-1,19 (0,71)	9	8,94 (3,20)	2,34 (0,97)	-3,53 [-6,06; -0,99], 0,0084	-1,14 [-2,01; -0,28]	0,0474
Nein	110	8,04 (4,02)	-1,38 (0,23)	55	7,96 (4,08)	-0,40 (0,31)	-0,99 [-1,71; -0,26], 0,0082	-0,42 [-0,74; -0,09]	
Pruritus NRS									
<4	79	5,78 (2,36)	-0,36 (0,24)	42	5,74 (2,04)	0,28 (0,31)	-0,63 [-1,36; 0,10], 0,0890	-0,30 [-0,68; 0,07]	0,0085
≥4	49	12,31 (3,43)	-3,54 (0,43)	22	12,41 (2,80)	-0,65 (0,64)	-2,89 [-4,43; -1,34], 0,0004	-0,94 [-1,47; -0,42]	
Gesamt Bilirubin 0,6									
<0,6 x ULN	59	7,21 (3,39)	-0,80 (0,31)	31	8,09 (4,17)	-0,44 (0,41)	-0,36 [-1,31; 0,58], 0,4474	-0,15 [-0,59; 0,28]	0,0049
≥0,6 x ULN	69	9,19 (4,69)	-1,78 (0,32)	33	8,11 (3,81)	0,64 (0,46)	-2,42 [-3,51; -1,32], <0,0001	-0,90 [-1,33; -0,47]	

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Monat 12									
Pruritus NRS									
<4	79	5,78 (2,36)	-0,34 (0,29)	42	5,74 (2,04)	0,73 (0,39)	-1,06 [-1,99; -0,14], 0,0250	-0,41 [-0,79; -0,04]	0,0123
≥4	49	12,31 (3,43)	-4,24 (0,41)	22	12,41 (2,80)	-1,03 (0,60)	-3,21 [-4,65; -1,77], <0,0001	-1,12 [-1,66; -0,59]	
Gesamt-Bilirubin 0,6									
<0,6 x ULN	59	7,21 (3,39)	-0,65 (0,38)	31	8,09 (4,17)	0,06 (0,53)	-0,71 [-1,95; 0,53], 0,2594	-0,24 [-0,67; 0,20]	0,0137
≥0,6 x ULN	69	9,19 (4,69)	-2,33 (0,32)	33	8,11 (3,81)	0,41 (0,45)	-2,74 [-3,82; -1,67], <0,0001	-1,04 [-1,48; -0,60]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar.									
a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.									
FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, NRS=Numerical Rating Scale, OCA=Obeticholsäure, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel									

Gesamt-Score

In dem Endpunkt 5-D Pruritus Gesamt-Score – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal vorherige Einnahme von OCA und/oder Fibraten ($p=0,0250$). Bei Patienten ohne vorherige Einnahme von OCA und/oder Fibraten zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -2,32 [-3,35; -1,29]; $p<0,0001$), der auch klinisch relevant (SMD [95%-KI]: -0,69 [-1,03; -0,35]) war. Bei Patienten mit vorheriger Einnahme von OCA und/oder Fibraten zu Baseline war der Effekt nicht statistisch signifikant. Entsprechend wird diese Effektmodifikation als nicht fazitrelevant erachtet.

In dem Endpunkt 5-D Pruritus Gesamt-Score – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Pruritus NRS ($p=0,0216$). Bei Patienten mit einer Pruritus NRS ≥ 4 zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -3,42 [-5,32; -1,51]; $p=0,0006$), der auch klinisch relevant (SMD [95%-KI]: -0,91 [-1,43; -0,38]) war. Bei Patienten mit einer Pruritus NRS < 4 zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

In dem Endpunkt 5-D Pruritus Gesamt-Score – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin 0,6 ($p=0,0099$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6$ zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -3,00 [-4,41; -1,59]; $p<0,0001$), der auch klinisch relevant (SMD [95%-KI]: -0,87 [-1,30; -0,44]) war. Bei Patienten mit einem Gesamt-Bilirubin-Wert $< 0,6$ zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

In dem Endpunkt 5-D Pruritus Gesamt-Score – stetige Analyse zeigte sich eine potenzielle Effektmodifikation zu Monat 12 für das Subgruppenmerkmal Gesamt-Bilirubin 0,6 ($p=0,0213$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6 \times \text{ULN}$ zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -3,56 [-4,97; -2,15]; $p<0,0001$), der auch klinisch relevant (SMD [95%-KI]: -1,04 [-1,48; -0,60]) war. Bei Patienten mit einem Gesamt-Bilirubin-Wert $< 0,6 \times \text{ULN}$ zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

Subdomäne Grad

In dem Endpunkt 5-D Pruritus Subdomäne Grad – stetige Analyse zeigte sich sowohl zu Monat 6 als auch zu Monat 12 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin 0,6 (Monat 6: $p=0,0138$, Monat 12: $0,0382$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6$ zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (Monat 6: LS-MWD [95%-KI]: -0,60 [-0,91; -0,29]; $p=0,0002$, Monat 12: LS-MWD [95%-KI]: -0,61 [-0,93; -0,29]; $p=0,0003$), der auch klinisch relevant (Monat 6: SMD [95%-KI]: -0,79 [-1,22; -0,36], Monat 12: SMD [95%-KI]: -0,79 [-1,22; -0,36]) war. Bei Patienten mit einem Gesamt-Bilirubin-Wert $< 0,6$ zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

Subdomäne Dauer

In dem Endpunkt 5-D Pruritus Subdomäne Dauer – stetige Analyse zeigte sich zu Monat 12 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Alter zum Screening ($p=0,0413$). Bei Patienten mit mindestens 65 Jahren zum Screening ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,89 [-1,50; -0,28]; $p=0,0055$), der auch klinisch relevant (SMD [95%-KI]: -0,94 [-1,65; -0,24]) war. Bei Patienten jünger als 65 Jahre zum Screening war der Effekt gleichgerichtet, jedoch, nicht statistisch signifikant. Diese Effektmodifikation wird als nicht fazitrelevant erachtet.

In dem Endpunkt 5-D Pruritus Subdomäne Dauer – stetige Analyse zeigte sich zu Monat 12 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Pruritus NRS ($p=0,0227$). Bei Patienten mit einer Pruritus NRS ≥ 4 zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,71 [-1,22; -0,20]; $p=0,0075$), der jedoch nicht klinisch relevant war. Bei Patienten mit einer Pruritus NRS < 4 zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

Subdomäne Beeinträchtigung

In dem Endpunkt 5-D Pruritus Subdomäne Beeinträchtigung – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Pruritus NRS ($p=0,0147$). Bei Patienten mit einer Pruritus NRS ≥ 4 zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,97 [-1,55; -0,40]; $p=0,0012$), der auch klinisch relevant (SMD [95%-KI]: -0,85 [-1,38; -0,33]) war. Bei Patienten mit einer Pruritus NRS < 4 zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

In dem Endpunkt 5-D Pruritus Subdomäne Beeinträchtigung – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin 0,6 ($p=0,0493$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6$ zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,72 [-1,12; -0,32]; $p=0,0006$), der auch klinisch relevant (SMD [95%-KI]: -0,73 [-1,16; -0,31]) war. Bei Patienten mit einem Gesamt-Bilirubin-Wert $< 0,6$ zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

Subdomäne Verlauf

In dem Endpunkt 5-D Pruritus Subdomäne Verlauf – stetige Analyse zeigte sich sowohl zu Monat 6 als auch zu Monat 12 eine potenzielle Effektmodifikation für das Subgruppenmerkmal vorherige Einnahme von OCA und/oder Fibraten (Monat 6: $p=0,0217$, Monat 12: $p=0,0414$). Bei Patienten ohne vorherige Einnahme zu Baseline ergab sich jeweils ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (Monat 6: LS-MWD [95%-KI]: -0,55 [-0,94; -0,15]; $p=0,0070$, Monat 12: LS-MWD [95%-KI]: -0,73 [-1,19; -0,27]; $p=0,0022$), zu beiden Zeitpunkten war der Effekt nicht klinisch relevant. Bei Patienten mit vorheriger Einnahme von OCA und/oder Fibraten zu Baseline war der Effekt zu beiden Zeitpunkten nicht statistisch signifikant. Diese Effektmodifikation wird als nicht fazitrelevant erachtet.

Subdomäne Lokalisation

In dem Endpunkt 5-D Pruritus Subdomäne Lokalisation – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin $\geq 0,6$ ($p=0,0080$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6$ zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,78 [-1,16; -0,39]; $p<0,0001$), der auch klinisch relevant (SMD [95%-KI]: -0,83 [-1,26; -0,40]) war. Bei Patienten mit einem Gesamt-Bilirubin-Wert $<0,6$ zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

Des Weiteren zeigte sich in dem Endpunkt 5-D Pruritus Subdomäne Lokalisation stetige Analyse zu Monat 12 eine Effektmodifikation für das Subgruppenmerkmal Pruritus NRS zu Baseline ($p=0,0025$). Bei Patienten mit einer Pruritus NRS ≥ 4 zu Baseline ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -1,03 [-1,61; -0,44]; $p=0,0009$), der auch klinisch relevant (SMD [95%-KI]: -0,89 [-1,41; -0,36]) war. Bei Patienten mit einer Pruritus NRS <4 Punkte zu Baseline war der Effekt gleichgerichtet, jedoch nicht statistisch signifikant.

Modifizierter Gesamt-Score

In dem Endpunkt 5-D Pruritus Modifizierter Gesamt-Score – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Zirrhose ($p=0,0474$). Bei Patienten mit Zirrhose ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -3,53 [-6,06; -0,99]; $p=0,0084$), der auch klinisch relevant (SMD [95%-KI]: -1,14 [-2,01; -0,28]) war. Bei Patienten ohne Zirrhose ergab sich ebenfalls ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -0,99 [-1,71; -0,26]; $p=0,0082$), der jedoch nicht klinisch relevant war. Diese Effektmodifikation wird als nicht fazitrelevant erachtet.

Auch in dem Endpunkt 5-D Pruritus Modifizierter Gesamt-Score – stetige Analyse zeigte sich sowohl zu Monat 6 als auch zu Monat 12 eine Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin $\geq 0,6$ (Monat 6: $p=0,0049$, Monat 12: $p=0,0137$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6 \times \text{ULN}$ zu Baseline ergab sich jeweils ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (Monat 6: LS-MWD [95%-KI]: -2,42 [-3,51; -1,32]; $p<0,0001$, Monat 12: LS-MWD [95%-KI]: -2,74 [-3,82; -1,67]; $p<0,0001$), der zu beiden Zeitpunkten auch klinisch relevant war (Monat 6: SMD [95%-KI]: -0,90 [-1,33; -0,47], Monat 12: SMD [95%-KI]: -1,04 [-1,48; -0,60]). Für Patienten mit einem Gesamt-Bilirubin-Wert $<0,6 \times \text{ULN}$ zu Baseline war der Effekt jeweils gleichgerichtet, jedoch nicht statistisch signifikant.

Die Effektmodifikation für das Subgruppenmerkmal Pruritus NRS zu Baseline (Monat 6: $p=0,0085$, Monat 12: $p=0,0123$) zeigte sich sowohl zu Monat 6 als auch zu Monat 12 auch in dem Endpunkt 5-D Pruritus Modifizierter Gesamt-Score – stetige Analyse. Bei Patienten mit einer Pruritus NRS ≥ 4 Punkte zu Baseline ergab sich jeweils ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (Monat 6: LS-MWD [95%-KI]: -2,89 [-4,43; -1,34]; $p=0,0004$, Monat 12: LS-MWD [95%-KI]: -3,21 [-4,65; -1,77]; $p<0,0001$), der auch klinisch

relevant (Monat 6: SMD [95%-KI]: -0,94 [-1,47; -0,42], Monat 12: SMD [95%-KI]: -1,12 [-1,66; -0,59]) war. Bei Patienten mit einer Pruritus NRS <4 Punkte zu Baseline war der Effekt zu beiden Zeitpunkten gleichgerichtet, wobei zu Monat 12 eine statistische Signifikanz vorlag (LS-MWD [95%-KI]: -1,06 [-1,99; -0,14]; $p=0,0250$), die nicht klinisch relevant war.

Für die Subgruppenmerkmale Gesamt-Bilirubin 0,6 und Pruritus NRS treten vermehrt statistisch signifikante Ereignisse zum 5-D Pruritus auf. Die unterschiedlichen Effektgrößen sind zwischen den Subgruppenausprägungen in der Art der Analyse begründet, außerdem ist ein höherer Bilirubin-Wert mit einer erhöhten Pruritus NRS assoziiert (vergleiche Abschnitt 4.3.1.3.2.8). Entsprechend wird die Effektmodifikation zur Pruritus NRS daher als fazitrelevant, alle weiteren Effektmodifikationen als nicht fazitrelevant eingestuft.

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

4.3.1.3.2.4 PGI-C & PGI-S – RCT

Tabelle 4-60: Ergebnisse für Subgruppenanalysen zum Endpunkt PGI-S – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12

Subgruppe	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI] p-Wert	SMD [95%-KI]	Interaktion
Monat 12									
Zirrhose									
Ja	18	2,44 (0,92)	-0,46 (0,19)	9	2,44 (1,13)	-1,13 (0,35)	0,68 [-0,14; 1,50], 0,0988	0,73 [-0,10; 1,56]	0,0463
Nein	107	2,08 (0,98)	-0,35 (0,07)	53	2,11 (0,92)	-0,22 (0,10)	-0,13 [-0,36; 0,10], 0,2611	-0,18 [-0,51; 0,15]	
ALP									
<350 U/L	91	1,96 (0,90)	-0,23 (0,07)	44	2,07 (0,95)	-0,30 (0,10)	0,07 [-0,17; 0,31], 0,5728	0,10 [-0,26; 0,46]	0,0221
≥350 U/L	34	2,60 (1,03)	-0,65 (0,14)	18	2,39 (0,92)	-0,12 (0,19)	-0,53 [-1,01; -0,06], 0,0271	-0,66 [-1,24; -0,07]	
Gesamt-Bilirubin 1									
≤1 x ULN	105	2,10 (0,98)	-0,34 (0,08)	58	2,17 (0,95)	-0,28 (0,10)	-0,06 [-0,29; 0,17], 0,6141	-0,08 [-0,40; 0,24]	0,0257
>1 x ULN	20	2,30 (0,98)	-0,45 (0,16)	4	2,00 (1,00)	0,75 (0,47)	-1,21 [-2,26; -0,16], 0,0261	-1,52 [-2,70; -0,35]	

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

Subgruppe	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI] p-Wert	SMD [95%-KI]	Interaktion
Gesamt-Bilirubin 0,6									
<0,6 x ULN	58	1,90 (0,84)	-0,08 (0,10)	31	2,16 (0,92)	-0,30 (0,13)	0,22 [-0,08; 0,52], 0,1520	0,28 [-0,16; 0,72]	0,0033
≥0,6 x ULN	67	2,33 (1,04)	-0,58 (0,09)	31	2,16 (0,99)	-0,16 (0,13)	-0,42 [-0,73; -0,11], 0,0085	-0,57 [-1,00; -0,13]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar.									
a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.									
ALP=Alkalische Phosphatase, FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, PGI-S= Patient Global Impression of Severity RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel									

Für den Endpunkt PGI-S – stetige Analyse zeigte sich zu Monat 12 eine potenzielle Effektmodifikation für das Merkmal Zirrhose ($p=0,0463$). In keiner der Subgruppenausprägungen zeigte sich ein statistisch signifikanter Unterschied. Entsprechend wird diese Effektmodifikation als nicht fazitrelevant erachtet.

Für den Endpunkt PGI-S – stetige Analyse zeigte sich zu Monat 12 eine weitere potenzielle Effektmodifikation für das Merkmal ALP zu Baseline ($p=0,0221$). Bei Patienten mit einem $ALP \geq 350$ U/L zeigte sich ein statistisch signifikanter Unterschied zugunsten von SEL (LS-MWD [95%-KI]: -0,53 [-1,01 -0,06], 0,0271). Der Unterschied war jedoch nicht klinisch relevant. Bei Patienten mit einem $ALP < 350$ U/L war der Unterschied nicht statistisch signifikant. Entsprechend wird diese Effektmodifikation als nicht fazitrelevant erachtet.

Für den Endpunkt PGI-S – stetige Analyse zeigte sich zu Monat 12 eine weitere potenzielle Effektmodifikation für das Merkmal Gesamt-Bilirubin 1 ($p=0,0257$). Bei Patienten mit einem Gesamt-Bilirubin-Wert >1 x ULN zu Baseline zeigte sich ein statistisch signifikanter Unterschied zugunsten von SEL (LS-MWD [95%-KI]: -1,21 [-2,26 -0,16], 0,0261), der auch klinisch relevant (SMD [95%-KI]: -1,52 [-2,70; 0,35]) war. Bei Patienten mit einem Gesamt-Bilirubin-Wert ≤ 1 x ULN war der Unterschied nicht statistisch signifikant. Des Weiteren zeigte sich auch für die Analyse des Gesamt-Bilirubins mit dem Cut-off 0,6 eine weitere potenzielle Effektmodifikation ($p=0,0033$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6$ x ULN zeigte sich ein statistisch signifikanter Unterschied zugunsten von SEL (LS-MWD [95%-KI]: -0,42 [-0,73 -0,11], 0,0085). Der Unterschied war jedoch nicht klinisch relevant. Bei Patienten mit einem Gesamt-Bilirubin-Wert $< 0,6$ x ULN war der Unterschied nicht statistisch signifikant. Beide Effektmodifikation werden als nicht fazitrelevant erachtet (vergleiche Abschnitt 4.3.1.3.2.8).

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

4.3.1.3.2.5 Fibrose – RCT

Tabelle 4-61: Ergebnisse für Subgruppenanalysen zum Endpunkt Lebersteifigkeit (FibroScan®) aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12

Subgruppe	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI] p-Wert	SMD [95%-KI]	Interaktion
Monat 6									
Vorherige Einnahme von OCA und/oder Fibraten									
Ja	17	11,37 (8,68)	1,44 (0,95)	12	7,89 (3,09)	-1,58 (1,36)	3,02 [-0,41; 6,45], 0,0815	0,69 [-0,07; 1,45]	0,0107
Nein	92	9,54 (5,54)	0,85 (0,50)	47	8,96 (4,43)	2,49 (0,65)	-1,64 [-3,12; -0,16], 0,0297	-0,35 [-0,70; 0,00]	
Gesamt-Bilirubin 0,6									
<0,6 x ULN	54	8,17 (5,82)	0,90 (0,54)	30	7,46 (3,15)	0,29 (0,63)	0,61 [-0,80; 2,02], 0,3933	0,16 [-0,29; 0,61]	0,0109
≥0,6 x ULN	55	11,43 (6,10)	0,92 (0,71)	29	10,02 (4,72)	3,83 (0,98)	-2,91 [-5,27; -0,55], 0,0163	-0,55 [-1,01; -0,09]	
Monat 12									
Alter zum Zeitpunkt der PBC Diagnose									
<50 Jahre	50	10,56 (7,19)	0,58 (0,64)	28	8,74 (3,68)	-0,09 (0,87)	0,68 [-1,40; 2,75], 0,5189	0,15 [-0,32; 0,61]	0,0387

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

Subgruppe	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI] p-Wert	SMD [95%-KI]	Interaktion
≥50 Jahre	59	9,21 (5,05)	0,38 (0,82)	31	8,74 (4,70)	3,05 (1,04)	-2,66 [-5,11; -0,21], 0,0336	-0,43 [-0,87; 0,01]	
Vorherige Einnahme von OCA und/oder Fibraten									
Ja	17	11,37 (8,68)	1,25 (0,89)	12	7,89 (3,09)	-0,95 (1,25)	2,20 [-1,00; 5,40], 0,1680	0,54 [-0,21; 1,29]	0,0283
Nein	92	9,54 (5,54)	0,54 (0,59)	47	8,96 (4,33)	2,28 (0,77)	-1,55 [-3,55; 0,07], 0,0600	-0,31 [-0,66; 0,04]	
Mono-/Kombitherapie									
Monotherapie	8	8,11 (2,84)	2,25 (0,25)	4	10,45 (4,53)	1,30 (0,66)	0,95 [-0,76; 2,66], 0,2309	0,93 [-0,35; 2,22]	0,0328
Kombinationstherapie	101	9,97 (6,33)	0,50 (0,61)	55	8,62 (4,17)	2,08 (0,77)	-1,58 [-3,43; 0,27], 0,0945	-0,26 [-0,59; 0,07]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar.									
a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline.									
FAS=Full-Analysis-Set, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, OCA=Obeticholsäure, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel									

Für den Endpunkt Lebersteifigkeit (FibroScan®) zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal vorherige Einnahme von OCA und/oder Fibraten ($p=0,0107$). Bei Patienten ohne vorherige Einnahme (LS-MWD [95%-KI]: -1,64 [-3,12; -0,16], 0,0297) zeigte sich ein statistisch signifikanter Unterschied zugunsten von SEL. Der Unterschied war jedoch nicht klinisch relevant (SMD [95%-KI]: -0,35 [-0,70; 0,00]). Bei Patienten mit vorheriger Einnahme von OCA und/oder Fibraten war der Unterschied nicht statistisch signifikant. Entsprechend wird diese Effektmodifikation als nicht fazitrelevant erachtet.

Für den Endpunkt Lebersteifigkeit (FibroScan®) zeigte sich zu Monat 6 eine weitere potenzielle Effektmodifikation für das Subgruppenmerkmal Gesamt-Bilirubin 0,6 ($p=0,0109$). Bei Patienten mit einem Gesamt-Bilirubin-Wert $\geq 0,6 \times \text{ULN}$ (LS-MWD [95%-KI]: -2,91 [-5,27; -0,55], 0,0163) zeigte sich ein statistisch signifikanter Unterschied zugunsten von SEL, welcher nicht klinisch relevant war (SMD [95%-KI]: -0,55 [-1,01; -0,09]). Bei Patienten mit einem Gesamt-Bilirubin-Wert $< 0,6 \times \text{ULN}$ zu Baseline war der Unterschied nicht statistisch signifikant. Entsprechend wird diese Effektmodifikation als nicht fazitrelevant erachtet (vergleiche Abschnitt 4.3.1.3.2.8).

Für den Endpunkt Lebersteifigkeit (FibroScan®) zeigte sich zu Monat 12 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Alter zum Zeitpunkt der PBC-Diagnose ($p=0,0387$). Bei Patienten, die mindestens 50 Jahre zur PBC-Diagnose alt waren, zeigte sich ein statistisch signifikanter Unterschied zuungunsten von SEL (LS-MWD [95%-KI]: -2,66 [-5,11; -0,21], 0,0336). Der Unterschied war jedoch nicht klinisch relevant (SMD [95%-KI]: -0,43 [-0,87; 0,01]). Bei Patienten jünger als 50 Jahre zur PBC-Diagnose war der Unterschied nicht statistisch signifikant. Entsprechend wird diese Effektmodifikation als nicht fazitrelevant erachtet.

Für den Endpunkt Lebersteifigkeit (FibroScan®) zeigte sich zu Monat 12 eine weitere potenzielle Effektmodifikation für das Merkmal vorherige Einnahme von OCA und/oder Fibrate ($p=0,0283$). In der Subgruppenausprägung mit vorheriger Einnahme war eine deutlich geringere Patientenzahl als in der Subgruppenkategorie ohne vorherige Einnahme von OCA und/oder Fibraten ($N=29$ vs. $N=139$), was die Aussagekraft der Analyse stark eingeschränkt. Da sich zusätzlich in keiner der Subgruppenausprägungen ein statistisch signifikanter Behandlungsunterschied zeigte, wird die Effektmodifikation als nicht fazitrelevant angesehen.

Für den Endpunkt Lebersteifigkeit zeigte sich zu Monat 12 eine weitere potenzielle Effektmodifikation für das Subgruppenmerkmal Mono-/Kombitherapie ($p=0,0328$). Keine der Subgruppenausprägungen war zwischen den Behandlungsgruppen statistisch signifikant unterschiedlich. Zusätzlich war die Patientenzahl in der Subgruppenkategorie Monotherapie deutlich geringer als in der Subgruppenkategorie Kombinationstherapie ($N=12$ vs. $N=156$). Dadurch ist die Aussagekraft der Analyse stark eingeschränkt und folglich als nicht fazitrelevant eingestuft.

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

4.3.1.3.2.6 Gesundheitsbezogene Lebensqualität PBC-40 – RCT

Tabelle 4-62: Ergebnisse Subgruppenanalysen zum Endpunkt für PBC-40 – stetige Analyse aus RCT mit dem zbAM (Studie RESPONSE, FAS) zu Monat 12

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Gesamt-Score									
Monat 12									
Region									
Nordamerika	49	92,78 (30,06)	-2,43 (2,46)	12	93,42 (32,95)	-6,10 (4,63)	3,68 [-6,28; 13,64], 0,4611	0,21 [-0,42; 0,85]	0,0233
Europa	39	91,88 (28,96)	-8,64 (2,46)	23	90,23 (25,87)	1,27 (3,18)	-9,91 [-17,76; -2,06], 0,0143	-0,64 [-1,17; -0,11]	
Restliche Welt	38	89,94 (29,38)	-6,34 (3,39)	28	91,55 (31,62)	-11,62 (3,85)	5,28 [-4,75; 15,31], 0,2964	0,25 [-0,24; 0,74]	
Kognitive Domäne									
Monat 6									
Mono-/Kombitherapie									
Monotherapie	8	14,00 (7,20)	-0,82 (0,70)	4	11,00 (3,83)	3,08 (1,23)	-3,91 [-7,85; 0,04], 0,0514	-1,69 [-3,15; -0,23]	0,0218
Kombinationstherapie	118	13,17 (5,41)	-0,36 (0,30)	59	12,96 (4,95)	-0,04 (0,42)	-0,32 [-1,32; 0,68], 0,5301	-0,10 [-0,41; 0,22]	

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Pruritus NRS									
<4	79	11,91 (5,32)	0,37 (0,38)	41	11,17 (4,56)	0,04 (0,50)	0,34 [-0,84; 1,51], 0,5710	0,10 [-0,28; 0,48]	0,0207
≥4	47	15,34 (5,18)	-1,37 (0,44)	22	15,89 (3,95)	0,61 (0,68)	-1,98 [-3,59; -0,36], 0,0172	-0,64 [-1,16; -0,12]	
GGT									
≤3 x ULN	32	13,15 (5,55)	-0,02 (0,75)	13	12,79 (4,81)	-1,70 (1,07)	1,67 [-0,53; 3,87], 0,1325	0,39 [-0,26; 1,04]	0,0314
>3 x ULN	94	13,24 (5,52)	-0,47 (0,32)	50	12,85 (4,95)	0,47 (0,45)	-0,94 [-2,02; 0,14], 0,0863	-0,30 [-0,64; 0,05]	
Monat 12									
Region									
Nordamerika	49	14,37 (6,18)	-0,57 (0,47)	12	13,04 (5,05)	-1,94 (0,90)	1,37 [-0,55; 3,30], 0,1583	0,42 [-0,22; 1,05]	0,0357
Europa	39	12,94 (5,08)	-0,68 (0,56)	23	12,73 (4,17)	1,20 (0,72)	-1,88 [-3,67; -0,09], 0,0403	-0,53 [-1,06; -0,01]	
Restliche Welt	38	12,03 (4,80)	-0,67 (0,66)	28	12,84 (5,51)	-1,11 (0,74)	0,44 [-1,52; 2,40], 0,6525	0,11 [-0,38; 0,60]	

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

Endpunkt	Seladelpar			Placebo			Seladelpar vs. Placebo		
	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	N ^a	Baseline MW (SD)	Differenz zu Baseline LS-MW (SE)	LS-MWD [95%-KI], p-Wert	SMD [95%-KI]	Interaktion
Fatigue Domäne									
Monat 12									
Region									
Nordamerika	49	28,12 (9,82)	-0,49 (1,01)	12	29,46 (11,80)	-0,59 (1,88)	0,10 [-3,93; 4,13], 0,9589	0,01 [-0,62; 0,65]	0,0307
Europa	39	28,26 (9,97)	-3,00 (1,17)	23	26,08 (10,32)	2,06 (1,52)	-5,06 [-8,85; -1,28], 0,0097	-0,68 [-1,21; -0,15]	
Restliche Welt	38	26,19 (10,41)	-2,43 (1,29)	28	27,59 (10,59)	-4,15 (1,46)	1,72 [-2,08; 5,53], 0,3690	0,22 [-0,27; 0,71]	
Soziale Domäne									
Monat 6									
Pruritus NRS									
<4	79	20,29 (8,18)	0,20 (0,56)	41	18,98 (6,40)	-2,05 (0,75)	2,26 [0,50; 4,01], 0,0122	0,46 [0,08; 0,84]	0,0288
≥4	47	26,98 (7,16)	-2,11 (0,76)	22	28,11 (8,29)	-0,75 (1,18)	-1,36 [-4,16; 1,44], 0,3343	-0,25 [-0,76; 0,26]	
Ein LS-MWD & SMD <0 bedeutet einen Vorteil für die Behandlung mit Seladelpar. a: N beschreibt die Anzahl der Patienten, die in das MMRM eingeschlossen wurden, d. h. alle Patienten mit Baseline-Werten und mindestens einer Visite nach Baseline. FAS=Full-Analysis-Set, GGT= Gamma-Glutamyltransferase, KI=Konfidenzintervall, LS-MW=Kleinste-Quadrate-Mittelwert, LS-MWD= Kleinste-Quadrate-Mittelwertdifferenz, MMRM=Mixed Effect Model Repeat Measurement, MW=Mittelwert, N=Anzahl der Patienten im jeweiligen Behandlungsarm, NRS=Numerical Rating Scale, PBC=Primär biliäre Cholangitis, RCT=Randomisierte kontrollierte Studie, SD=Standardabweichung, SE=Standardfehler, SMD=Standardisierte Mittelwertdifferenz, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel									

Gesamt-Score

In dem Endpunkt PBC-40 Gesamt-Score – stetige Analyse zeigte sich eine Effektmodifikation für das Subgruppenmerkmal Region ($p=0,0233$). Bei Patienten aus Europa ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -9,91 [-17,76; -2,06]; $p=0,0143$), der jedoch nicht klinisch relevant war. Bei Patienten aus Nordamerika und der restlichen Welt war der Unterschied nicht statistisch signifikant.

Kognitive Domäne

Für den Endpunkt PBC-40 Kognitive Domäne – stetige Analyse zeigte sich zu Monat 6 eine potenzielle Effektmodifikation für das Subgruppenmerkmal Mono-/Kombitherapie. Da sich in keiner der Subgruppenausprägungen ein signifikanter Behandlungsunterschied zeigte, wird die Effektmodifikation als nicht fazitrelevant angesehen.

Für den Endpunkt PBC-40 Kognitive Domäne – stetige Analyse zeigte sich zu Monat 6 eine Effektmodifikation durch das Subgruppenmerkmal Pruritus NRS. Bei Patienten mit einer Pruritus NRS ≥ 4 zu Baseline zeigte sich ein statistisch signifikanter Vorteil zugunsten von SEL (LS-MWD [95%-KI]: -1,98 [-3,59; -0,36]; $p=0,0172$), der jedoch nicht klinisch relevant war. Die Effektmodifikation zur Pruritus NRS wird als fazitrelevant eingestuft.

Für den Endpunkt PBC-40 Kognitive Domäne – stetige Analyse zeigte sich zu Monat 6 eine weitere Effektmodifikation für das Subgruppenmerkmal GGT. Da sich in keiner der Subgruppenausprägungen ein signifikanter Behandlungsunterschied zeigte, wird die Effektmodifikation als nicht fazitrelevant angesehen.

Für den Endpunkt PBC-40 Kognitive Domäne – stetige Analyse zeigte sich zu Monat 12 ebenso eine Effektmodifikation für das Subgruppenmerkmal Region ($p=0,0357$). Bei Patienten aus Europa ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -1,88 [-3,67; -0,09]; $p=0,0403$), der jedoch nicht klinisch relevant war. Bei Patienten aus Nordamerika und der restlichen Welt war der Unterschied nicht statistisch signifikant.

Fatigue Domäne

Auch in dem Endpunkt PBC-40 Fatigue Domäne – stetige Analyse zeigte sich eine Effektmodifikation für das Subgruppenmerkmal Region ($p=0,0307$). Bei Patienten aus Europa ergab sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter Vorteil für SEL (LS-MWD [95%-KI]: -5,06 [-8,85; -1,28]; $p=0,0097$), der jedoch nicht klinisch relevant war. Bei Patienten aus Nordamerika und der restlichen Welt war der Unterschied nicht statistisch signifikant. Zusammengekommen (Gesamt-Score, Kognitive Domäne sowie Fatigue Domäne) wurde der Effekt jeweils durch die Veränderung im Placebo-Arm getrieben. So wies der Placebo-Arm in der Subgruppenausprägung Europa in allen drei Endpunkten einen positiven LS-MW auf, gegensätzlich zu den anderen beiden Subgruppenausprägungen. Da sich in keiner der Subgruppen ein klinisch relevanter Unterschied zeigte und sich die Aussagen zum Zusatznutzen zwischen den Regionen nicht unterscheiden, wird die Effektmodifikation als nicht fazitrelevant angesehen.

Soziale Domäne

Für den Endpunkt PBC-40 Soziale Domäne – stetige Analyse zeigte sich zu Monat 6 eine Effektmodifikation für das Subgruppenmerkmal Pruritus NRS. Bei Patienten mit einer Pruritus NRS <4 zu Baseline zeigte sich ein statistisch signifikanter Unterschied zuungunsten von SEL (LS-MWD [95%-KI]: 2,26 [0,50; 4,01]; $p=0,0122$), der jedoch nicht klinisch relevant war. Die Effektmodifikation zur Pruritus NRS wird als fazitrelevant eingestuft.

4.3.1.3.2.7 Unerwünschte Ereignisse – RCT

Tabelle 4-63: Ergebnisse Subgruppenanalysen zum Endpunkt für Kardiovaskuläre Toxizität (UESI) aus RCT mit dem zbAM (Studie RESPONSE, SAS)

Subgruppe	Seladelpar	Placebo	Seladelpar vs. Placebo	
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI], p-Wert, OR [95%-KI], p-Wert, RD [95%-KI], p-Wert	Interaktion
GGT				
≤3 x ULN	0/33 (0,0)	2/14 (14,3)	0,09 [0,00; 1,73], 0,1097 0,07 [0,00; 1,66], 0,1014 -0,14 [-0,33; 0,04], 0,1266	0,0458
>3 x ULN	13/95 (13,7)	3/51 (5,9)	2,33 [0,69; 7,79], 0,1709 2,54 [0,69; 9,35], 0,1621 0,08 [-0,02; 0,17], 0,1059	
Ein RR <1 bedeutet einen Vorteil für die Behandlung mit Seladelpar. GGT=Gamma-Glutamyltransferase, KI=Konfidenzintervall, N=Anzahl der Patienten im jeweiligen Behandlungsarm, n=Anzahl der Patienten mit Ereignis, OR=Odds Ratio, RCT=Randomisierte kontrollierte Studie, RD=Risikodifferenz, RR=Relatives Risiko, SAS=Safety-Analysis-Set, UESI=Unerwünschtes Ereignis von speziellem Interesse, ULN=Oberer Normwert, zbAM=Zu bewertendes Arzneimittel				

Für den Endpunkt Kardiovaskuläre Toxizität (UESI) ergab sich zu Monat 12 eine Effektmodifikation für das Subgruppenmerkmal GGT ($p=0,0458$). Da sich in keiner der Subgruppenausprägungen ein signifikanter Behandlungsunterschied zeigte, wird die Effektmodifikation als nicht fazitrelevant angesehen.

Es traten keine weiteren statistisch signifikanten Interaktionen für die Subgruppenanalysen der UE auf.

4.3.1.3.2.8 Fazit Subgruppenanalysen

Zu beachten ist, dass unter Annahme der Unabhängigkeit im Mittel 5% der durchgeführten Interaktionstests zu einem falsch positiven Ergebnis ($p<0,05$) führen. Es wurden mehr als 570 Subgruppenanalysen durchgeführt. Bei einem Signifikanzniveau von $\alpha=0,05$ ist somit mit

ca. 29 falsch positiven Ergebnissen (Fehler 1. Art) zu rechnen. Insgesamt wiesen 41 Subgruppenanalysen einen statistisch signifikanten Interaktionstest auf.

Auffällig ist das gehäufte Auftreten der Effektmodifikationen zu den Subgruppenanalysen der Merkmale Gesamt-Bilirubin (beide Cut-offs 0,6 und 1) sowie Pruritus NRS. Zum Gesamt-Bilirubin 1 traten diese in den Endpunkten ALP – stetige Analyse und PGI-S auf; für Gesamt-Bilirubin 0,6 in Pruritus NRS, 5-D Pruritus (Gesamt-Score Subdomäne Grad Subdomäne Beeinträchtigung, Subdomäne Lokalisation und modifizierter Gesamt-Score), PGI-S sowie bei der Lebersteifigkeit (FibroScan®). Die Effektmodifikationen zur Subgruppe Pruritus NRS sind in den Endpunkten Pruritus NRS, 5-D Pruritus (Gesamt-Score, Subdomäne Dauer, Subdomäne Beeinträchtigung, Subdomäne Lokalisation und modifizierter Gesamt-Score) und dem PBC-40 (Kognitive Domäne und Soziale Domäne) aufgetreten. Über die Endpunkte hinweg zeigten sich insbesondere Vorteile für Patienten unter SEL-Behandlung mit einem höheren Gesamt-Bilirubin-Wert zu Baseline ($\geq 0,6 \times \text{ULN}$ bzw. $\geq 1 \times \text{ULN}$) bzw. einem erhöhten Pruritus (≥ 4 Punkte in der Pruritus NRS) zu den Monaten 6 und 12. Die unterschiedlichen Effektgrößen sind zwischen den Subgruppenausprägungen in der Art der Analyse begründet. Da die Patienten in der Subgruppe Pruritus NRS < 4 Punkte zu Baseline bereits definitionsbedingt geringere Werte auf der Pruritus NRS aufweisen, können sie sich hier weniger verbessern als Patienten, deren Pruritus NRS-Werte zu Baseline einen moderaten bis starken Pruritus zeigen. Des Weiteren sind erhöhte Werte des Gesamt-Bilirubins mit Pruritus assoziiert. Da die Gesamt-Bilirubin Werte der Patienten in der Subgruppe $< 0,6 \times \text{ULN}$ zu Baseline bereits nah am Normalbereich liegen und damit mit keinem Pruritus zu assoziieren sind, ist hier der Effekt kleiner als bei Patienten deren Gesamt-Bilirubin Werte mit einem stärkeren Pruritus assoziiert wird. Entsprechend werden die Effektmodifikationen zum Gesamt-Bilirubin (beide Cut-off: 0,6 und 1) als nicht fazitrelevant eingestuft. Dahingegen wird die Effektmodifikation zur Pruritus NRS als fazitrelevant eingestuft und entsprechend in der Bewertung des Zusatznutzens herangezogen.

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).

RESPONSE

A Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Efficacy and Safety of Seladelpar in Patients with Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA)

Quellen:

- Studienbericht [51]
- Registereinträge [48-50]
- Publikationen [47]

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

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

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

Nicht zutreffend.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

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

Nicht zutreffend.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-64: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in drei Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

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

Tabelle 4-65: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs- therapie 1>	<Vergleichs- therapie 2>	<Vergleichs- therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2>	•		•	
	<Studie 3>	•		•	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

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

Tabelle 4-66: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-67: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

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

Tabelle 4-68: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere inkonsistente Ergebnisse.*

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

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt für den ein indirekter Vergleich vorgenommen wird fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht zutreffend.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie***

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

Nicht zutreffend.

4.3.2.2.2 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 4-69: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-70: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	<Mortalität>	<Gesundheits-bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-71: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-72: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie

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

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-73: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-74: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind,

gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.2 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.

Evidenzklassifizierung

Die Vollständigkeit des Studienpools wurde durch eine umfassende systematische bibliografische Literaturrecherche und Studienregistersuche bestätigt (siehe Abschnitte 4.2.3.2, 4.2.3.3, Anhang 4-A und Anhang 4-B). Außerdem sind alle in die Nutzenbewertung

eingeschlossenen Endpunkte valide und patientenrelevante Endpunkte, die auch von der EMA als solche herangezogen wurden. Die relevanten Endpunkte wurden in allen für die Zusatznutzenableitung herangezogenen Studien gemäß aktuellen Leitlinien und Standards erfasst und bewertet.

Für die Bewertung des Zusatznutzens von SEL gegenüber Placebo wurde eine RCT des pharmazeutischen Unternehmers (RESPONSE) mit dem zbAM eingeschlossen. Bei der Studie handelt es sich um eine internationale, multizentrische, randomisierte klinische Phase-III-Studie. Es wurden jeweils 128 Patienten auf den SEL-Arm und 65 Placebo-Arm randomisiert. Die Phase-III-Studie RESPONSE entspricht aufgrund ihres randomisierten, kontrollierten Designs der Evidenzstufe Ib [11].

Studienqualität/ Verzerrungspotenzial

Die Studie RESPONSE wurde verblindet durchgeführt und ist somit niedrig verzerrt. Es gibt keine Hinweise auf ergebnisabhängige Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Aspekte.

Die in die Nutzenbewertung eingeschlossenen Endpunkte sind valide und patientenrelevant, sie wurden adäquat unter Verwendung des ITT-Prinzips ausgewertet. Auch hier ergeben sich keine Hinweise auf ergebnisabhängige Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Aspekte, sodass auch auf Endpunktebene das Verzerrungspotenzial niedrig ist.

Wahrscheinlichkeit des Zusatznutzens (Aussagesicherheit)

Unter Berücksichtigung der hohen Qualität der Studie, die der Evidenzstufe 1b entspricht, des niedrigen Verzerrungspotenzials auf Studien- und Endpunktebene sowie der validen Endpunkte kann auf Basis der Studie RESPONSE ein Hinweis auf einen Zusatznutzen abgeleitet werden.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Die PBC ist eine chronisch-entzündliche, autoimmun-vermittelte, cholestatische Lebererkrankung, die durch eine gestörte Gallenflussregulation, intrahepatische Entzündung und eine fortschreitende Fibrose gekennzeichnet ist, die unbehandelt zu einer biliären Zirrhose im Endstadium führen kann [14].

Initial ist die PBC oft über mehrere Jahre asymptomatisch, Symptome entwickeln sich meist erst im Verlauf der Erkrankung [4]. Die Auswirkung der PBC auf die Patienten zeigt sich sowohl durch das Fortschreiten der Lebererkrankung bis zum Endstadium (Zirrhose und die Notwendigkeit einer Lebertransplantation) als auch in den teils stark ausgeprägten, belastenden Begleitsymptomen [14]. Dabei korrelieren die Symptome nicht immer mit dem Grad der Cholestase oder Fibrose, obwohl Patienten mit einer schwereren Erkrankung im Allgemeinen zu mehr Symptomen neigen [27]. Die klinische Präsentation der PBC kann variieren [4]. Cholestase, die Schädigung der Gallensekretion und/oder des Gallenflusses, kann asymptomatisch sein oder sich in Form von Fatigue, Pruritus, unspezifischen Oberbauchbeschwerden und Ikterus äußern [14]. Auch eine Sicca-Symptomatik (trockene Augen und trockener Mund) und Arthralgien werden häufig berichtet. Asymptomatische Patienten werden häufig zufällig aufgrund erhöhter Leberwerte entdeckt [4, 14]. Zu den frühen biochemischen Markern gehören erhöhte ALP- und GGT-Werte im Serum, gefolgt von konjugierter Hyperbilirubinämie in fortgeschritteneren Stadien [14].

An Fatigue leiden bis zu 85% der Patienten mit PBC [4, 23, 57], darunter berichten mehr als 40% von einer moderaten bis schweren Ausprägung. Diese ist häufig assoziiert mit Schlafstörungen, ausgeprägter Tagesmüdigkeit und autonomer Dysfunktion. Betroffene Patienten beschreiben die Fatigue als physische Erschöpfung mit der Unfähigkeit Alltagsaktivitäten zu bewältigen, sowie als verminderte geistige und physische Leistungsfähigkeit [4]. Chronischer Pruritus, ein weiteres schwerwiegendes Symptom der PBC kann bei bis zu 80% der Patienten auftreten und wird von betroffenen Patienten bei schwerer Ausprägung als quälend und unerträglich empfunden. Der Pruritus selbst führt zu keinen spezifischen Hautveränderungen, es kann aber sekundär durch Kratzen zur Entwicklung von Hautverletzungen unterschiedlicher Art (Exkoriationen, Hämatome, Ulzerationen, Lichenifikationen, Pigmentierungsstörungen und Narben) sowie Infektionen kommen, die mit einer erheblichen Beeinträchtigung der Lebensqualität einhergehen [4, 23, 24]. Die gesteigerte

nächtliche Intensität von Pruritus führt zu teils schwerem Schlafentzug, Abgeschlagenheit und Depression und kann bei extremem, therapierefraktärem Pruritus sogar Suizidalität zur Folge haben [4].

Die Lebensqualität von Patienten mit PBC wird sowohl durch die Erkrankung selbst als auch durch sozioökonomische Umstände beeinträchtigt. Neben den primären Symptomen Fatigue und Pruritus sind auch soziale (Folge-)Probleme, emotionale und kognitive (Folge-)Symptome, und Depression mit schlechterer Lebensqualität assoziiert. Verglichen mit anderen chronischen Erkrankungen wie Diabetes mellitus empfanden Patienten mit PBC ihre Erkrankung als größere Belastung und wiesen somit eine reduzierte Lebensqualität auf [4].

Es existiert keine ursächliche Therapie der PBC. Ziel der lebenslangen Behandlung ist es, die Gallentoxizität zu reduzieren, die Entzündung zu unterdrücken und die Fibrose zu begrenzen, um die Leberfunktion zu erhalten, sowie die mit der Krankheit verbundenen Symptome insbesondere Pruritus zu lindern und die Lebensqualität der Patienten zu verbessern. Als Standardtherapie der PBC wird UDCA meist lebenslang verabreicht und ein Großteil der Patienten mit PBC zeigt ein gutes Ansprechen darauf. Neben der Reduktion der prognostisch wichtigen und für die Beurteilung des Therapieansprechens etablierten Laborparameter ALP und Bilirubin wurde unter UDCA auch die Leberhistologie positiv beeinflusst und die histologische Progression verzögert [4]. Unabhängig vom Stadium der Erkrankung verbesserte sich unter der Therapie mit UDCA zudem das transplantationsfreie Überleben. Biochemische Ansprechraten liegen in großen internationalen Kohorten zwischen 50% (Paris-II-Kriterien) und 83% (GLOBE Score) [4]. Als Risikofaktoren für einen ungünstigen Verlauf wird unter anderem ein unzureichendes Ansprechen auf UDCA genannt [4, 52]. Entsprechend besteht insbesondere bei Patienten im vorliegenden Anwendungsgebiet, die nicht ausreichend auf UDCA alleine ansprechen, oder die UDCA nicht vertragen, ein erhöhtes Risiko einer Krankheitsprogression.

Bei diesen Patienten mit nicht ausreichendem Therapieansprechen auf die UDCA-Therapie steigt das Risiko für Krankheitsprogression, Mortalität und Notwendigkeit einer Lebertransplantation [4]. Es wird eine Zweitlinientherapie in Kombination mit UDCA unter Berücksichtigung des Stadiums der Erkrankung, des Alters des Patienten, der Symptomlast und den zusätzlichen therapeutischen Effekten der Substanzen empfohlen [4]. Als Zweitlinientherapie stehen aktuell die zugelassenen Substanzen SEL und Elafibranor zur Verfügung. OCA steht in der EU zur Behandlung der PBC nicht mehr zur Verfügung¹⁷. Mit SEL steht im hier relevanten Anwendungsgebiet nun eine weitere Therapieoption zur Verfügung. SEL ist ein potenter und hochselektiver PPAR (Peroxisom-Proliferator-aktivierter Rezeptor)- δ -Agonist mit entzündungshemmenden und anticholestatischen Eigenschaften [59,

¹⁷ Die Europäische Kommission hat nach Prüfung der verfügbaren Daten durch den Ausschuss für Humanarzneimittel die bedingte Zulassung von OCA am 30. August 2024 in der EU widerrufen, da der Nutzen des Arzneimittels seine Risiken nicht überwiegt. Nach einer vorübergehenden Aussetzung dieser Entscheidung vom 4. September 2024 ist der Widerruf der bedingten Zulassung von OCA durch die Europäische Kommission seit dem 26. November 2024 wieder in Kraft [58].

60], der an mehreren Stellen die Gallensäure-Homöostase, darunter die Cholesterol-Biosynthese und die Aufnahme von Cholesterol aus der Nahrung beeinflusst [59, 61].

Bei SEL handelt es sich um ein Arzneimittel zur Behandlung einer seltenen Erkrankung (Orphan Drug) nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden [2]. Der Orphan Drug-Status wurde am 5. Dezember 2024 durch die EMA bestätigt. Gemäß § 35a Abs. 1 Satz 11 SGB V gilt der medizinische Zusatznutzen für Orphan Drugs durch die Zulassung als belegt [3]. Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im Verhältnis zur zVT müssen nach § 35a Abs. 1 Satz 12 SGB V nur nach Aufforderung durch den G-BA vorgelegt werden, wenn der Umsatz des Arzneimittels in den letzten zwölf Kalendermonaten einen Betrag von 30 Millionen Euro übersteigt [3]. Es ist das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien für die Anzahl der Patienten und Patientengruppen nachzuweisen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht. Es bedarf keiner Ausführungen zu einer zVT.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Im Folgenden werden die Ergebnisse zum medizinischen Zusatznutzen von SEL für alle relevanten Endpunkte beschrieben. Der Nutzenbewertung werden die Ergebnisse der RCT RESPONSE zugrunde gelegt. Erhoben wurden Ergebnisse zu einer Vielzahl von Monaten, dargestellt werden die Ergebnisse zu Monat 6 und Monat 12 (alle weiteren sind dem Anhang 4-G zu entnehmen). Sowohl bei Monat 6 als auch bei Monat 12 handelt es sich jeweils um eine Momentaufnahme des aktuellen Erkrankungszustands der Patienten. Speziell hinsichtlich der Symptomatik und Lebensqualität ist der Verlauf bei der PBC nicht linear progredient, sondern kann Tagesform-abhängigen Schwankungen unterliegen. Durch die gleichwertige Betrachtung beider Erhebungszeitpunkte kann eine umfassendere Betrachtung der Ergebnisse, unter Berücksichtigung möglicher Veränderungen, gewährleistet werden. Daher werden beide Zeitpunkte als gleich relevant angesehen und zur Ableitung des Zusatznutzens herangezogen.

Das Ausmaß des Zusatznutzens wird anhand der folgenden patientenrelevanten Endpunkte abgeleitet:

- Mortalität
- Morbidität
 - Biochemisches Ansprechen
 - Symptomatik
- Gesundheitsbezogene Lebensqualität anhand patientenrelevanter Fragebögen
- UE

Tabelle 4-75: Ausmaß des Zusatznutzen von SEL gegenüber Placebo auf Endpunktebene (Studie RESPONSE)

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Mortalität		
Mortalität	Es ist kein Patient verstorben.	Kein Zusatznutzen/Schaden ableitbar
Morbidität – Biochemisches Ansprechen		
ALP-Normalisierung (ALP $\leq 1.0 \times$ ULN)		
Monat 6	RR: 16,40 [2,32; 116,20], 0,0051	
Monat 12	RR: 14,80 [2,08; 105,44], 0,0071	
Kombiniertes biochemisches Ansprechen		
Monat 6	RR: 3,60 [2,16; 6,01], <0,0001	
Monat 12	RR: 3,09 [1,87; 5,10], <0,0001	
ALP – Stetige Analyse		
Monat 6	LS-MWD: -122,56 [-142,15; -102,98], <0,0001 SMD: -1,70 [-2,05; -1,35]	Ausmaß: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis
Monat 12	LS-MWD: -117,02 [-144,20; -89,84], <0,0001 SMD: -1,23 [-1,56; -0,90]	
Gesamt-Bilirubin – Stetige Analyse		
Monat 6	LS-MWD: -0,07 [-0,15; 0,01], 0,0896 SMD: -0,26 [-0,56; -0,05]	
Monat 12	LS-MWD: -0,02 [-0,14; 0,11], 0,7777 SMD: -0,04 [-0,35; 0,26]	
Morbidität – Symptomatik ^a		
Pruritus NRS - Stetige Analyse		
Monat 6	LS-MWD: -0,90 [-1,35; -0,45], <0,0001 SMD: -0,58 [-0,89; -0,28]	<u>Gesamtpopulation:</u> Ausmaß: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis
Monat 12	LS-MWD: -0,83 [-1,35; -0,31], 0,0018 SMD: -0,47 [-0,78; -0,17]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 12	LS-MWD: -1,77 [-2,97; -0,57], 0,0045 SMD: -0,76 [-1,28; -0,24]	
5-D Pruritus Veränderung – Gesamt-Score		
Monat 6	LS-MWD: -1,76 [-2,72; 0,81], 0,0003	<u>Für Patienten mit Pruritus NRS ≥ 4:</u> Ausmaß: Erheblicher Zusatznutzen Wahrscheinlichkeit:

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Monat 12	SMD: -0,53 [-0,83; -0,22] LS-MWD: -2,31 [-3,40; -1,23], <0,0001 SMD: -0,62 [-0,93; -0,31]	Hinweis
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -3,42 [-5,32; -1,51], 0,0006 SMD: -0,91 [-1,43; -0,38]	
5-D Pruritus Veränderung – Subdomäne Grad		
Monat 6	LS-MWD: -0,35 [-0,56; -0,15], 0,0008 SMD: -0,49 [-0,80; -0,19]	
Monat 12	LS-MWD: -0,38 [-0,61; -0,15], 0,0016 SMD: -0,47 [-0,77; -0,16]	
5-D Pruritus Veränderung – Subdomäne Dauer		
Monat 6	LS-MWD: -0,18 [-0,39; 0,02], 0,0731 SMD: -0,26 [-0,56; 0,04]	
Monat 12	LS-MWD: -0,40 [-0,63; -0,16], 0,0011 SMD: -0,49 [-0,79; -0,19]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 12	LS-MWD: -0,71 [-1,22; -0,20], 0,0075 SMD: -0,70 [-1,21; -0,18]	
5-D Pruritus Veränderung – Subdomäne Beeinträchtigung		
Monat 6	LS-MWD: -0,47 [-0,73; -0,20], 0,0006 SMD: -0,50 [-0,81; -0,20]	
Monat 12	LS-MWD: -0,60 [-0,91; -0,29], 0,0002 SMD: -0,57 [-0,87; -0,26]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -0,97 [-1,55; -0,40], 0,0012 SMD: -0,85 [-1,38; -0,33]	
5-D Pruritus Veränderung – Subdomäne Verlauf		
Monat 6	LS-MWD: -0,31 [-0,68; 0,06], 0,0981 SMD: -0,25 [-0,55; 0,06]	
Monat 12	LS-MWD: -0,47 [-0,90; -0,04], 0,0340 SMD: -0,32 [-0,62; -0,02]	
5-D Pruritus Veränderung – Subdomäne Lokalisation		
Monat 6	LS-MWD: -0,41 [-0,68; -0,14], 0,0034	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Monat 12	SMD: -0,43 [-0,73; -0,13] LS-MWD: -0,38 [-0,67; -0,10], 0,0085 SMD: -0,39 [-0,69; -0,09]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 12	LS-MWD: -1,03 [-1,61; -0,44], 0,0009 SMD: -0,89 [-1,41; -0,36]	
5-D Pruritus Veränderung – Modifizierter Gesamt-Score		
Monat 6	LS-MWD: -1,41 [-2,14; -0,68], 0,0002 SMD: -0,55 [-0,86; -0,25]	
Monat 12	LS-MWD: -1,79 [-2,60; -0,97], <0,0001 SMD: -0,63 [-0,94; -0,33]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -2,89 [-4,43; -1,34], 0,0004 SMD: -0,94 [-1,47; -0,42]	
Monat 12	LS-MWD: -3,21 [-4,65; -1,77], <0,0001 SMD: -1,12 [-1,66; -0,59]	
PGI-S - Stetige Analyse		
Monat 6	LS-MWD: -0,42 [-0,61; -0,23], <0,0001 SMD: -0,64 [-0,95; -0,32]	
Monat 12	LS-MWD: -0,10 [-0,32; 0,11], 0,3453 SMD: -0,14 [-0,44; 0,17]	
PGI-C - Stetige Analyse		
Monat 6	LS-MWD: -0,45 [-0,88; -0,02], 0,0391 SMD: -0,31 [-0,61; -0,01]	
Monat 12	LS-MWD: -0,41 [-0,85; 0,03], 0,0673 SMD: -0,27 [-0,58; 0,03]	
Fibrose - Lebersteifigkeit (FibroScan®)		
Monat 6	LS-MWD: -1,15 [-2,67; 0,36], 0,1354 SMD: -0,22 [-0,54; 0,09]	
Monat 12	LS-MWD: -1,37 [-3,09; 0,35], 0,1174 SMD: -0,24 [-0,55; 0,08]	
Fibrose - Erweiterter Leberfibrosetest (ELF-Score)		
Monat 6	LS-MWD: 0,10 [-0,03; 0,24], 0,1264 SMD: 0,23 [-0,08; 0,53]	
Monat 12	LS-MWD: 0,08 [-0,11; 0,26], 0,4215 SMD: 0,12 [-0,18; 0,42]	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Hospitalisierungen aufgrund der Erkrankung Monat 12	RR: 1,06 [0,27; 4,17], 0,9384	
Gesundheitsbezogene Lebensqualität ^a		
PBC-40: Veränderung - Gesamt-Score Monat 6 Monat 12	LS-MWD: -0,92 [-5,30; 3,46], 0,6787 SMD: -0,06 [-0,36; 0,24] LS-MWD: 0,33 [-4,98; 5,64], 0,9019 SMD: 0,02 [-0,28; 0,32]	Kein Zusatznutzen/Schaden ableitbar
PBC-40: Veränderung - Kognitive Domäne Monat 6 Monat 12	LS-MWD: -0,43 [-1,38; 0,52], 0,3750 SMD: -0,13 [-0,44; 0,17] LS-MWD: -0,13 [-1,20; 0,94], 0,8106 SMD: -0,04 [-0,34; 0,27]	
Subgruppenanalyse Pruritus NRS ≥ 4 zu Monat 6	LS-MWD: -1,98 [-3,59; -0,36], 0,0172 SMD: -0,64 [-1,16; -0,12]	
PBC-40: Veränderung - Emotionale Domäne Monat 6 Monat 12	LS-MWD: 0,21 [-0,40; 0,81], 0,5013 SMD: 0,10 [-0,20; 0,40] LS-MWD: 0,25 [-0,40; 0,90], 0,4488 SMD: 0,11 [-0,19; 0,41]	
PBC-40: Veränderung - Fatigue Domäne Monat 6 Monat 12	LS-MWD: -1,19 [-3,07; 0,69], 0,2131 SMD: -0,18 [-0,49; 0,12] LS-MWD: -0,47 [-2,66; 1,71], 0,6692 SMD: -0,06 [-0,37; 0,24]	
PBC-40: Veränderung - Pruritus Domäne Monat 6 Monat 12	LS-MWD: -0,93 [-1,68; -0,18], 0,0159 SMD: -0,35 [-0,66; -0,05] LS-MWD: -0,83 [-1,68; 0,02], 0,0544 SMD: -0,28 [-0,59; 0,02]	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
PBC-40: Veränderung - Soziale Domäne Monat 6 Monat 12 PBC-40: Veränderung - Symptomatische Domäne Monat 6 Monat 12	LS-MWD: 0,95 [-0,58; 2,47], 0,2216 SMD: 0,18 [-0,12; 0,49] LS-MWD: 1,05 [-0,47; 2,56], 0,1735 SMD: 0,20 [-0,10; 0,50] LS-MWD: 0,17 [-0,64; 0,97], 0,6827 SMD: 0,06 [-0,24; 0,36] LS-MWD: 0,09 [-0,87; 1,06], 0,8475 SMD: 0,03 [-0,27; 0,33]	
Unerwünschte Ereignisse		
UE	RR: 1,02 [0,91; 1,16], 0,6976	Ausmaß: Geringer Zusatznutzen Wahrscheinlichkeit: Hinweis
SUE	RR: 1,14 [0,37; 3,57], 0,8186	
Schwere UE (CTCAE-Grad ≥ 3)	RR: 1,42 [0,54; 3,78], 0,4799	
Therapieabbrüche aufgrund von UE	RR: 0,68 [0,16; 2,94], 0,6023	
UE, die zum Tod führten	Es führte kein UE zum Tod.	
UE, SUE und schwere UE von speziellem Interesse	UE von speziellem Interesse Pruritus-assoziiert: RR: 0,36 [0,14; 0,89], 0,0273 SUE von speziellem Interesse Es traten keine statistisch signifikanten Ergebnisse auf ^c Schwere UE von speziellem Interesse Es traten keine statistisch signifikanten Ergebnisse auf ^d	
UE, SUE und schwere UE ohne erkrankungsbezogene Ereignisse	Jegliche UE ohne erkrankungsbezogene Ereignisse RR: 1,04 [0,92; 1,19], 0,5144 SUE ohne erkrankungsbezogene Ereignisse RR: 1,14 [0,37; 3,57], 0,8186 Schwere UE ohne erkrankungsbezogene Ereignisse RR: 1,78 [0,61; 5,18], 0,2923	

Endpunkte	Effektschätzer [95%-KI], p-Wert	Ausmaß des Zusatznutzen Wahrscheinlichkeit des Zusatznutzen
Jegliche UE nach SOC und PT ^b	SOC Erkrankungen der Haut und des Unterhautgewebes RR: 0,68 [0,37; 1,23], 0,2021 PT Pruritus RR: 0,30 [0,12; 0,80], 0,0160	
SUE nach SOC und PT ^b	Keine SUE nach SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen	
Schwere UE nach SOC und PT ^b	Keine schweren UE nach SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen	
a: Schwerwiegende bzw. schwere Endpunkte b: Aus Gründen der Übersichtlichkeit werden an dieser Stelle nur fazitrelevante Ergebnisse (signifikanter und relevanter Unterschied in Gesamtpopulation oder Subgruppe) dargestellt. c: Kardiovaskuläre Toxizität: 1 (0,8%) Patient im SEL-Arm; Ischämische Herzkrankheit: 1 (0,8%) Patient im SEL-Arm; Lebertoxizität: 1 (0,8%) Patient im SEL-Arm d: Kardiovaskuläre Toxizität: 1 (0,8%) Patient im SEL-Arm; Ischämische Herzkrankheit: 1 (0,8%) Patient im SEL-Arm; Lebertoxizität: 1 (0,8%) Patient im SEL-Arm ALP=Alkalische Phosphatase, CTCAE= Common Terminology Criteria for Adverse Events, ELF=Fortgeschrittene Leberfibrose, KI=Konfidenzintervall, LS-MWD=Kleinste-Quadrate-Mittelwertdifferenz, NRS=Numerical Rating Scale, PBC=Primär biliäre Cholangitis, PGI-S=Patient Global Impression of Severity, PGI-C=Patient Global Impression of Change, PT=Preferred Term, RR=Relatives Risiko, SEL=Seladelpar, SMD=Standardisierte Mittelwertdifferenz, SOC=System Organ Class, SUE=Schwerwiegendes unerwünschtes Ereignis, UE=Unerwünschtes Ereignis, ULN=Oberer Normwert		

Im Folgenden werden die Ergebnisse zum Zusatznutzen von SEL gegenüber Placebo beschrieben.

Mortalität

In der Studie RESPONSE ist kein Patient verstorben.

Insgesamt wird **kein Zusatznutzen oder Schaden** abgeleitet.

Morbidität

Die Morbidität wird basierend auf den Endpunkten biochemisches Ansprechen, Pruritus, Fibrose und Hospitalisierung beschrieben.

Biochemisches Ansprechen

ALP-Normalisierung

Es zeigten sich statistisch signifikante Ergebnisse für die ALP-Normalisierung zugunsten von SEL. Dieser statistisch signifikante Vorteil zeigte sich sowohl zu Monat 6 (RR [95%-KI]: 16,40 [2,32; 116,20], p=0,0051) als auch zu Monat 12 (RR [95%-KI]: 14,80 [2,08; 105,44], p=0,0071) und belegt die nachhaltige Normalisierung.

Kombiniertes biochemisches Ansprechen

Für den kombinierten Endpunkt biochemisches Ansprechen zeigte sich ebenfalls ein statistisch signifikanter Vorteil zugunsten von SEL. Der Vorteil war sowohl zu Monat 6 (RR [95%-KI]: 3,60 [2,16; 6,01]; $p < 0,0001$) als auch zu Monat 12 (RR [95%-KI]: 3,09 [1,87; 5,10]; $p < 0,0001$) statistisch signifikant. Zu Monat 6 sprachen 66,4% der Patienten unter SEL an, unter Placebo 18,5%. Dieser Unterschied im biochemischen Ansprechen bestätigt sich zu Monat 12 mit Ansprechraten von 61,7% unter SEL und 20,0% unter Placebo.

ALP und Gesamt-Bilirubin – Stetige Analysen

Auch in der stetigen Analyse setzt sich der Vorteil von SEL gegenüber Placebo fort: Zu beiden berichteten Zeitpunkten zeigte sich ein statistisch signifikanter Unterschied (Monat 6 und 12: $p < 0,0001$). Und auch die stetige Analyse zum Gesamt-Bilirubin zeigt den Vorteil auf (SMD [95%-KI] zu Monat 6: -0,26 [-0,56; -0,05]; zu Monat 12: -0,04 [-0,35; 0,26]).

Fazit biochemisches Ansprechen

Für Patienten mit PBC ist das primäre Therapieziel eine Verlangsamung der Krankheitsprogression, um die damit einhergehende höhere Symptombelastung mit potenziell schwerwiegenden Folgekomplikationen (Fibrose, Zirrhose, Lebertransplantation und Tod) zu vermeiden. Unabhängig von Risikofaktoren wie Alter und Krankheitsstadium steigt die transplantationsfreie Überlebenszeit, je niedriger der ALP-Wert ist [15]. Dabei ist die Normalisierung der ALP mit dem niedrigsten Risiko für eine Lebertransplantation oder Tod verbunden [16], sowie auch eine Senkung der ALP-Werte mit einem verbesserten transplantationsfreien Überleben oder Tod assoziiert ist [15]. Weiter konnte gezeigt werden, dass bei Patienten mit PBC das Erreichen einer ALP-Normalisierung nach einem Jahr (unter UDCA-Therapie) mit einem deutlich längeren komplikationsfreien Überleben assoziiert ist [17]. Bereits eine Veränderung des ALP-Wertes von über 10% im Krankheitsverlauf wird als adäquates Maß zur Erfassung einer klinisch relevanten Verbesserung erachtet [19]. Eine anhaltende Erhöhung der ALP bei PBC unter UDCA ist hingegen mit einem gewissen Grad an persistierender Krankheitsaktivität verbunden [13].

Für das Therapieziel einer vollständigen Kontrolle der Krankheitsaktivität bei PBC ist auch eine Normalisierung des Bilirubins erforderlich [13]. Im Hinblick auf das Gesamt-Bilirubin ist die mit fortschreitender PBC auftretende erhöhte Bilirubin-Konzentration mit einer sinkenden Überlebenswahrscheinlichkeit assoziiert [14, 15, 19]. Der Gesamt-Bilirubinwert kann daher als wichtiger Prädiktor für einen ungünstigen Verlauf der PBC und das individuelle Risiko einer Krankheitsprogression angesehen werden [14]. Folglich kann die ALP-Normalisierung als geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC herangezogen werden [15], ebenso die Einzelkomponenten des kombinierten biochemischen Ansprechens (ALP und Gesamt-Bilirubin).

Die hier gezeigten statistisch signifikanten und anhaltenden Vorteile zugunsten von SEL in allen Teilkomponenten des biochemischen Ansprechens sowohl nach 6 als auch nach

12 Monaten belegen, dass durch SEL im vorliegenden Anwendungsgebiet das Fortschreiten der Lebererkrankung nachhaltig verhindert, respektive verlangsamt werden kann.

Es ergibt sich für den Endpunkt biochemisches Ansprechen ein **Hinweis** auf einen **beträchtlichen Zusatznutzen** von SEL in der Gesamtpopulation.

Symptomatik

Pruritus NRS

Hinsichtlich der Pruritus NRS zeigten sich ausschließlich Vorteile durch die Behandlung mit SEL. Sowohl zu Monat 6 als auch zu Monat 12 zeigte sich ein stärkerer mittlerer Abfall der Pruritus NRS im SEL-Arm als im Placebo-Arm. Zu beiden Monaten sind diese Vorteile statistisch signifikant (Monat 6: $p=0,0001$; Monat 12: $p=0,0018$) und zu Monat 6 zusätzlich klinisch relevant (SMD [95%-KI]: -0,58 [-0,89; -0,28]).

Zudem zeigt sich in dem Endpunkt Pruritus NRS – stetige Analyse eine Effektmodifikation für das Subgruppenmerkmal Pruritus NRS zu Baseline ($p=0,0256$). Bei Patienten mit einer Pruritus NRS ≥ 4 Punkte zu Baseline ergibt sich ein dem Haupteffekt gleichgerichteter statistisch signifikanter ($p=0,0045$) sowie klinisch relevanter Vorteil für SEL zu Monat 12 (SMD [95%-KI]: -0,76 [-1,28; -0,24]).

In der Gesamtschau zeigen sich für den Endpunkt Pruritus NRS nur Vorteile unter der Behandlung mit SEL, sowohl in der Gesamtpopulation als in der Teilpopulation mit einer Pruritus NRS ≥ 4 Punkte zu Baseline.

5-D Pruritus

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus im Gesamt-Score. Der Unterschied zwischen SEL und Placebo war zu beiden Zeitpunkten statistisch signifikant ($p=0,0003$ und $p<0,0001$) und klinisch relevant zugunsten von SEL (SMD [95%-KI] zu Monat 6: -0,53 [-0,83; -0,22]; zu Monat 12: -0,62 [-0,93; -0,31]).

Die Vorteile einer Behandlung mit SEL gegenüber Placebo, statistisch signifikant und auch klinisch relevant, zeigten sich in allen Subdomänen sowie im modifizierten Gesamt-Score zu allen Zeitpunkten:

- Subdomäne Grad: Zu beiden Zeitpunkten führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus in der Subdomäne Grad. Der Vorteil der Behandlung mit SEL gegenüber Placebo war zu beiden Zeitpunkten statistisch signifikant ($p=0,0008$ und $p=0,0016$).
- Subdomäne Dauer: Auch in der Subdomäne Dauer führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall zu beiden berichteten Zeitpunkten, wobei dieser Unterschied zwischen SEL und Placebo zu Monat 12 statistisch signifikant war ($p=0,0011$).

- Subdomäne Beeinträchtigung: Zu beiden dargestellten Erhebungszeitpunkten zeigte sich ein statistisch signifikanter ($p=0,0006$ und $p=0,0002$) als auch klinisch relevanter (SMD [95%-KI] zu Monat 6: $-0,50 [-0,81; -0,20]$; zu Monat 12: $-0,57 [-0,87; -0,26]$) Vorteil für Patienten unter SEL-Behandlung, der entsprechend mit einem stärkeren mittleren Abfall des 5-D Pruritus in der Subdomäne Beeinträchtigung einher ging.
- Subdomäne Verlauf: Auch in der Subdomäne Verlauf führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall zu beiden Zeitpunkten. Der Unterschied zwischen SEL und Placebo war zu Monat 12 statistisch signifikant ($p=0,0340$).
- Subdomäne Lokalisation: Ebenfalls statistisch signifikant sind die Ergebnisse zur Subdomäne Lokalisation zu Monat 6 und Monat 12 ($p=0,0034$ und $p=0,0085$). Die Behandlung mit SEL führte im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des 5-D Pruritus in der Subdomäne Lokalisation.
- Modifizierter Gesamt-Score: Neben statistisch signifikanten zeigten sich klinisch relevante Vorteile zu Monat 6 und zu Monat 12 für die Behandlung mit SEL im Vergleich zu Placebo mit einem stärkeren mittleren Abfall des 5-D Pruritus im modifizierten Gesamt-Score (statistische Signifikanz: $p=0,0002$ und $p<0,0001$; klinische Relevanz: SMD [95%-KI] zu Monat 6: $-0,55 [-0,86; -0,25]$; zu Monat 12: $-0,63 [-0,94; -0,33]$).

Neben den Haupteffekten trat über die einzelnen Endpunkte und beide Zeitpunkte hinweg eine Effektmodifikation für das Subgruppenmerkmal Pruritus NRS zu Baseline auf. Es wurde jeweils für Patienten mit einer Pruritus NRS ≥ 4 Punkte zu Baseline eine gegenüber der Gesamtpopulation nochmals stärkere Verbesserung in den jeweiligen Endpunkten erreicht. Die Effektmodifikation war statistisch signifikant für den Gesamt-Score zu Monat 6, in der Subdomäne Dauer zu Monat 12, in der Subdomäne Beeinträchtigung zu Monat 6, in der Subdomäne Lokalisation zu Monat 12 sowie im modifizierten Gesamt-Score zu Monat 6 und 12. Die ausführlichen Ergebnisse zur Ableitung des Zusatznutzens sind der Tabelle 4-75 zu entnehmen.

Insgesamt zeigen sich damit nur Vorteile für die mit SEL behandelten Patienten im Endpunkt 5-D Pruritus. In der Gesamtpopulation sind die Vorteile bereits beträchtlichen Ausmaßes und gehen in der Teilpopulation ≥ 4 Punkte in der Pruritus NRS zu Baseline nochmal darüber hinaus.

PGI-C & PGI-S

Sowohl zu Monat 6 als auch zu Monat 12 führte die Behandlung mit SEL im Vergleich zu Placebo zu einem stärkeren mittleren Abfall des PGI-S. Der Unterschied zwischen SEL und Placebo war zu Monat 6 sowohl statistisch signifikant ($p<0,0001$) als auch klinisch relevant (SMD [95%-KI]: $-0,64 [-0,95; -0,32]$). Auch für den PGI-C zeigt sich der statistisch signifikante stärkere mittlere Abfall von SEL gegenüber Placebo zu Monat 6 ($p=0,0391$).

Die PGI-S- und PGI-C-Skalen sind valide Instrumente, anhand derer sowohl eine Einschätzung des Schweregrades der Erkrankung eines Patienten und seines klinischen Gesundheitszustands (PGI-S-Skala) als auch der Verbesserung bzw. Veränderung des klinischen Gesundheitszustands des Patienten im zeitlichen Verlauf (PGI-C) abgefragt werden kann. Entsprechend sind die positiven Ergebnisse von SEL in der Ableitung des Zusatznutzens heranzuziehen.

Fibrose

Die Veränderung der Lebersteifigkeit erfasst mittels FibroScan® ist unter SEL und Placebo vergleichbar, sowohl zu Monat 6 als auch zu Monat 12. Dies wird auch durch die Ergebnisse des erweiterten Leberfibrosetests (ELF-Score) bestätigt.

Hospitalisierung

Bei der betrachteten Patientenpopulation handelt es sich um eine verhältnismäßig „gesunde“ Kohorte dahingehend, dass die Erkrankung, respektive die Symptome in der Regel keine Einweisung in ein Krankenhaus erfordern. Innerhalb der 12-monatigen Behandlung wurden insgesamt 9 Patienten aufgrund der Erkrankung hospitalisiert. Der Anteil der Patienten, die aufgrund ihrer Erkrankung hospitalisiert wurden, ist mit 4,7% unter SEL und 4,6% unter Placebo in beiden Armen vergleichbar. Entsprechend zeigt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.

Fazit Symptomatik

Chronischer Pruritus ist ein schwerwiegendes Symptom der PBC, dessen Ursache bislang nicht vollständig geklärt ist [24, 25]. Er kann bei bis zu 80% der Patienten auftreten und wird von betroffenen Patienten bei schwerer Ausprägung als quälend und unerträglich empfunden. Der Pruritus selbst führt zu keinen spezifischen Hautveränderungen, es kann aber sekundär durch Kratzen zur Entwicklung von Hautverletzungen unterschiedlicher Art (Exkoriationen, Hämatome, Ulzerationen, Lichenifikationen, Pigmentierungsstörungen und Narben) sowie Infektionen kommen, die mit einer erheblichen Beeinträchtigung der Lebensqualität einhergehen [4, 23, 24]. Die gesteigerte nächtliche Intensität von Pruritus führt zu teils schwerem Schlafentzug, Abgeschlagenheit und Depression und kann bei extremem, therapierefraktärem Pruritus sogar Suizidalität zur Folge haben [4]. Im Verlauf der PBC-Erkrankung kann Pruritus mit variabler Intensität auftreten. Therapierefraktärer, unerträglicher Pruritus stellt bei Patienten mit PBC in seltenen Fällen eine Indikation zur Lebertransplantation dar [24, 27].

In der Studie RESPONSE wird unter der Behandlung mit SEL eine deutlich höhere und anhaltende Reduktion dieses sehr belastenden Symptoms erreicht als unter Placebo, das sowohl für die Pruritus NRS als auch für den 5-D Pruritus in statistisch signifikanten und klinisch relevanten Vorteilen von SEL gegenüber Placebo resultiert. Bei Betrachtung des zeitlichen Verlaufs zeigt sich, dass es unter SEL zu einem kontinuierlichen Rückgang von Pruritus kommt, während die Patienten unter der Behandlung mit Placebo Pruritus in insgesamt höherer und variabler Intensität spürbar wahrnehmen. Angesichts des beobachteten kontinuierlichen Rückgangs unter der Behandlung mit SEL kann davon ausgegangen werden, dass SEL Pruritus

nicht nur lindert, sondern auch dessen Entstehung hemmt. Zu berücksichtigen ist zudem, dass sich bei Patienten mit einer bereits erhöhten Belastung durch Pruritus zu Baseline (Subgruppe Pruritus NRS ≥ 4 Punkte zu Baseline, das entspricht Patienten mit moderat bis schwerwiegendem Pruritus) eine deutliche und nachhaltige Reduktion unter der Behandlung mit SEL zeigt und diese damit besonders deutlich von der Behandlung mit SEL profitieren.

Unterstützend zeigen auch die Ergebnisse des PGI-S und PGI-C den Vorteil der Patienten unter SEL-Behandlung im Hinblick auf die positive Entwicklung unter der Erkrankung hinsichtlich Schweregrad und Veränderung.

Damit ergibt sich in der Gesamtschau für die Symptomatik ein **Hinweis** auf einen **beträchtlichen Zusatznutzen** von SEL in der Gesamtpopulation. Zudem wird ein **Hinweis** auf einen **erheblichen Zusatznutzen** von SEL für Patienten mit einer Pruritus NRS ≥ 4 Punkte zu Baseline abgeleitet.

Gesundheitsbezogene Lebensqualität anhand des PBC-40

Der Fragebogen PBC-40 ist ein etablierter und validierter Fragebogen für die PBC und enthält Domänen zur Morbidität sowie der gesundheitsbezogenen Lebensqualität [44]. Auch der G-BA sieht darin ein „weitestgehend validiertes und krankheitsspezifisches Instrument zur Messung der Lebensqualität bei Patienten mit PBC“ [45].

Es zeigte sich zu Monat 6 in der Domäne Pruritus ein statistisch signifikanter Vorteil für die Behandlung mit SEL ($p=0,0159$).

Bei der betrachteten Patientenpopulation der Patienten mit PBC im frühen Stadium (Child-Pugh Stadium A) handelt es sich um Patienten mit einer bereits zu Beginn relativ hohen Lebensqualität. Eine Verbesserung ist entsprechend weniger wahrscheinlich, respektive möglich als bei Patienten mit primär geringer Lebensqualität. Unter dem Ziel die Symptome zu lindern und das Fortschreiten der Krankheit zu verlangsamen, ist bereits die hier gezeigte gleichbleibende Lebensqualität von hoher Patientenrelevanz.

Insgesamt wird **kein Zusatznutzen oder Schaden** im Hinblick auf die Lebensqualität für die Gesamtpopulation abgeleitet.

Unerwünschte Ereignisse

UE

Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungen.

SUE

Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungen.

Schwere UE

Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungen.

Therapieabbrüche aufgrund von UE

Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungen.

UE, die zum Tod führten

Es führte kein UE zum Tod.

Jegliche UESI

Hinsichtlich der UESI jeglichen Schweregrades zeigte sich für das UESI Pruritus-assoziierte UE ein statistisch signifikanter Vorteil zugunsten der Patienten unter der Behandlung mit SEL (RR [95%-KI]: 0,36 [0,14; 0,89], $p=0,0273$). Es traten keine weiteren statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen SEL vs. Placebo auf.

Schwerwiegende und schwere UESI

Es zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen hinsichtlich der schwerwiegenden und schweren UESI. Es trat jeweils ein SUE und ein schweres UE in dem Endpunkt kardiovaskuläre Toxizität, ischämische Herzkrankheit und Lebertoxizität im SEL-Arm auf.

UE, SUE und schwere UE ohne erkrankungsbezogene Ereignisse

Es zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen SEL vs. Placebo hinsichtlich jeglicher UE-Ereignisse, SUE, und schwerer UE ohne erkrankungsbezogene Ereignisse.

Jegliche UE nach SOC und PT

Für die UE nach SOC und PT jeglichen Schweregrades zeigte sich erneut der statistisch signifikante Vorteil im PT Pruritus für die Patienten im SEL-Arm (RR [95%-KI]: 0,30 [0,12; 0,80]; $p=0,0160$). Alle weiteren SOC und PT zeigten keinen statistisch signifikanten Unterschied.

SUE nach SOC und PT

Es traten unter den SUE keine SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen auf.

Schwere UE nach SOC und PT

Unter den schweren UE traten keine SOC und PT mit mindestens 5% Inzidenz in einer der Behandlungsgruppen auf.

Die ausführliche Darstellung der Effektschätzer der UE ist der Tabelle 4-75 zu entnehmen.

Fazit UE

Für fast alle betrachteten Endpunkte zu den Gesamtraten UE traten keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen auf – dies ist bei einer verblindeten Studie im Vergleich zu Placebo per se bemerkenswert. In dem UESI Pruritus (RR [95%-KI]: 0,36 [0,14; 0,89], $p=0,0273$) als auch in dem PT Pruritus jeglichen Schweregrades

zeigten sich sogar statistisch signifikante Vorteile für die Patienten im SEL-Arm (RR [95%-KI]: 0,30 [0,12; 0,80]; $p=0,0160$). Dem gegenüber stehen keine Nachteile unter der Behandlung mit SEL. Bei der Beurteilung der UE kommt als wesentlicher Faktor noch hinzu, dass in der Studie der Vergleich gegenüber Placebo betrachtet wird. Placebo umfasste primär symptomatische Therapieansätze, die aber im Wesentlichen genauso auch im SEL-Arm angewendet wurden. Somit wurde die Gabe von SEL schlussendlich mit keiner Therapie verglichen, womit der nicht-statistisch signifikante Unterschied hinsichtlich des Auftretens von UE bei Patienten unter der Behandlung mit SEL besonders hervorzuheben ist. Denn bei einer Behandlung mit einer aktiven Substanz sind generell mehr UE zu erwarten als bei Einnahme einer symptomatisch orientierten Therapie, respektive keiner Therapie.

Aus den genannten Gründen ergibt sich ein **Hinweis auf einen geringen Zusatznutzen** in den UE für die Gesamtpopulation.

Gesamtfazit

Die PBC ist eine chronisch-entzündliche, autoimmun-vermittelte, cholestatische Lebererkrankung, die unbehandelt zu Fibrose und einer biliären Zirrhose im Endstadium fortschreitet [14]. Das Vorhandensein einer Leberzirrhose stellt auch bei PBC-Patienten den entscheidenden Risikofaktor für die Entwicklung eines HCC dar [4].

Die Lebensqualität von Patienten mit PBC wird sowohl durch die Erkrankung selbst als auch durch sozioökonomische Umstände beeinträchtigt. Neben den bekannten Symptomen Fatigue und Pruritus sind auch soziale (Folge-)Probleme, emotionale und kognitive (Folge-)Symptome und Depression mit schlechterer Lebensqualität assoziiert. Verglichen mit anderen chronischen Erkrankungen wie Diabetes mellitus empfanden Patienten mit PBC ihre Erkrankung als größere Belastung und wiesen somit eine reduzierte Lebensqualität auf [4].

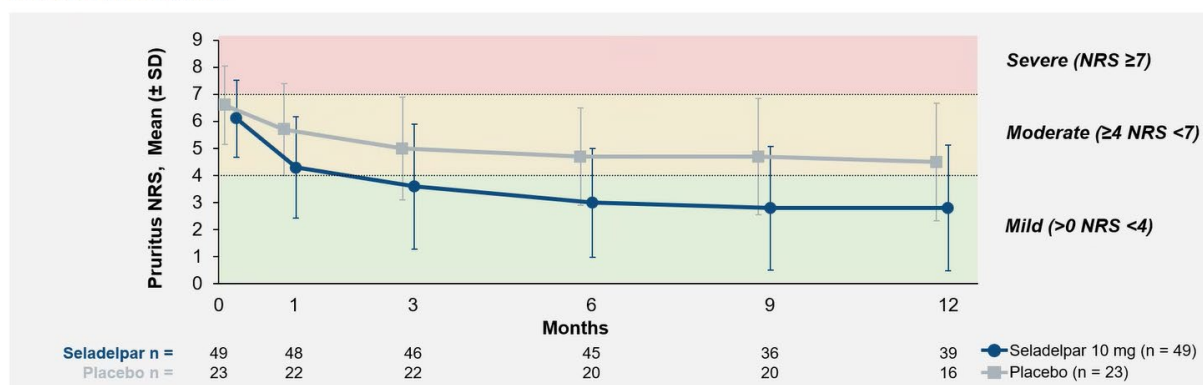
Da es sich bei der PBC um eine seltene Erkrankung handelt, stellt die geringe Anzahl Betroffener und somit potenzieller Studienteilnehmer naturgemäß eine besondere Herausforderung bei der Durchführung aussagekräftiger Studien dar. Mit der Studie RESPONSE konnte trotzdem eine RCT durchgeführt und somit eine Studie von hoher Evidenzgüte und Aussagekraft für die Nutzenbewertung vorgelegt werden. Die Vorteile der Behandlung mit SEL spiegeln sich in deutlichen Effekten mit statistischer Signifikanz und teilweise klinischer Relevanz wider, die trotz des vergleichsweise kleinen Patientenkollektivs gezeigt werden konnten. Die Ergebnisse der Studie RESPONSE belegen die Wirksamkeit von SEL im Hinblick auf das biochemische Ansprechen, die Symptomatik, Lebensqualität sowie das Eintreten unerwünschter Ereignisse. Die beobachtete Verringerung des Leberwertes alkalische Phosphatase (ALP) sowie der respektiven Normalisierung ebendieser unter SEL zeigt, dass ein Fortschreiten der Lebererkrankung verlangsamt und eventuell sogar verhindert und somit der Zustand und die Funktion dieses lebensnotwendigen Organs erhalten werden kann. Damit verbessert sich auch die Prognose für die Patienten im weiteren Krankheitsverlauf, welcher potenziell zu Leberfibrose, Zirrhose und letztendlich zu einer Lebererkrankung im Endstadium bis hin zum Tod führt.

Patienten mit PBC nehmen die Auswirkungen des Pruritus auf ihr tägliches Leben unmittelbar wahr, was mit einer erheblichen Beeinträchtigung der Lebensqualität einhergeht. Die Intensität des Pruritus ist mit Beeinträchtigungen in vielen Lebensbereichen, wie reduzierter Schlafqualität und Arbeitsproduktivität, sowie Depressionen und Suizidalität verbunden. Pruritus wird aufgrund der direkten Spürbarkeit durch den Patienten verbunden mit starken Einschränkungen des alltäglichen Lebens des Patienten als patientenrelevanter Endpunkt betrachtet. Hinzu kommt, dass Pruritus meist nicht zufriedenstellend therapiert werden kann [4]. Auftreten und Intensität von Pruritus korrelieren nicht mit dem Schweregrad der Lebererkrankung, können patientenindividuell sehr unterschiedlich sein [25] und treten unabhängig vom Ausmaß der Cholestase – oft schubweise – auf. Unter SEL konnte in der Studie RESPONSE eine deutliche Verbesserung des sehr belastenden Pruritus gezeigt werden. Sowohl für die Pruritus NRS als auch für den 5-D Pruritus zeigten sich ausschließlich statistisch signifikante Vorteile für die Behandlung mit SEL. Aufgrund der deutlichen Verringerung der Pruritus-Inzidenz über den gesamten Behandlungszeitraum wird gerade auch im Vergleich zu Placebo davon ausgegangen, dass SEL Pruritus nicht nur lindert, sondern auch einen antipruritogenen Effekt aufweist, das heißt auch dessen de novo-Entstehung hemmt. Das bedeutet, dass schlussendlich alle Patienten, also Patienten mit, aber auch ohne Pruritus von einer Therapie mit SEL im Hinblick auf bestehenden sowie neu auftretenden Pruritus beträchtlich profitieren.

Zu berücksichtigen ist zudem, dass sich bei Patienten mit einer bereits erhöhten Belastung durch Pruritus zu Baseline (Subgruppe Pruritus NRS ≥ 4 zu Baseline) eine noch deutlichere und nachhaltigere Reduktion unter der Behandlung mit SEL zeigte (siehe auch Abbildung 30). Neben dem gezeigten Vorteil in der Gesamtpopulation wird deutlich, dass speziell Patienten mit einer höheren Pruritus-Belastung nochmal stärker, und damit erheblich, von der Behandlung mit SEL profitieren.

Mean Pruritus NRS Over Time in Patients With Moderate to Severe Pruritus at Baseline

NRS ≥ 4 at Baseline



Seladelpar reduced the mean itch intensity from moderate to mild by month 12

Abbildung 30: Mittlere Pruritus NRS, Patienten mit mittlerem bis schweren Pruritus zu Baseline, Subgruppe NRS ≥ 4 Punkte zu Baseline

n=Anzahl der Patienten mit Ereignis, NRS=Numerical Rating Scale, SD=Standardabweichung

Des Weiteren ist bemerkenswert, dass unter der Behandlung mit SEL statistisch signifikante Verbesserungen in den UE zu beobachten waren, was SEL zu einem äußerst verträglichen Arzneimittel macht. Es ist zu betonen, dass dieser Vorteil gegenüber Placebo, also keiner Therapie, gezeigt werden konnte.

Diese positiven Effekte zeigen, dass SEL einen bislang ungedeckten therapeutischen Bedarf für Patienten mit einer PBC im vorliegenden Anwendungsgebiet deckt und dabei den Gesamtzustand der Patienten zudem noch deutlich verbessert. SEL trägt maßgeblich zur Reduktion der klinischen Manifestationen der PBC bei und führt so zu einem guten Gesundheitszustand und potenziell zu einer erhöhten Lebenserwartung. Vor dem Hintergrund einer schweren, potenziell lebensverkürzenden chronischen Erkrankung mit einer progredienten Akkumulation der Symptomlast im weiteren Verlauf ist für SEL ein **beträchtlicher Zusatznutzen in der Gesamtpopulation** festzustellen. Da sich in der Population der PBC-Patienten mit moderat-schwerem Pruritus zu Baseline (Pruritus NRS ≥ 4 Punkte) der Vorteil von SEL besonders deutlich zeigt, wird für diese Teilpopulation ein **erheblicher Zusatznutzen** festgestellt.

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 4-76: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit PBC, die nicht ausreichend auf UDCA alleine ansprechen oder die UDCA nicht vertragen	Hinweis auf einen beträchtlichen Zusatznutzen
Patienten mit PBC und moderat-schwerem Pruritus (Pruritus NRS ≥ 4 Punkte) zu Baseline	Hinweis auf einen erheblichen Zusatznutzen
NRS=Numerical Rating Scale, PBC=Primär biliäre Cholangitis, UDCA=Ursodeoxycholsäure	

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen

indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

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

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

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹⁸, Molenberghs 2010¹⁹). 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.

¹⁸ Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.

¹⁹ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006²⁰) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%- Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006²¹) 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.

In der vorliegenden Indikation wird der Surrogatparameter biochemisches Ansprechen zur Ableitung des Zusatznutzens herangezogen.

Biochemisches Ansprechen

Im vorliegenden Anwendungsgebiet basieren derzeit alle Ansprechkriterien auf eine Therapie (vorrangig mit UDCA) auf Grenzwerten für den Grad der Erhöhung der ALP und des Bilirubins. Diese Parameter sind auf Populationsebene prädiktiv für das Risiko einer Progression der Erkrankung bis zur Transplantation oder zum Tod [13].

Da sich Leberkomplikationen wie die Notwendigkeit einer Lebertransplantation und leberbedingte Mortalität erst nach längerer Zeit entwickeln, würde der direkte Nachweis der

²⁰ Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

²¹ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

Auswirkungen der Therapie auf klinische Endpunkte eine längere Nachbeobachtungszeit erfordern. Das Fortschreiten der Erkrankung wird durch die Therapie so stark verlangsamt, dass eine unverhältnismäßig lange Beobachtungszeit in klinischen Studien nötig wäre, um klinische Endpunkte wie Morbidität und Mortalität adäquat zu erfassen, und die Ereignishäufigkeiten bei normaler Studiendauer von 48 und 96 Wochen sind nicht ausreichend. Klinische Endpunkt-Studien sind vor diesem Hintergrund kaum mehr möglich: Zum einen ist es ethisch nicht vertretbar, einen Patienten bis zu einem klinischen Ereignis auf einem Regime zu belassen, wenn dies gegebenenfalls durch die Initiierung einer Therapie verhindert werden könnte, zum anderen stellen sowohl die erforderliche Studiendauer als auch die notwendige Patientenzahl praktische Hindernisse dar. In der Indikation HIV ist eine klare Korrelation zwischen den Surrogatparametern Viruslast und CD4-Zellzahl und den entsprechenden Morbiditäts- und Mortalitäts-Endpunkten in klinischen Studien vielfach validiert worden. Entsprechend werden Zulassungsstudien im HIV-Bereich seit 1997 im Grunde ausnahmslos unter Verwendung dieser validierten Surrogate durchgeführt. Auch im Anwendungsgebiet der PBC ist es ethisch nicht vertretbar, einen Patienten bis zu einem klinischen Ereignis zu beobachten, obwohl im Verumarm der Studie eine wirksame Therapie verfügbar wäre, um dieses zu verhindern, dies insbesondere vor dem Hintergrund, dass im Vergleichsarm keine wirksame Therapie zur Verfügung steht.

Im Rahmen des biochemischen Ansprechens werden die ALP-Normalisierung (definiert als $ALP \leq 1 \times ULN$), das kombinierte biochemische Ansprechen (definiert als $ALP < 1,67 \times ULN$, Gesamt-Bilirubin $\leq ULN$ und ALP-Reduktion von $\geq 15\%$) sowie die stetigen Analysen zu ALP und Gesamt-Bilirubin als Surrogate für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod herangezogen. Ergänzend werden die Veränderung der Parameter ALT, AST, direktes und indirektes Bilirubin sowie GGT und 5'-Nucleotidase dargestellt.

ALP-Normalisierung

ALP ist für viele Stoffwechselprozesse im Körper essentiell und findet sich unter anderem in Knochenzellen und im Lebergewebe. Bei Verdacht auf Lebererkrankungen wird die Konzentration der ALP im Blut bestimmt. Die PBC ist eine nicht-eitrige Entzündung der kleinen intrahepatischen Gallengänge, wodurch es zu einem Stau von Gallenflüssigkeit in der Leber (Cholestase) und damit einer vermehrten Freisetzung von ALP aus den betroffenen Gallengangs-Epithelzellen kommt [14]. Ein erhöhter ALP-Wert wird typischerweise im gesamten Verlauf der PBC-Erkrankung beobachtet [14, 15].

Studiendaten belegen, dass unabhängig von Risikofaktoren wie Alter und Krankheitsstadium die transplantationsfreie Überlebenszeit steigt, je niedriger der ALP-Wert ist [15]. Dabei ist die Normalisierung der ALP, definiert als $ALP \leq 1 \times ULN$, mit dem niedrigsten Risiko für eine Lebertransplantation oder Tod verbunden: In einer Patientenkohorte der GLOBAL PBC Study Group mit UDCA-behandelten und -unbehandelten Patienten zeigten sich bei Patienten, die – unabhängig von der UDCA-Therapie – nach einem Jahr einen normalen ALP-Wert ($\leq 1 \times ULN$) aufwiesen, 10-Jahres-Überlebensraten von 93,2%. Patienten mit einem ALP-Wert zwischen

1,0-1,67 x ULN nach einem Jahr hatten im weiteren Verlauf 10-Jahres-Überlebensraten von 86,1% [16]. Weitere Daten zeigen ebenfalls, dass bei Patienten mit PBC das Erreichen einer ALP-Normalisierung nach einem Jahr (unter UDCA-Therapie) mit einem deutlich längeren komplikationsfreien Überleben assoziiert ist [17]. Eine weitere Studie belegt, dass jede anhaltende Erhöhung der ALP bei PBC unter einer UDCA-Therapie mit einem gewissen Grad an persistierender Krankheitsaktivität, gemessen auf Basis des Entzündungs- und Stoffwechselproteoms im Serum, verbunden ist und belegt damit die Bedeutung biochemischer Marker wie der ALP bei PBC, bei denen eine lineare Beziehung zum Krankheitsverlauf besteht. Diese anhaltende Krankheitsaktivität birgt das Potenzial für ein stilles Fortschreiten der Krankheit und stellt damit eine mechanistische Erklärung für die Ergebnisse der GLOBAL PBC Study Group dar [13]. Daher wird der ALP-Wert unter PBC-Behandlung als entscheidender Prädiktor für das Langzeitüberleben und das Risiko einer Krankheitsprogression angesehen. Entsprechend sollte das Behandlungsziel die Normalisierung der ALP, respektive der Leberwerte sein, um eine vollständige Kontrolle der Krankheitsaktivität bei der PBC zu erreichen [13]. Es ist mittlerweile allgemein anerkannt, dass jede Form von Abnormalität bei den Leberwerten mit einer anhaltenden Leberentzündung, einem erhöhten Risiko für Fibrose und schlechteren klinischen Ergebnissen in Verbindung steht [13].

Zusammengefasst ist der nachweislich ausschlaggebende Faktor für die Krankheitsprogression bei der PBC die entzündliche Krankheitsaktivität in der Leber, die sich in den erhöhten ALP-Werten widerspiegelt. Folglich kann die ALP-Normalisierung als geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC herangezogen werden [15]. Die ALP-Normalisierung wird in der vorliegenden Nutzenbewertung als nicht-schwerer Endpunkt eingestuft.

Kombiniertes biochemisches Ansprechen

Das kombinierte biochemische Ansprechen ist definiert als ALP $<1,67$ x ULN, Gesamt-Bilirubin $\leq$ ULN und ALP-Reduktion von $\geq 15\%$. Eine Überschreitung dieser Grenzwerte für ALP und Bilirubin wird aktuell als klinische Indikation für eine Zweitlinientherapie herangezogen [18].

Wie bereits beschrieben, handelt es sich bei der ALP-Normalisierung aufgrund der Korrelation mit dem Grad der Leberschädigung um ein geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC. Ergänzend zur Bedeutung einer ALP-Normalisierung ist auch eine Senkung des ALP-Spiegels mit einem verbesserten transplantationsfreien Überleben oder Tod assoziiert: Studiendaten belegen, dass unabhängig von Risikofaktoren wie Alter und Krankheitsstadium die transplantationsfreie Überlebenszeit steigt, je niedriger der ALP-Wert ist. In einer Kohorte mit PBC-Patienten zeigte sich die Assoziation zwischen ALP-Werten und dem Risiko einer Lebertransplantation oder Tod sowohl für Baseline-ALP-Werte, als auch im Rahmen der Nachbeobachtungszeit von bis zu 5 Jahren und unabhängig von der UDCA-Therapie. Die transplantationsfreien Überlebensraten waren nach 5, 10 und 15 Jahren bei Patienten mit ALP $\leq 2,0$ x ULN deutlich höher als bei

Patienten mit höheren ALP-Werten ($>2,0 \times \text{ULN}$) [15]. Folglich kann der ALP-Wert unter PBC-Behandlung (beispielsweise mit UDCA) als entscheidender Prädiktor für das Langzeitüberleben und das Risiko einer Krankheitsprogression angesehen werden [15]. Im Rahmen klinischer Studien wird bezüglich des kombinierten biochemischen Ansprechens ein Cut-off von $1,67 \times \text{ULN}$ als geeignet angesehen [15]. Eine Senkung des ALP-Wertes $<1,67 \times \text{ULN}$ und damit unter den genannten Schwellenwert von $2,0 \times \text{ULN}$ stellt daher analog zur ALP-Normalisierung ein valides Surrogat dar.

Neben der Senkung des ALP-Wertes wird eine Veränderung des ALP-Wertes von über 10% im Krankheitsverlauf und damit oberhalb möglicher natürlicher Schwankungen als relevant im Rahmen klinischer Studien erachtet [19]. Die hier betrachtete ALP-Reduktion um $\geq 15\%$ kann daher als adäquates Maß zur Erfassung einer klinisch relevanten Verbesserung angesehen werden.

Neben der ALP gilt auch das Gesamt-Bilirubin als etablierter prognostischer Faktor der PBC [14, 16, 19]. Bei der PBC ist besonders das Gesamt-Bilirubin von klinischer Relevanz. Es setzt sich zusammen aus dem direkten (wasserlösliches) und dem indirekten (unkonjugiertes, wasserunlösliches) Bilirubin. Ein Anstieg des Gesamt-Bilirubins erfolgt bei der PBC primär dadurch, dass es zu einer gestörten Sekretion des direkten Bilirubins über die Gallengänge kommt [20]. Primär dient der Gesamt-Bilirubinwert als wichtiges Surrogat für das Progressionsrisiko, respektive die klinischen Endpunkte Lebertransplantation oder Tod [14, 15, 19] und wird sowohl bei Erstdiagnose als auch im Krankheitsverlauf zur Ermittlung des Krankheitsstadiums der PBC herangezogen. Die mit fortschreitender PBC auftretende erhöhte Bilirubin-Konzentration ist mit einer sinkenden Überlebenswahrscheinlichkeit assoziiert [14, 15, 19]. Das Risiko für eine Lebertransplantation oder Tod steigt bei Gesamt-Bilirubinwerten jenseits des Schwellenwertes von $0,6 \times \text{ULN}$ an. Eine Senkung des Bilirubinwertes $<0,6 \times \text{ULN}$ nach einem Jahr unter UDCA-Therapie führte zu einer Verbesserung der 10-Jahres-Überlebensrate um 11%. Folglich ist das Erreichen von Gesamt-Bilirubinwerten $\leq 0,6 \times \text{ULN}$ bei Patienten mit PBC mit dem geringsten Risiko für eine Lebertransplantation oder Tod verbunden [16]. Untersuchungen in einem kanadischen Register (mit einer zusätzlichen Validierungskohorte) zum Entzündungs- und Stoffwechselproteom im Serum von PBC-Patienten unter UDCA-Therapie zeigen ebenfalls, dass für das Therapieziel einer vollständigen Kontrolle der Krankheitsaktivität bei PBC sowohl eine ALP-Normalisierung als auch eine Normalisierung des Bilirubins erforderlich ist [13]. Basierend auf den Daten der GLOBE-Kohorte zeigt sich weiter, dass erhöhte Gesamt-Bilirubinwerte (sowohl zu Baseline als auch nach 5 Jahren) mit einer schlechteren Prognose assoziiert sind. Die transplantationsfreien Überlebensraten nach 5, 10 und 15 Jahren waren bei Patienten mit normalen Gesamt-Bilirubinwerten ($\leq 1 \times \text{ULN}$) signifikant höher als bei Patienten mit erhöhten Gesamt-Bilirubinwerten [15]. Obschon die Bilirubinwerte meist relativ spät im Krankheitsverlauf ansteigen, ist der prädiktive Wert nicht nur auf das späte Krankheitsstadium beschränkt und selbst bei Patienten mit normalen Werten verbessert sich die Prognose, wenn die Werte sinken [15]. Bei anderen Lebererkrankungen hat sich der Gesamt-Bilirubinwert ebenfalls als prädiktiv für den klinischen Verlauf bestätigt und ist in gängigen Prognosemodellen in Kombination mit weiteren Leberwerten enthalten [15, 19]. Als

Schwellenwert für das Gesamt-Bilirubin wird in der Literatur der im Rahmen des kombinierten biochemischen Ansprechens verwendete Cut-off von $\leq 1,0 \times \text{ULN}$ als adäquat angesehen [15]. Der Gesamt-Bilirubinwert kann daher als wichtiger Prädiktor für einen ungünstigen Verlauf der PBC und das individuelle Risiko einer Krankheitsprogression angesehen werden [14].

Untersuchungen aus dem kanadischen Register (mit einer zusätzlichen Validierungskohorte) zeigen unterstützend dazu, dass der Verlust des biochemischen Ansprechens zu einem beliebigen Zeitpunkt mit einem erhöhten Risiko für Lebertransplantation oder Tod korreliert – umgekehrt ist das (Wieder-)Erreichen eines biochemischen Ansprechens im Verlauf der Krankheit zu jedem Zeitpunkt mit besseren klinischen Ergebnissen verbunden. Ein weiterer Beleg dafür, dass das Erreichen eines nachhaltigen biochemischen Ansprechens auf die Therapie ein wichtiges Ziel während des gesamten Krankheitsverlaufs darstellt [18].

Folglich werden die Einzelkomponenten des kombinierten biochemischen Ansprechens (ALP und Gesamt-Bilirubin) ebenfalls als geeignetes Surrogat für die klinischen Endpunkte zur Messung der Morbidität/Krankheitsaktivität und Krankheitsschwere, respektive der Endpunkte Lebertransplantation oder Tod bei der PBC herangezogen. Es wird davon ausgegangen, dass sowohl eine Abnahme der ALP- als auch der Gesamt-Bilirubinwerte mit einer besseren Prognose für den Patienten assoziiert sind, da niedrigere Werte eindeutig mit einem besseren transplantationsfreien Überleben verbunden sind. Für die vorliegende Nutzenbewertung wird das kombinierte biochemische Ansprechen sowie die stetigen Analysen von ALP und Gesamt-Bilirubin als nicht-schwere Endpunkte eingestuft.

Ergänzend werden die Veränderungen der ALT, AST, des direkten und indirekten Bilirubins sowie der GGT und der 5'-Nucleotidase dargestellt. Diese Laborparameter korrelieren ebenfalls mit dem Schweregrad der PBC-Erkrankung und können in Kombination mit ALP und/oder Bilirubin zur Beurteilung der Krankheitsaktivität herangezogen werden. Beispielsweise wird in der klinischen Praxis der hepatische Ursprung eines erhöhten ALP-Wertes meist durch eine gleichzeitige Erhöhung der GGT (oder der 5'-Nukleotidase) sowie des direkten Bilirubins bestätigt [4, 14, 15, 21, 22].

Zusammengefasst ist die positive Auswirkung der hier relevanten Surrogatparameter (ALP-Normalisierung und kombiniertes biochemisches Ansprechen) im Hinblick auf die langfristige Vermeidung der Folgekomplikation und somit die klinischen Endpunkte Lebertransplantation und Tod qualitativ belegt. Biochemisches Ansprechen stellt somit in der vorliegenden Indikation einen patientenrelevanten Endpunkt dar.

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] Gilead Sciences Ireland UC. Fachinformation Seladelpar Gilead 10 mg Hartkapsel. Stand der Information: Februar 2025.
- [2] European Parliament, Council of the European Union. Regulation (EC) No 141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products. 2000.
- [3] Bundesministerium der Justiz. Das Fünfte Buch Sozialgesetzbuch – Gesetzliche Krankenversicherung – (Artikel 1 des Gesetzes vom 20. Dezember 1988, BGBl. I S. 2477, 2482), das zuletzt durch Artikel 10 des Gesetzes vom 18. Dezember 2024 (BGBl. 2024 I Nr. 423) geändert worden ist. 2024. URL: https://www.gesetze-im-internet.de/sgb_5/_35a.html (aufgerufen am: 5. Februar 2025).
- [4] Sebode M, Bantel H, Baumann U, Bufler P, Dold L, Hohenester S, et al. S3-Leitlinie „Seltene Lebererkrankungen (LeiSe LebEr) – autoimmune Lebererkrankungen von der Pädiatrie bis zum Erwachsenenalter“ der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS). Februar 2025 – AWMF-Registernummer: 021 – 027 2025.
- [5] Gemeinsamer Bundesausschuss. 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 – Emtricitabin/Rilpivirin/Tenofoviridisoproxil (neues Anwendungsgebiet). 19. Juni 2014.
- [6] Gemeinsamer Bundesausschuss. 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 – Bictgravir/Emtricitabin/Tenofovirafenamid. 20. Dezember 2018.
- [7] European Medicines Agency. Reflection paper on regulatory requirements for the development of medicinal products for primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC). 4. Dezember 2023.
- [8] Gemeinsamer Bundesausschuss. 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 – Dapagliflozin. 6. Juni 2013.
- [9] Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V) und Anlage XIIa – Kombinationen von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Evinacumab (homozygote familiäre Hypercholesterinämie, ≥ 12 Jahre). 4. Juli 2024.
- [10] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Evinacumab (homozygote familiäre Hypercholesterinämie, Erwachsene und Jugendliche ab 12 Jahre). Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 1762. 8. April 2024.
- [11] Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses (Fassung vom 18. Dezember 2008, zuletzt geändert durch den Beschluss vom 19. September 2024). Stand: 3. Januar 2025.
- [12] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden. Version 7.0 vom 19. September 2023.

- [13] Jones DEJ, Wetten A, Barron-Millar B, Ogle L, Mells G, Flack S, et al. The relationship between disease activity and UDCA response criteria in primary biliary cholangitis: A cohort study. *EBioMedicine*. 2022 Jun;80:104068.
- [14] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol*. 2017 Jul;67(1):145-72.
- [15] Lammers WJ, van Buuren HR, Hirschfield GM, Janssen HL, Invernizzi P, Mason AL, et al. Levels of alkaline phosphatase and bilirubin are surrogate end points of outcomes of patients with primary biliary cirrhosis: an international follow-up study. *Gastroenterology*. 2014 Dec;147(6):1338-49 e5.
- [16] Murillo Perez CF, Harms MH, Lindor KD, van Buuren HR, Hirschfield GM, Corpechot C, et al. Goals of Treatment for Improved Survival in Primary Biliary Cholangitis: Treatment Target Should Be Bilirubin Within the Normal Range and Normalization of Alkaline Phosphatase. *Am J Gastroenterol*. 2020 Jul;115(7):1066-74.
- [17] Cancado GGL, Fucuta PDS, Gomes NMF, Couto CA, Cancado ELR, Terrabuio DRB, et al. Alkaline phosphatase and liver fibrosis at diagnosis are associated with deep response to ursodeoxycholic acid in primary biliary cholangitis. *Clin Res Hepatol Gastroenterol*. 2024 Oct;48(8):102453.
- [18] Roberts SB, Choi WJ, Worobetz L, Vincent C, Flemming JA, Cheung A, et al. Loss of biochemical response at any time worsens outcomes in UDCA-treated patients with primary biliary cholangitis. *JHEP Rep*. 2024 Oct;6(10):101168.
- [19] Silveira MG, Brunt EM, Heathcote J, Gores GJ, Lindor KD, Mayo MJ. American Association for the Study of Liver Diseases endpoints conference: design and endpoints for clinical trials in primary biliary cirrhosis. *Hepatology*. 2010 Jul;52(1):349-59.
- [20] Reshetnyak VI, Maev IV. Pathophysiology of biochemical signs of primary biliary cholangitis. *Exploration of Digestive Diseases*. 2023;2(4):149-71.
- [21] Hirschfield GM, Shiffman ML, Gulamhusein A, Kowdley KV, Vierling JM, Levy C, et al. Seladelpar efficacy and safety at 3 months in patients with primary biliary cholangitis: ENHANCE, a phase 3, randomized, placebo-controlled study. *Hepatology*. 2023 Aug 1;78(2):397-415.
- [22] Lammers WJ, Hirschfield GM, Corpechot C, Nevens F, Lindor KD, Janssen HL, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology*. 2015 Dec;149(7):1804-12 e4.
- [23] Schregel I, Steinmann S, Schramm C. Autoimmune Lebererkrankungen – ein Update (Teil 2). Primär biliäre Cholangitis (PBC) und primär sklerosierende Cholangitis (PSC). Thieme. 2022.
- [24] Tanaka A. Current understanding of primary biliary cholangitis. *Clin Mol Hepatol*. 2021 Jan;27(1):1-21.
- [25] Hassan A, Hurtado Diaz De Leon I, Tapper EB. Symptom burden in chronic liver disease. *Gastroenterol Rep (Oxf)*. 2024;12:goae078.
- [26] Smith H, Fettiplace J, Von Maltzahn R, Das S, McLaughlin MM, Jones D. More Than Just an Itch: Impact of Cholestatic Pruritus in Primary Biliary Cholangitis (PBC) on Health-Related Quality of Life (HRQoL). S107. 2022.
- [27] Trivella J, John BV, Levy C. Primary biliary cholangitis: Epidemiology, prognosis, and treatment. *Hepatol Commun*. 2023 Jun 1;7(6).

- [28] Phan NQ, Blome C, Fritz F, Gerss J, Reich A, Ebata T, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritus. *Acta Derm Venereol.* 2012 Sep;92(5):502-7.
- [29] Mathur VS, Lindberg J, Germain M, Block G, Tumlin J, Smith M, et al. A longitudinal study of uremic pruritus in hemodialysis patients. *Clin J Am Soc Nephrol.* 2010 Aug;5(8):1410-9.
- [30] Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. *Br J Dermatol.* 2010 Mar;162(3):587-93.
- [31] Eremenco S, Chen WH, Blum SI, Bush EN, Bushnell DM, DeBusk K, et al. Comparing patient global impression of severity and patient global impression of change to evaluate test-retest reliability of depression, non-small cell lung cancer, and asthma measures. *Qual Life Res.* 2022 Dec;31(12):3501-12.
- [32] Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V). Olipudase alfa (Mangel an saurer Sphingomyelinase (ASMD) Typ A/B oder Typ B). 16. März 2023.
- [33] AstraZeneca GmbH. Dossier zur Nutzenbewertung gemäß § 35a SGB V. Modul 4A. Dapagliflozin (Forxiga® 5 und 10 mg Filmtabletten). Zur Behandlung einer symptomatischen, chronischen Herzinsuffizienz mit reduzierter Ejektionsfraktion bei erwachsenen Patienten. 30. November 2020.
- [34] Snyder ES, Tao P, Svetnik V, Lines C, Herring WJ. Use of the single-item Patient Global Impression-Severity scale as a self-reported assessment of insomnia severity. *J Sleep Res.* 2021 Feb;30(1):e13141.
- [35] Rampakakis E, Ste-Marie PA, Sampalis JS, Karellis A, Shir Y, Fitzcharles MA. Real-life assessment of the validity of patient global impression of change in fibromyalgia. *RMD Open.* 2015;1(1):e000146.
- [36] Viktrup L, Hayes RP, Wang P, Shen W. Construct validation of patient global impression of severity (PGI-S) and improvement (PGI-I) questionnaires in the treatment of men with lower urinary tract symptoms secondary to benign prostatic hyperplasia. *BMC Urol.* 2012 Nov 7;12:30.
- [37] Somnay K, Wadgaonkar P, Sridhar N, Roshni P, Rao N, Wadgaonkar R. Liver Fibrosis Leading to Cirrhosis: Basic Mechanisms and Clinical Perspectives. *Biomedicines.* 2024 Sep 30;12(10).
- [38] Horwich BH, Han H. Diagnosis and Treatment of Primary Biliary Cholangitis: A Patient-Friendly Summary of the 2018 AASLD Practice Guidance. *Clin Liver Dis (Hoboken).* 2021 Nov;18(5):255-9.
- [39] Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses zur Richtlinie „Ambulante Behandlung im Krankenhaus nach § 116b SGB V“, Anlage 2 Nr. 6. Diagnostik und Versorgung von Patientinnen und Patienten mit biliärer Zirrhose. 17. September 2009.
- [40] Gomez-Dominguez E, Mendoza J, Garcia-Buey L, Trapero M, Gisbert JP, Jones EA, et al. Transient elastography to assess hepatic fibrosis in primary biliary cirrhosis. *Aliment Pharmacol Ther.* 2008 Mar 1;27(5):441-7.
- [41] Floreani A, Cazzagon N, Martines D, Cavalletto L, Baldo V, Chemello L. Performance and utility of transient elastography and noninvasive markers of liver fibrosis in primary biliary cirrhosis. *Dig Liver Dis.* 2011 Nov;43(11):887-92.

- [42] Friedrich-Rust M, Zeuzem S. [Transient elastography (FibroScan) for the non-invasive assessment of liver fibrosis: current status and perspectives]. *Z Gastroenterol.* 2007 May;45(5):387-94.
- [43] Guechot J, Trocme C, Renversez JC, Sturm N, Zarski JP, Group AHEFS. Independent validation of the Enhanced Liver Fibrosis (ELF) score in the ANRS HC EP 23 Fibrostar cohort of patients with chronic hepatitis C. *Clin Chem Lab Med.* 2012 Jan 13;50(4):693-9.
- [44] Jacoby A, Rannard A, Buck D, Bhala N, Newton JL, James OF, et al. Development, validation, and evaluation of the PBC-40, a disease specific health related quality of life measure for primary biliary cirrhosis. *Gut.* 2005 Nov;54(11):1622-9.
- [45] Gemeinsamer Bundesausschuss. Zusammenfassende Dokumentation über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Obeticholsäure. 6. Juli 2017.
- [46] Corpechot C, Chazouilleres O, Poupon R. Early primary biliary cirrhosis: biochemical response to treatment and prediction of long-term outcome. *J Hepatol.* 2011 Dec;55(6):1361-7.
- [47] Hirschfield GM, Bowlus CL, Mayo MJ, Kremer AE, Vierling JM, Kowdley KV, et al. A Phase 3 Trial of Seladelpar in Primary Biliary Cholangitis. *N Engl J Med.* 2024 Feb 29;390(9):783-94.
- [48] ClinicalTrials.gov. NCT04620733 – Titel: RESPONSE: A Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Efficacy and Safety of Seladelpar in Patients With Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA). 2024. URL: <https://www.clinicaltrials.gov/study/NCT04620733> (aufgerufen am: 28. Januar 2025).
- [49] EU-CTR. 2020-004348-27 – Titel: RESPONSE: A Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Efficacy and Safety of Seladelpar in Patients With Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA). 0000. URL: <https://www.clinicaltrialsregister.eu/ctr-search/search?query=2020-004348-27> (aufgerufen am: 28. Januar 2025).
- [50] WHO-ICTRP. NCT04620733 – Titel: RESPONSE: A Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Efficacy and Safety of Seladelpar in Patients With Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA). 2023. URL: <https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04620733> (aufgerufen am: 28. Januar 2025).
- [51] Gilead Sciences Inc. RESPONSE: A Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Efficacy and Safety of Seladelpar in Patients with Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA). 2024.
- [52] Hirschfield GM, Dyson JK, Alexander GJM, Chapman MH, Collier J, Hubscher S, et al. The British Society of Gastroenterology/UK-PBC primary biliary cholangitis treatment and management guidelines. *Gut.* 2018 Sep;67(9):1568-94.
- [53] Statistisches Bundesamt. Gesundheitszustand und -relevantes Verhalten: Körpermaße nach Altersgruppen und Geschlecht – Insgesamt. Körpermaße der Bevölkerung nach Altersgruppen 2021 (Endergebnisse). Stand: 27. März 2023. URL: <https://www.destatis.de/DE/Themen/Gesellschaft->

- Umwelt/Gesundheit/Gesundheitszustand-Relevantes-Verhalten/Tabellen/liste-koerpermasse.html (aufgerufen am: 5. Februar 2025).
- [54] Wilde AB, Lieb C, Leicht E, Greverath LM, Steinhagen LM, Wald de Chamorro N, et al. Real-World Clinical Management of Patients with Primary Biliary Cholangitis-A Retrospective Multicenter Study from Germany. *J Clin Med*. 2021 Mar 4;10(5).
 - [55] Agraria Pharma GmbH. Fachinformation. Vasosan® P (Colestyramin). Stand der Information: Januar 2023.
 - [56] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol*. 2009 Aug;51(2):237-67.
 - [57] Lindor KD, Bowlus CL, Boyer J, Levy C, Mayo M. Primary Biliary Cholangitis: 2018 Practice Guidance from the American Association for the Study of Liver Diseases. *Hepatology*. 2019 Jan;69(1):394-419.
 - [58] European Medicines Agency. Revocation of conditional marketing authorisation for Ocaliva. Benefits of Ocaliva no longer considered to outweigh its risks. 3. Dezember 2024.
 - [59] Kremer AE, Mayo MJ, Hirschfield G, Levy C, Bowlus CL, Jones DE, et al. Seladelpar improved measures of pruritus, sleep, and fatigue and decreased serum bile acids in patients with primary biliary cholangitis. *Liver Int*. 2022 Jan;42(1):112-23.
 - [60] Colapietro F, Gershwin ME, Lleo A. PPAR agonists for the treatment of primary biliary cholangitis: Old and new tales. *J Transl Autoimmun*. 2023;6:100188.
 - [61] Jones D, Boudes PF, Swain MG, Bowlus CL, Galambos MR, Bacon BR, et al. Seladelpar (MBX-8025), a selective PPAR-delta agonist, in patients with primary biliary cholangitis with an inadequate response to ursodeoxycholic acid: a double-blind, randomised, placebo-controlled, phase 2, proof-of-concept study. *Lancet Gastroenterol Hepatol*. 2017 Oct;2(10):716-26.
 - [62] Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc*. 2006 Oct;94(4):451-5.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ²²] – 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

²² 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

Eine bibliografische Literaturrecherche zum zbAM SEL unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, wurde durchgeführt. Die Suchstrategien sind in Tabelle 4-77, Tabelle 4-78 und Tabelle 4-79 dargestellt. Die Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Die Suche wurde für Abschnitt 4.3.1.1.2 durchgeführt.

Datenbankname	EMBASE Classic+EMBASE
Suchoberfläche	Ovid
Datum der Suche	15.01.2025
Zeitsegment	1947 bis 14.01.2025
Suchfilter	Filter für RCT nach Wong 2006-Strategy minimizing difference between sensitivity and specificity [62]

Tabelle 4-77 (Anhang): Ovid: Suchstrategie EMBASE Classic+EMBASE 1947 bis heute: 14.01.2025

#	Suchbegriffe	Ergebnis
1.	exp seladelpar/	236
2.	(seladelpar* or MBX-8025* or MBX8025* or RWJ-800025* or RWJ800025* or 928821-40-3 or 851528-79-5).mp.	244
3.	(928821-40-3 or 851528-79-5).rn.	201
4.	1 or 2 or 3	244
5.	exp primary biliary cirrhosis/ or exp primary biliary cholangitis/	15.098
6.	(cholangit* or cirrhos*).mp.	310.539
7.	bili*.mp.	364.436
8.	6 and 7	68.704
9.	5 or 8	68.704
10.	4 and 9	147
11.	(random* or double-blind*).tw. or placebo*.mp.	2.455.597
12.	10 and 11	60

Datenbankname	Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions
Suchoberfläche	Ovid
Datum der Suche	15.01.2025
Zeitsegment	1946 bis 14.01.2025
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity

Tabelle 4-78 (Anhang): Ovid: Suchstrategie: MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions 1946 bis heute: 14.01.2025

#	Suchbegriffe	Ergebnis
1.	(seladelpar* or MBX-8025* or MBX8025* or RWJ-800025* or RWJ800025* or 928821-40-3 or 851528-79-5).mp.	50
2.	(928821-40-3 or 851528-79-5).rn.	0
3.	1 or 2	50
4.	exp Liver Cirrhosis, Biliary/ or exp Cholangitis/	20.005
5.	(cholangit* or cirrhos*).mp.	178.684
6.	bili*.mp.	189.848
7.	5 and 6	30.884
8.	4 or 7	36.814
9.	3 and 8	38
10.	(randomi#ed or placebo).mp. or randomi#ed controlled trial.pt.	1.232.383
11.	9 and 10	13

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials
Suchoberfläche	Ovid
Datum der Suche	15.01.2025
Zeitsegment	bis Dezember 2024
Suchfilter	Kein Suchfilter verwendet

Tabelle 4-79 (Anhang): Ovid Suchstrategie: EBM Reviews – Cochrane Central Register of Controlled Trials bis Dezember 2024

#	Suchbegriffe	Ergebnis
1.	(seladelpar* or MBX-8025* or MBX8025* or RWJ-800025* or RWJ800025* or 928821-40-3 or 851528-79-5).mp.	62
2.	exp Liver Cirrhosis, Biliary/ or exp Cholangitis/	666
3.	(cholangit* or cirrhos*).mp.	13.397
4.	bili*.mp.	19.376
5.	3 and 4	2.717
6.	2 or 5	2.834
7.	1 and 6	53

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

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

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

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

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

Die Suche nach registrierten Studien mit dem zbAM SEL erfolgte anhand spezifischer Stichwörter. Die Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Die Suche in dem Portal ClinicalTrials.gov wurde über die Advanced Search-Funktion durchgeführt.

Die Suche wurde für Abschnitt 4.3.1.1.3 durchgeführt.

Studienregister	ClinicalTrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	15.01.2025
Suchstrategie	Seladelpar OR MBX-8025 OR MBX8025 [Intervention/Treatment]
Treffer	13

Studienregister	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	15.01.2025
Suchstrategie	Seladelpar OR MBX-8025 OR MBX8025 OR (MBX 8025) OR (928821-40-3)
Treffer	8

Studienregister	WHO-ICTRP
Internetadresse	https://trialsearch.who.int/AdvSearch.aspx
Datum der Suche	15.01.2025
Suchstrategie	Seladelpar OR MBX-8025 OR MBX8025 OR MBX 8025 OR 928821-40-3
Treffer	19

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

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

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

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

Nicht zutreffend.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

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

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-80 (Anhang): Übersicht der ausgeschlossenen Studien aus der Studienregistersuche zum zbAM

Register	Trefferzahl entsprechend Anhang 4-B	Ausgeschlossene Registereinträge (Anhang 4-D)	Eingeschlossene Registereinträge (Tabelle 4-5)
ClinicalTrials.gov	13	12 (Nr. 1 – 12)	1
EU-CTR	8	7 (Nr. 13 – 19)	1
WHO-ICTRP	19	18 (Nr. 20 – 37)	1
Summe	$\Sigma=40$	$\Sigma=37$	$\Sigma=3$

Nr.	Studien-ID	Titel	Ausschlussgrund
ClinicalTrials.gov			
1.	NCT00701883	Multicenter Randomized Double-Blind Placebo and Active Comparator Controlled Parallel-group, Multicenter Study to Evaluate the Safety, Tolerability, Pharmacokinetics and Efficacy of MBX-8025 in Moderately Obese Hyperlipidemic Patients With/Out Concomitant Atorvastatin". ClinicalTrials.gov. 7-08. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT00701883	A1 andere Population
2.	NCT02472535	A 12-week, Open-label, Dose-escalating, Phase 2 Study to Evaluate the Effects of MBX-8025 in Patients With Homozygous Familial Hypercholesterolemia (HoFH). ClinicalTrials.gov. 5-04. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT02472535	A5 anderer Studientyp

Nr.	Studien-ID	Titel	Ausschlussgrund
3.	NCT02609048	A 12-week, Double-blind, Randomized, Placebo-controlled, Phase 2 Study, to Evaluate the Effects of Two Doses of MBX-8025 in Subjects With Primary Biliary Cirrhosis (PBC) and an Inadequate Response to Ursodeoxycholic Acid (UDCA). ClinicalTrials.gov. 5-11. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT02609048	A2 andere Intervention
4.	NCT02955602	An 8-week, Dose Ranging, Open Label, Randomized, Phase 2 Study With a 44-week Extension, to Evaluate the Safety and Efficacy of MBX-8025 in Subjects With Primary Biliary Cholangitis (PBC) and an Inadequate Response to or Intolerance to Ursodeoxycholic Acid (UDCA). ClinicalTrials.gov. 1-28. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT02955602	A3 andere Vergleichs-therapie
5.	NCT03301506	ASSURE: An Open Label Long-Term Study to Evaluate the Safety and Tolerability of Seladelpar in Subjects With Primary Biliary Cholangitis (PBC). ClinicalTrials.gov. 2-12. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT03301506	A5 anderer Studientyp
6.	NCT03369002	A Phase 1, Open Label, Non-Randomized Study to Evaluate the Pharmacokinetics, Safety, and Tolerability of Oral Seladelpar in Subjects With Varying Degrees of Hepatic Impairment and Healthy Matched Control Subjects With Normal Hepatic Function. ClinicalTrials.gov. 1-27. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT03369002	A5 anderer Studientyp
7.	NCT03551522	A Phase 2, Double-Blind, Randomized, Placebo-Controlled Study Followed by an Open-Label Extension Period to Evaluate the Activity of Seladelpar in Subjects With Nonalcoholic Steatohepatitis (NASH). ClinicalTrials.gov. 4-30. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT03551522	A1 andere Population
8.	NCT03602560	A 52-week, Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Safety and Efficacy of Seladelpar in Subjects With Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA). ClinicalTrials.gov. 0-01. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT03602560	A6 andere Studiendauer

Nr.	Studien-ID	Titel	Ausschlussgrund
9.	NCT04024813	A Phase 2, Randomized, Double Blind, Placebo Controlled, Multiple Center Study to Evaluate the Safety, Tolerability, and Efficacy of Seladelpar Administered for 24 Weeks in Adult Patients With Primary Sclerosing Cholangitis (PSC). ClinicalTrials.gov. 1-12. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT04024813	A1 andere Population
10.	NCT04950764	The Effect of Hepatic Impairment on The Pharmacokinetics of Seladelpar: An Open-Label Study Following Oral Dosing of Seladelpar to Subjects With Primary Biliary Cholangitis (PBC) and Hepatic Impairment. ClinicalTrials.gov. 9-17. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT04950764	A5 anderer Studientyp
11.	NCT06051617	AFFIRM: A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Effect of Seladelpar on Clinical Outcomes in Patients With Primary Biliary Cholangitis (PBC) and Compensated Cirrhosis. ClinicalTrials.gov. 9-07. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT06051617	A7 anderer Publikationstyp
12.	NCT06060665	IDEAL: A 52-week, Double-blind, Placebo-controlled, Randomized, Phase 3 Study Intended to Determine the Effects of Seladelpar on Normalization of Alkaline Phosphatase Levels in Subjects With Primary Biliary Cholangitis (PBC) and an Incomplete Response or Intolerance to Ursodeoxycholic Acid (UDCA). ClinicalTrials.gov. 9-05. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrials.gov/study/NCT06060665	A7 anderer Publikationstyp
EU-CTR			
13.	2015-002698-39	A 12-week, double-blind, randomized, placebo-controlled, Phase 2 study to evaluate the effects of two doses of MBX-8025 in subjects with Primary Biliary Cirrhosis (PBC) and an inadequate response to ursodeoxycholic acid (UDCA).. EU-CTR. 2015. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2015-002698-39	A2 andere Intervention
14.	2017-003910-16	An Open Label Long-Term Study to Evaluate the Safety and Tolerability of Seladelpar in Subjects with Primary Biliary Cholangitis (PBC). EU-CTR. 2017. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2017-003910-16	A5 anderer Studientyp

Nr.	Studien-ID	Titel	Ausschlussgrund
15.	2020-005198-29	ASSURE: An Open Label Long-Term Study to Evaluate the Safety and Tolerability of Seladelpar in Subjects with Primary Biliary Cholangitis (PBC). EU-CTR. 2021. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2020-005198-29	A5 anderer Studientyp
16.	2018-001171-20	A 52-week, placebo-controlled, randomized, Phase 3 study to evaluate the safety and efficacy of seladelpar in subjects with primary biliary cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA).. EU-CTR. 2018. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2018-001171-20	A6 andere Studiendauer
17.	2019-001760-30	A Phase 2, Randomized, Double-Blind, Placebo-Controlled Multiple Center Study to Evaluate the Safety, Tolerability, and Efficacy of Seladelpar Administered for 24 Weeks in Adult Patients with Primary Sclerosing Cholangitis (PSC). EU-CTR. 2019. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2019-001760-30	A1 andere Population
18.	2014-004856-68	A 12-week, open-label, dose-escalating, phase 2 study to evaluate the effects of MBX-8025 in patients with Homozygous Familial Hypercholesterolemia (HoFH). EU-CTR. 2015. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2014-004856-68	A1 andere Population
19.	2016-002996-91	An 8-week, dose ranging, open label, randomized, Phase 2 study with a 44-week extension, to evaluate the safety and efficacy of MBX-8025 in subjects with Primary Biliary Cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA). EU-CTR. 2016. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=2016-002996-91	A3 andere Vergleichs-therapie
WHO-ICTRP			
20.	NCT00701883	Multicenter Randomized Double-Blind Placebo and Active Comparator Controlled Parallel-Group, Multicenter Study to Evaluate the Safety, Tolerability, Pharmacokinetics and Efficacy of MBX-8025 in Moderately Obese Hyperlipidemic Patients With/Out Concomitant Atorvastatin". ICTRP. 2015. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00701883	A1 andere Population

Nr.	Studien-ID	Titel	Ausschlussgrund
21.	NCT02472535	A 12-week, Open-label, Dose-escalating, Phase 2 Study to Evaluate the Effects of MBX-8025 in Patients With Homozygous Familial Hypercholesterolemia (HoFH). ICTRP. 2016. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02472535	A5 anderer Studientyp
22.	EUCTR2015-002698-39-GB	A 12-week, double-blind, randomized, placebo-controlled, Phase 2 study to evaluate the effects of two doses of MBX-8025 in subjects with Primary Biliary Cirrhosis (PBC) and an inadequate response to ursodeoxycholic acid (UDCA).. ICTRP. 2020. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-002698-39-GB	A2 andere Intervention
23.	EUCTR2020-005198-29-PL	ASSURE: An Open Label Long-Term Study to Evaluate the Safety and Tolerability of Seladelpar in Subjects with Primary Biliary Cholangitis (PBC) - ASSURE. ICTRP. 2024. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2020-005198-29-PL	A5 anderer Studientyp
24.	NCT03369002	A Phase 1, Open Label, Non-Randomized Study to Evaluate the Pharmacokinetics, Safety, and Tolerability of Oral Seladelpar in Subjects With Varying Degrees of Hepatic Impairment and Healthy Matched Control Subjects With Normal Hepatic Function. ICTRP. 2020. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03369002	A5 anderer Studientyp
25.	NCT03602560	A 52-week, Placebo-controlled, Randomized, Phase 3 Study to Evaluate the Safety and Efficacy of Seladelpar in Subjects With Primary Biliary Cholangitis (PBC) and an Inadequate Response to or an Intolerance to Ursodeoxycholic Acid (UDCA). ICTRP. 2021. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03602560	A6 andere Studiendauer
26.	EUCTR2018-001171-20-AT	A 52-week, placebo-controlled, randomized, Phase 3 study to evaluate the safety and efficacy of seladelpar in subjects with primary biliary cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA).. ICTRP. 2020. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-001171-20-AT	A6 andere Studiendauer

Nr.	Studien-ID	Titel	Ausschlussgrund
27.	EUCTR2018-001171-20-FR	A 52-week, placebo-controlled, randomized, Phase 3 study to evaluate the safety and efficacy of seladelpar in subjects with primary biliary cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA).. ICTRP. 2020. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-001171-20-FR	A6 andere Studiendauer
28.	EUCTR2018-001171-20-GR	A 52-week, placebo-controlled, randomized, Phase 3 study to evaluate the safety and efficacy of seladelpar in subjects with primary biliary cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA).. ICTRP. 2020. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-001171-20-GR	A6 andere Studiendauer
29.	EUCTR2018-001171-20-PL	A 52-week, placebo-controlled, randomized, Phase 3 study to evaluate the safety and efficacy of seladelpar in subjects with primary biliary cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA).. ICTRP. 2020. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-001171-20-PL	A6 andere Studiendauer
30.	NCT04024813	A Phase 2, Randomized, Double Blind, Placebo Controlled, Multiple Center Study to Evaluate the Safety, Tolerability, and Efficacy of Seladelpar Administered for 24 Weeks in Adult Patients With Primary Sclerosing Cholangitis (PSC). ICTRP. 2021. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04024813	A1 andere Population
31.	NCT04950764	The Effect of Hepatic Impairment on The Pharmacokinetics of Seladelpar: An Open-Label Study Following Oral Dosing of Seladelpar to Subjects With Primary Biliary Cholangitis (PBC) and Hepatic Impairment. ICTRP. 2024. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04950764	A5 anderer Studientyp
32.	NCT06051617	AFFIRM: A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Effect of Seladelpar on Clinical Outcomes in Patients With Primary Biliary Cholangitis (PBC) and Compensated Cirrhosis. ICTRP. 2025. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT06051617	A7 anderer Publikationstyp

Nr.	Studien-ID	Titel	Ausschlussgrund
33.	NCT06060665	IDEAL: A 52-week, Double-blind, Placebo-controlled, Randomized, Phase 3 Study Intended to Determine the Effects of Seladelpar on Normalization of Alkaline Phosphatase Levels in Subjects With Primary Biliary Cholangitis (PBC) and an Incomplete Response or Intolerance to Ursodeoxycholic Acid (UDCA). ICTRP. 2025. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT06060665	A7 anderer Publikationstyp
34.	EUCTR2014-004856-68-NO	A 12-week, open-label, dose-escalating, phase 2 study to evaluate the effects of MBX-8025 in patients with Homozygous Familial Hypercholesterolemia (HoFH). ICTRP. 2016. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-004856-68-NO	A1 andere Population
35.	EUCTR2016-002996-91-DE	An 8-week, dose ranging, open label, randomized, Phase 2 study with a 44-week extension, to evaluate the safety and efficacy of MBX-8025 in subjects with Primary Biliary Cholangitis (PBC) and an inadequate response to or intolerance to ursodeoxycholic acid (UDCA). ICTRP. 2019. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2016-002996-91-DE	A3 andere Vergleichstherapie
36.	NCT04250155	A Phase Ia/Ib, Open-Label, Multicenter, Global, Dose-Escalation Study to Evaluate the Safety and Pharmacokinetics of XmAb24306 as a Single Agent and in Combination With Atezolizumab in Patients With Locally Advanced or Metastatic Solid Tumors. ICTRP. 2024. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04250155	A1 andere Population
37.	NL-OMON42038	A 12-week, open-label, dose-escalating, phase 2 study to evaluate the effects of MBX-8025 in patients with Homozygous Familial Hypercholesterolemia (HoFH) (study CB8025-21427) - CB8025-21427 (CymaBay). ICTRP. 2024. [Zugriffsdatum: 15.01.2025]. Verfügbar unter: https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NL-OMON42038	A1 andere Population

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E : Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-81 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 4-81 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-81 (Anhang): Studiendesign und -methodik für Studie RESPONSE

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primäre Fragestellung: • Bewertung der Behandlungseffekte von Seladelpar auf die zusammengesetzte biochemische Verbesserung der Cholestase-Marker basierend auf alkalischer Phosphatase (ALP) und Gesamt-Bilirubin nach 12 Monaten Behandlung im Vergleich zu Placebo • Bewertung der Sicherheit von Seladelpar über einen Zeitraum von 12 Monaten im Vergleich zu Placebo Wichtige sekundäre Fragestellungen: • Bewertung der Wirkung von Seladelpar auf die Normalisierung der ALP-Werte nach 12 Monaten Behandlung im Vergleich zu Placebo • Bewertung der Wirkung von Seladelpar auf Pruritus nach 6 Monaten Behandlung im Vergleich zu Placebo bei Patienten mit anfänglich mäßigem bis schweren Pruritus Weitere sekundäre Fragestellungen: • Bewertung der Wirkung von Seladelpar auf andere Cholestase-Messwerte, Stoffwechselmarker und PBC-Prognosekriterien • Bewertung der Wirkung von Seladelpar auf die Lebensqualität • Bewertung der Wirkung von Seladelpar auf PBC-assoziierte klinische Endpunkte Explorative Fragestellungen: • Bewertung der Wirkung von Seladelpar auf die Leberhistologie, zusätzliche Messinstrumente zur Lebensqualität, Biomarker für Cholestase und Entzündung, Lipid- und Autoantikörperprofile, Gallensäuresynthese, Leberfibrose und Leberschäden • Bewertung der Plasmakonzentrationen von Seladelpar und seinen Metaboliten Hypothesentests: Die Fehlerwahrscheinlichkeit vom Typ I von 5% wird für die gesamte Studie eingehalten, indem eine hierarchische, festgelegte Sequenz für die primären und wichtigsten sekundären Wirksamkeitsanalysen wie folgt verwendet wird: Wenn die primäre Wirksamkeitsanalyse für Seladelpar 10 mg im Vergleich zu Placebo bei einem zweiseitigen Signifikanzniveau von 0,05 positiv ausfällt, werden die

Item ^a	Charakteristikum	Studieninformation
		2 wichtigsten sekundären Endpunkte hierarchisch in folgender Reihenfolge analysiert: • Normalisierung der ALP nach 12 Monaten (Seladelpar 10 mg vs. Placebo): Falls negativ bei einem zweiseitigen Signifikanzniveau von 0,05, werden keine weiteren Inferenztests durchgeführt. Falls positiv, wird das Testen fortgesetzt. Veränderungen in der Pruritus NRS vom Ausgangswert bis Monat 6 (Seladelpar 10 mg vs. Placebo) werden bei einem zweiseitigen Signifikanzniveau von 0,05 bewertet.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, randomisierte, multizentrische Phase 3-Parallelgruppenstudie bei männlichen und weiblichen Patienten mit PBC. 193 Patienten wurden im Verhältnis 2:1 (SEL: Placebo) in die Behandlungsgruppen randomisiert: • Gruppe A: SEL 10 mg für 12 Monate • Gruppe B: Placebo für 12 Monate Die Gesamtdauer des Behandlungszeitraums beträgt 12 Monate. Dieser folgen 17 Tage Nachbeobachtung. Die Gesamtdauer beträgt 16 Monate (einschließlich der Screening-, Run-in-, Behandlungs- und Nachbeobachtungsphase). Patienten hatten außerdem die Möglichkeit im Anschluss an die Behandlungsdauer in eine offene Langzeitnachbeobachtungsstudie (ASSURE) eingeschlossen zu werden. Für diese Patienten entfällt die 17-tägige Nachbeobachtungsphase.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<i>Protokoll Version 3.0 (30. Juni 2021)</i> Abschnitt 4.2 Ausschlusskriterien • Das Ausschlusskriterium 4a wurde überarbeitet, um klarzustellen, dass der INR-Ausgangswert in Verbindung mit den aktuellen Dosisanpassungen der Antikoagulanzen bewertet wird. • Das Ausschlusskriterium 15 (Immunsuppressive Therapien [z. B. Cyclosporin, Tacrolimus, Anti-TNF oder andere immunsuppressive Biologika]) wurde hinzugefügt. • Das Ausschlusskriterium 16 (Andere Medikamente, die die Leber- oder Magen-Darm-Funktion beeinflussen, wie z. B. die Absorption von Arzneimitteln, können verboten sein und sollten im Einzelfall mit dem medizinischen Monitor besprochen werden.) wurde hinzugefügt. Abschnitt 7.3 Verbotene Begleitmedikation

Item ^a	Charakteristikum	Studieninformation
		- Wurde analog Ausschlusskriterium 15 und 16 angepasst. Abschnitt 9.3.4 Maßnahmen im Zusammenhang mit der Studienmedikation - Um die Definitionen an die Standarddefinitionen des Clinical Data Interchange Standards Consortiums anzupassen, wurden die Begriffe und Definitionen für Maßnahmen im Zusammenhang mit der Studienmedikation angepasst. <i>Protokoll Version 4.0 (9. Februar 2022)</i> Abschnitt 4.1 Einschlusskriterium 5e - Basierend auf einer Diskussion mit FDA wurde das Protokoll geändert, um die Einschlusskriterien für die eGFR zu aktualisieren. Einschlusskriterium 5e wurde von einer eGFR von >60 mL/min/1,73 m² auf eine eGFR von ≥45 mL/min/1,73 m² angepasst. Abschnitt 4.2 Ausschlusskriterien - Zum Ausschlusskriterium 2 wurde „andere aktive Infektion“ hinzugefügt. - Bei dem Ausschlusskriterium 9 wurde die Zeit für die Behandlung mit Obeticholsäure und Fibraten (z. B. Bezafibrat, Fenofibrat, Elafibranor, Lanifibranor, Pemafibrat, Saroglitazar) von 3 Monaten zu 6 Wochen vor Screening abgeändert. - Zum Ausschlusskriterium 16 wurde „Medikamenten oder Roux-en-Y-Magenbypass-Verfahren“ hinzugefügt. - Ausschlusskriterium 17 (Aktive COVID-19-Infektion während des Screenings) wurde hinzugefügt. Abschnitt 5.1 Studiendauer - Die Safety Follow-up-Visite wurde von einem Monat nach der letzten Dosis zu zwei Wochen (14 Tage ± 3) nach der letzten Dosis geändert. Erklärung: Der Zeitraum für die das Safety-Follow-up wurde auf der Grundlage der langfristigen Sicherheit von SEL in PBC und der SEL-Halbwertszeit verkürzt. Nach einer Einzeldosis variierte die terminale Halbwertszeit (t_{1/2}) von SEL zwischen 5 und 8 Stunden. In einer Studie mit mehreren aufsteigenden Dosen von 50, 100 und 200 mg schwankte die t_{1/2} von SEL an Tag 21 zwischen 9 und 14 Stunden. Es gab keine Hinweise auf eine Akkumulation zwischen Tag 1 und Tag 21, wobei die Zunahme der Exposition relativ

Item ^a	Charakteristikum	Studieninformation
		dosisproportional war. Auf der Grundlage dieser Daten ist eine Nachbeobachtungszeit von 14 bis 17 Tagen ausreichend (z. B. mehr als 5 Halbwertszeiten), um sicherzustellen, dass SEL vollständig aus dem System ausgeschieden ist. <i>Protokoll Version 4.1 DE (14. Februar 2022); genannte Änderungen zu Protokoll Version 4</i> Abschnitt 5.1 Studienüberblick - Die Anzahl der Studienzentren wurde auf 180 erhöht. Abschnitt 8.1.1 Screening - Ergänzungen bezüglich Screeningzeitraum. Abschnitt 8.1.7 Follow Up - Ergänzungen bezüglich jährlicher Nachuntersuchungen. Abschnitt 8.2.5 Labortest - HIV-Testung wurde aufgenommen. Abschnitt 10.1.4 Tabelle 6 - Der Ausschluss einer Pankreatitis soll per Computertomographie oder eine Magnetresonanztomographie durchgeführt worden sein.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien Die Probanden müssen alle folgenden Kriterien erfüllen, um für die Teilnahme an der Studie in Frage zu kommen: - Sie müssen eine schriftliche Einwilligung nach Aufklärung (unterschrieben und datiert) und alle nach örtlichem Recht erforderlichen Genehmigungen erteilt haben. 18 bis einschließlich 75 Jahre alt - Männlich oder weiblich mit einer PBC-Diagnose, die auf zwei der folgenden Kriterien beruht: - ALP-Wert über 1,0 x ULN seit mindestens 6 Monaten - Positiver AMA-Titer (>1:40 in der Immunfluoreszenz oder M2 positiv in ELISA oder positive PBC-spezifische ANA. - UDCA in den letzten 12 Monaten (stabile Dosis seit >3 Monaten vor der Untersuchung) ODER Unverträglichkeit von UDCA (letzte UDCA-Dosis >3 Monate vor der Untersuchung). - Vom Zentrallabor beim Screening gemessene Laborparameter

Item ^a	Charakteristikum	Studieninformation
		a. ALP $\geq 1,67 \times \text{ULN}$ b. AST $\leq 3 \times \text{ULN}$ c. ALT $\leq 3 \times \text{ULN}$ d. Gesamt-Bilirubin $\leq 2 \times \text{ULN}$ e. eGFR $> 45 \text{ mL/min/1,73 m}^2$ (berechnet nach der Gleichung der Modification of Diet in Renal Disease Study). f. INR unter $1,1 \times \text{ULN}$ Bei Patienten, die eine Antikoagulationstherapie erhalten, muss der INR-Wert in dem Bereich gehalten werden, der für die Prophylaxe ihrer spezifischen Erkrankung erforderlich ist. g. Thrombozytenzahl $\geq 100 \times 10^3/\mu\text{L}$. Hinweis: Prothrombinzeit, INR und Thrombozyten können beim Screening-Besuch vor Ort bestimmt werden, wenn der Prüfer dies nach Rücksprache mit dem medizinischen Betreuer für notwendig hält, falls die zentral gemessenen Proben als ungültig erachtet werden. 6. Frauen im gebärfähigen Alter müssen während der Studie und für mindestens 90 Tage nach der letzten Dosis mindestens ein Barriere-Kontrazeptivum und eine zweite wirksame Verhütungsmethode anwenden. Männliche Probanden, die mit weiblichen Partnern im gebärfähigen Alter sexuell aktiv sind, müssen eine Barriere-Kontrazeption anwenden, und ihre weiblichen Partner müssen während der Studie und mindestens 90 Tage nach der letzten Dosis eine zweite wirksame Verhütungsmethode anwenden. Ausschlusskriterien Die Probanden dürfen keines der folgenden Kriterien erfüllen, um für eine Studienteilnahme in Frage zu kommen: 1. Frühere Exposition gegenüber SEL (MBX-8025) 2. Eine andere Erkrankung als PBC, die nach Ansicht des Prüfarztes eine uneingeschränkte Teilnahme an der Studie ausschließen würde (z. B. Krebs) oder die Ergebnisse verfälschen würde (z. B. Morbus Paget, eine andere aktive Infektion) 3. Fortgeschrittene PBC im Sinne der Rotterdam-Kriterien (Albumin unter der unteren Grenze

Item ^a	Charakteristikum	Studieninformation
		der Norm UND Gesamt-Bilirubin über 1,0 x ULN) 4. Vorliegen einer klinisch bedeutsamen dekompenzierten Lebererkrankung, einschließlich: a. Lebertransplantation in der Vorgeschichte, aktuelle Aufnahme in die Lebertransplantationsliste oder aktueller MELD-Wert ≥ 12. Bei Probanden, die blutgerinnungshemmende Medikamente einnehmen, wird bei der Berechnung des MELD-Scores der INR-Ausgangswert in Verbindung mit der aktuellen Dosisanpassung der blutgerinnungshemmenden Medikamente berücksichtigt. Dies wird in Absprache mit dem medizinischen Betreuer geschehen. b. Komplikationen der portalen Hypertension, einschließlich bekannter Ösophagusvarizen, früherer Varizenblutungen oder damit zusammenhängende Eingriffe (z. B. Anlage eines transjugulären intrahepatischen portosystemischen Shunts), Aszites und hepatischer Enzephalopathie. c. Zirrhose mit Komplikationen, einschließlich Vorgeschichte oder Vorliegen einer spontanen bakteriellen Peritonitis, eines hepatozellulären Karzinoms oder eines hepatorenenalen Syndroms. 5. Andere chronische Lebererkrankungen: a. Akute Merkmale einer Autoimmunhepatitis, diagnostiziert mittels Immunserologie, Leberbiochemie oder historisch bestätigter Leberhistologie durch den Prüfer. b. Primäre sklerosierende Cholangitis, bestimmt durch das Vorliegen diagnostischer cholangiografischer Befunde. c. Vorgeschichte oder klinische Anzeichen einer alkoholischen Lebererkrankung. d. Vorgeschichte oder klinische Anzeichen auf Alpha-1-Antitrypsin-Mangel. e. Vorgeschichte einer durch Biopsie betätigten nicht-alkoholischen Steatohepatitis. f. Vorgeschichte oder Anzeichen des Gilbert-Syndroms mit erhöhtem Gesamt-Bilirubin. g. Vorgeschichte oder Anzeichen einer Hämochromatose. h. Hepatitis B, definiert als Vorhandensein von Hepatitis B-Oberflächen-Antigen.

Item ^a	Charakteristikum	Studieninformation
		i. Hepatitis C, definiert als das Vorhandensein von Hepatitis C-Virus-Ribonukleinsäure. j. Vorgeschichte, Nachweis oder starker Verdachte auf hepatobiliäre Malignität aufgrund von Bildgebung, Screening-Laborwerten und/oder klinischen Symptomen. 6. Bekannte Vorgeschichte von HIV oder positiver Antikörpertest beim Screening. 7. Klinisch bedeutsamer Alkoholkonsum, definiert als mehr als 2 Getränkeinheiten pro Tag (entspricht 20 g) bei Frauen und 3 Getränkeinheiten pro Tage (entspricht 30 g) bei Männern, oder die Unfähigkeit, den Alkoholkonsum zuverlässig zu quantifizieren. 8. Eine diagnostizierte oder bösartige Erkrankung in der Vorgeschichte, aktiv oder innerhalb von 2 Jahren, oder eine laufende Untersuchung auf eine bösartige Erkrankung; eine lokale Behandlung von Plattenepithelkarzinom oder nicht-invasiven Basalzellkarzinom und Gebärmutterhalskrebs in situ ist zulässig, wenn sie vor dem Screening angemessen behandelt wurden. 9. Behandlung mit OCA und Fibraten (z. B. Bezafibrat, Fenofibrat, Elafibranor, Lanifibranor, Pemafibrat, Saroglitazar [PPAR-Agonisten]) während der letzten 6 Wochen vor dem Screening. 10. Behandlung mit Colchizin, Methotrexat, Azathioprin oder langfristigen systematischen Kortikosteroiden (>2 Wochen) während der letzten 2 Monate vor dem Screening. 11. Die Behandlung mit pruritusstillenden Medikamenten (z. B. Colestyramin, Naltrexon, Rifampicin, Sertralin oder ein experimenteller Ansatz) muss innerhalb eines Monats vor dem Screening in einer stabilen Dosis erfolgen. 12. Behandlung mit einer anderen Prüftherapie oder -gerät innerhalb von 30 Tagen oder innerhalb von 5 Halbwertszeiten, je nachdem, was länger ist, vor dem Screening. 13. Bei Frauen: Schwangerschaft oder Stillen 14. Jeder andere Zustand, der nach dem Urteil des Prüfers die Sicherheit des Probanden oder die Qualität der klinischen Studie beeinträchtigen würde. 15. Immunsuppressive Therapien (z. B. Cyclosporin, Tacrolimus, Anti-

Item ^a	Charakteristikum	Studieninformation
		Tumornekrosefaktor oder andere immunsuppressive Biologika). 16. Andere Medikamente, die sich auf die Leber- oder Magen-Darm-Funktionen auswirken, wie z. B. die Absorption von Medikamenten oder Roux-en-Y-Magenbypass-Verfahren, können verboten sein und sollten von Fall zu Fall mit dem medizinischen Betreuer besprochen werden. 17. Aktive COVID-19-Infektion während des Screenings.
4b	Studienorganisation und Ort der Studiendurchführung	90 Studienzentren in 24 Ländern. Die Probanden, die an dieser Studie teilnahmen, kamen aus den folgenden Ländern: Argentinien, Australien, Österreich, Belgien, Kanada, Chile, Tschechische Republik, Frankreich, Deutschland, Griechenland, Ungarn, Israel, Italien, Republik Korea, Mexiko, Neuseeland, Polen, Rumänien, Russland, Spanien, Schweiz, Türkei, Vereinigtes Königreich und Vereinigte Staaten von Amerika.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<u>Studienmedikament:</u> Probanden erhalten SEL 10 mg (oder SEL 5 mg, falls heruntertitriert) oder ein passendes Placebo einmal täglich, oral, etwa zur gleichen Tageszeit <u>Heruntertitration:</u> Bei Probanden, die spezifische Sicherheitsüberwachungskriterien erfüllen (Abschnitt 10 im Protokoll) oder bei denen Probleme mit der Verträglichkeit auftreten, kann eine Dosisreduktion vorgenommen werden. Probanden, die ursprünglich 10 mg SEL erhalten haben, werden verblindet auf 5 mg heruntertitriert. Bei Probanden, die ursprünglich Placebo erhalten haben, erfolgt eine verblindete Heruntertitration und sie bleiben in der Placebo-Gruppe. Probanden, bei denen eine klinisch bedeutsame UE auftritt, das nach dem klinischen Urteil des Prüfarztes eine Dosisreduktion rechtfertigt, kommen ebenfalls für eine ähnliche Dosisreduktion in Frage. Die Dosisreduktion muss von dem medizinischen Monitor genehmigt werden und wird verblindet durchgeführt. <u>UDCA:</u> Während der Studie wird das Studienmedikament bei Probanden, die UDCA vertragen, als Zusatz zur UDCA-Therapie verabreicht; bei Patienten mit UDCA-Unverträglichkeit wird das Studienmedikament als Monotherapie verabreicht.

Item ^a	Charakteristikum	Studieninformation
		UDCA wird oral ein- oder mehrmals täglich in der Dosis von vor Studienbeginn und nach dem klinischen Urteil des Prüfarztes.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäre Endpunkte: - Das kombinierte Ansprechen zu 12 Monaten. Das kombinierte Ansprechen ist definiert als a. $ALP < 1,67 \times ULN$ b. $\geq 15\%$ Verringerung in ALP c. Gesamt-Bilirubin $\leq 1,0 \times ULN$ - Bewertung von unerwünschten Ereignissen unter Behandlung sowie der Biochemie und Hämatologie Wichtige Sekundäre Endpunkte: - Anteil der Probanden mit $ALP \leq 1,0 \times ULN$ nach 12 Monaten (d. h. Normalisierung) - Veränderung der wöchentlich gemittelten Pruritus NRS nach 6 Monaten bei Probanden mit Baseline-NRS ≥ 4 Andere Sekundäre Endpunkte: - Kombiniertes Ansprechen zu 6 Monaten (Das kombinierte Ansprechen ist definiert als a. $ALP < 1,67 \times ULN$, b. $\geq 15\%$ Verringerung in ALP, c. Gesamt-Bilirubin $\leq 1,0 \times ULN$) - Anteil der Probanden mit $ALP \leq 1,0 \times ULN$ nach 6 Monaten - Anteil der Probanden mit $ALP < 1,67 \times ULN$ und $ALP < 1,5 \times ULN$ nach 6 und 12 Monaten - Absolute und relative Veränderungen von ALP nach 3, 6 und 12 Monaten - Anteil der Probanden mit einem Rückgang der NRS ≥ 2, NRS ≥ 3 oder NRS ≥ 4 bei jedem Besuch bei Probanden mit NRS ≥ 4 zu Beginn Baseline - Veränderungen des NRS-Wertes für Pruritus nach 3 und 12 Monaten bei Probanden mit einem NRS-Wert ≥ 4 zu Beginn der Baseline - Veränderung des Lebensqualitätsmaßes unter Verwendung des PBC-Fragebogens PBC-40 bei jedem Besuch (Gesamtwert und Bereichswert) im Vergleich zu Baseline - Veränderung gegenüber Baseline im Vereinigten Königreich - Primäre biliäre Zirrhose und globale PBC Study Group Risikowerte bei jedem Besuch - Absolute und relative Veränderungen von ALT, AST, GGT, Bilirubin (gesamt, direkt und indirektes Bilirubin und 5'-Nukleotidase bei jedem Besuch

Item ^a	Charakteristikum	Studieninformation
		- Das erste Auftreten von klinischen PBC-Ergebnisse im Sinne der folgenden Definitionen: - Gesamttod - Lebertransplantation - MELD-Score ≥ 15 bei mindestens 2 aufeinanderfolgenden Besuchen - Behandlungsbedürftiger Aszites - Hospitalisierung wegen eines Neuauftretens oder Wiederauftretens einer der folgenden Erkrankungen: - Varizenblutung - Hepatische Enzephalopathie (wie definiert durch einen West-Haven-Score ≥ 2) - Spontane bakterielle Peritonitis (bestätigt durch Kultur aus diagnostischer Parazentese) Explorative Endpunkte: - Histologische Veränderungen der Leber auf der Grundlage der pathologischen Untersuchung von Biopsiegeweben - Ergebnisse der PBC-Ansprechkriterien (Barcelona, Paris I und II, Toronto I und II, Rotterdam) - Änderungen gegenüber dem Ausgangswert in der Pruritus NRS basierend auf zusätzlichen Schwellenwerten für Verbesserungen und dem Ausgangswert des Juckreizstatus - Änderungen gegenüber dem Ausgangswert im PBC-40 QoL Juckreiz-Domäne und der 5-D Pruritus-Skala, PGI-C und PGI-S - Absolute und relative Veränderungen bei Lipiden, Gallensäuren, Sterolen und Biomarkern der Gallensäuresynthese: 7α-Hydroxy-4-Cholesten-3-on (C4) und Fibroblasten-Wachstumsfaktor 19 - Plasmakonzentrationen von Seladelpar und seinen Metaboliten (M1, M2, M3) - Absolute und relative Veränderungen der Entzündungsmarker/Immunreaktivität (z. B. hochsensibles C-reaktives Protein, Fibrinogen, Haptoglobin, Tumornekrosefaktor-alpha und Antikörper [siehe Veränderungen in Abschnitt 12]) - Absolute und relative Veränderungen der Marker der verstärkten Leberfibrose (ELF) und der Leber Lebersteifigkeit mittels FibroScan[®] - Absolute und relative Veränderungen der Marker für Leberschäden: CK18 (M65) und miR-122 (siehe Veränderungen in Abschnitt 12)
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<i>Protokoll Version 2.0 (1. Dezember 2020)</i> Abschnitt 9.1

Item ^a	Charakteristikum	Studieninformation
		Es wurde gelöscht, dass für die Meldung von UE UDCA nicht als Studienmedikament betrachtet wird. Der Text wurde entfernt, um Missverständnisse bei der Sicherheitsberichterstattung zu vermeiden. <i>Protokoll Version 3.0 (30. Juni 2021)</i> Um die Definitionen an die Standarddefinitionen des Clinical Data Interchange Standards Consortiums anzupassen, wurden die Definitionen für Endpunktabhängige UE angepasst.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Für die Schätzung des Stichprobenumfangs wird die Ansprechquote der Placebo-Gruppe auf 20% geschätzt. Die Ansprechquote von SEL 10 mg wird auf 55% geschätzt. Bei Verwendung eines zweiseitigen Tests auf Gleichheit der binominalen Proportionen auf der Grundlage des exakten Tests von Fisher auf dem Signifikanzniveau 0,05 bietet eine Stichprobengröße von 180 randomisierten Probanden eine Aussagekraft von >90%, um einen Unterschied zwischen der SEL 10 mg Gruppe und Placebo-Gruppe festzustellen. Bei der wichtigsten sekundären Wirksamkeitsanalyse, der Normalisierung des ALP-Wertes, wird eine Placebo-Ansprechrare und eine Seladelpar-Ansprechrare von 2,5% bzw. 25,5% erwartet. Eine Stichprobengröße von 180 randomisierten Probanden bietet eine Power von >90%, um einen Unterschied zwischen der Seladelpar- und der Placebo-Gruppe zu erkennen, basierend auf einem exakten Fisher-Test mit einem Signifikanzniveau von 0,05. Die Ansprechrate für den sekundären Wirksamkeitsendpunkt ALP-Normalisierung wird in der Placebo-Gruppe auf 2,5% und in der SEL 10 mg-Gruppe auf 25,5% geschätzt. Eine Stichprobengröße von 180 randomisierten Probanden bietet eine Power von >90%, um einen Unterschied zwischen der Seladelpar- und der Placebo-Gruppe zu erkennen, basierend auf einem exakten Fisher-Test mit einem Signifikanzniveau von 0,05. Die Berechnung des Stichprobenumfangs der sekundären Wirksamkeitsanalyse zur Veränderung der wöchentlich gemittelten Pruritus NRS im Monat 6 gegenüber Baseline basiert auf einem zweiseitigen t-Test mit einem Signifikanzniveau von 0,05. Die Standardabweichung wird auf 2 geschätzt. Unter diesen Annahmen sowie bei einer Gesamtzahl von 48 randomisierten Probanden mit einer NRS von =4 zu Studienbeginn bietet dieser Test eine Aussagekraft von >80%, um einen Behandlungsunterschied von =2 zwischen SEL 10 mg-Gruppe und Placebo-Gruppe festzustellen.

Item ^a	Charakteristikum	Studieninformation
		Die angegebenen Power-Berechnungen basieren auf früheren Ergebnissen der Studie CB8025-31735 sowie einer angenommenen Dropout-Rate von etwa 10%.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Interimanalysen durchgeführt.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung wird zentral über ein IWRS am Tag 1 der Studie durchgeführt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	2:1 (SEL: Placebo) Stratifizierungsfaktoren: • ALP-Level (ALP <350 U/L vs. ALP ≥350 U/L) • Pruritus (NRS <4 vs. ≥ 4 NRS)
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	Durchführung der Zuteilung: siehe 8a Es handelt sich um eine doppelblinde Studie.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierung wird zentral über ein IWRS am Tag 1 der Studie durchgeführt. Der Randomisierungsplan wird von einem nicht-verblindeten Statistiker erstellt, unabhängig vom Studienteam.
11	Verblindung	Es handelt sich um eine doppelblinde RCT. In der Zeit von der Randomisierung bis zum Database lock sind die für die Studienaufsicht verantwortlichen Mitglieder des Sponsor-Studienteams, die Probanden, die Prüfer und das gesamte Personal des Prüfzentrums hinsichtlich der Behandlungszuweisung verblindet. Ein DSMB ist berechtigt, unverblindete Daten zu überprüfen, wie in der DSMB-Charta festgelegt. Die Kriterien für eine Notfall-Entblindung durch den Prüfarzt sind im Prüfplan festgelegt. Wenn eine solche Entblindung vorgenommen wird, muss sie klar begründet werden und durch einen Kommentar in der Quellendokumentation erläutert werden, zusammen mit dem Datum, an dem die Entblindung gebrochen wurde, und die Identität der Person, die die Entblindung genehmigt hat. Das Datum und die Uhrzeit, zu denen der Prüfer die Verblindung für einen einzelnen Probanden

Item ^a	Charakteristikum	Studieninformation
		aufgehoben hat, werden vom IWRS dokumentiert, und dem Sponsor wird eine automatische Benachrichtigung über den Vorfall zugesandt. Das Pharmakovigilanz-Team des Auftragsforschungsinstituts muss möglicherweise auch die Verblindung für Zwecke der behördlichen Berichterstattung unterbrechen.
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Ja b) Ja c) Ja
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Placebo wurde passend zu SEL oral verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5mg SEL, 10 mg SEL, Placebo) Verhalten bei Heruntertitration: • Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert. • Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert. • Die Heruntertitration wurde verblindet durchgeführt.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<u>Analysesets:</u> • Alle gescreenten Probanden (All Subjects Screened Analysis Set) umfasst alle Personen, die für die Teilnahme an der Studie gescreent wurden, unabhängig davon, ob sie tatsächlich in die Studie aufgenommen wurden. • ITT-Analyseset: umfasst alle Probanden, die randomisiert in die Studie aufgenommen wurden und mindestens eine Dosis des Studienmedikaments erhalten haben. • Moderater bis schwerer Pruritus Numerical Rating Scale (MSPN): umfasst Probanden aus der ITT-Analyseset, die einen Ausgangswert von ≥ 4 Pruritus NRS aufwiesen. • Per-Protokoll-Analyseset: umfasst Probanden aus der ITT-Analysegruppe, die mindestens eine post-Baseline Erhebung von ALP und Gesamt-Bilirubin aufweisen und keine Protokollverletzung haben, die als beeinflussend für die Wirksamkeitsanalyse betrachtet wird.

Item ^a	Charakteristikum	Studieninformation
		• Biopsie-Analyseset: umfasst alle Probanden, bei denen eine Biopsie zu Studienbeginn oder nach 12 Monaten/Ende der Behandlung durchgeführt wurde. • Sicherheits-Analyseset: umfasst alle Probanden, die mindestens eine Dosis des Studienmedikaments erhalten haben. Probanden werden in der Gruppe gemäß der tatsächlich erhaltenen Behandlung berücksichtigt, falls diese von der zugewiesenen Behandlung abweicht. • Pharmakokinetik-Analyseset: umfasst alle Probanden, die an der Sammlung von pharmakokinetischen-Proben teilnehmen. <u>Analysemethoden:</u> <u>Alle statistischen Tests werden, sofern nicht anders angegeben, nach den Baseline-Werten wie folgt stratifiziert:</u> • Alterskategorien (Alter bei Screening: <65, ≥65 Jahre; Alter bei PBC-Diagnose <50, ≥50 Jahre) • Geschlecht (weiblich, männlich) • Ethnische Herkunft (Weiß, Schwarz, Asiatisch, Andere) • Region (Nordamerika, Europa, Rest der Welt) • Ausgangs-ALP (<350 U/L, ≥350 U/L) • Gesamt-Bilirubin (<0,6 × ULN, ≥0,6 × ULN) • Pruritus NRS (<4, ≥4) • UDCA-Gebrauch vs. UDCA-Intoleranz • Vorherige Anwendung von OCA und/oder Fibraten (ja, nein) • Zirrhose (ja, nein) • Gesamt-Bilirubin (≤1 × ULN, >1 × ULN) Analyse zum primären Endpunkt: Alle Patienten ohne Messung, Patienten, welche die Behandlung vor dem festgelegten Zeitpunkt für die Evaluation abgebrochen haben oder alle Patienten mit sonstigen fehlenden Daten werden als Nicht-Responder betrachtet. Die Inzidenz der Antwort nach 12 Monaten wird zusammengefasst; Analysen für den zusammengesetzten Endpunkt werden unter Verwendung eines CMH-Tests durchgeführt. Die CMH-Analyse wird nach dem Randomisierungs-Stratum stratifiziert und auf dem ITT-Analyseset durchgeführt. Die Risikodifferenz und das 95%-KI unter Verwendung von Miettinen und Nurminen werden ebenfalls angegeben. Die statistische Signifikanz des Unterschieds

Item ^a	Charakteristikum	Studieninformation
		zwischen Placebo und Seladelpar wird als zweiseitiger p-Wert $\leq 0,05$ definiert. Wenn die Annahmen für den CMH-Test nicht erfüllt sind, wird eine Poolung der Strata angewendet. Wenn das Mantel-Fleiss (1980)-Kriterium kleiner als 5 ist, wird das Stratum entfernt. Das Pruritus-Stratum wird zuerst entfernt, und die Daten werden unter Verwendung eines CMH-Tests stratifiziert nach dem Baseline-ALP zur Randomisierung analysiert. Darüber hinaus werden sowohl das Pruritus- als auch das ALP-Stratum entfernt, und in diesem Fall wird ein Pearson-χ^2-Test verwendet, mit einem 95%-KI (Wald). Wenn beim Pearson-χ^2-Test mindestens eine erwartete Zellanzahl unter 5 liegt, wird stattdessen der exakte Fisher-Test und das exakte 95%-KI (Santner und Snell, 1980) für Inferenzanalysen verwendet. Für den CMH-Test wird der Breslow-Day-Test verwendet, um die Annahme der Homogenität der Behandlungseffekte über die Strata hinweg zu überprüfen. Falls diese Annahme fehlschlägt, werden der Anteil der Responder in jeder Behandlung zusammen mit den entsprechenden KI und der Vergleich zwischen den Behandlungen für jedes Stratum separat berichtet. Analysen zu den wichtigen sekundären Endpunkten: Die wichtigste sekundäre Wirksamkeitsanalyse für den Anteil der Probanden, die zu Monat 12 eine Normalisierung der ALP erreichen, wird im ITT-Analyseset durchgeführt und folgt dem gleichen Ansatz wie für die primäre Wirksamkeitsanalyse. Die Veränderung vom Ausgangswert in der wöchentlich gemittelten Pruritus NRS nach 6 Monaten wird unter Verwendung eines MMRM für die Patienten im MSPN-Analyseset analysiert. Das Modell enthält Terme für den Ausgangswert der NRS, das Ausgangsstratum der Randomisierung (ALP-Wert < 350 U/L versus ALP-Wert ≥ 350 U/L), die Behandlungsgruppe, die Woche und die Wechselwirkung zwischen Behandlung und Woche. Falls ein Zeitpunkt fehlt, wird dieser als Durchschnitt der beiden benachbarten wöchentlichen Durchschnitte imputiert (höchstens eine Woche auseinander); andernfalls wird der benachbarte wöchentliche Durchschnitt, der vorliegt, zur Imputation verwendet. Ein Proband, der vor oder während Woche 24 die Behandlung abbricht, erhält keinen imputierten Wert für Woche 26. Nach Monat 6 gesammelte Daten werden nicht zur Imputation verwendet. Die Interaktion zwischen Behandlung und Ausgangswert der NRS wird untersucht und als Term hinzugefügt,

Item ^a	Charakteristikum	Studieninformation
		wenn eine Interaktion festgestellt wird (p-Wert <0,05). Die LS-Mittelwerte für die NRS-Veränderung vom Ausgangswert nach Behandlung und die zugehörigen Standardfehler, die LS-Mittelwerte für den Unterschied zwischen den Behandlungsgruppen sowie die zugehörigen zweiseitigen 95%-Konfidenzintervalle und zweiseitigen p-Werte werden aus dem MMRM-Modell abgeleitet. Die Signifikanz dieser wichtigsten sekundären Wirksamkeitsanalyse basiert auf dem Unterschied der Behandlungen in Woche 26 (d. h. Monat 6). Zuerst wird für das MMRM ein unstrukturierter Kovarianzmatrixansatz angenommen. Falls das Modell nicht konvergiert, werden die folgenden Kovarianzstrukturen in der Reihenfolge angewendet, bis das Modell konvergiert: heterogene Toeplitz, heterogene Compound Symmetry und dann Compound Symmetry. Die Kenward-Roger-Korrektur für die Freiheitsgrade des Nenners wird angewendet. Beobachtete Werte und Veränderungen vom Ausgangswert in der wöchentlich gemittelten Pruritus NRS werden am Ausgangswert und für jede Woche nach dem Ausgangswert zusammengefasst, wie in Tabelle 2 angegeben. Die Imputation fehlender wöchentlicher Durchschnitte erfolgt wie für das MMRM-Modell beschrieben. Nach Monat 6 gesammelte Daten werden nicht für die Imputation verwendet, und Besuche nach Monat 6 werden ebenfalls nicht imputiert. Analysen zu weiteren sekundären Endpunkten: Die Analysen zu weiteren sekundären Endpunkten orientieren sich entsprechend an den oben beschriebenen Analysen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<u>Präspezifizierte Subgruppen:</u> • Alter (Alter beim Screening <65, ≥65 Jahre; Alter bei PBC-Diagnose <50, ≥50 Jahre) • Geschlecht (weiblich, männlich) • Ethnie (Weiß, Schwarz, Asiatisch, Andere) • Region (Nordamerika, Europa, Rest der Welt) • Ausgangswert ALP (<350 U/L, ≥350 U/L) • Gesamt-Bilirubin (<0,6 × ULN, ≥0,6 × ULN) • Pruritus NRS (<4, ≥4) • UDCA-Einsatz vs. UDCA-Unverträglichkeit • Frühere Verwendung von OCA und/oder Fibraten (ja, nein) • Zirrhose (ja, nein) • Gesamt-Bilirubin (≤1 × ULN, >1 × ULN) Zur Adjustierung siehe Item 12a.
Resultate		

Item ^a	Charakteristikum	Studieninformation
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart
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	SEL: a. 128 b. 128 c. 128 Placebo: d. 65 e. 65 f. 65
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	SEL: • Unerwünschtes Ereignis: 3 • Lost to Follow-up: 1 • Protokollabweichung: 1 • Einwilligung zurückgezogen: 5 • Andere: 1 Placebo: • Unerwünschtes Ereignis: 4 • Lost to Follow-up: 1 • Protokollabweichung: 1 • Einwilligung zurückgezogen: 2
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Randomisierung des ersten Patienten in die Studie: 21. April 2023 Letzter Patient letzte Visite: 11. August 2023
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.
a: nach CONSORT 2010. ALP=Alkalische Phosphatase, ALT=Alanin-Aminotransferase, AMA=Antimitochondriale Antikörper, ANA=Antinukleäre Antikörper, AST=Aspartat-Aminotransferase, CMH=Cochran-Mantel-Haenszel, COVID-19=Coronavirus-Krankheit 2019, DSMB=Datenüberwachungskomitee, eGFR=Geschätzte Glomeruläre Filtrationsrate, ELF=Fortgeschrittene Leberfibrose, ELISA=Enzyme Linked Immunosorbent Assay, FDA=Food and Drug Administration, GGT=Gamma-Glutamyltransferase, HIV=Humanes Immundefizienzvirus, INR=International Normalized Ratio, ITT=Intention-to-Treat, IWRS=Interaktives Webdialogsystem, KI=Konfidenzintervall, MELD=Model for End-Stage Liver Disease, MMRM=Mixed Effect Model Repeat Measurement, MSPN=Moderater bis schwerer Pruritus Numerical Rating Scale, NRS=Numerical Rating Scale, OCA=Obeticholsäure, PBC=Primär biliäre Cholangitis, PGI-C=Patient Global Impression of Change, PGI-S=Patient Global Impression of Severity, PPAR=Peroxisom-Proliferator-aktivierter Rezeptor, RCT=Randomisierte kontrollierte Studie, SEL=Seladelpar, UDCA=Ursodeoxycholsäure, UE=Unerwünschtes Ereignis, ULN=Oberer Normwert		

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

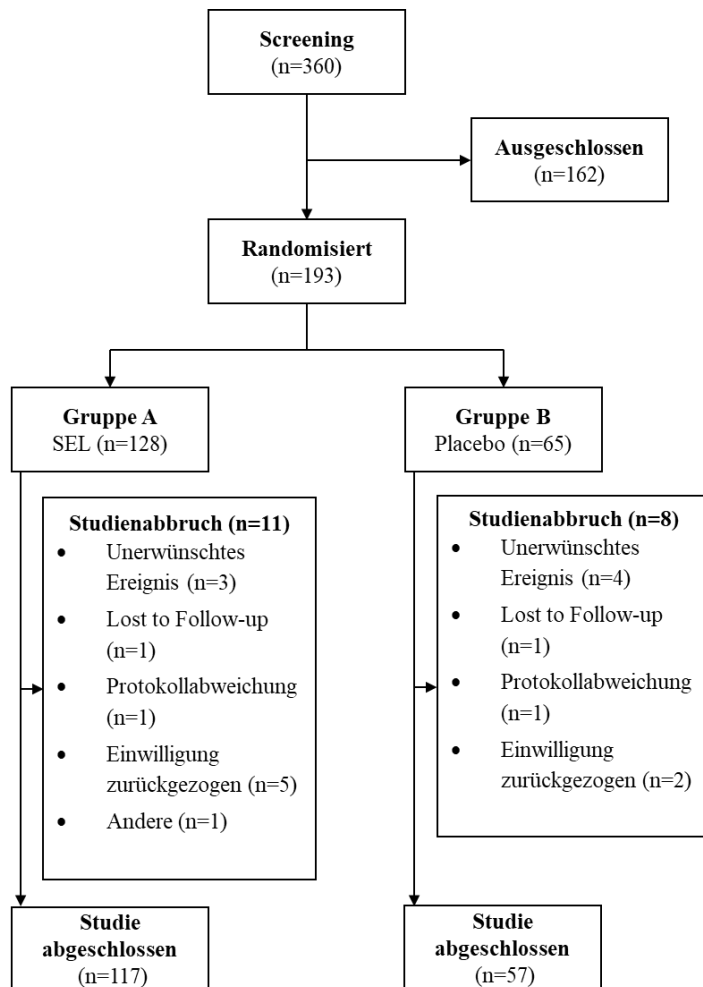


Abbildung 31 (Anhang): Flow-Chart der Studie RESPONSE

n=Anzahl der Patienten mit Ereignis, SEL=Seladelpar

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. Quellspezifische 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 4-82 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie RESPONSE.

Studie: RESPONSE

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht	A
Statistischer Analyseplan	B
Studienprotokoll	C

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

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

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

C: Abschnitt 6.4

Die Randomisierung wurde zentral über ein IWRS durchgeführt.

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

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

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

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

A, B, C

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ ja ☐ unklar ☐ nein

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

3. Verblindung von Patienten und behandelnden Personen**Patient:**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

behandelnde bzw. weiterbehandelnde Personen:

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

A, B, C

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

A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Da es keine Hinweise auf Aspekte gibt, die Ergebnisse auf Studienebene verzerren, wird das Verzerrungspotenzial auf Studienebene als niedrig bewertet.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**Endpunkt:** Biochemisches Ansprechen**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: Pruritus NRS

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

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

A, B, C

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

A, B, C

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

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: 5-D Pruritus**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

A, B, C

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

☒ ja ☐ nein

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

A, B, C

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: PGI-C und PGI-S**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: Fibrose

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

A, B, C

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

☒ ja ☐ nein

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

A, B, C

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: Hospitalisierung

5. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

6. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: PBC-40

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen

Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

A, B, C

Seladelpar und Placebo wurden in der gleichen Darreichungsform verabreicht. Die unterschiedlichen Dosierungen von SEL und auch Placebo werden jeweils in einer einzelnen Hartkapsel verabreicht (mögliche Hartkapseln: 5 mg SEL, 10 mg SEL, Placebo). Verhalten bei Heruntertitration:

- Probanden, die 10 mg SEL erhalten haben, wurden auf 5 mg SEL heruntertitriert.
- Probanden, die Placebo erhalten, wurden auf Placebo heruntertitriert.

Die Heruntertitration wurde verblindet durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

A, B, C

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

☒ ja ☐ nein

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

A, B, C

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

A, B, C

Aus den oben genannten Punkten ergeben sich keine Anhaltspunkte für ein erhöhtes Verzerrungspotenzial.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte auf Studienebene:**Einstufung als randomisierte Studie**

☐ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein:** Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

☐ **ja:** Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

☐ **unklar:** Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

☐ **nein:** Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

☐ **ja:** Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.☐ **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

☐ **unklar:** Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.☐ **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:**☐ **ja:** Die Patienten waren verblindet.☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.
- Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.
- Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.
- Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- Subgruppen
- Zeitpunkte/-räume
- Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)
- Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)
- Cut-off-points bei Dichotomisierung
- statistischer Verfahren

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).
- Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.

Zulässige Gründe sind:

- erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung
- Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe
- geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben
- Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.
- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

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

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

☐ **ja**

☐ **nein**

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

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

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- ☐ **hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ☐ **ja:** Der Endpunkt wurde verblindet erhoben.
- ☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.
- ☐ **nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ☐ **ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

☐ **unklar:** Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

☐ **nein:** Keines der unter „ja“ genannten drei Merkmale trifft zu.

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

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

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

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

☐ **ja**

☐ **nein**

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

Anhang 4-G: Ergänzende Unterlagen

Anhang 4-G liegt als separater Anhang vor.