

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Obeticholsäure (OCA; OCALIVA®)

Intercept Pharma Deutschland GmbH

Modul 4 A

für die Behandlung der primären biliären Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt) in Verbindung mit Ursodesoxycholsäure (UDCA) bei Erwachsenen, die unzureichend auf UDCA ansprechen, oder als Monotherapie bei Erwachsenen, die UDCA nicht tolerieren können.

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	11
Abkürzungsverzeichnis.....	15
4.1 Zusammenfassung der Inhalte von Modul 4.....	19
4.2 Methodik	26
4.2.1 Fragestellung	26
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	27
4.2.3 Informationsbeschaffung	31
4.2.3.1 Studien des pharmazeutischen Unternehmers	32
4.2.3.2 Bibliografische Literaturrecherche	32
4.2.3.3 Suche in Studienregistern	33
4.2.3.4 Selektion relevanter Studien	34
4.2.4 Bewertung der Aussagekraft der Nachweise.....	35
4.2.5 Informationssynthese und -analyse	36
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	36
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	38
4.2.5.3 Meta-Analysen.....	43
4.2.5.4 Sensitivitätsanalysen	45
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	46
4.2.5.6 Indirekte Vergleiche	47
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	49
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	49
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	49
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	49
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	51
4.3.1.1.3 Studien aus der Suche in Studienregistern	54
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	55
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	56
4.3.1.2.1 Studiendesign und Studienpopulationen	56
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	65
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	66
4.3.1.3.1 Endpunkt - RCT	67
4.3.1.3.1.1 Mortalität: Gesamtmortalität – RCT.....	67
4.3.1.3.1.2 Morbidität: AP und Bilirubin – RCT.....	71
4.3.1.3.1.3 Morbidität: Fibrose – RCT	97
4.3.1.3.1.4 Morbidität: Osteoporose – RCT	106
4.3.1.3.1.5 Morbidität: PBC-40 – RCT	110
4.3.1.3.1.6 Morbidität: Pruritus – RCT.....	117
4.3.1.3.1.7 Gesundheitsbezogene Lebensqualität: PBC-40 – RCT.....	128

4.3.1.3.1.8	Verträglichkeit – RCT	133
4.3.1.3.2	Subgruppenanalysen – RCT.....	140
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	141
4.3.2	Weitere Unterlagen.....	145
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	145
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	145
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	145
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	146
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	146
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	148
4.3.2.2	Nicht randomisierte vergleichende Studien	149
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	149
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	149
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	150
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	150
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien.....	151
4.3.2.3	Weitere Untersuchungen.....	152
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	152
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	153
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	158
4.3.2.3.3.1	Mortalität: Gesamtmortalität – weitere Untersuchungen	158
4.3.2.3.3.2	Morbidität: AP und Bilirubin – weitere Untersuchungen	159
4.3.2.3.3.3	Morbidität: Fibrose – weitere Untersuchungen	175
4.3.2.3.3.4	Morbidität: Osteoporose – weitere Untersuchungen	178
4.3.2.3.3.5	Morbidität: PBC-40– weitere Untersuchungen	179
4.3.2.3.3.6	Morbidität: Pruritus – weitere Untersuchungen	182
4.3.2.3.3.7	Gesundheitsbezogene Lebensqualität: PBC-40 – weitere Untersuchungen	186
4.3.2.3.3.8	Verträglichkeit – weitere Untersuchungen	188
4.3.2.3.3.9	Subgruppenanalysen – weitere Untersuchungen	191
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	192
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	193
4.4.1	Beurteilung der Aussagekraft der Nachweise	193
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	194
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	198
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	199
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	199
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	199
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	199
4.5.4	Verwendung von Surrogatendpunkten	199
4.6	Liste der eingeschlossenen Studien.....	211

4.7 Referenzliste.....	211
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche	216
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....	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).....	226
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT	229
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	246

Tabellenverzeichnis

	Seite
Tabelle 4-1: Übersicht der Ein- und Ausschlusskriterien für Studien zur Bewertung des Ausmaßes des Zusatznutzens von Obeticholsäure	30
Tabelle 4-2: Übersicht über die erhobenen demographischen Parameter und Baseline-Patientencharakteristika der POISE-Studie.....	39
Tabelle 4-3: Wichtige Response-Kriterien für UDCA in der Literatur	42
Tabelle 4-4: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	50
Tabelle 4-5: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	51
Tabelle 4-6: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	54
Tabelle 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	56
Tabelle 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	57
Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	58
Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel - Patienten mit Kombinationstherapie.....	59
Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel - Patienten mit Monotherapie.....	61
Tabelle 4-12: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-13: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel.....	66
Tabelle 4-14: Operationalisierung von Gesamtmortalität	68
Tabelle 4-15: Bewertung des Verzerrungspotenzials für Gesamtmortalität in RCT mit dem zu bewertenden Arzneimittel	68
Tabelle 4-16: Ergebnisse für Gesamtmortalität aus RCT mit dem zu bewertenden Arzneimittel - Patienten mit Kombinationstherapie.....	69
Tabelle 4-17: Ergebnisse für Gesamtmortalität aus RCT mit dem zu bewertenden Arzneimittel - Patienten mit Monotherapie.....	69
Tabelle 4-18: Operationalisierung von AP und Bilirubin	71
Tabelle 4-19: Bewertung des Verzerrungspotenzials für AP und Bilirubin in RCT mit dem zu bewertenden Arzneimittel	72
Tabelle 4-20: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP <1,67x ULN, Gesamt-Bilirubin ≤ULN und AP-Reduktion von ≥15%) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	73

Tabelle 4-21: Ergebnisse für die Einzelkomponenten des primären Endpunkts aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	74
Tabelle 4-22: Ergebnisse für den Anteil Patienten mit AP-Reduktion zu Baseline aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	76
Tabelle 4-23: Ergebnisse für den Anteil Patienten mit $AP \leq ULN$ aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	79
Tabelle 4-24: Ergebnisse für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	80
Tabelle 4-25: Ergänzende Ergebnisse gemäß der Kriterien nach PARIS, MAYO, ROTTERDAM, TORONTO und BARCELONA aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	83
Tabelle 4-26: Ergebnisse für den primären Endpunkt (Anteil Patienten mit $AP < 1,67 \times ULN$, Gesamt-Bilirubin $\leq ULN$ und AP-Reduktion von $\geq 15\%$) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie.....	94
Tabelle 4-27: Ergebnisse für die Einzelkomponenten des primären Endpunkts aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	94
Tabelle 4-28: Ergebnisse für den Anteil Patienten mit AP-Reduktion zu Baseline aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie.....	95
Tabelle 4-29: Ergebnisse für den Anteil Patienten mit $AP \leq ULN$ aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie.....	96
Tabelle 4-30: Ergebnisse für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	96
Tabelle 4-31: Operationalisierung von Fibrose (Morbidität)	98
Tabelle 4-32: Bewertung des Verzerrungspotenzials für Fibrose (Morbidität) in RCT mit dem zu bewertenden Arzneimittel.....	98
Tabelle 4-33: Ergebnisse für ELF-Score aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	99
Tabelle 4-34: Ergebnisse für Fibrose anhand des Fibroscan® (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	102
Tabelle 4-35: Ergebnisse der Ansprecheranalysen für Fibrose anhand des Fibroscan® (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	103
Tabelle 4-36: Ergebnisse für ELF-Score aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	104
Tabelle 4-37: Ergebnisse für Fibrose anhand des Fibroscan® (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	104
Tabelle 4-38: Ergebnisse der Ansprecheranalysen für Fibrose anhand des Fibroscan® (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie.....	105
Tabelle 4-39: Operationalisierung von Osteoporose (Morbidität).....	106
Tabelle 4-40: Bewertung des Verzerrungspotenzials für Osteoporose (Morbidität) in RCT mit dem zu bewertenden Arzneimittel	106

Tabelle 4-41: Ergebnisse für Osteoporose anhand des DEXA-Scan (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel in der Sicherheitspopulation – Patienten mit Kombinationstherapie	107
Tabelle 4-42: Ergebnisse für Osteoporose anhand des DEXA-Scan (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel in der Sicherheitspopulation – Patienten mit Monotherapie	109
Tabelle 4-43: Operationalisierung von PBC-40 (Morbidity)	110
Tabelle 4-44: Bewertung des Verzerrungspotenzials für PBC-40 (Morbidity) in RCT mit dem zu bewertenden Arzneimittel.....	111
Tabelle 4-45: Rücklaufquoten für PBC-40 (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	111
Tabelle 4-46: Ergebnisse für PBC-40 (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	112
Tabelle 4-47: Ergebnisse für PBC-40 (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	115
Tabelle 4-48: Operationalisierung von Pruritus (Morbidity)	117
Tabelle 4-49: Bewertung des Verzerrungspotenzials für 5-D und Pruritus-VAS (Morbidity) in RCT mit dem zu bewertenden Arzneimittel.....	118
Tabelle 4-50: Ergebnisse für das UE Pruritus (Inzidenzraten) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	119
Tabelle 4-51: Ergebnisse für das UE Pruritus (Time-to-first-event) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	119
Tabelle 4-52: Rücklaufquoten für 5-D Gesamtscore und Pruritus-VAS (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	120
Tabelle 4-53: Rücklaufquoten für 5-D Einzelskalen aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	120
Tabelle 4-54: Ergebnisse für 5-D und Pruritus-VAS (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	121
Tabelle 4-55: Ergebnisse für Pruritus (Inzidenzraten) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	125
Tabelle 4-56: Ergebnisse für 5-D und Pruritus-VAS (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie.....	126
Tabelle 4-57: Operationalisierung von PBC-40 (Lebensqualität).....	128
Tabelle 4-58: Bewertung des Verzerrungspotenzials für PBC-40 (Lebensqualität) in RCT mit dem zu bewertenden Arzneimittel	128
Tabelle 4-59: Rücklaufquoten für PBC-40 (Lebensqualität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	129
Tabelle 4-60: Ergebnisse für PBC-40 (Lebensqualität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	129
Tabelle 4-61: Ergebnisse für PBC-40 (Lebensqualität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	131

Tabelle 4-62: Operationalisierung von Verträglichkeit	133
Tabelle 4-63: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel	134
Tabelle 4-64: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	135
Tabelle 4-65: Ergänzende Analyse der schwere UE mit dem einzelnen UE Pruritus aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	136
Tabelle 4-66: Ergänzende Ergebnisse für UE mit einer Inzidenz $\geq 5\%$ in mindestens einer Behandlungsgruppe aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	136
Tabelle 4-67: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	137
Tabelle 4-68: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	138
Tabelle 4-69: Ergänzende Ergebnisse für UE mit einer Inzidenz $\geq 40\%$ in mindestens einer Behandlungsgruppe aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie	139
Tabelle 4-70: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	146
Tabelle 4-71: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden	146
Tabelle 4-72: Operationalisierung von <Endpunkt xxx>	147
Tabelle 4-73: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	147
Tabelle 4-74: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche	148
Tabelle 4-75: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	150
Tabelle 4-76: Operationalisierung von <Endpunkt xxx>	150
Tabelle 4-77: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	151
Tabelle 4-78: Charakterisierung der eingeschlossenen Studien – weitere Unterlagen mit dem zu bewertenden Arzneimittel	154
Tabelle 4-79: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel	155
Tabelle 4-80: Charakterisierung der Studienpopulationen – Langzeitsicherheitsstudie mit dem zu bewertenden Arzneimittel	156
Tabelle 4-81: Operationalisierung von Gesamtmortalität - weitere Untersuchungen	158
Tabelle 4-82: Ergebnisse der Obeticholsäure Patienten für Gesamtmortalität aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden	159
Tabelle 4-83: Ergebnisse der Obeticholsäure Patienten für Gesamtmortalität aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	159

Tabelle 4-84: Operationalisierung von AP und Bilirubin – weitere Untersuchungen	160
Tabelle 4-85: Ergebnisse der Obeticholsäure Patienten für den primären Endpunkt (AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin ≤ULN) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden	161
Tabelle 4-86: Ergebnisse der Obeticholsäure Patienten für die Einzelkomponenten des primären Endpunkts aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden	162
Tabelle 4-87: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP-Reduktion zu Baseline aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden	163
Tabelle 4-88: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP ≤ ULN aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	164
Tabelle 4-89: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden	164
Tabelle 4-90: Ergänzende Ergebnisse der Obeticholsäure Patienten für AP gemäß der Kriterien nach PARIS, MAYO, ROTTERDAM, TORONTO und BARCELONA aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	165
Tabelle 4-91: Ergebnisse der Obeticholsäure Patienten für den primären Endpunkt (AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin ≤ULN) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe.....	168
Tabelle 4-92: Ergebnisse der Obeticholsäure Patienten für die Einzelkomponenten des primären Endpunkts aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe.....	169
Tabelle 4-93: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP-Reduktion zu Baseline aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe.....	169
Tabelle 4-94: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP ≤ ULN aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	171
Tabelle 4-95: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	171
Tabelle 4-96: Ergänzende Ergebnisse der Obeticholsäure Patienten für AP gemäß der Kriterien nach PARIS, MAYO, ROTTERDAM, TORONTO und BARCELONA aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	171
Tabelle 4-97: Operationalisierung von Fibrose: Fibroscan® und ELF-Score - weitere Untersuchungen.....	175

Tabelle 4-98: Ergebnisse der Obeticholsäure Patienten für ELF-Score aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	176
Tabelle 4-99: Ergebnisse der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden	177
Tabelle 4-100: Ergebnisse der Ansprecheranalysen der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	177
Tabelle 4-101: Ergebnisse der Obeticholsäure Patienten für ELF-Score aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe.....	177
Tabelle 4-102: Ergebnisse der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe.....	178
Tabelle 4-103: Ergebnisse der Ansprecheranalysen der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	178
Tabelle 4-104: Operationalisierung von PBC-40 (Morbidity) - weitere Untersuchungen....	179
Tabelle 4-105: Ergebnisse der Obeticholsäure Patienten für PBC-40 (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	180
Tabelle 4-106: Ergebnisse der Obeticholsäure Patienten für PBC-40 (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	181
Tabelle 4-107: Operationalisierung von Pruritus (Morbidity) - weitere Untersuchungen....	182
Tabelle 4-108: Ergebnisse für Pruritus (Inzidenzraten) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	183
Tabelle 4-109: Ergebnisse für Pruritus (Time-to-first-event) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	183
Tabelle 4-110: Ergebnisse der Obeticholsäure Patienten für 5-D und Pruritus-VAS aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	184
Tabelle 4-111: Ergebnisse für Pruritus (Inzidenzraten) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	184
Tabelle 4-112: Ergebnisse für Pruritus (Time-to-first-event) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	185
Tabelle 4-113: Ergebnisse der Obeticholsäure Patienten für 5-D und Pruritus-VAS aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	185
Tabelle 4-114: Operationalisierung von PBC-40 (Lebensqualität) - weitere Untersuchungen	186

Tabelle 4-115: Ergebnisse für PBC-40 (Lebensqualität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	187
Tabelle 4-116: Ergebnisse für PBC-40 (Lebensqualität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	187
Tabelle 4-117: Operationalisierung von Verträglichkeit - weitere Untersuchungen	188
Tabelle 4-118: Ergebnisse für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	189
Tabelle 4-119: Ergebnisse der ergänzenden Analysen für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	190
Tabelle 4-120: Ergebnisse für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe	190
Tabelle 4-121: Ergebnisse der ergänzenden Analysen für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden.....	191
Tabelle 4-122: Ausmaß des Zusatznutzens von Obeticholsäure auf Endpunktebene - Patienten mit Kombinationstherapie	195
Tabelle 4-123: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	198
Tabelle 4-124: Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in Cochrane Register of Controlled Trials.....	217
Tabelle 4-125: Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in MEDLINE.....	219
Tabelle 4-126: Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in EMBASE	221
Tabelle 4-127 (Anhang): Aus der Bewertung ausgeschlossene RCT aus der Suche in Studienregistern mit Ausschlussgrund	226
Tabelle 4-128 (Anhang): Studiendesign und -methodik für POISE-Studie (747-301).....	230
Tabelle 4-129 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie <Studienbezeichnung>.....	247

Abbildungsverzeichnis**Seite**

Abbildung 4-1: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP <1,67x ULN, Gesamt-Bilirubin ≤ULN und AP-Reduktion von ≥15%) aus RCT und Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	24
Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	53
Abbildung 4-3: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP <1,67x ULN, Gesamt-Bilirubin ≤ULN und AP-Reduktion von ≥15%) aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	74
Abbildung 4-4: Ergebnisse für den Anteil Patienten mit AP <1,67x ULN aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	75
Abbildung 4-5: Ergebnisse für den Anteil Patienten mit Gesamt-Bilirubin ≤ULN aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	76
Abbildung 4-6: Ergebnisse für den Anteil Patienten mit AP-Reduktion ≥10% aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	77
Abbildung 4-7: Ergebnisse für den Anteil Patienten mit AP-Reduktion ≥15% aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	78
Abbildung 4-8: Ergebnisse für den Anteil Patienten mit AP-Reduktion ≥20% aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	79
Abbildung 4-9: Ergebnisse für den Anteil Patienten mit AP-Reduktion ≥40% aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	79
Abbildung 4-10: Ergebnisse für den Anteil Patienten mit AP≤ULN aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	80
Abbildung 4-11: Ergebnisse für den Endpunkt AP, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	82
Abbildung 4-12: Ergebnisse für den Endpunkt Gesamt-Bilirubin, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	82
Abbildung 4-13: Ergebnisse für den Endpunkt Direktes Bilirubin, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	83

Abbildung 4-14: Ergänzende Ergebnisse für Anteil Patienten, die das Kriterium PARIS I erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	86
Abbildung 4-15: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium PARIS I erfüllt haben, Patienten aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie	86
Abbildung 4-16: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium PARIS I erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline erfüllt haben – Patienten mit Kombinationstherapie	87
Abbildung 4-17: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium PARIS II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	87
Abbildung 4-18: Ergänzende Ergebnisse den Anteil Patienten, die das Kriterium PARIS II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie	88
Abbildung 4-19: Ergänzende Ergebnisse den Anteil Patienten, die das Kriterium MAYO II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	88
Abbildung 4-20: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium MAYO II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie	89
Abbildung 4-21: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium ROTTERDAM erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	89
Abbildung 4-22: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium ROTTERDAM erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie	90
Abbildung 4-23: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium ROTTERDAM erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline erfüllt haben – Patienten mit Kombinationstherapie	90
Abbildung 4-24: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium TORONTO II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	91
Abbildung 4-25: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium TORONTO II erfüllt haben, Patienten aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie	91

Abbildung 4-26: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium BARCELONA erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	92
Abbildung 4-27: Ergänzende Analyse der Patienten, die x von sechs möglichen Kriterien für AP und Bilirubin (Paris I, Paris II, Mayo II, Rotterdam, Toronto II, Barcelona) erfüllt haben – Patienten mit Kombinationstherapie aus RCT; Obeticholsäure+UDCA versus UDCA+Placebo.....	93
Abbildung 4-28: Ergebnisse für ELF-Score, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	100
Abbildung 4-29: Ergebnisse für HA (ng/mL), absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	101
Abbildung 4-30: Ergebnisse für P3NP (ng/mL), absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	101
Abbildung 4-31: Ergebnisse für TIMP-1 (ng/mL), absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	102
Abbildung 4-32: Ergebnisse für PBC-40 generelle Symptome, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	113
Abbildung 4-33: Ergebnisse für PBC-40 Jucken, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	113
Abbildung 4-34: Ergebnisse für PBC-40 Fatigue, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	114
Abbildung 4-35: Ergebnisse für PBC-40 kognitive Funktion, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	114
Abbildung 4-36: Ergebnisse für den 5-D Gesamtscore, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	122
Abbildung 4-37: Ergebnisse für 5-D Dauer, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	122
Abbildung 4-38: Ergebnisse für 5-D Schweregrad, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	123
Abbildung 4-39: Ergebnisse für 5-D Richtung, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie.....	123

Abbildung 4-40: Ergebnisse für 5-D Behinderung, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	124
Abbildung 4-41: Ergebnisse für 5-D Verteilung, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	124
Abbildung 4-42: Ergebnisse für 5-D Pruritus-VAS, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	125
Abbildung 4-43: Ergebnisse für PBC-40, soziale Domäne, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	130
Abbildung 4-44: Ergebnisse für PBC-40, emotionale Domäne, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie	131
Abbildung 4-45: Serum-Bilirubin gegen Zeit der Nachbeobachtung (39)	202
Abbildung 4-46: Mayo Score gegen die Zeit der Nachbeobachtung , getrennt nach UDCA Ansprechern und Nicht-Ansprechern (23)	203
Abbildung 4-47: UK-PBC Vorhersagemodell für Tod auf Grund Leberveränderungen oder Lebertransplantation in 15 Jahren bei unterschiedlicher AP nach 1 Jahre Therapie mit UDCA.....	204
Abbildung 4-48: Hansen et al (2015): Kaplan-Meier-Kurve für den Vergleich der AP-Werte (7)	206
Abbildung 4-49: Hansen et al (2015): Kaplan-Meier-Kurve für den Vergleich der Bilirubin-Werte (7).....	207
Abbildung 4-50: Lammers et al (2014): Kaplan-Meier-Kurve für den Vergleich der Bilirubin-Werte (40).....	208
Abbildung 4-51: GLOBE Vorhersagemodell für Lebertransplantat-freies Überleben	209
Abbildung 4-52: Lammers et al (2013): Forestplot der Hazard Ratios im Vergleich verschiedener Szenarien für den kombinierten Endpunkt aus AP und Bilirubin - Subgruppenanalysen (9).....	210
Abbildung 4-53: Flow-Chart der Patienten der POISE-Studie	245

Abkürzungsverzeichnis

Abkürzung	Bedeutung
Abs.	Absatz
AM-NutzenV	Arzneimittel-Nutzenverordnung
AP (auch: ALP)	Alkalische Phosphatase
AST	Asparat-Aminotransferase
ALT	Alanin-Aminotransferase
BMI	Body Mass Index
bzgl.	bezüglich
bspw.	beispielsweise
bzw.	beziehungsweise
CHMP	Komitee für medizinische Produkte zur Verwendung im Menschen (Committee for Medicinal Products for Human Use)
CMH	Cochran-Mantel-Haenszel Test
COMP	Komitee für medizinische Produkte für seltene Erkrankungen (Committee for Orphan Medicinal Products)
CONSORT	Konsolidierte Standards für das Berichten von Studien (Consolidated Standards of Reporting Trials)
CSR	klinischer Studienreport
d.h.	das heißt
DIMDI	Deutsches Institut für Medizinische Dokumentation
EG	Europäische Gemeinschaft
ELF	Fortgeschrittene Leber-Fibrose (Enhanced liver fibrosis)
EMA	Europäische Zulassungsbehörde für Arzneimittel (European Medicines Agency)
EMBASE	Medizinische Datenbank (Excerpta Medical Database)
etc.	et cetera
evtl.	eventuell
G	Gramm
G-BA	Gemeinsamer Bundesausschuss
GGT	Gamma-Glutamyl-Transferase
HR	Hazard Ratio
HS	Halbsatz
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

INR	Internationales normiertes Verhältnis
ITT	Intention to treat
IxRS	Interaktives Sprach- und Internet-Antwortsystem (Interactive Voice-response and Web-response system)
Kg	Kilogramm
KI	Konfidenzintervall
L	Liter
LS	Kleinste Quadrate (Least-square)
LTSE	Langzeit-Sicherheitsstudie
m ²	Quadratmeter
MEDLINE	Medizinisches Literatur Analyse und Zugriffssystem (Medical Literature Analysis and Retrieval System Online)
MedDRA	Medizinisches Lexikon für regulatorische Aktivitäten (Medical Dictionary for Regulatory Activities)
Mg	Milligramm
µmol	Mikromol
MTC	Netzwerkmetaanalyse (Mixed Treatment Comparison)
MW	Mittelwert
MWD	Mittelwertdifferenz
N	Anzahl Patienten mit Ereignis
N	Anzahl Patienten insgesamt
n.b.	nicht berechenbar (Not estimable)
OCA	Obeticholsäure (Obeticholic Acid)
PBC	primäre biliäre Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt)
PBC-40	Standardisierter Fragebogen zur gesundheitsbezogenen Lebensqualität bei PBC mit 40 Items
Pbo	Placebo
RCT	Randomisierte kontrollierte Studie (Randomized Controlled Trial)
RD	Risikodifferenz
RR	Relatives Risiko
RRR	Invertiertes relatives Risiko
OCA	Obeticholsäure
OR	Odds Ratio

SAP	Statistischer Analyseplan
SD	Standardabweichung
SE	Standardfehler
SGB	Sozialgesetzbuch
SMD	Standardisierte Mittelwertdifferenz
STE	Surrogat-Grenzwerteffekt (Surrogate Threshold Effects)
STROBE	Stärkung des Berichtens von Beobachtungsstudien in der Epidemiologie (Strengthening the Reporting of Observational Studies in Epidemiology)
SUE	Schwerwiegendes unerwünschtes Ereignis
TE	Transiente Elastographie
TREND	Transparentes Berichten von nicht-randomisierten Studien (Transparent Reporting of Evaluations with Non-Randomized Design)
U	Einheit
UDCA	Ursodesoxycholsäure (Ursodeoxycholic Acid)
UE	Unerwünschtes Ereignis
ULN	Obere Grenze des Normbereichs
VAS	Visuelle Analogskala (Visual Analogue Scale)
vs.	versus
WHO	Welt-Gesundheitsbehörde (World Health Organization)
z.B.	zum Beispiel
ZVT	zweckmäßige Vergleichstherapie

Anmerkung:

Zur einfacheren Lesbarkeit wurde meist die männliche Form (z.B. „Ansprecher“, „Patient“, „Arzt“ usw.) verwendet. Damit ist stets auch die weibliche Form gemeint.

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

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

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

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Die im vorliegenden Dossier zu beantwortende Fragestellung lautet:

Wie ist das Ausmaß des medizinischen Zusatznutzens von Obeticholsäure (Obeticholic Acid, im Folgenden teils: OCA) in der Behandlung von erwachsenen Patienten mit primärer biliärer Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt) in Kombination mit Ursodesoxycholsäure (Ursodeoxycholic acid, im Folgenden: UDCA) bei Erwachsenen, die unzureichend auf UDCA ansprechen, oder als Monotherapie bei Erwachsenen, die UDCA nicht tolerieren, hinsichtlich der patientenrelevanten Endpunkte Mortalität, Morbidität, Lebensqualität und Verträglichkeit – im Vergleich zur Therapie mit UDCA+Placebo bzw. zum abwartenden Beobachten (Placebo – zu bewerten?

Obeticholsäure erfüllt gemäß der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 die Bedingungen zur Zulassung als Arzneimittel zur Behandlung eines seltenen Leidens, einem sogenannten „Orphan Drug“ (1). Der Orphan Drug-Status wurde am 14.11.2016 von der EU-Kommission auf Grund der Orphan Designation EMEA/OD/073/09 und EU/3/10/753 erteilt (2).

Gemäß den Vorgaben des G-BA sind für Orphan Drugs nur die Abschnitte 4.4.2 und 4.4.3 zwingend auszufüllen. Der G-BA stellt dem pharmazeutischen Hersteller jedoch frei, auch darüber hinaus Angaben in anderen Abschnitten zu machen, sofern dies nach Ermessen des Herstellers der strukturierten Darstellung der Ergebnisse zum Nachweis des Ausmaßes des Zusatznutzens dient. Intercept stellt im vorliegenden Dossier auch über die oben genannten Abschnitte hinausgehende Informationen zu Obeticholsäure dar.

Für Orphan Drugs gilt der medizinische Zusatznutzen gemäß § 35a Abs. 1 Satz 10 SGB V bereits durch die Zulassung als belegt. Daher ist das Erbringen des Nachweises für einen medizinischen Zusatznutzen von Obeticholsäure gegenüber einer ZVT gemäß § 35a Abs. 1 Satz 10 1. HS SGB V nicht erforderlich.

Da der Zusatznutzen als belegt gilt, wird im Folgenden über die formalen Anforderungen hinaus das Ausmaß des Zusatznutzens von Obeticholsäure im Rahmen der Vorgaben gemäß Verfahrensordnung des G-BA anhand von Daten aus einer randomisierten klinischen Studie (RCT) hergeleitet.

Insgesamt umfasste das klinische Programm für Obeticholsäure, das bei der EMA zur Zulassung eingereicht wurde, zwei Phase-II- und eine Phase-III-Studie.

Für Obeticholsäure im Anwendungsgebiet wurden zwei Phase-II- (Studien 747-201 und 747-202) und eine Phase-III-Studie (747-301, POISE) doppelblind mit Studienzentren in Nordamerika und Europa mit Beobachtungszeiten von 18 Wochen bis 14 Monaten durchgeführt. In jeder der drei Studien wurde eine offene, unkontrollierte

Langzeitnachsbeobachtungsphase zur Evaluierung der Wirksamkeit und Sicherheit von Obeticholsäure bei PBC angeschlossen. In den beiden Phase-II-Studien wurden Obeticholsäure-Dosen zwischen 10 mg und 50 mg pro Tag verabreicht. In der Phase-III-Studie wurde in einem Studienarm eine Obeticholsäure-Dosis von 10 mg verwendet, im zweiten Arm eine Startdosis von 5 mg, die auf 10 mg titriert wurde, wenn die Patienten den primären Endpunkt der Studie sechs Monate nach Behandlungsbeginn noch nicht erreicht hatten.

Die Kontrollgruppe in der Phase-II-Studie 747-201 erhielt als Vergleichstherapie Placebo, in der Phase-II-Studie 747-202 UDCA+Placebo, sowie in der Phase-III-Studie POISE zum einen für Patienten mit Kombinationstherapie UDCA+Placebo, zum anderen für Patienten mit Unverträglichkeit von UDCA abwartendes Beobachten unter Verwendung eines Placebo.

Primäre Endpunkte der drei Studien waren die Änderungen der alkalischen Phosphatase (AP, Phase-II) bzw. die Ansprecherraten im kombinierten Endpunkt aus $AP < 1.67 \times \text{Normbereich (ULN)}$ bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und des Gesamt-Bilirubins $\leq \text{ULN}$ in der Phase III.

In allen drei Studien erreichte Obeticholsäure den primären Endpunkt, so dass eine signifikante Überlegenheit von Obeticholsäure gegenüber der jeweiligen Kontrollgruppe gezeigt werden konnte. Vorteile zeigten sich zudem in allen drei Studien in anderen Endpunkten wie den biochemischen Kenngrößen GGT, AST und ALT.

Die Langzeitnachsbeobachtung von bis zu 3 Jahren bestätigt das in der POISE-Studie beobachtete Sicherheitsprofil für Obeticholsäure.

Das Anwendungsgebiet für Obeticholsäure ist:

OCALIVA® wird angewendet für die Behandlung der primären biliären Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt) in Kombination mit Ursodesoxycholsäure (UDCA) bei Erwachsenen, die unzureichend auf UDCA ansprechen, oder als Monotherapie bei Erwachsenen, die UDCA nicht tolerieren können (3).

Datenquellen

Zur Ermittlung der Evidenz aus Studien zu Obeticholsäure für die Behandlung der primären biliären Cholangitis wurden bibliographische Literaturrecherchen in den relevanten Datenbanken Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medical Database (EMBASE) und Cochrane Library mittels vorab definierter Schlagwörter und Suchstrategien durchgeführt. Zudem wurden Studienregister systematisch mittels vorab definierter Schlagwörter durchsucht. Weitere Daten wurden firmeninternen Datenquellen entnommen.

Ein-/Ausschlusskriterien für Studien

Kriterien für den Ein- und Ausschluss von Studien wurden gemäß den Vorgaben des G-BA in Abschnitt 4.2.2 gewählt.

Die Selektion der relevanten Studien aus der bibliographischen Recherche und der Recherche in den Studienregistern wurde gemäß dieser Kriterien von zwei Reviewern unabhängig voneinander durchgeführt. Diskrepanzen zwischen den beiden Reviewern wurden im Konsens gelöst.

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

Die aus der Literaturrecherche identifizierte randomisierte, kontrollierte Studie (Randomised Controlled Trial; RCT) wird gemäß der ITEMS 1 bis 14 des Consolidated Standards of Reporting Trials (CONSORT) Statements beurteilt. Neben der Checkliste des CONSORT Statements werden auch die weiteren Ausführungen zu den einzelnen ITEMS zur besseren Beurteilung der internen und externen Validität der Studie herangezogen. Das Ziel der Studie, die Interventionen, die Zielgröße(n) und die Studienpopulation sollten mittels entsprechender Kriterien klar definiert sein. Ausschlaggebend für eine Bewertung der Studie bezüglich der unverzerrten Aussagekraft sind beispielsweise die angewandten Methoden zur Fallzahlbestimmung, die Art und Weise der Randomisierung (insbesondere die Generierung der Randomisierungssequenz und die verdeckte Zuteilung der Patienten zu den Gruppen), die Verblindung, die adäquate Anwendung von statistischen Analysemethoden und der Umgang mit Studienaustritten und Protokollverletzern. Weiterhin wird die Darstellung der Ergebnisse, die Vollständigkeit berichteter Ergebnisse und der Patientenfluss zur Beurteilung eines potentiellen Verzerrungspotenzials betrachtet.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Zur Bewertung des Ausmaßes des medizinischen Zusatznutzens von Obeticholsäure für die Behandlung der primären biliären Cholangitis wurde eine randomisierte, placebokontrollierte, multizentrische, doppelblinde Phase-III-Studie, die POISE-Studie, herangezogen.

Im vorliegenden Dossier zur Nutzenbewertung von Obeticholsäure werden folgende patientenrelevante Endpunkte ausgewertet:

Mortalität

- Gesamtüberleben

Morbidität

- AP und Bilirubin

- Fibrose
- Osteoporose
- PBC-40
- Pruritus

Gesundheitsbezogene Lebensqualität

- PBC-40

Verträglichkeit

- Gesamtrate der unerwünschten Ereignisse (UE) Schwerwiegende UE
- Schwere UE ohne Pruritus
- UE, die zum Therapieabbruch führten

Die Patientenrelevanz der oben genannten Endpunkte begründet sich dabei wie folgt:

Das **Gesamtüberleben** wird in RCTs außerhalb der Onkologie als Ereignisrate im Studienzeitraum im Rahmen der Sicherheitsendpunkte erhoben. Die Gesamtmortalität wird im Rahmen des AMNOG als herausragendes Therapieziel angesehen.

In der Endpunktdimension **Morbidität** werden folgende Endpunkte dargestellt.

Die Endpunkte AP und Bilirubin werden in diesem Dossier in verschiedenen Operationalisierungen dargestellt, da sich der primäre Wirksamkeitsendpunkt aus diesen zusammensetzt. Die patientenrelevanten Endpunkte Gesamtüberleben und Lebertransplantation sind bei PBC in klinischen Studien nur schwer erhebbar, weshalb AP und Bilirubin als biochemische Kenngrößen für die Morbidität der Patienten empfohlen werden (4). Beide Kenngrößen wurden in großen PBC-Registern validiert („UK-PBC“ (5),(6) sowie „GLOBE“ (7),(8),(9), für Details siehe auch Abschnitt 4.5.4). FDA (10) und EMA (11) bestätigen beide die Validität dieser biochemischen Kenngrößen für PBC. Diese Surrogatparameter gelten daher als validiert.

Der Endpunkt Pruritus wird anhand des unerwünschten Ereignisses Pruritus dargestellt. Zusätzlich werden zwei validierte Instrumente herangezogen: Der Pruritus-Fragebogen 5-D ist ein validiertes Instrument zur Messung des Pruritus (12). Die Pruritus-Visuelle Analogskala (Pruritus-VAS) ist für Patienten mit chronischem Pruritus validiert (13).

Der Endpunkt Fibrose wird anhand des ELF-Scores und der transienten Elastografie dargestellt. Der ELF-Score wurde validiert (14): Die Autoren bestätigen mit der Studie die Aussagekraft des Tests als Index zur Prädiktion einer Leber-Fibrose oder -Zirrhose. Die transiente

Elastografie (TE; z.B. Fibroscan®) ist ein validiertes, nicht-invasives Verfahren zur Bestimmung des quantitativen Ausmaßes der Fibrosierung der Leber (15).

Die Osteoporose wird anhand des DEXA-Scans dargestellt, einer nicht-invasiven Methode zur Bestimmung der Osteoporose bei postmenopausalen Frauen.

Der Fragebogen PBC-40 ist ein etablierter und validierter Fragebogen für die PBC und enthält Domänen zur **Morbidität** sowie der **gesundheitsbezogenen Lebensqualität** (16).

Im vorliegenden Nutzendossier wird die **Verträglichkeit** anhand der Endpunkte Gesamtrate UE, schwerwiegende UE (SUE), schwere UE ohne Pruritus, UE, die zum Therapieabbruch führten, dargestellt.

UE werden grundsätzlich als patientenrelevant betrachtet (17). Pruritus wird als UE von speziellem Interesse als patientenrelevantes Symptom separat in der Morbidität dargestellt. Um eine Doppelzählung des Symptoms Pruritus zu vermeiden, wird als UE-Endpunkt "schwere UE ohne Pruritus" dargestellt.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Die doppelblinde, randomisierte, kontrollierte POISE-Studie („747-301“, NCT 01473524) steht nach der G-BA-Verfahrensordnung auf Stufe Ib der Evidenzhierarchie und ist damit zur Herleitung des Ausmaßes des Zusatznutzens sowohl für Patienten mit Kombinationstherapie als auch für Patienten mit Monotherapie geeignet.

Patienten mit Kombinationstherapie

Für den Endpunkt Mortalität ließ sich kein Zusatznutzen für Obeticholsäure nachweisen.

In der Endpunktdimension Morbidität wurde der Zusatznutzen anhand der Endpunktkomplexe AP und Bilirubin, Fibrose, Osteoporose und PBC-40 hergeleitet. Für den Endpunktkomplex AP und Bilirubin wurde der primäre Endpunkt der POISE-Studie als Hauptoperationalisierung herangezogen (Anteil Patienten mit AP $<1.67 \times$ ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin $\leq$ ULN). Wie oben dargestellt, werden diese biochemischen Kenngrößen als valide Surrogate für die Leber-bedingte Mortalität und die Lebertransplantation angesehen. Der primäre Endpunkt wurde in 46% der Patienten der OCA+UDCA-Gruppe erreicht im Vergleich zu 10% in der Kontrollgruppe ($p < 0,0001$) erreicht. Dies entspricht einem relativen Risiko von 0,22 (95% KI: 0,11; 0,47). Da die biochemischen Kenngrößen AP und Bilirubin mit dem Grad der Leberschädigung korrelieren, handelt es sich um einen validen Surrogatendpunkt für eine patientenrelevante Morbiditäts- und Mortalitätsveränderung. Wie in Abschnitt 4.5.4 ausführlich beschrieben, ist die Verbesserung der Surrogatparameter AP und Bilirubin gemäß der Analysen des UK-PBC- und des GLOBE-Registers prädiktiv für die Krankheitsbelastung und insbesondere die Überlebensdauer und die Zeit bis zu einer Lebertransplantation.

Es zeigt sich mit dem durchgängigen Erreichen des primären Endpunkts in der POISE-Studie eine patientenrelevante, spürbare Linderung der Erkrankung, wodurch ein beträchtlicher Zusatznutzen gegenüber der Kontrollgruppe belegt ist.

Die Einzelkomponenten und die übrigen Operationalisierungen des Endpunktes AP und Bilirubin unterstützen den beträchtlichen Zusatznutzen.

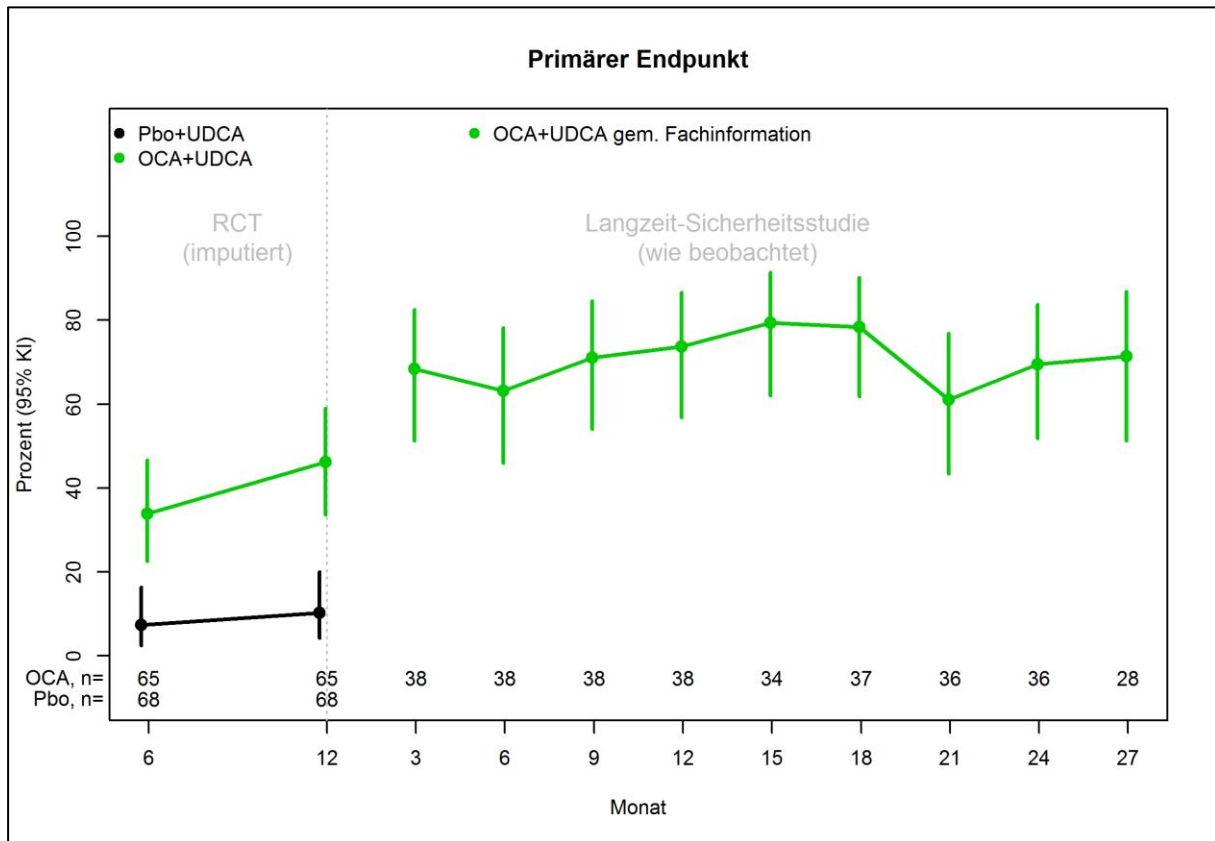


Abbildung 4-1: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP $<1,67 \times$ ULN, Gesamt-Bilirubin $\leq$ ULN und AP-Reduktion von $\geq 15\%$) aus RCT und Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Bezüglich der übrigen Morbiditätsendpunkte zeigten sich keine Vor- oder Nachteile für Obeticholsäure gegenüber der Kontrollgruppe UDCA+Placebo. Dies ist unter anderem der sehr langsam fortschreitenden Erkrankung und dem Zeitpunkt der Behandlung geschuldet, da die klinischen Ereignisse Leberversagen oder Leber-bedingte Mortalität sowie andere Krankheitssymptome erst im Verlauf der Erkrankung entstehen können. Die frühzeitige Behandlung der PBC hat das Ziel, diese schwerwiegenden Ereignisse zu vermeiden. Eine Erhebung solcher Ereignisse in einer klinischen Studie erfordert eine Beobachtungszeit von mehreren Jahren / Jahrzehnten und kann zum Zeitpunkt der Zulassung von Obeticholsäure daher noch nicht vorliegen.

Bei den Verträglichkeitsendpunkten bestand kein Unterschied in den UE, den SUE, den schweren UE ohne Pruritus und den UE, die zum Therapieabbruch führten. Pruritus ist ein

krankheitsbedingtes Grundsymptom der PBC. In den verschiedenen Operationalisierungen der POISE-Studie zeigten sich keine signifikanten Unterschiede gegenüber der Kontrollgruppe. Die Ergebnisse der Verträglichkeitsendpunkte mindern somit nicht den beträchtlichen Zusatznutzen von Obeticholsäure.

In der einarmigen Langzeit-Sicherheitsstudie (LTSE) bestätigte sich der in der randomisierten Therapiephase nachgewiesene Zusatznutzen von Obeticholsäure. Die Langzeitdaten erlauben dabei eine Abschätzung der langfristigen Sicherheit und Wirksamkeit von Obeticholsäure in dieser chronischen und progressiven Erkrankung. Bei den Endpunkten Mortalität, Morbidität und gesundheitsbezogene Lebensqualität zeigten sich Ergebnisse, die auf dem Niveau des Monats 12 der RCT-Phase der POISE-Studie oder darüber lagen. Insbesondere bei AP und Bilirubin bestätigt sich damit eine langfristige Wirksamkeit, die über den gesamten Beobachtungszeitraum (POISE, 12 Monate, und LTSE, 27 Monate mit ausreichender Evidenz) bis zu 39 Monaten nachgewiesen ist. Hinsichtlich Sicherheits-Endpunkten liegen keine Auffälligkeiten vor. Insbesondere die Rate an Patienten, die die Langzeit-Sicherheitsstudie wegen UE abbrechen, lag dabei mit 5% sehr niedrig. Die in der LTSE beobachteten Effekte sind daher zu den Ergebnissen der POISE-Studie gleichgerichtet und bestätigen den dort gefundenen beträchtlichen Zusatznutzen.

In der Gesamtschau ergibt sich damit für Obeticholsäure bei Patienten mit Kombinationstherapie ein beträchtlicher Zusatznutzen.

Patienten mit Monotherapie

Für die Patienten mit Monotherapie lagen insgesamt nur Daten von 5 Patienten aus der Obeticholsäure-Behandlungsgruppe und 5 Patienten aus der Placebo-Behandlungsgruppe vor.

Eine Nichttoleranz auf UDCA stellt eine seltene Subentität dar, die im Rahmen der Zulassungsstudie nur in sehr begrenzten Zahlen eingeschlossen werden konnte. In dieser kleinen Population konnten jedoch durchaus Effekte nachgewiesen werden, die allerdings – wie auch dargestellt – eine Quantifizierung nicht zulassen. Die Richtung und das Ausmaß der Effekte sprechen jedoch gerade für die Richtigkeit der Nutzenfiktion bei Orphan Drugs.

Auf Grund der geringen, nicht ausreichend belastbaren Evidenzlage wurde auf die Herleitung eines Zusatznutzens auf Endpunktebene verzichtet. Gemäß § 35a Abs. 1 Satz 10 1. HS SGB V gilt der medizinische Zusatznutzen in dieser Teilpopulation durch die Zulassung als belegt.

In der Gesamtschau ergibt sich damit für Obeticholsäure bei Patienten mit Monotherapie ein nicht quantifizierbarer Zusatznutzen.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

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

Wie ist das Ausmaß des medizinischen Zusatznutzens von Obeticholsäure (Obeticholic Acid, im Folgenden teils: OCA) in der Behandlung von erwachsenen Patienten mit primärer biliärer Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt) in Verbindung mit Ursodesoxycholsäure (Ursodeoxycholic acid, im Folgenden: UDCA) bei Erwachsenen, die unzureichend auf UDCA ansprechen, oder als Monotherapie bei Erwachsenen, die UDCA nicht tolerieren können, hinsichtlich der patientenrelevanten Endpunkte Mortalität, Morbidität, Lebensqualität und Verträglichkeit – im Vergleich zur Monotherapie mit UDCA+Placebo bzw. zum abwartenden Beobachten (Placebo) – zu bewerten?

Das zugelassene **Anwendungsgebiet** und die zugelassene **Patientenpopulation** für Obeticholsäure ergeben sich aus der Fachinformation:

OICALIVA® wird angewendet für die Behandlung der primären biliären Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt) in Verbindung mit Ursodesoxycholsäure (UDCA) bei Erwachsenen, die unzureichend auf UDCA ansprechen, oder als Monotherapie bei Erwachsenen, die UDCA nicht tolerieren können (3).

Obeticholsäure erfüllt gemäß der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 die Bedingungen zur Zulassung als Arzneimittel zur Behandlung eines seltenen Leidens, einem sogenannten „Orphan Drug“ (1).

Da der Zusatznutzen als belegt gilt, wird im Folgenden über die formalen Anforderungen hinaus das Ausmaß des Zusatznutzens von Obeticholsäure im Rahmen der Vorgaben gemäß Verfahrensordnung des G-BA anhand von Daten aus einer randomisierten klinischen Studie (RCT) hergeleitet.

Insgesamt umfasste das klinische Programm für Obeticholsäure, das bei der EMA zur Zulassung eingereicht wurde, die folgenden **Studientypen**: zwei Phase-II-Studien (Studien 747-201 und 747-202) und eine Phase-III-Studie (747-301, POISE) doppelblind mit Studienzentren in Nordamerika und Europa mit Beobachtungszeiten von 18 Wochen bis 14 Monaten durchgeführt. In jeder der drei Studien wurde eine offene, unkontrollierte Langzeitnachbeobachtungsphase zur Evaluierung der Wirksamkeit und Sicherheit von Obeticholsäure bei PBC angeschlossen. Als **Intervention** wurden in den beiden Phase-II-Studien Obeticholsäure in Dosen zwischen 10 mg und 50 mg pro Tag verabreicht. In der Phase-III-Studie wurde in einem Studienarm eine Obeticholsäure Dosis von 10 mg verwendet, im zweiten Arm eine Startdosis von 5 mg, die auf 10 mg titriert wurde, wenn die Patienten den primären Endpunkt der Studie 6 Monate nach Behandlungsbeginn noch nicht erreicht hatten.

Die Kontrollgruppe in der Phase-II-Studie 747-201 erhielt als **Vergleichstherapie** Placebo, in der Phase-II-Studie 747-202 UDCA+Placebo, sowie in der Phase-III-Studie POISE zum einen für Patienten mit Kombinationstherapie UDCA+Placebo, zum anderen für Patienten mit Unverträglichkeit von UDCA abwartendes Beobachten unter Verwendung eines Placebo.

Primäre **Endpunkte** der drei Studien waren die Änderungen der alkalischen Phosphatase (AP, Phase-II) bzw. die Ansprecherraten im kombinierten Endpunkt aus $AP < 1.67 \times \text{Normbereich (ULN)}$ bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und des Gesamt-Bilirubins $\leq \text{ULN}$ in der Phase-III.

In allen drei Studien erreichte Obeticholsäure den primären Endpunkt, so dass eine signifikante Überlegenheit von Obeticholsäure gegenüber der jeweiligen Kontrollgruppe gezeigt werden konnte. Vorteile zeigten sich zudem in allen drei Studien in anderen Endpunkten wie den biochemischen Kenngrößen GGT, AST und ALT. Die Langzeitnachbeobachtung von bis zu 4 Jahren bestätigt das in der POISE-Studie beobachtete Sicherheitsprofil für Obeticholsäure.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

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

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

Nach der oben formulierten Fragestellung zur Bewertung des medizinischen Nutzens und Zusatznutzens von Obeticholsäure bzw. Obeticholsäure in Kombination mit UDCA wurden folgende Kriterien für den Einschluss von Studien in die Nutzenbewertung berücksichtigt:

Patientenpopulation

Die Patientenpopulation umfasst erwachsene Patienten mit primärer biliärer Cholangitis mit nicht ausreichendem Ansprechen auf UDCA oder bei Unverträglichkeit von UDCA.

Intervention

Die Intervention ist Obeticholsäure (OCALIVA®), ein selektiver und potenter Agonist für den nukleären Farnesoid X-Rezeptor (FXR).

Vergleichstherapie

Obeticholsäure hat seit dem 27.07.2010 den Status als Medikament zur Behandlung eines seltenen Leidens für die Behandlung der primären biliären Cholangitis (18) und ist gemäß Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates über Arzneimittel für seltene Leiden (1) seit dem 12.12.2016 zugelassen.

Somit gilt der medizinische Zusatznutzen von Obeticholsäure laut § 35a Abs. 1 Satz 10 SGB V durch die Zulassung als belegt. Im vorliegenden Dossier wird das Ausmaß des Zusatznutzens für die Teilpopulationen (Obeticholsäure in Kombination mit UDCA, Obeticholsäure als Monotherapie) anhand der Phase-III-Studie POISE gegenüber den in der Studie verwendeten Vergleichsgruppen hergeleitet.

Endpunkte

Als Endpunkte zum Nachweis des Zusatznutzens sollen gemäß der Verfahrensordnung des G-BA (5. Kapitel § 3 Abs. 1 und § 5 Abs. 2) patientenrelevante Endpunkte wie Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit herangezogen werden (19). Für das vorliegende Nutzendossier werden deshalb die Endpunkte Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit definiert.

Mortalität

- Gesamtüberleben

Morbidität

- AP und Bilirubin

- Fibrose
- Osteoporose
- PBC-40
- Pruritus

Gesundheitsbezogene Lebensqualität

- PBC-40

Verträglichkeit

- Gesamtrate der unerwünschten Ereignisse (UE)
- Schwerwiegende UE
- Schwere UE ohne PruritusUE, die zum Therapieabbruch führten

Studientyp

Das Ausmaß des medizinischen Zusatznutzens von Obeticholsäure kann mittels prospektiver RCT am besten bewertet werden. In der Selektion gilt dieser Studientyp als Einschlusskriterium.

Für die Beantwortung der oben formulierten Fragestellung ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien (Randomized Controlled Trials, RCTs) möglich und praktisch durchführbar. Da sie die höchste Ergebnissicherheit aufweisen, sollen nur RCTs in die Nutzenbewertung einbezogen werden. Studien, für die weder eine Volltext-Publikation noch ein Studienbericht oder ein ausführlicher Ergebnisbericht aus einem Studienregister verfügbar sind, werden aus der Bewertung ausgeschlossen, da bei Fehlen umfassender Informationen Verzerrungspotenzial und Ergebnissicherheit der Studie nicht bewertet werden können. Studien, für die noch keine Ergebnisse vorliegen, sowie tierexperimentelle Studien und In-vitro-Untersuchungen bleiben ebenfalls unberücksichtigt.

Studiendauer

Da es sich bei der PBC um eine chronische Erkrankung handelt, wird die Suche auf Studien mit einer Mindestdauer von 24 Wochen eingeschränkt.

Ein- und Ausschlusskriterien

Tabelle 4-1: Übersicht der Ein- und Ausschlusskriterien für Studien zur Bewertung des Ausmaßes des Zusatznutzens von Obeticholsäure

Kriterium	Einschlussgrund	Kürzel	Ausschlussgrund	Kürzel
Studienpopulation	Erwachsene Patienten mit primärer biliärer Cholangitis, die ein inadäquates Ansprechen auf Ursodesoxycholsäure (UDCA) aufweisen (Kombinationstherapie mit Obeticholsäure), oder Erwachsene, die intolerant auf UDCA sind (Monotherapie mit Obeticholsäure).	E1	Erwachsene Patienten ohne primäre biliäre Cholangitis.	A1
Indikation	primäre biliäre Cholangitis (auch unter der Bezeichnung primäre biliäre Zirrhose bekannt)	E2	Patienten, die nicht an einer primären biliären Cholangitis leiden	A2
Endpunkte	Patientenrelevante Endpunkte wie Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit: <u>Mortalität</u> • Gesamtüberleben <u>Morbidität</u> • AP und Bilirubin • Pruritus (UE Pruritus, Pruritus-Fragebogen 5-D und Pruritus-VAS) • Enhanced liver fibrosis (ELF) Score und dessen Einzelkomponenten • Items zu Symptomen des Fragebogens PBC-40 <u>Gesundheitsbezogene Lebensqualität</u> • Items zur Lebensqualität des Fragebogens PBC-40 <u>Verträglichkeit</u> • Gesamtrate der unerwünschten Ereignisse (UE) • Schwerwiegende UE (SUE) • Schwere UE ohne Pruritus • UE, die zum	E3	Keine patientenrelevanten Endpunkte, die zur Bewertung der Mortalität, der Morbidität, der gesundheitsbezogenen Lebensqualität oder der Verträglichkeit herangezogen werden können	A3

	Therapieabbruch führten			
Intervention	Obeticholsäure (OCALIVA®)	E4	Keine Intervention mit Obeticholsäure (OCALIVA®)	A4
Dosierung/Applikation	Kontinuierliche Gabe von Obeticholsäure in der Startdosis von 5 mg einmal täglich. Auf der Grundlage der Vertraglichkeit sollte die Dosis nach 6 Monaten auf 10 mg einmal täglich erhöht werden.	E5	Dosierung bzw. Applikation abweichend vom Dosisschema laut Fachinformation.	A5
Publikationstyp	Vollpublikation oder Studien-/Ergebnisbericht	E6	Studien, für die weder eine Volltext-Publikation noch ein Studienbericht oder ein ausführlicher Ergebnisbericht in einem Studienregister verfügbar sind. Zusätzliche Analysen zu bereits identifizierten Studien	A6
Studiendesign	RCT	E7	Keine RCT, experimentelle Studien, In-vitro-Untersuchung	A7
Studiendauer	Mindestens 24 Wochen	E8	Weniger als 24 Wochen	A8
Vergleichstherapie	Keine Einschränkungen	E9		A9
Veröffentlichungszeitraum	Keine Einschränkungen	E10		A10
Sprache	Keine Einschränkungen	E11		A11
Dublette	Publikation mit Zusatzinformation	E12	Publikation ohne Zusatzinformation	A12
Studienstatus	Abgeschlossene Studien, für die Ergebnisse vorliegen	E13	Noch nicht begonnene, laufende oder abgeschlossene Studien, für die noch keine Ergebnisse vorliegen	A13
Abkürzungen: bzw.: beziehungsweise; mg; Milligramm; RCT: Randomisierte kontrollierte Studie.				

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

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

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

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

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

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

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

Die bibliografische Literaturrecherche erfolgte am 31.10.2016 systematisch in den Datenbanken MEDLINE, EMBASE und der Cochrane Library für die im vorliegenden Dossier formulierte Fragestellung. Für alle Datenbanken wurde die Suchoberfläche des Deutschen

Instituts für medizinische Dokumentation und Information (DIMDI), abrufbar unter <https://www.dimdi.de/static/de/db/index.htm> (Link überprüft am 14.11.2016) verwendet. In Anhang 4-A ist die jeweilige Suchstrategie mit den in der Recherche verwendeten Schlagwörtern, eingeteilt nach Intervention und Indikation dargestellt.

Für jede Datenbank wurde eine eigene adaptierte Suchstrategie verwendet. Die vollständigen Suchstrategien sind im Anhang 4-A dokumentiert. Die Suche wurde dabei so sensitiv wie möglich gehalten. So wurden keinerlei Einschränkungen bezüglich des Zeitpunktes der Veröffentlichung und der Sprache getroffen.

Um die bibliografische Literaturrecherche auf RCTs zu beschränken, wurde in den Datenbanken MEDLINE und EMBASE ein Suchfilter nach Wong verwendet (20). Der Suchfilter dient der Optimierung zwischen Sensitivität und Spezifität und wurde entsprechend der jeweiligen Datenbank adaptiert.

4.2.3.3 Suche in Studienregistern

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

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

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

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

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

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

Die Recherche erfolgte am 17.10.2016 in den Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov) (<http://www.clinicaltrials.gov> – Link überprüft am 14.11.2016), Europäisches Register klinischer Studien (EU-CTR, <http://www.clinicaltrialsregister.eu> – Link überprüft am 14.11.2016), Klinische Prüfungen PharmNet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm> – Link überprüft am 14.11.2016) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/> – Link überprüft am 14.11.2016). Die Suchstrategie und die vollständigen Rechercheergebnisse sind in Anhang 4-B hinterlegt.

4.2.3.4 Selektion relevanter Studien

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

Die Ergebnisse der bibliographischen Literaturrecherche wurden in das Literaturarchivierungsprogramm Endnote X7 importiert und gespeichert. Doppelte Literaturstellen wurden gekennzeichnet und nicht weiter berücksichtigt. Die verbliebenen Literaturstellen wurden anhand ihrer Titel und der Zusammenfassung (Abstract) von zwei Reviewern unabhängig voneinander auf Übereinstimmung mit den für die Recherche definierten Einschlusskriterien überprüft (siehe Abschnitt 4.2.2). Falls die Zusammenfassung einen Hinweis auf Übereinstimmung lieferte, wurden die Studien im Volltext bewertet. Diskrepanzen zwischen den beiden Reviewern wurden im Konsens gelöst.

Die Ergebnisse der Recherche in den Studienregistern wurden auf doppelte Treffer untersucht. Doppelte Studien wurden gekennzeichnet und entfernt. Die Studien wurden gemäß ihren Einträgen in den Registern von zwei Reviewern unabhängig voneinander auf Übereinstimmung mit den für die Recherche definierten Einschlusskriterien überprüft. Diskrepanzen zwischen den beiden Reviewern wurden im Konsens gelöst. Die Ausschlussgründe werden pro Studie in Anhang 4-D angeführt.

4.2.4 Bewertung der Aussagekraft der Nachweise

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

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

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

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

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

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

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

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

Zur Bewertung von Verzerrungspotenzialen wurden Angaben der identifizierten Studien gemäß den Kriterien im Anhang 4-F extrahiert. Zur Extraktion wurden die Studienberichte der herstellereigenen Studien herangezogen.

Bei der Beurteilung der Methodik der Studie wurden die für das Studiendesign gültigen Kriterien herangezogen, um mögliche Verzerrungspotenziale der Studie, die die Aussagekraft der Studienergebnisse einschränken könnten, abzuschätzen.

Die aus der Recherche resultierenden Studien sollten RCTs sein. Um Verzerrungen auf Studienebene hinsichtlich der Selektion der Studienpopulation auszuschließen, müssen in der Studie adäquate Methoden zur Erzeugung der Randomisierungssequenz angewandt und sichergestellt worden sein, damit die Sequenz den Behandlern verborgen blieb.

Um Verzerrungspotenziale auf Endpunktebene zu bewerten, wurde der Umgang mit Drop-outs in den statistischen Analysen bewertet. Eine Analyse der Daten der Intention-to-treat Population (ITT) wurde dabei als Methode mit einem geringen Verzerrungspotenzial angesehen. Für eine RCT mit hoher Evidenz ist in der Regel eine Verblindung sowohl der Patienten als auch der Behandler Voraussetzung. Weiterhin sollte nicht nur die Methodik der Studie, sondern auch die Art und Weise der Darstellung der Ergebnisse kritisch betrachtet werden. Alle im Studienprotokoll festgelegten Zielgrößen müssen vollständig berichtet sein.

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)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

Die selektierten RCT werden gemäß den ITEMS 1 bis 14 des CONSORT Statements beurteilt. Dabei werden sowohl die Checkliste (21) als auch die weiteren Ausführungen zu den einzelnen ITEMS (22) zur besseren Beurteilung der internen und externen Validität der Studien herangezogen. Das Ziel der Studie, die Interventionen, die Zielgröße(n) und die Studienpopulation sind mittels entsprechender Kriterien klar definiert (ITEM 3-6). Ausschlaggebend für eine Bewertung der Studien bezüglich der unverzerrten Aussagekraft sind die angewandten Methoden zur Fallzahlbestimmung, die Art und Weise der Randomisierung, insbesondere die Generierung der Randomisierungssequenz und die verdeckte Zuteilung der Patienten zu den Gruppen, die Verblindung und die adäquate Anwendung von statistischen Analysemethoden (ITEM 7, 8-12). Da letztere Punkte eine zentrale Größe in der Beurteilung des Verzerrungspotenzials der Aussagekraft von Studien haben, werden hierzu zusätzlich die vom IQWiG herausgegebenen Methoden herangezogen (17).

Die Darstellung der Ergebnisse wird zunächst beurteilt nach den Angaben zu der Zahl an geplanten und tatsächlich eingeschlossenen zu untersuchenden Patienten, zu relevanten Basis-Patientencharakteristika und der grafischen Darstellung des Patientenflusses. Studienabbrüche und evtl. Therapiewechsel sind klar und transparent beschrieben (ITEM 13 -16). Die Ergebnisse selbst sind anschaulich dargestellt und beschrieben (17),(22). Zur Beurteilung der Wirksamkeit ist es erforderlich, dass Angaben der statistischen Analysen vorliegen, die über die Angaben eines p-Wertes hinausgehen, wie z. B. KI der Effektschätzer. Weiterhin ist eine Überprüfung auf Vollständigkeit der berichteten Ergebnisse in Bezug auf die geplanten Zielgrößen empfehlenswert. Als weiteres wichtiges Ergebnis sollten alle relevanten unerwünschten Ereignisse (UE) dargestellt sein (ITEM 18). Abschließend wird beurteilt, inwiefern die dargestellten Kriterien eingehalten wurden. Daraus wird das Verzerrungspotenzial für die berichteten Endpunkte und für die Studie abgeleitet.

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

² 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.

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

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

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

Patientencharakteristika

Die Patientenpopulation umfasst erwachsene Patienten mit primärer biliärer Cholangitis mit nicht ausreichendem Ansprechen auf UDCA oder bei Unverträglichkeit von UDCA.

Die nachfolgende Tabelle 4-2 gibt einen Überblick über die demographischen Parameter und Baseline-Patientencharakteristika der POISE-Studie.

Tabelle 4-2: Übersicht über die erhobenen demographischen Parameter und Baseline-Patientencharakteristika der POISE-Studie

Kategorie	Ausprägung
Geschlecht	Männlich, weiblich
Alter	Jahre
Altersklasse	≤65, >65 Jahre
Region	Europa, Nordamerika, Australien
Ethnische Gruppe	Kaukasier, Nicht-Kaukasier
Körpergewicht	Kilogramm (kg)
Body Mass Index (BMI)	kg/m ²
BMI-Klasse	<30 kg/m ² , ≥30 kg/m ²
Leber-Biopsie vor Studienbeginn	Ja, Nein
UDCA Behandlung zu Baseline	Ja, Nein
Pruritus in der Anamnese	Ja, Nein
Intensität des letzten Pruritus-Ereignisses mit Pruritus in der Anamnese	Mild, Moderat, Schwer, unbekannt
Pruritus zu Baseline	Ja (Mild, Moderat, Schwer), Nein
Fatigue in der Anamnese	Ja, Nein
Gesamt Intensität der PBC-bedingten Fatigue	Mild, Moderat, Schwer
Alter bei PBC Diagnose	Jahre
Altersklasse bei PBC Diagnose	<50, ≥50 Jahre
Dauer der PBC	Jahre
Dauer der PBC-Klasse	≤7,5, >7,5 Jahre
AP zu Baseline	U/L
Gesamt Bilirubin zu Baseline	µmol/L
Konjugiertes Bilirubin zu Baseline	µmol/L
GGT zu Baseline	U/L
ALT zu Baseline	U/L
AST zu Baseline	U/L
Albumin zu Baseline	g/L
INR zu Baseline	Verhältnis
Abkürzungen: kg = Kilogramm, U = Einheit, L = Liter, µmol = mikromol, g = Gramm, AP = Alkalische Phosphatase, GGT = Gamma-Glutamyl Transferase, ALT = Alanin Aminotransferase, AST = Aspartat Aminotransferase, INR = Internationales normalisiertes Verhältnis	

Patientenrelevante Endpunkte

Im vorliegenden Dossier zur Nutzenbewertung von Obeticholsäure werden folgende patientenrelevante Endpunkte ausgewertet:

Mortalität

- Gesamtüberleben

Morbidität

- AP und Bilirubin
- Fibrose
- Osteoporose
- PBC-40
- Pruritus

Gesundheitsbezogene Lebensqualität

- PBC-40

Verträglichkeit

- Gesamtrate der unerwünschten Ereignisse (UE)
- Schwerwiegende UE
- Schwere UE ohne PruritusUE, die zum Therapieabbruch führten

Die Patientenrelevanz der oben genannten Endpunkte begründet sich dabei wie folgt:

Das **Gesamtüberleben** wird in RCTs außerhalb der Onkologie als Ereignisrate im Studienzeitraum im Rahmen der Sicherheitsendpunkte erhoben. Die Gesamtmortalität wird im Rahmen des AMNOG als herausragendes Therapieziel angesehen.

In der Endpunktdimension **Morbidität** werden folgende Endpunkte dargestellt.

Die Endpunkte AP und Bilirubin werden in diesem Dossier in verschiedenen Operationalisierungen dargestellt, da sich der primäre Wirksamkeitsendpunkt aus diesen zusammensetzt. AP und Bilirubin sind als Surrogate für die in klinischen Studien nur schwer erhebbaren Endpunkte Mortalität und Lebertransplantation als biochemische Kenngrößen empfohlen. (4) Im Weiteren wurden beide Kenngrößen in großen Registern validiert („UK-PBC“ (5), (6) sowie „GLOBE“ (7),(8),(9). FDA (10) und EMA (11) bestätigen beide die

Validität dieser biochemischen Kenngrößen für PBC. Eine ausführliche Darstellung der Validität der AP und des Bilirubin findet sich in Abschnitt 4.5.4.

Tabelle 4-3: Wichtige Response-Kriterien für UDCA in der Literatur

Name der Studie	Endpunkt Kriterium	Zeitpunkt für den Endpunkt	Validierungs-Endpunkt	Patientenzahl N Etablierungs-Kohorte (Nicht-Ansprechen n;n%)	Kommentar	Quelle
Barcelona*	>40% Reduktion oder Normalisierung der AP	1 Jahr	Tod oder Lebertransplantation über im Mittel 6,8 Jahre	N=192 (75; 39%)	Patienten unter Therapie mit UDCA; Überleben für Ansprecher wie Normalbevölkerung	(23)
Mayo I	AP < 2x ULN (und/oder: Mayo Risiko Score < 4,5)	6 Monate	Kurzzeit-Response nach 2 Jahren; (Langzeit für Mayo Risiko Score: Tod und Lebertransplantation)	N=180 (4/36, 11%)	Patienten in RCT mit UDCA Therapie; Langzeitergebnisse nicht nach Ansprechern dargestellt	(24)
Mayo II *	AP ≤ 1,67x ULN und Bilirubin ≤ 1mg/dL	1 Jahr	Verschiedene hepatische Komplikationen wie Varizen, Aszites und Enzephalopathie nach 36 Monaten	N=73 (nicht berichtet)	Auch Darstellung verschiedener AP Kriterien, besonders AP<2x ULN und Bilirubin≤1mg/dL	(25)
Paris I *	AP ≤ 3x ULN und AST ≤ 2x ULN und Bilirubin gesamt ≤ 1mg/dL	1 Jahr	Tod oder Lebertransplantation über im Mittel 6,1 Jahre	N=292 (113; 38%)	Patienten unter Therapie mit UDCA	(26)
Paris II *	AP ≤ 1,5x ULN und AST ≤ 1,5x und Bilirubin gesamt ≤ 1mg/dL	1 Jahr	Tod oder Lebertransplantation oder Komplikationen der Zirrhose (Aszites, Varizenblutung, Enzephalopathie, HCC) oder histologisch gesicherte Zirrhoseentwicklung über im Mittel 7,1 Jahre	N=292 (86/165; 52%)	Selbe Patientenkohorte wie Paris I	(27)
Rotterdam *	Normales Bilirubin (≤ULN) und normales Albumin wenn ein oder beide Parameter abnormal waren; Normalisierung von Bilirubin oder Albumin wenn beide Parameter abnormal waren	1 Jahr	Tod oder Lebertransplantation über im Mittel 9,7 Jahre	N=375 (75/311; 24%)	Kriterium fokussiert auf in späteren Stadien auftretende hepatische Dysfunktion; auch Darstellung der Paris I und Barcelona Kriterien	(28)
Toronto I bzw. II *	I: AP ≤ 1,67 ULN und Bilirubin ≤ 1 mg/dL II: AP ≤ 1,76 ULN	2 Jahre	Histologische Progression über im Mittel 9,4 Jahre I: ≥1 Stufe Progression II: ≥2 Stufen Progression	N=69 (25/64; 39%)	Patienten 2 Jahre unter Therapie mit UDCA	(29)

* in Modul 4 verwendete Response-Kriterien; ULN: oberer Normwert

Der Endpunkt Pruritus wird anhand des unerwünschten Ereignisses Pruritus und anhand des Fragebogens 5-D sowie der Pruritus-VAS dargestellt. Der Pruritus-Fragebogen 5-D ist ein Instrument zur Messung des Pruritus. Er wurde an 234 Patienten mit chronischem Pruritus validiert. Darunter waren 63 Patienten mit einer Lebererkrankung. Der 5-D wurde dabei als

verlässliches und multidimensionales Messinstrument für Pruritus etabliert und validiert (12). Die Pruritus Visuelle Analogskala (Pruritus-VAS) wurde in einer Studie von Phan et al. an 471 Patienten mit chronischem Pruritus validiert (13). Sie wurde als verlässliches und validiertes Messinstrument für den Pruritus etabliert.

Der Endpunkt Fibrose wird anhand des ELF-Scores und der transienten Elastografie dargestellt. Der ELF-Score wurde in der Studie von Guéchet et al. in 512 Patienten mit chronischer Hepatitis C gegenüber anderen als den für den ELF-Score verwendeten Reagenzien validiert (14). Die Autoren bestätigen mit der Studie die Aussagekraft des Tests als Index zur Prädiktion einer Leber-Fibrose oder Zirrhose bei Patienten mit chronischer Hepatitis C. Daher wird dieser Test als valides Surrogat für die patientenrelevanten Endpunkte Zirrhose und Fibrose angesehen.

Die Fibrosierung der Leber wurde zudem mittels transientser Elastografie (Fibroscan®) in den Zentren erhoben, in denen das Medical Device (Echosens, Paris, Frankreich) verfügbar war. Der Fibroscan ist ein validiertes, nicht-invasives Verfahren (15) im Rahmen der TE zur Bestimmung der Leberfibrose. Zusätzlich zur metrischen Auswertung wurde die Präsenz von Fibrose mit einer Grenze für die Lebersteifigkeit $\geq 17,3$ kPa (30) als Ansprecheranalyse auf Basis des Fibroscan® durchgeführt.

Die Osteoporose wird anhand des DEXA-Scans dargestellt, eine nicht-invasive Methode zur Bestimmung der Osteoporose bei postmenopausalen Frauen.

Der Fragebogen PBC-40 enthält Domänen zur **Morbidität** und **gesundheitsbezogenen Lebensqualität** und wurde in einer Arbeit von Jacoby et al. 2005 vorgestellt (16). Der Fragebogen besteht aus 40 Fragen in 6 Domänen und wurde an 900 PBC-Patienten gegen andere Instrumente zur Messung der Lebensqualität validiert. Er ist ein etablierter Fragebogen in der PBC und stellt damit einen validen und patientenrelevanten Endpunkt dar.

Im vorliegenden Nutzendossier wird die Verträglichkeit anhand der Endpunkte UE, SUE, schwere UE ohne Pruritus und UE, die zum Therapieabbruch führten, dargestellt.

UE werden grundsätzlich als patientenrelevant betrachtet, entsprechend dem IQWiG-Methodenpapier als „ein Aspekt der therapiebedingten Morbidität“ (17). Pruritus wird als UE im Rahmen seiner Sonderstellung als PBC-Grundsymptom in der Endpunktdimension Morbidität separat dargestellt. Um eine Doppelzählung der des Symptoms Pruritus zu vermeiden, wird als UE Endpunkt "schwere UE ohne Pruritus" aufgeführt.

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 Effekten⁴ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁵ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität⁶ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Da nur eine Studie in die Nutzenbewertung eingeschlossen wurde, konnten keine Metaanalysen durchgeführt werden.

⁴ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

⁵ 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.

⁶ 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.

Da keine Meta-Analysen durchgeführt wurden, entfällt die Durchführung von Sensitivitätsanalysen zur Untersuchung der Stabilität der Ergebnisse bezüglich methodischer Faktoren im Rahmen meta-analytischer Untersuchungen. Statt dessen wurden Sensitivitätsanalysen zur Überprüfung der Robustheit der Ergebnisse auf Ebene der Einzelstudie POISE durchgeführt.

Pruritus stellt das wichtigste unerwünschte Ereignis im Rahmen der Behandlung mit Obeticholsäure dar. Pruritus wurde deshalb im Rahmen des klinischen Studienreports (CSR) der POISE-Studie ausführlich analysiert und im Modul 4 als separater Endpunkt analog den Analysen im CSR dargestellt. Um den Effekt des Symptoms Pruritus für die Bewertung der Sicherheit darzustellen, wurde im Dossier für die schweren UE eine Sensitivitätsanalyse durchgeführt, in der das Ereignis Pruritus für die Bestimmung der Anzahl von Patienten mit schweren UE unberücksichtigt blieb.

Weitere Sensitivitätsanalysen wurden nicht durchgeführt.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

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

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

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

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

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

Gemäß den Vorgaben des G-BA betreffend Format, Gliederung sowie der einzureichenden Unterlagen im Rahmen der Erstellung und Einreichung eines Dossiers zur Nutzenbewertung sind in Modul 4 eines Nutzendossiers für ein Orphan Drug die Abschnitte 4.4.2 und 4.4.3 vom pharmazeutischen Unternehmer auszufüllen (siehe Abschnitt 4.2.1). Im vorliegenden Nutzendossier werden keine Subgruppenanalysen dargestellt und entsprechend nicht zum Nachweis des Ausmaßes des Zusatznutzens herangezogen, da diese Angaben gemäß 1.3 der Anlage II.1 zum 5. Kapitel der VerfO nicht zu machen sind.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethoden dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁷. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹.

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vorliegen.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

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

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

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

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

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

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

¹¹ B. Schöttker, D. Lüthmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

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

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

Indirekte Vergleiche wurden nicht durchgeführt.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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

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

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

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

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
POISE (Studie 747-301)	ja	ja	abgeschlossen	März 2012 bis Dezember 2013	OCA (10 mg)+UDCA oder OCA (10 mg), OCA (Titration)+UDCA oder OCA (Titration), UDCA+Placebo oder Placebo
Studie 747-201	nein	ja	abgeschlossen	Januar 2008 bis September 2010	OCA 10 mg, OCA 50 mg, Placebo
Studie 747-202	nein	ja	abgeschlossen	November 2007 bis September 2009	OCA 10 mg + UDCA, OCA 25 mg + UDCA, OCA 50 mg + UDCA, UDCA+Placebo
COBALT (Studie 747-302)	nein	ja	laufend	Dezember 2014 bis April 2023 (geschätzt)	OCA (Titration) Placebo
Studie 747-205	nein	ja	abgeschlossen	Dezember 2013 bis August 2014	OCA (10 mg)
Studie 747-401	nein	ja	in Planung	in Planung	Liegt noch nicht vor

Geben Sie an, welchen Stand die Information in

Tabelle 4-4 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.

Stand der

Tabelle 4-4 „Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel“ ist der 31.10.2016.

Geben Sie in der nachfolgenden Tabelle an, welche der in

Tabelle 4-4 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-5: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
Studie 747-201	Dosierung bzw. Applikation abweichend vom Dosisschema laut Fachinformation (Obeticholsäure ohne Titrationsphase mit 5 mg für 6 Monate); Studiendauer <24 Wochen
Studie 747-202	Dosierung bzw. Applikation abweichend vom Dosisschema laut Fachinformation (Obeticholsäure ohne Titrationsphase mit 5 mg für 6 Monate); Studiendauer <24 Wochen
COBALT (Studie 747-302)	Studie läuft noch bis 2023 (geschätzt), es liegen noch keine Ergebnisse vor.
Studie 747-205	Einarmige Studie, keine Vergleichsgruppe. Dosierung bzw. Applikation abweichend vom Dosisschema laut Fachinformation (Obeticholsäure ohne Titrationsphase mit 5 mg für 6 Monate); Studiendauer <24 Wochen
Studie 747-401	Studie ist in der Planungsphase, es liegen noch keine Ergebnisse vor.

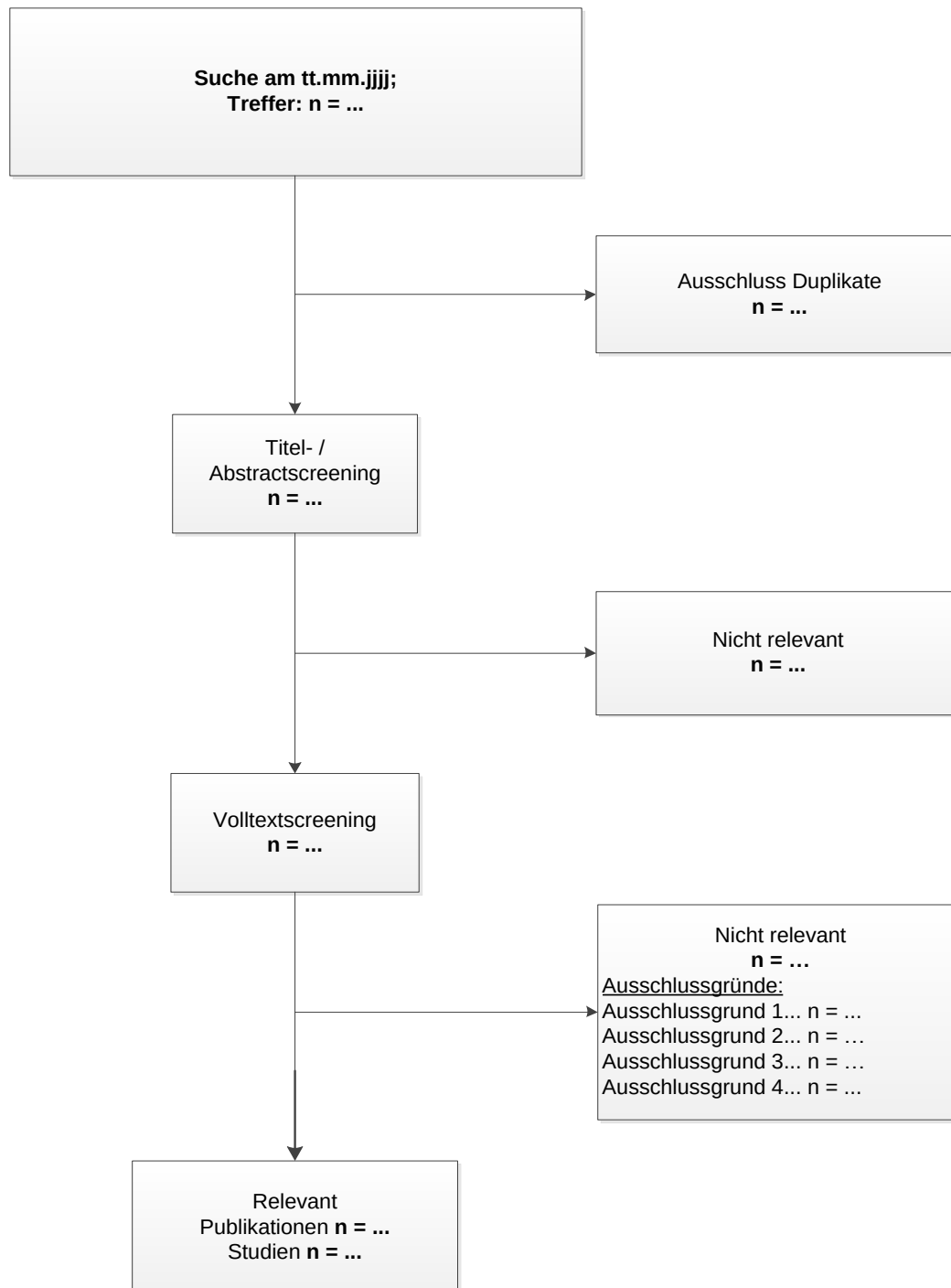
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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



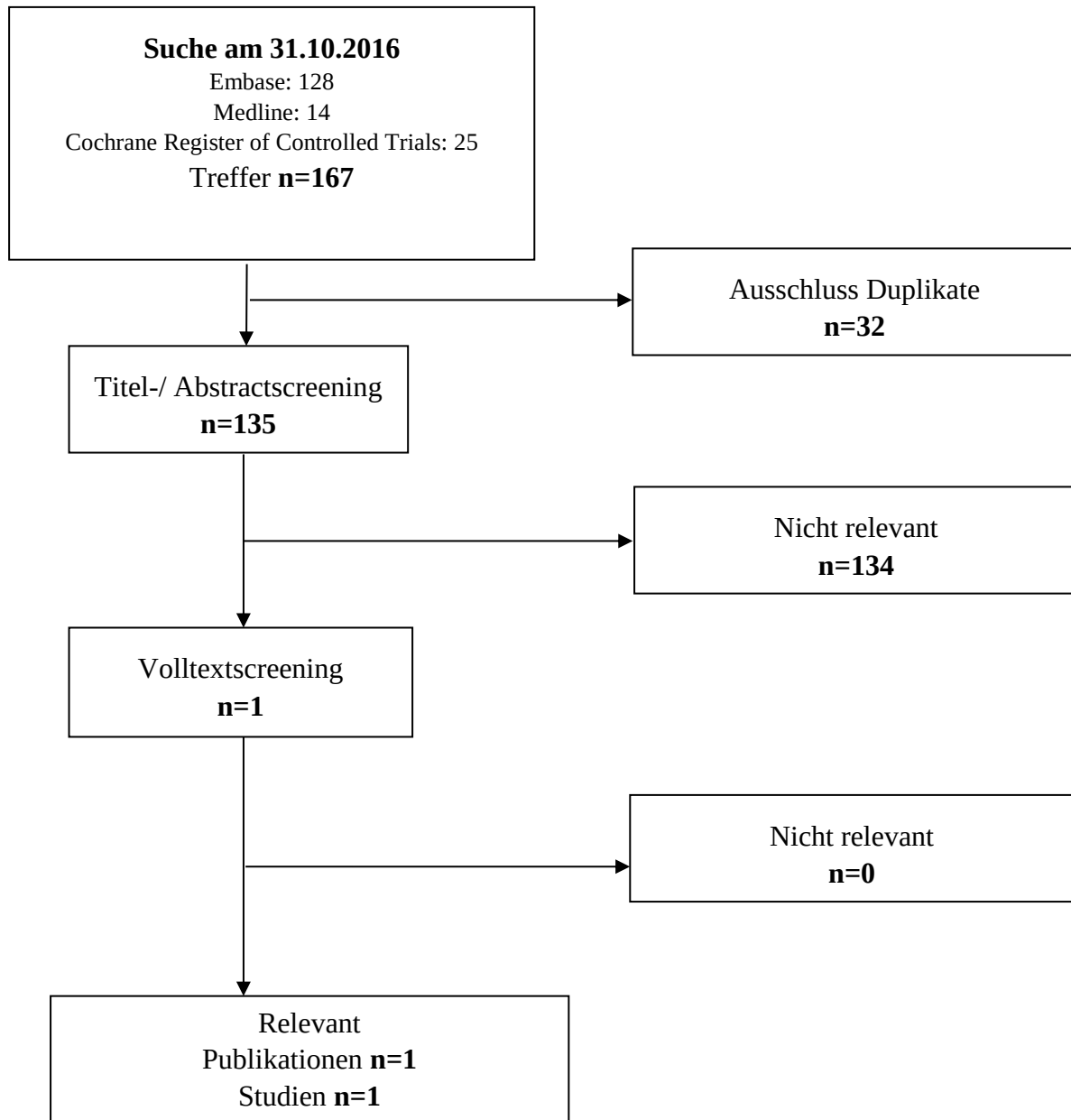


Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Im Rahmen der bibliographischen Literaturrecherche wurden folgende Datenbanken durchsucht: MEDLINE, EMBASE, die Cochrane Database of Systematic Reviews sowie das Cochrane Central Register of Clinical Trials. Die Suche wurde am 31.10.2016 durchgeführt. Die Suchstrategien sind in Anhang 4-A dargestellt. Insgesamt wurden 167 Treffer identifiziert (MEDLINE: 14, EMBASE: 128, Cochrane Library: 25). Nach Ausschluss von Duplikaten verblieben 135 Treffer. Die Auswahl relevanter Studien aus diesem Pool erfolgte durch zwei unabhängige Reviewer anhand der in Abschnitt 4.2.2 beschriebenen Kriterien und führte zum Ausschluss von 134 Treffern. Der verbleibende eine Eintrag dokumentiert die Ergebnisse der Studie POISE, die bereits in Abschnitt 4.3.1.1.1 aufgelistet wurde und der vorliegenden

Nutzenbewertung zu Grunde liegt. Der Ablauf der Recherche ist im Flussdiagramm in Abbildung 4-2 dokumentiert.

4.3.1.1.3 Studien aus der Suche in Studienregistern

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

Tabelle 4-4) 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 4-6: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein)	Status (abgeschlossen / abgebrochen / laufend)
Phase 3 Study of Obeticholic Acid in Patients With Primary Biliary Cirrhosis (POISE)	Clinicaltrials.gov (31), EUCTR (32), ICTRP (33), Pharm Net Bund (34)	ja (747-301)	ja	abgeschlossen
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

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 erfolgte in den Studienregistern ClinicalTrials.gov, EUCTR, ICTRP und Pharm Net Bund und wurde am 17.10.2016 durchgeführt. Die dabei verwendeten Suchstrategien sind in Anhang 4-B dargestellt. Nach Ausschluss der nicht relevanten Studien verblieb nur die POISE-Studie, die bereits in Abschnitt 4.3.1.1.1 aufgelistet wurde und der vorliegenden Nutzenbewertung zu Grunde liegt. Eine Liste der ausgeschlossenen Studien findet sich in Anhang 4-D.

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

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

Tabelle 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebokontrolliert						
POISE-Studie 747-301	ja	ja	nein	ja [(35)]	NCT01473524 Clinicaltrials.gov (31) ICTRP (33) EUCTR (32) Pharm Net Bund (34)	ja [(36)]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
keine						
a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden. b: Studie, für die der Unternehmer Sponsor war. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

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

4.3.1.2.1 Studiendesign und Studienpopulationen

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

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

Tabelle 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
POISE (Studie 747-301)	RCT, doppelblind, parallel, 1:1:1 Randomisierung; Patienten im Kontrollarm ^a konnten in einer Extensionsphase weiterbeobachtet werden und dort optional mit OCA offen behandelt werden, Weiterbehandlung der Patienten der OCA- Arme	Erwachsene mit primärer biliärer Cholangitis	OCA 10 mg (n= 73) OCA Titration (n= 71, davon 70 mit mindestens einer Dosis Obeticholsäure) Kontrollarm ^a (n= 73)	Behandlung: 12 Monate Nachbeobachtung: 4 Wochen nach Ende der Behandlung für die Erhebung der Sicherheit Extensionsphase: Möglichkeit des Cross-over zur offenen Behandlung mit OCA für Patienten des Placeboarms und Weiterbehandlung der Patienten der OCA-Arme	Europa (Deutschland, Vereinigtes Königreich, Polen, Niederlande, Italien, Spanien, Österreich, Belgien, Frankreich, Schweden), Nordamerika (USA, Kanada), Australien 3/2012 – 12/2013	<u>Primärer Endpunkt:</u> Kombinierter Endpunkt als Ansprecher-Endpunkt mit den Komponenten AP < 1,67x Obere Normgrenze (ULN), Gesamt Bilirubin ≤ ULN und AP-Reduktion um ≥ 15% von Baseline zum Monat 12 <u>Sekundäre Endpunkte:</u> AP, Bilirubin, Biochemische Kriterien zum Behandlungserfolg, erkrankungsspezifische Symptome (PBC-40), Sicherheitsparameter
a) Der Kontrollarm umfaßte für Patienten mit Kombinationstherapie UDCA + Placebo, für Patienten mit Monotherapie Placebo						

Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Therapie	Kontrollarm	OCA Titration	OCA 10 mg	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run- in-Phase etc.</i>
POISE (Studie 747-301)	Kombina- tions- therapie	UDCA Basistherapie plus Placebo Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück	UDCA Basistherapie plus OCA 5 mg Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück Nach Monat 6 Möglichkeit der Titration zu 10 mg einmal täglich	UDCA Basistherapie plus OCA 10 mg Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück	Patienten im Kontrollarm konnten in einer Extensionsphase weiterbeobachtet werden und dort optional mit OCA offen behandelt werden, Weiterbehandlung der Patienten der OCA-Arme
	Mono- therapie	Placebo Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück	OCA 5 mg Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück Nach Monat 6 Möglichkeit der Titration zu 10 mg einmal täglich	OCA 10 mg Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück	Patienten im Kontrollarm konnten in einer Extensionsphase weiterbeobachtet werden und dort optional mit OCA offen behandelt werden, Weiterbehandlung der Patienten der OCA-Arme

Die POISE-Studie deckt zwei Therapieschemata ab: Zum einen die Kombination von Obeticholsäure mit UDCA, zum anderen Obeticholsäure als Monotherapie im Vergleich zum beobachtenden Abwarten (Placebo). Bei Patienten, die noch mit UDCA behandelt werden konnten, wurde Obeticholsäure+UDCA gegen UDCA+Placebo verglichen. Für Patienten, die für eine Therapie mit UDCA nicht oder nicht mehr zur Verfügung standen, wurde Obeticholsäure als Monotherapie gegen Placebo als Platzhalter für ein beobachtendes Abwarten verglichen.

Die Indikation für Obeticholsäure bei PBC sieht eine Titration der Behandlung vor. Somit ist der Behandlungsarm "Obeticholsäure 10 mg" mit einer festen Dosis ohne Titration nicht zulassungskonform und wird im Folgenden nicht weiter dargestellt.

In den folgenden Tabellen sind die Demographie und Baseline-Charakteristika der beiden Teilpopulationen der Patienten mit Kombinationstherapie und Monotherapie für den relevanten Behandlungsarm "Obeticholsäure Titration" sowie den Kontrollarm dargestellt. Hier zeigen sich keine relevanten Unterschiede zwischen den beiden Behandlungsgruppen.

Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel - Patienten mit Kombinationstherapie

	OCA+UDCA ^a (N=65)	UDCA+Placebo (N=68)
Geschlecht, n (%)		
Männlich	5 (8%)	5 (7%)
Weiblich	60 (92%)	63 (93%)
Alter (Jahre)		
n	65	68
MW (SD)	55,2 (9,42)	55,2 (10,28)
Median	54	55,5
Minimum; Maximum	29; 81	34; 79
Altersklasse, n (%)		
<65 Jahre	58 (89%)	56 (82%)
≥65 Jahre	7 (11%)	12 (18%)
Region, n (%)		
Europa	42 (65%)	44 (65%)
Nordamerika	18 (28%)	21 (31%)
Australien	5 (8%)	3 (4%)
Ethnische Gruppe, n (%)		
Kaukasier	63 (97%)	61 (90%)
Nicht-Kaukasier	2 (3%)	7 (10%)
Körpergewicht (kg)		
n	65	68
MW (SD)	68,74 (13,253)	70,36 (13,577)
Median	66,15	71,61
Minimum; Maximum	46,7; 101,8	41,0; 106,0
Body Mass Index (BMI) (kg/m²)		
n	65	68
MW (SD)	25,87 (5,028)	26,19 (4,517)
Median	24,76	25,99
Minimum; Maximum	17,7; 40,7	16,4; 37,6
BMI-Klasse, n (%)		
<30 kg/m ²	53 (82%)	53 (78%)
≥30 kg/m ²	12 (18%)	15 (22%)
Leber Biopsie vor Studienbeginn, n (%)		
Ja	11 (17%)	7 (10%)
Nein	54 (83%)	61 (90%)
UDCA Behandlung zu Baseline, n (%)		
Ja	65 (100%)	68 (100%)
Nein	0	0
Pruritus in der Anamnese, n (%)		
Ja	42 (65%)	45 (66%)
Nein	23 (35%)	23 (34%)
Intensität des letzten Pruritus-Ereignisses mit Pruritus in der Anamnese, n (%)		
Mild	28 (67%)	29 (64%)
Moderat	11 (26%)	14 (31%)
Schwer	2 (5%)	1 (2%)
unbekannt	1 (2%)	1 (2%)
Pruritus zu Baseline, n (%)		
Ja	35 (54%)	45 (66%)
Nein	30 (46%)	23 (34%)
Fatigue in der Anamnese, n (%)		
Ja	33 (51%)	45 (66%)
Nein	32 (49%)	23 (34%)
Gesamt-Intensität der PBC-bedingten Fatigue, n (%)		
Mild	14 (22%)	26 (38%)
Moderat	14 (22%)	14 (21%)
Schwer	5 (8%)	3 (4%)
Alter bei PBC Diagnose (Jahre)		

	OCA+UDCA ^a (N=65)	UDCA+Placebo (N=68)
n	65	68
MW (SD)	46,7 (10,47)	47,1 (9,57)
Median	47	47
Minimum; Maximum	25; 77	31; 74
Altersklasse bei PBC Diagnose, n (%)		
<50 Jahre	37 (57%)	43 (63%)
≥50 Jahre	28 (43%)	25 (37%)
Dauer der PBC (Jahre)		
n	65	68
MW (SD)	8,6 (5,76)	8,3 (5,35)
Median	7,3	7,4
Minimum; Maximum	0,8; 27,0	0,9; 21,8
Dauer der PBC – Klassen, n (%)		
≤7,5 Jahre	33 (51%)	37 (54%)
>7,5 Jahre	32 (49%)	31 (46%)
AP zu Baseline (U/L)		
n	65	68
MW (SD)	313,57 (94,843)	324,60 (114,713)
Median	279,47	310,18
Minimum; Maximum	186,8; 566,3	143,8; 745,9
Totales Bilirubin zu Baseline (µmol/L)		
n	65	68
MW (SD)	10,296 (5,699)	11,695 (7,438)
Median	8,550	8,935
Minimum; Maximum	2,05; 36,42	2,31; 37,39
Konjugiertes Bilirubin zu Baseline (µmol/L)		
n	65	68
MW (SD)	4,522 (4,670)	5,448 (6,417)
Median	3,335	3,206
Minimum; Maximum	1,54; 35,23	1,54; 33,94
GGT zu Baseline (U/L)		
n	65	68
MW (SD)	244,01 (165,517)	286,25 (436,022)
Median	181,95	187,8
Minimum; Maximum	43,1; 727,8	45,2; 3263,5
ALT zu Baseline (U/L)		
n	65	68
MW (SD)	61,18 (39,425)	55,37 (30,872)
Median	52,15	47,95
Minimum; Maximum	17,1; 245,5	17,1; 174,9
AST zu Baseline		
n	65	68
MW (SD)	51,77 (26,070)	47,89 (22,849)
Median	47,9	42,85
Minimum; Maximum	20,1; 173,4	20,6; 156,5
Albumin zu Baseline (g/L)		
n	65	68
MW (SD)	43,23 (2,914)	42,68 (3,155)
Median	43	43
Minimum; Maximum	34,0; 50,5	32,5; 48,0
INR zu Baseline		
n	64	67
MW (SD)	1,07 (0,443)	0,98 (0,080)
Median	1	1
Minimum; Maximum	0,9; 3,5	0,7; 1,2
a) Behandlungsarm "Obeticholsäure Titration"		
kg = Kilogramm, U = Einheit, L = Liter, µmol = mikromol, g = Gramm,		
AP = Alkalische Phosphatase, INR = Internationales normalisiertes Verhältnis		

Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel - Patienten mit Monotherapie

	OCA ^a (N=5)	Placebo (N=5)
Geschlecht, n (%)		
Männlich	0	0
Weiblich	5 (100%)	5 (100%)
Alter (Jahre)		
n	5	5
MW (SD)	63,2 (20,44)	60,2 (3,49)
Median	66	60
Minimum; Maximum	30; 83	57; 66
Altersklasse, n (%)		
<65 Jahre	2 (40%)	4 (80%)
≥ 65 Jahre	3 (60%)	1 (20%)
Region, n (%)		
Europa	3 (60%)	5 (100%)
Nordamerika	2 (40%)	0
Australien	0	0
Ethnische Gruppe, n (%)		
Kaukasier	4 (80%)	5 (100%)
Nicht-Kaukasier	1 (20%)	0
Körpergewicht (kg)		
n	5	5
MW (SD)	60,44 (9,291)	67,40 (9,381)
Median	56,25	69,55
Minimum; Maximum	52,0; 74,6	53,4; 79,0
Body Mass Index (BMI) (kg/m²)		
n	5	5
MW (SD)	24,23 (2,856)	25,59 (2,515)
Median	23,43	25,26
Minimum; Maximum	21,3; 27,7	22,8; 29,4
BMI-Klasse, n (%)		
<30 kg/m ²	5 (100%)	5 (100%)
≥30 kg/m ²	0	0
Leber Biopsie vor Studienbeginn, n (%)		
Ja	2 (40%)	0
Nein	3 (60%)	5 (100%)
UDCA Behandlung zu Baseline, n (%)		
Ja	0	0
Nein	5 (100%)	5 (100%)
Pruritus in der Anamnese, n (%)		
Ja	3 (60%)	2 (40%)
Nein	2 (40%)	3 (60%)
Intensität des letzten Pruritus-Ereignisses mit Pruritus in der Anamnese, n (%)		
Mild	1 (33%)	2 (100%)
Moderat	2 (67%)	0
Schwer	0	0
unbekannt	0	0
Pruritus zu Baseline, n (%)		
Ja	2 (40%)	2 (40%)
Nein	3 (60%)	3 (60%)
Fatigue in der Anamnese, n (%)		
Ja	5 (100%)	4 (80%)
Nein	0	1 (20%)
Gesamt-Intensität der PBC-bedingten Fatigue, n (%)		
Mild	3 (60%)	2 (40%)
Moderat	2 (40%)	2 (40%)
Schwer	0	0
Alter bei PBC Diagnose (Jahre)		

	OCA^a (N=5)	Placebo (N=5)
n	5	5
MW (SD)	58,6 (20,61)	50,6 (4,39)
Median	54	50
Minimum; Maximum	30; 82	45; 55
Altersklasse bei PBC Diagnose, n (%)		
<50 Jahre	1 (20%)	2 (40%)
≥50 Jahre	4 (80%)	3 (60%)
Dauer der PBC (Jahre)		
n	5	5
MW (SD)	4,6 (5,45)	9,5 (6,45)
Median	1,3	8,3
Minimum; Maximum	0,3; 11,6	4,5; 20,6
Dauer der PBC – Klassen, n (%)		
≤7,5 Jahre	3 (60%)	2 (40%)
>7,5 Jahre	2 (40%)	3 (60%)
AP zu Baseline (U/L)		
n	5	5
MW (SD)	485,72 (233,981)	366,73 (124,998)
Median	539,4	362,85
Minimum; Maximum	257,2; 811,0	232,0; 557,3
Totales Bilirubin zu Baseline (µmol/L)		
n	5	5
MW (SD)	8,841 (3,029)	12,597 (3,539)
Median	9,148	12,569
Minimum; Maximum	5,73; 12,65	7,44; 17,36
Konjugiertes Bilirubin zu Baseline (µmol/L)		
n	5	5
MW (SD)	2,787 (1,135)	5,763 (2,265)
Median	2,48	4,959
Minimum; Maximum	1,54; 4,53	3,08; 8,64
GGT zu Baseline (U/L)		
n	5	5
MW (SD)	367,39 (159,051)	626,94 (561,231)
Median	298,4	358,45
Minimum; Maximum	231,6; 606,1	285,1; 1619,4
ALT zu Baseline (U/L)		
n	5	5
MW (SD)	66,51 (37,210)	64,39 (21,896)
Median	52	63,45
Minimum; Maximum	32,8; 123,9	45,9; 100,2
AST zu Baseline		
n	5	5
MW (SD)	58,49 (10,244)	61,03 (11,171)
Median	54,65	62,9
Minimum; Maximum	48,9; 75,6	44,0; 71,8
Albumin zu Baseline (g/L)		
n	5	5
MW (SD)	40,69 (4,575)	44,00 (2,850)
Median	41,45	43,5
Minimum; Maximum	33,0; 45,0	40,0; 47,0
INR zu Baseline		
n	5	5
MW (SD)	0,98 (0,045)	0,97 (0,104)
Median	0,95	0,95
Minimum; Maximum	1,0; 1,1	0,9; 1,1
a) Behandlungsarm "OCA Titration"		
kg = Kilogramm, U = Einheit, L = Liter, µmol = mikromol, g = Gramm,		
AP = Alkalische Phosphatase, INR = Internationales normalisiertes Verhältnis		

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

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

Für die Bewertung des Ausmaßes des Zusatznutzens von Obeticholsäure wurde die POISE-Studie herangezogen. Diese wird im Folgenden hinsichtlich Studiendesign, Randomisierung, Endpunkten, statistischen Methoden, Patientencharakteristika und Übertragbarkeit auf den deutschen Versorgungskontext beschrieben.

Studiendesign

Die POISE-Studie ist eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie, die Obeticholsäure zur Behandlung von Patienten mit PBC gegen eine Kontrollgruppe vergleicht. Dabei deckt die POISE-Studie zwei Therapieschemata ab: Zum einen die Kombination von Obeticholsäure mit UDCA, zum anderen Obeticholsäure als Monotherapie im Vergleich zum beobachtenden Abwarten. In Patienten, die noch mit UDCA behandelt werden konnten, wurde Obeticholsäure+UDCA gegen UDCA+Placebo verglichen. Bei Patienten, die für eine Therapie mit UDCA nicht oder nicht mehr zur Verfügung standen, wurde Obeticholsäure als Monotherapie gegen Placebo als Platzhalter für ein beobachtendes Abwarten verglichen. Die Studie selbst unterteilt sich in zwei Studienabschnitte: 1. Randomisierungsphase, 2. Extensionsphase. Vor der Randomisierungsphase wurde zunächst die Eignung der Patienten für den Einschluss in die Studie anhand der im Studienprotokoll präspezifizierten Ein- und Ausschlusskriterien festgestellt (Screening-Zeitraum). Im Anschluss wurden die aktuell bestehenden Krankheitscharakteristika erhoben (Baseline-Zeitraum) und erneut die Eignung zur Studienteilnahme geprüft. In der Randomisierungsphase wurden die Patienten mittels verblindeter 1:1:1-Randomisierung einem der drei Studienarme, dem Obeticholsäure 10 mg-Arm, dem Obeticholsäure Titrationsarm bzw. dem Placeboarm zugeteilt. Die gesamte Randomisierungsphase begann mit der verblindeten, zufälligen Zuteilung des ersten Patienten und endete zum Monat 12 der Behandlungsphase. Die einzelnen Patienten verblieben bis zu diesem Zeitpunkt verblindet in der Randomisierungsphase bis entweder Monat 12 erreicht war, oder aber ein unerwünschtes Ereignis eintrat, das zum Studienabbruch führte bzw. wenn die Einwilligung zur Teilnahme widerrufen wurde. Nach Monat 12 konnten die Patienten in die Extensionsphase aufgenommen werden. Die Extensionsphase besteht aus einer open-label Behandlungsphase mit Obeticholsäure in Form einer Eingangsphase mit 5 mg täglich gefolgt von einer anschließenden Titration bis maximal 25 mg täglich sowie einem Follow-up Zeitraum (35).

Randomisierung

Die Randomisierung in der POISE-Studie erfolgte zufällig in einem 1:1:1 Verhältnis entweder zur Behandlung mit Obeticholsäure 10 mg oder Obeticholsäure Titration oder zur Kontrollgruppe, stratifiziert nach

- AP $>3\times$ ULN und/oder AST $>2\times$ ULN und/oder Bilirubin $>ULN$ zu Baseline und intolerant für UDCA
- AP $\leq 3\times$ ULN und AST $\leq 2\times$ ULN und Bilirubin $\leq ULN$ zu Baseline und intolerant für UDCA
- AP $>3\times$ ULN und/oder AST $>2\times$ ULN und/oder Bilirubin $>ULN$ zu Baseline und nicht intolerant für UDCA
- AP $\leq 3\times$ ULN und AST $\leq 2\times$ ULN und Bilirubin $\leq ULN$ zu Baseline und nicht intolerant für UDCA.

Das Randomisierungsschema wurde von einem unabhängigen Statistiker erstellt. Die Randomisierung wurde zentral anhand eines interaktiven Sprach- und Webdialogsystems (Interactive Voice and Web Response System, IxRS) durchgeführt (35).

Endpunkte

Primärer Endpunkt war der kombinierte Endpunkt als Ansprecher-Endpunkt mit den Komponenten AP $< 1,67\times$ Obere Normgrenze (ULN), Gesamt Bilirubin $\leq ULN$ bei AP-Reduktion um $\geq 15\%$ von Baseline zum Monat 12. Sekundäre Endpunkte umfaßten AP, Bilirubin, biochemische Kriterien zum Behandlungserfolg, erkrankungsspezifische Symptome (PBC-40) und Verträglichkeit.

Patientenpopulation

Eingeschlossen wurden erwachsene Patienten mit primärer biliärer Cholangitis mit einer Randomisierung von 1:1:1 in die Gruppen Obeticholsäure 10 mg, Obeticholsäure Titration und Placebo, jeweils in Kombination oder ohne UDCA, abhängig vom Toleranzstatus der Patienten zu UDCA. Die Patienten wurden eingeschlossen in Europa (Deutschland, Vereinigtes Königreich, Polen, Niederlande, Italien, Spanien, Österreich, Belgien, Frankreich, Schweden), Nordamerika (USA, Kanada) und Australien.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Studie wurde in Studienzentren in Europa, Nordamerika und Australien durchgeführt.

Die zur Herleitung des Zusatznutzens berücksichtigten Studienarme bilden die in Deutschland gültige Zulassung von Obeticholsäure ab.

Es liegen keine Hinweise auf biodynamische oder kinetische Unterschiede zwischen den einzelnen Bevölkerungsgruppen und in Bezug auf Deutschland vor, die sich relevant auf die Studienergebnisse auswirken würden. Zudem wurden überwiegend Patienten kaukasischer Abstammung eingeschlossen, die Mehrzahl der Patienten stammt aus Europa. Für alle

Studienzentren ist davon auszugehen, dass die medizinischen Versorgungsstandards der Länder, in denen Patienten behandelt wurden (Europa, Nordamerika, Australien), mit denen in Deutschland vergleichbar sind.

Zudem ist besonders hervorzuheben, dass 10 Studienzentren der insgesamt 59 Studienzentren in Deutschland waren. Insgesamt stammten über alle Studienarme 13% der Patienten aus Deutschland, nur aus den USA (25%) und Italien (15%) stammten noch mehr Patienten (Quelle: CSR, S. 301 (35)).

Daher kann insgesamt davon ausgegangen werden, dass die Ergebnisse unter Berücksichtigung der Unsicherheit, die mit der Übertragbarkeit klinischer Daten assoziiert sind, auf den deutschen Versorgungskontext übertragbar sind.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

Tabelle 4-12: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
POISE	Ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der POISE-Studie wurde die Randomisierung adäquat durchgeführt und die Gruppenzuteilung verdeckt über ein Sprachdialogsystem (IxRS) durchgeführt. Patienten und behandelnde Personen waren während der randomisierten Phase verblindet. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan (SAP) definiert waren oder – wie bei den Verträglichkeitsendpunkten – gemäß Verfahrensordnung berichtet werden mussten (z.B. die Berechnung der Effektmaße für den Vergleich der Behandlungsgruppen). Es lagen keine sonstigen Aspekte für ein erhöhtes Verzerrungspotenzial vor, so dass das Verzerrungspotenzial

insgesamt als niedrig eingestuft wurde. Siehe auch Tabelle 4-12 in Anhang 4-E für eine detaillierte Bewertung des Verzerrungspotenzials.

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 4-13: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität Gesamt- mortalität	Morbidität AP und Bilirubin	Morbidität Fibrose	Morbidität Osteoporose	Morbidität PBC-40
POISE	ja	ja	ja	ja	ja
AP = Alkalische Phosphatase, ELF-Score = Enhanced liver fibrosis Score, PBC = primäre biliäre Cholangitis					

Studie	Morbidität Pruritus	Lebensqualität PBC-40
POISE	ja	ja
PBC = primäre biliäre Cholangitis		

Studie	Verträglichkeit Gesamtrate UE ^a	Verträglichkeit schwerwiegende UE	Verträglichkeit schwere UE ohne Pruritus	Verträglichkeit UE, die zum Therapieabbruch führten
POISE	ja ^a	ja	ja	ja
a: UE = unerwünschtes Ereignis, die Gesamtrate UE wird nur deskriptiv dargestellt				

4.3.1.3.1 Endpunkt - RCT

4.3.1.3.1.1 Mortalität: Gesamtmortalität – RCT

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

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

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

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

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

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

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

Tabelle 4-14: Operationalisierung von Gesamtmortalität

Studie	Operationalisierung
POISE	Gesamtrate der Todesfälle zwischen Studienbeginn und Studienende basierend auf der Sicherheitspopulation. Analysiert wurden die Inzidenzraten mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.

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 4-15: Bewertung des Verzerrungspotenzials für Gesamtmortalität in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Der Endpunkt „Gesamtmortalität“ wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte nach dem ITT-Prinzip auf Basis der Sicherheitspopulation. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) weisen auf eine hohe methodische Qualität des Endpunktes hin. Das Verzerrungspotenzial für den Endpunkt „Gesamtmortalität“ wurde daher mit niedrig bewertet (siehe Tabelle 4-15 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Tabelle 4-16: Ergebnisse für Gesamtmortalität aus RCT mit dem zu bewertenden Arzneimittel - Patienten mit Kombinationstherapie

Todesfälle/N (%) OCA+UDCA	Todesfälle/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
1/65 (1,54)	0/68 (0)	3,14 [0,13; 75,63] ^(a)	0,32 [0,01; 7,69] ^(a)	3,19 [0,13; 79,63] ^(a)	-0,02 [-0,05; 0,01] ^(a)	0,3063
(a) Modell ohne Stratifikations-Faktoren						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test						

Bei den Patienten mit Kombinationstherapie trat ein Todesfall auf. Es zeigten sich keine signifikanten Unterschiede zwischen den Behandlungsgruppen.

Patienten mit Monotherapie

Tabelle 4-17: Ergebnisse für Gesamtmortalität aus RCT mit dem zu bewertenden Arzneimittel - Patienten mit Monotherapie

Todesfälle/N (%) OCA	Todesfälle/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
0/5 (0)	0/5 (0)	1 [0,02; 42,65] ^(a)	1 [0,02; 42,65] ^(a)	1 [0,02; 59,99] ^(a)	0 [-0,01; 0,01] ^(a)	NA
(a) Modell ohne Stratifikations-Faktoren						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test NA = nicht berechenbar						

In der Teilpopulation der Patienten mit Monotherapie traten in beiden Behandlungsgruppen keine Todesfälle auf.

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

Für beide Teilpopulationen lag genau eine Studie mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

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

4.3.1.3.1.2 Morbidität: AP und Bilirubin – RCT

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

Tabelle 4-18: Operationalisierung von AP und Bilirubin

Studie	Operationalisierung
Primärer Endpunkt	
POISE	Der primäre Endpunkt ist definiert als dichotomer kombinierter Endpunkt aus AP $<1.67 \times$ ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin $\leq$ULN zu Monat 3, 6 und 12. Analysiert wurden die Inzidenzraten auf Basis der ITT-Population mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.
AP-Reduktion zu Baseline	
POISE	Der dichotome Endpunkt ist definiert als Reduktion im AP von Baseline zu Monat 3, 6 und 12 $>10\%$, $>15\%$, $>20\%$ und $>40\%$. Analysiert wurden die Inzidenzraten auf Basis der ITT-Population mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.
AP $\leq$ ULN	
POISE	Der dichotome Endpunkt ist definiert als Reduktion der AP von Baseline zu Monat 3, 6 und 12 $\leq$ der oberen Grenze des Normbereichs (upper limit of normal). Analysiert wurden die Inzidenzraten auf Basis der ITT-Population mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.
Änderung zu Baseline in den biochemischen Kenngrößen AP und Bilirubin	
POISE	Für die metrischen Endpunkte "Änderung zu Baseline in den biochemischen Kenngrößen " AP und Bilirubin wurde die mittlere Änderung der biochemischen Kenngrößen AP und Bilirubin (Gesamt und direktes) zu Baseline mittels gemischtem Modell für Wiederholungsmessungen (MMRM) berechnet. Als Basis diente die ITT-Population. Dabei wurden die Faktoren "AP $> 3 \times$ ULN und/oder AST $> 2 \times$ ULN und/oder Bilirubin $> ULN$ zu Baseline" und "Intoleranz gegenüber UDCA" als Strata sowie der Baselinewert als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde auf Basis der Kleinste-Quadrate-(Least-Square-) Methode berechnet und die LS-Mittelwerte pro Behandlungsarm und die Differenz der LS-Mittelwerte dargestellt. Zudem wurde die standardisierte Mittelwertdifferenz (SMD) mit 95% Konfidenzintervall berechnet. Die Analyse erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline vorgelegen hatte.

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 4-19: Bewertung des Verzerrungspotenzials für AP und Bilirubin in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Primärer Endpunkt (Anteil Patienten mit AP <1.67x ULN, Gesamt-Bilirubin ≤ULN und AP-Reduktion von ≥15% vom Ausgangswert)						
POISE	niedrig	ja	ja	ja	nein	hoch
AP-Reduktion zu Baseline						
POISE	niedrig	ja	ja	ja	nein	hoch
AP ≤ ULN						
POISE	niedrig	ja	ja	ja	nein	hoch
Änderung zu Baseline in den biochemischen Kenngrößen AP und Bilirubin						
POISE	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Endpunkte zu „AP“ und „Bilirubin“ wurden in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis der ITT-Population. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da beide Endpunkte als Surrogate für die klinischen Endpunkte Mortalität und Lebertransplantation dienen, aber formal nicht validiert sind, wurde das Verzerrungspotenzial für diese Endpunkte mit hoch bewertet (siehe Tabelle 4-19 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

In der folgenden Tabelle wurden die Ergebnisse des primären Endpunkts der POISE-Studie dargestellt. Die darauf folgende Abbildung zeigt den zeitlichen Verlauf des primären

Endpunkts für die Zeitpunkte des RCT und der Langzeit-Sicherheitsstudie zum jährlichen Datenschnitt des 31. Januars (hier: 31.01.2016), für den adäquate Daten für die Nachverfolgung bis Monat 27 vorliegen.

Tabelle 4-20: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP $<1,67 \times$ ULN, Gesamt-Bilirubin $\leq$ ULN und AP-Reduktion von $\geq 15\%$) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
6	22/65 (33,85)	5/68 (7,35)	4,64 [1,89; 11,38]	0,22 [0,09; 0,53]	7 [2,41; 20,38]	-0,24 [-0,38; -0,1]	0,0001
12	30/65 (46,15)	7/68 (10,29)	4,45 [2,12; 9,34]	0,22 [0,11; 0,47]	8,22 [3,18; 21,21]	-0,34 [-0,49; -0,2]	$<0,0001$
CMH = Cochran-Mantel-Haenszel Test RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test							

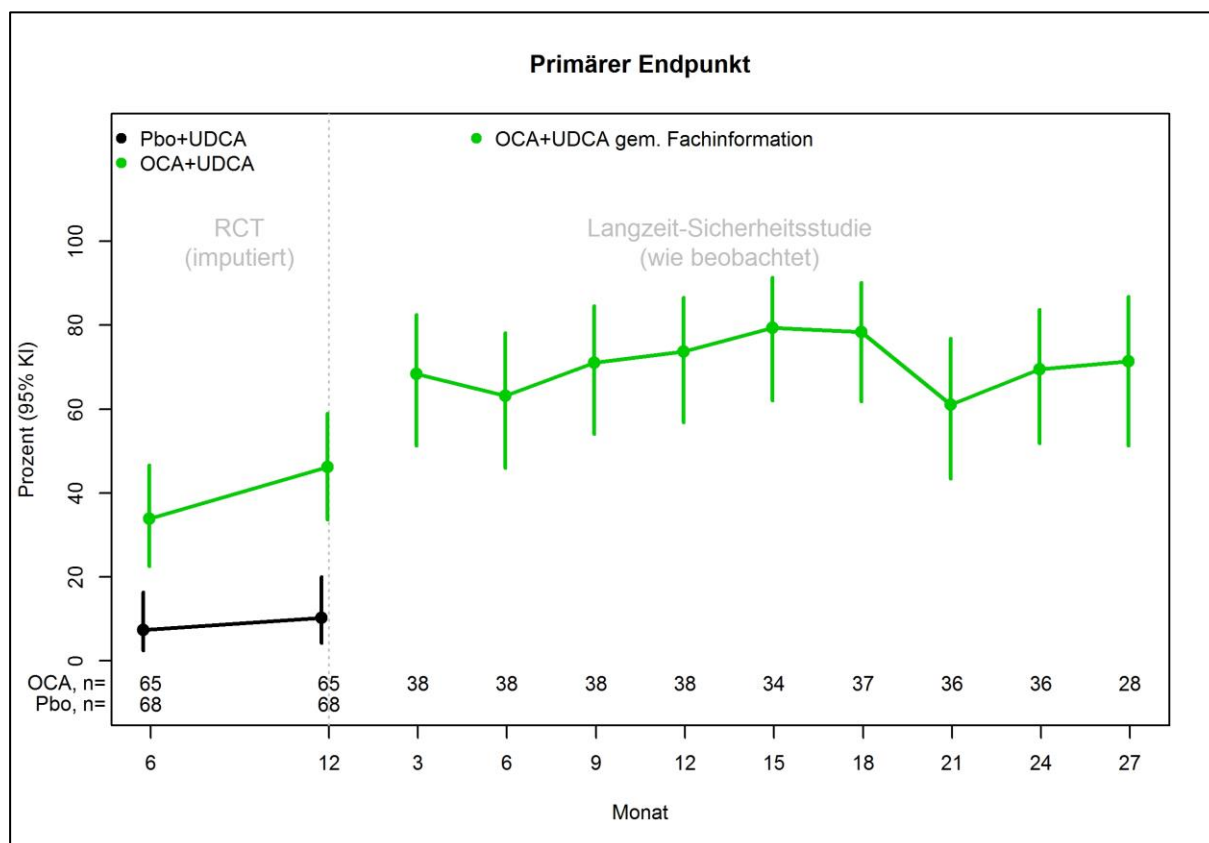


Abbildung 4-3: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP <1,67x ULN, Gesamt-Bilirubin ≤ULN und AP-Reduktion von ≥15%) aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Für den primären Endpunkt zeigte sich ein statistisch signifikanter Vorteil bezüglich der Ansprech-Kriterien für Obeticholsäure+UDCA gegenüber UDCA+Placebo. Dieser Vorteil bleibt auch in der einarmigen Langzeit-Sicherheitsstudie bestehen, die in Abschnitt 4.3.2.3 ausführlich dargestellt ist.

Tabelle 4-21: Ergebnisse für die Einzelkomponenten des primären Endpunkts aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
AP < 1.67xULN							
6	23/65 (35,38)	8/68 (11,76)	3,07 [1,5; 6,27]	0,33 [0,16; 0,67]	4,39 [1,75; 11]	-0,21 [-0,35; -0,06]	0,0011
12	31/65	12/68	2,69	0,37	4,64	-0,28	0,0002

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
	(47,69)	(17,65)	[1,53; 4,71]	[0,21; 0,65]	[2,04; 10,54]	[-0,43; -0,13]	
Bilirubin <= ULN							
6	61/65 (93,85)	55/68 (80,88)	1,09 [0,96; 1,24]	0,91 [0,81; 1,04]	3,92 [1,16; 13,24]	-0,09 [-0,19; 0,02]	0,0219
12	57/65 (87,69)	54/68 (79,41)	1,08 [0,96; 1,22]	0,92 [0,82; 1,04]	1,99 [0,73; 5,4]	-0,07 [-0,18; 0,03]	0,1782

RR = Relatives Risiko
 RRR = Invertiertes relatives Risiko
 OR = Odds Ratio
 RD = Risikodifferenz
 KI = Konfidenzintervall
 Pbo = Placebo
 CMH = Cochran-Mantel-Haenszel Test
 ULN = Upper limit of normal (Obere Grenze des Normalbereichs)
 AP-Reduktion >=15% siehe Tabelle 4-22.

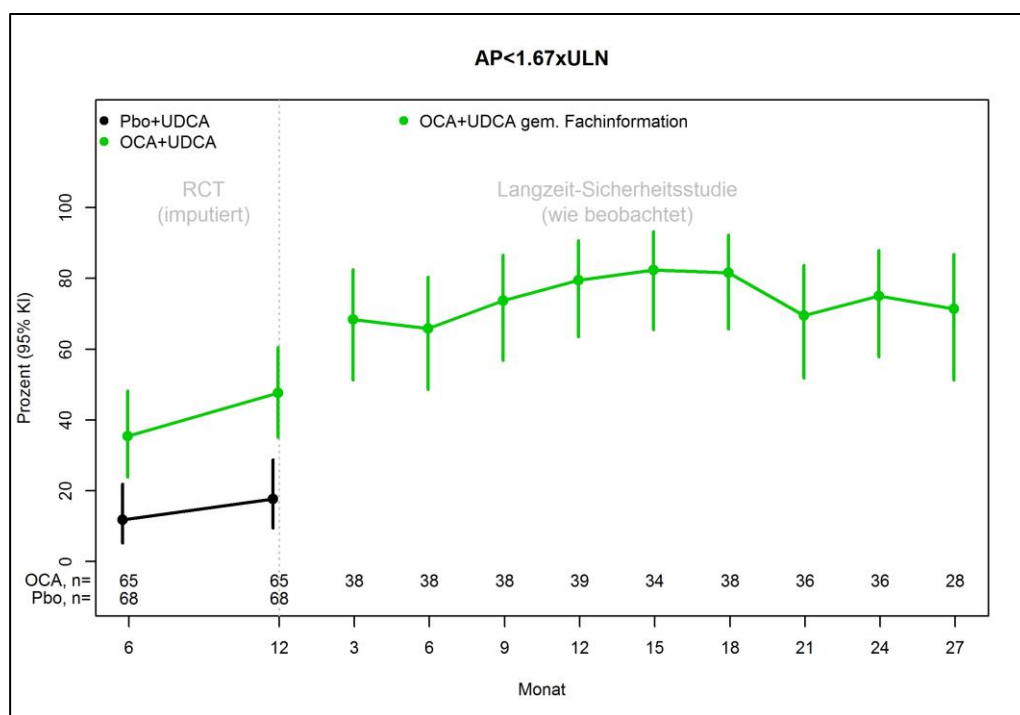


Abbildung 4-4: Ergebnisse für den Anteil Patienten mit AP <1,67x ULN aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

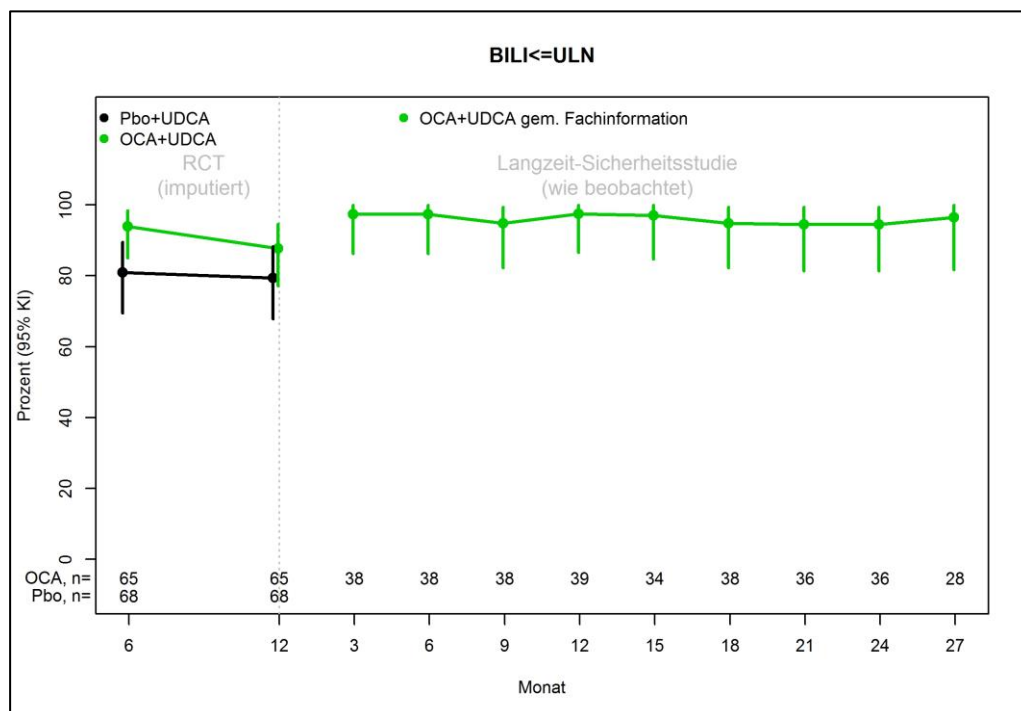


Abbildung 4-5: Ergebnisse für den Anteil Patienten mit Gesamt-Bilirubin $\leq$ ULN aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Auch in den drei Einzel-Komponenten des primären Endpunktes zeigten sich die statistisch signifikanten Vorteile von Obeticholsäure+UDCA gegenüber UDCA+Placebo. Diese Ergebnisse werden unterstützt durch die im Studienprotokoll als sekundäre Analysen definierten weiteren Ansprech-Kriterien nach prozentualer Reduktion von AP im Vergleich zu Baseline. In allen Analysen fand sich ein statistisch signifikanter Vorteil für Obeticholsäure+UDCA gegenüber UDCA+Placebo. Dieser Vorteil bleibt auch in der einarmigen Langzeit-Sicherheitsstudie bestehen.

Tabelle 4-22: Ergebnisse für den Anteil Patienten mit AP-Reduktion zu Baseline aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
AP-Reduktion $\geq$ 10%							
6	50/65 (76,92)	32/68 (47,06)	1,62 [1,22; 2,15]	0,62 [0,46; 0,82]	3,82 [1,79; 8,11]	-0,3 [-0,45; -0,14]	0,0004
12	51/65 (78,46)	29/68 (42,65)	1,84 [1,36; 2,49]	0,54 [0,4; 0,74]	4,91 [2,29; 10,52]	-0,36 [-0,51; -0,2]	<,0001
AP-Reduktion $\geq$ 15%							
6	47/65 (72,31)	18/68 (26,47)	2,7 [1,77; 4,14]	0,37 [0,24; 0,57]	7,45 [3,44; 16,17]	-0,46 [-0,61; -0,31]	<,0001

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
12	50/65 (76,92)	21/68 (30,88)	2,48 [1,7; 3,63]	0,4 [0,28; 0,59]	7,55 [3,47; 16,42]	-0,46 [-0,61; -0,31]	<,0001
AP-Reduktion $\geq 20\%$							
6	42/65 (64,62)	11/68 (16,18)	3,93 [2,22; 6,95]	0,25 [0,14; 0,45]	10,1 [4,34; 23,51]	-0,48 [-0,63; -0,34]	<,0001
12	45/65 (69,23)	17/68 (25)	2,72 [1,74; 4,24]	0,37 [0,24; 0,57]	7,07 [3,25; 15,38]	-0,44 [-0,59; -0,29]	<,0001
AP-Reduktion $\geq 40\%$							
6	17/65 (26,15)	0/68 (0)	36,59 [2,25; 596,21] ^(a)	0,03 [0; 0,45] ^(a)	49,43 [2,9; 841,94] ^(a)	-0,26 [-0,37; -0,15] ^(a)	<,0001
12	18/65 (27,69)	1/68 (1,47)	18,57 [2,57; 134,21]	0,05 [0,01; 0,39]	29,46 [3,7; 234,5]	-0,26 [-0,37; -0,15] ^(a)	<,0001
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test							

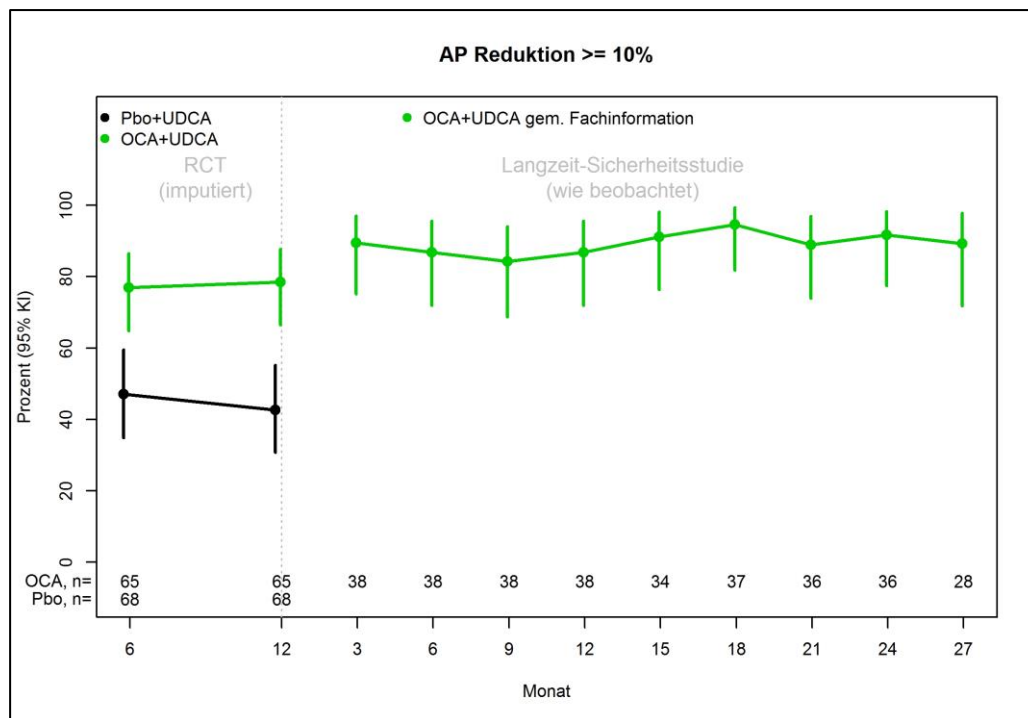


Abbildung 4-6: Ergebnisse für den Anteil Patienten mit AP-Reduktion $\geq 10\%$ aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

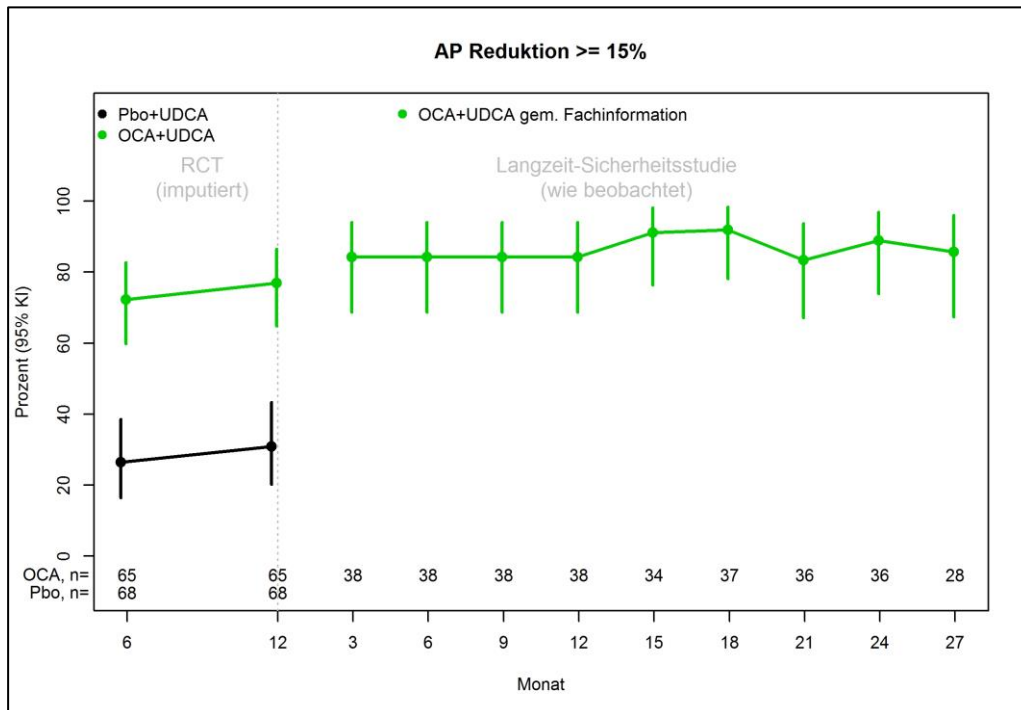


Abbildung 4-7: Ergebnisse für den Anteil Patienten mit AP-Reduktion $\geq 15\%$ aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

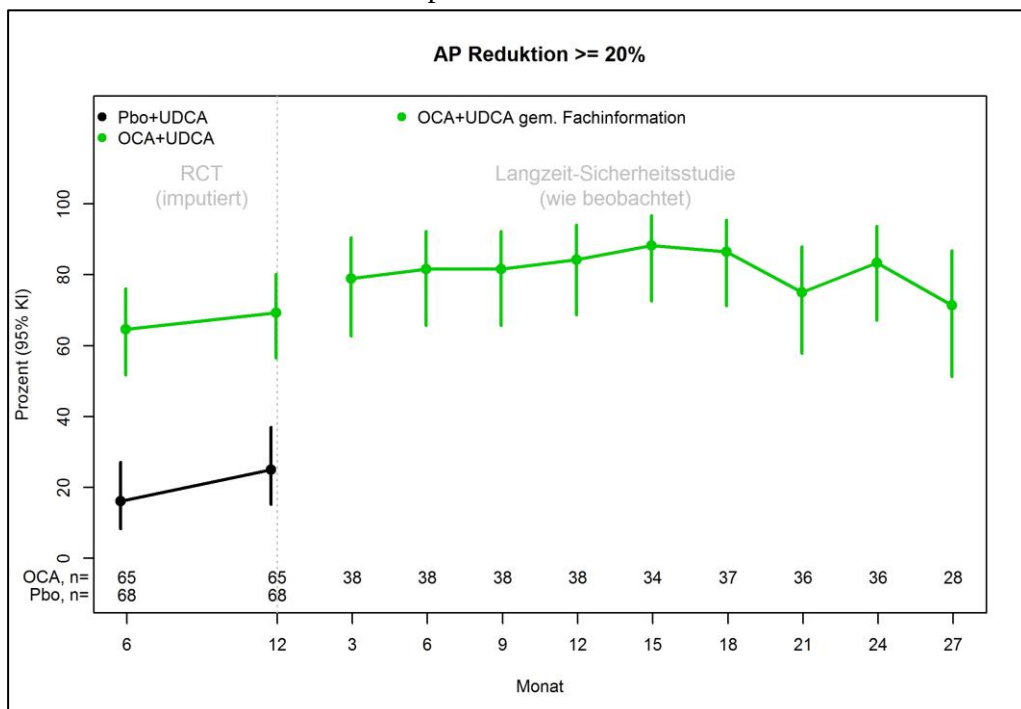


Abbildung 4-8: Ergebnisse für den Anteil Patienten mit AP-Reduktion $\geq 20\%$ aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

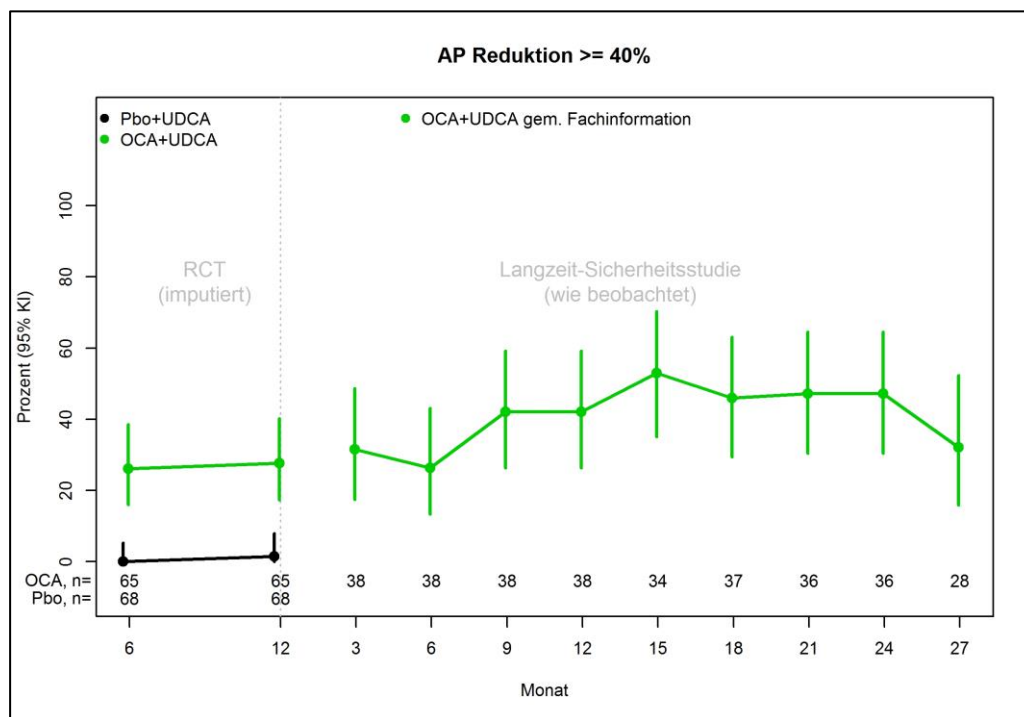


Abbildung 4-9: Ergebnisse für den Anteil Patienten mit AP-Reduktion $\geq 40\%$ aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Tabelle 4-23: Ergebnisse für den Anteil Patienten mit AP $\leq$ ULN aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
6	1/65 (1,54)	0/68 (0)	3,14 [0,13; 75,63] ^(a)	0,32 [0,01; 7,69] ^(a)	3,19 [0,13; 79,63] ^(a)	-0,02 [-0,05; 0,01] ^(a)	0,3063
12	1/65 (1,54)	0/68 (0)	3,14 [0,13; 75,63] ^(a)	0,32 [0,01; 7,69] ^(a)	3,19 [0,13; 79,63] ^(a)	-0,02 [-0,05; 0,01] ^(a)	0,3063
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test							

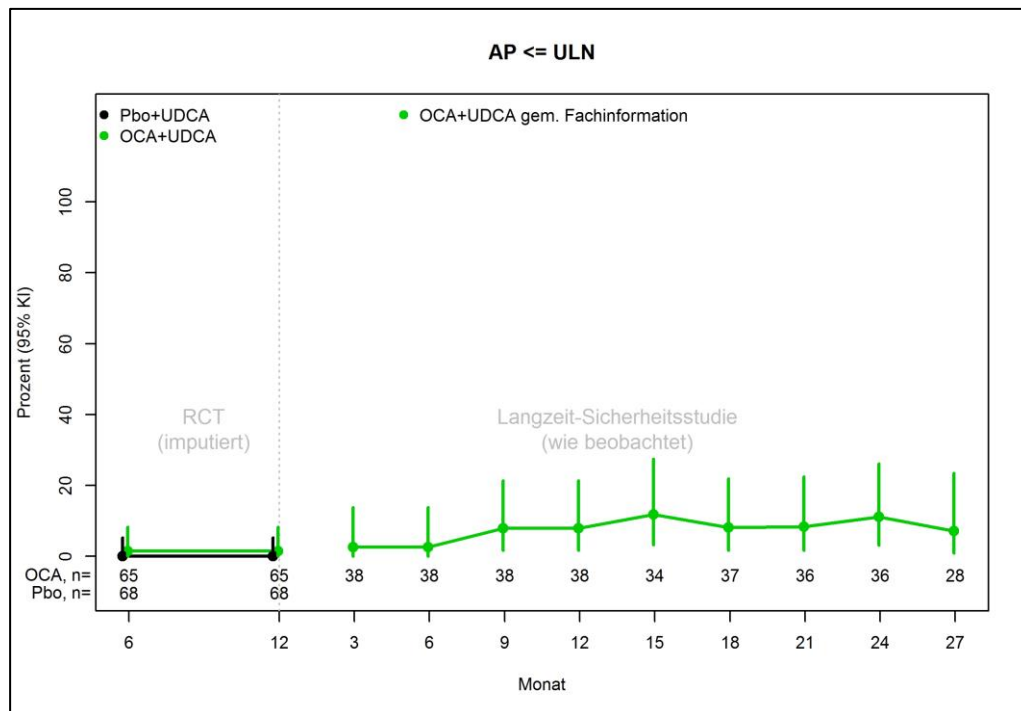


Abbildung 4-10: Ergebnisse für den Anteil Patienten mit $AP \leq ULN$ aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Die ebenfalls a priori definierte Analyse für Patienten mit $AP \leq ULN$ nach 6 und 12 Monaten zeigte keine Unterschiede, da nur ein Patient der Obeticholsäure+UDCA Gruppe dieses Kriterium erreichte, aber kein Patient der UDCA+Placebo-Gruppe. Dieser Vorteil bleibt auch in der einarmigen Langzeit-Sicherheitsstudie bestehen

Tabelle 4-24: Ergebnisse für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
AP, abs. Veränderung							
64	314 (95,5)	-85,24 (4)	68	324,6 (114,7)	-13,2 (3,92)	-72,04 [-82,45; -61,63] <,0001	-2,35 [-2,8; -1,91]
Gesamt-Bilirubin, abs. Veränderung							
64	10,4 (5,7)	-0,8 (0,19)	68	11,7 (7,4)	0,26 (0,18)	-1,06 [-1,56; -0,57] <,0001	-0,73 [-1,08; -0,37]
Direktes Bilirubin, abs. Veränderung							
64	4,6 (4,7)	-0,65 (0,13)	68	5,4 (6,4)	0,35 (0,13)	-1,00 [-1,34; -0,66] <,0001	-0,99 [-1,35; -0,63]
MW = Mittelwert							

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

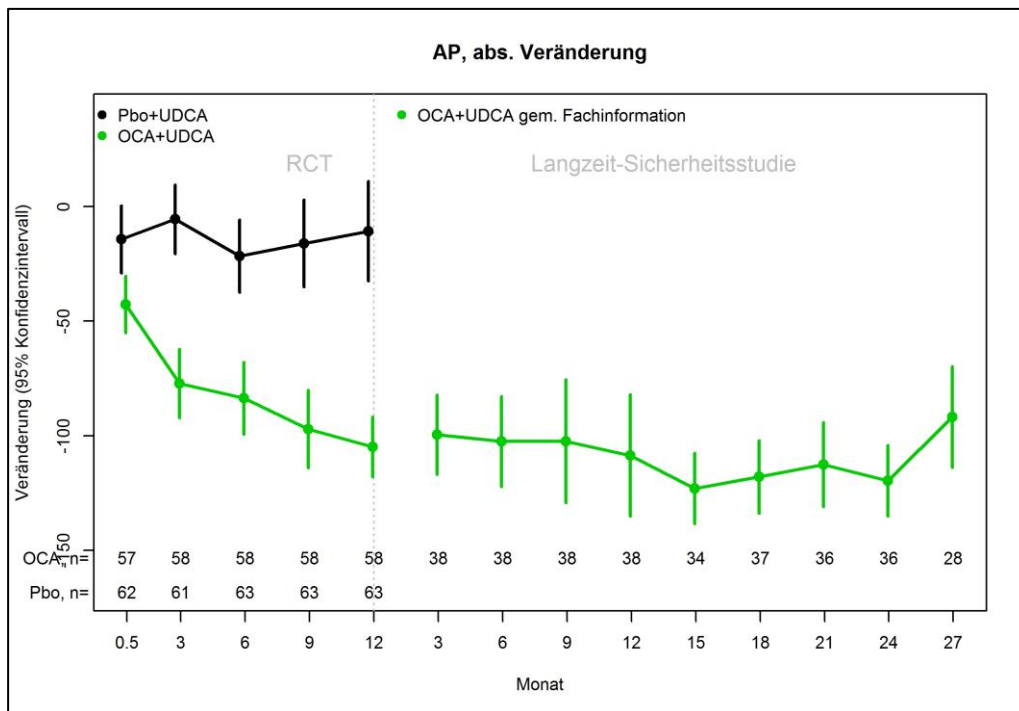


Abbildung 4-11: Ergebnisse für den Endpunkt AP, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

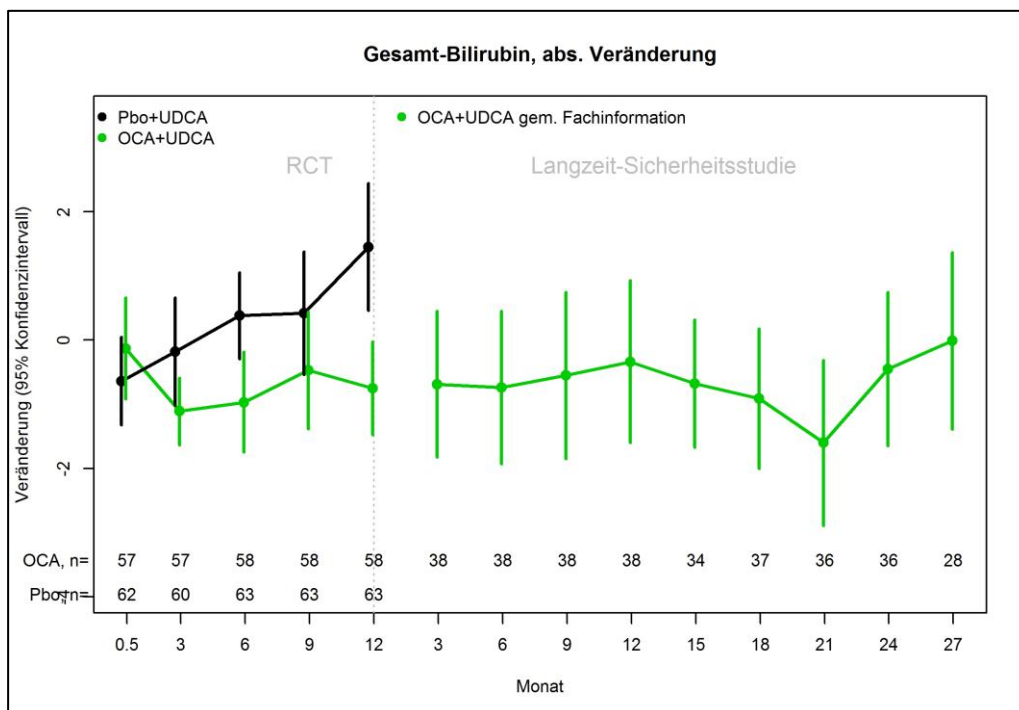


Abbildung 4-12: Ergebnisse für den Endpunkt Gesamt-Bilirubin, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

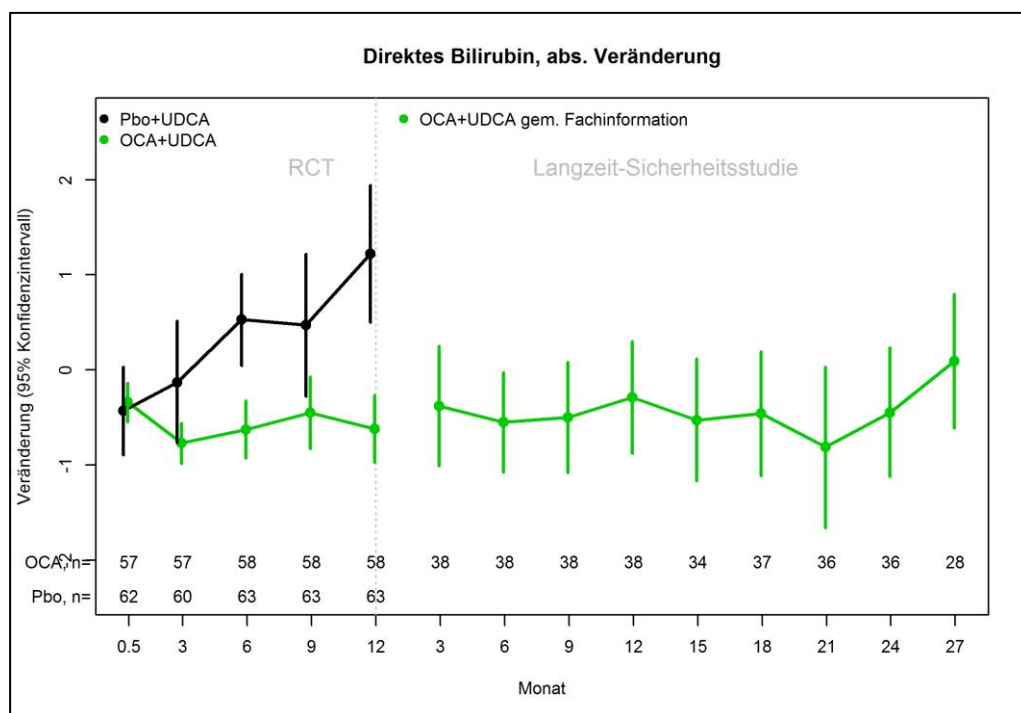


Abbildung 4-13: Ergebnisse für den Endpunkt Direktes Bilirubin, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Die zusätzlich durchgeführten und a priori geplanten Analysen bezüglich der Änderung zu Baseline für die AP und das Gesamt-, sowie auch das direkte Bilirubin, zeigten in der SMD statistisch signifikante und klinisch relevante Vorteile für Obeticholsäure+UDCA unter Berücksichtigung der Relevanzschwelle von -0,2 bzw. 0,2, wie sie unter anderem vom IQWiG definiert wurde (17). Dieser Vorteil bleibt auch in der einarmigen Langzeit-Sicherheitsstudie bestehen.

Ergänzend werden die Analysen der Kriterien nach PARIS I und II, MAYO II, Rotterdam, Toronto II und Barcelona als weiterführende Analysen unterstützend dargestellt. Deren Definitionen sind in Abschnitt 4.2.5.2 ausführlich beschrieben (siehe insbesondere Tabelle 4-3).

Tabelle 4-25: Ergänzende Ergebnisse gemäß der Kriterien nach PARIS, MAYO, ROTTERDAM, TORONTO und BARCELONA aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Paris I normal							
6	48/65 (73.85)	39/68 (57.35)	1,14 [0,92; 1,41]	0,88 [0,71; 1,09]	2,48 [1,09; 5,65]	-0,14 [-0,29; 0,01]	0,0292

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
12	52/65 (80)	35/68 (51.47)	1,38 [1,1; 1,73]	0,72 [0,58; 0,91]	5,86 [2,26; 15,2]	-0,27 [-0,41; -0,13]	0,0001
Paris I normal, abnormal zu Baseline							
6	17/31 (54.84)	6/30 (20)	2,85 [1,3; 6,24]	0,35 [0,16; 0,77]	5,05 [1,59; 16,04]	-0,35 [-0,57; -0,12]	0,0051
12	20/31 (64.52)	6/30 (20)	3,07 [1,41; 6,65]	0,33 [0,15; 0,71]	7,21 [2,22; 23,48]	-0,43 [-0,65; -0,21]	0,0007
Paris I normal, normal zu Baseline							
6	31/34 (91.18)	33/38 (86.84)	1,05 [0,89; 1,23] ^(a)	0,95 [0,81; 1,12] ^(a)	2,35 [0,42; 13]	-0,04 [-0,19; 0,1] ^(a)	0,3205
12	32/34 (94.12)	29/38 (76.32)	1,23 [1,01; 1,5] ^(a)	0,81 [0,67; 0,99] ^(a)	9,93 [1,18; 83,25]	-0,18 [-0,33; -0,02] ^(a)	0,0132
Paris II normal							
6	12/65 (18.46)	3/68 (4.41)	4,15 [1,25; 13,83]	0,24 [0,07; 0,8]	5,23 [1,37; 19,93]	-0,14 [-0,25; -0,03] ^(a)	0,0093
12	17/65 (26.15)	3/68 (4.41)	5,86 [1,82; 18,91]	0,17 [0,05; 0,55]	8,12 [2,22; 29,7]	-0,22 [-0,33; -0,1] ^(a)	0,0004
Paris II normal, abnormal zu Baseline							
6	12/65 (18.46)	3/68 (4.41)	4,15 [1,25; 13,83]	0,24 [0,07; 0,8]	5,23 [1,37; 19,93]	-0,14 [-0,25; -0,03] ^(a)	0,0093
12	17/65 (26.15)	3/68 (4.41)	5,86 [1,82; 18,91]	0,17 [0,05; 0,55]	8,12 [2,22; 29,7]	-0,22 [-0,33; -0,1] ^(a)	0,0004
Mayo II normal							
6	22/65 (33.85)	8/68 (11.76)	2,93 [1,42; 6,03]	0,34 [0,17; 0,7]	4,06 [1,62; 10,18]	-0,19 [-0,34; -0,05]	0,0020
12	31/65 (47.69)	11/68 (16.18)	2,88 [1,61; 5,17]	0,35 [0,19; 0,62]	5,31 [2,28; 12,37]	-0,3 [-0,45; -0,15]	0,0001
Mayo II normal, abnormal zu Baseline							
6	21/64 (32.81)	8/68 (11.76)	2,85 [1,38; 5,89]	0,35 [0,17; 0,73]	3,88 [1,54; 9,75]	-0,19 [-0,33; -0,04]	0,0030
12	30/64 (46.88)	11/68 (16.18)	2,84 [1,58; 5,12]	0,35 [0,2; 0,63]	5,14 [2,2; 12]	-0,29 [-0,44; -0,15]	0,0001
Rotterdam normal							
6	48/65 (73.85)	46/68 (67.65)	1 [0,81; 1,24]	1 [0,8; 1,23]	1,37 [0,64; 2,95]	-0,03 [-0,18; 0,12]	0,4262
12	47/65	41/68	1,08	0,92	1,84	-0,09	0,1240

Monat	Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
	(72.31)	(60.29)	[0,87; 1,35]	[0,74; 1,15]	[0,85; 4]	[-0,25; 0,06]	
Rotterdam normal, abnormal zu Baseline							
6	3/11 (27.27)	2/16 (12.5)	2,05 [0,36; 11,82]	0,49 [0,08; 2,81]	2,41 [0,3; 19,14]	-0,15 [-0,45; 0,16]	0,4170
12	2/11 (18.18)	1/16 (6.25)	2,99 [0,29; 30,73]	0,33 [0,03; 3,44]	3,5 [0,25; 48,77]	-0,12 [-0,38; 0,14] ^(a)	0,3537
Rotterdam normal, normal zu Baseline							
6	45/54 (83.33)	44/52 (84.62)	1 [0,84; 1,18]	1 [0,85; 1,19]	0,87 [0,3; 2,47]	0,01 [-0,13; 0,15]	0,7897
12	45/54 (83.33)	40/52 (76.92)	1,07 [0,89; 1,28]	0,94 [0,78; 1,13]	1,69 [0,63; 4,57]	-0,06 [-0,21; 0,09]	0,3009
Toronto II normal							
6	32/65 (49.23)	13/68 (19.12)	2,57 [1,52; 4,37]	0,39 [0,23; 0,66]	4,7 [2,07; 10,67]	-0,27 [-0,43; -0,12]	0,0001
12	35/65 (53.85)	14/68 (20.59)	2,59 [1,57; 4,28]	0,39 [0,23; 0,64]	5,27 [2,33; 11,92]	-0,31 [-0,46; -0,16]	<,0001
Toronto II normal, abnormal zu Baseline							
6	29/62 (46.77)	10/65 (15.38)	3,04 [1,65; 5,59]	0,33 [0,18; 0,61]	5,75 [2,36; 13,96]	-0,29 [-0,44; -0,14]	0,0001
12	32/62 (51.61)	11/65 (16.92)	3,03 [1,71; 5,36]	0,33 [0,19; 0,59]	6,39 [2,66; 15,35]	-0,33 [-0,48; -0,18]	<,0001
Barcelona normal							
6	17/65 (26.15)	0/68 (0)	36,59 [2,25; 596,21] ^(a)	0,03 [0; 0,45] ^(a)	49,43 [2,9; 841,94] ^(a)	-0,26 [-0,37; -0,15] ^(a)	<,0001
12	18/65 (27.69)	1/68 (1.47)	18,57 [2,57; 134,21]	0,05 [0,01; 0,39]	29,46 [3,7; 234,5]	-0,26 [-0,37; -0,15] ^(a)	<,0001
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test							

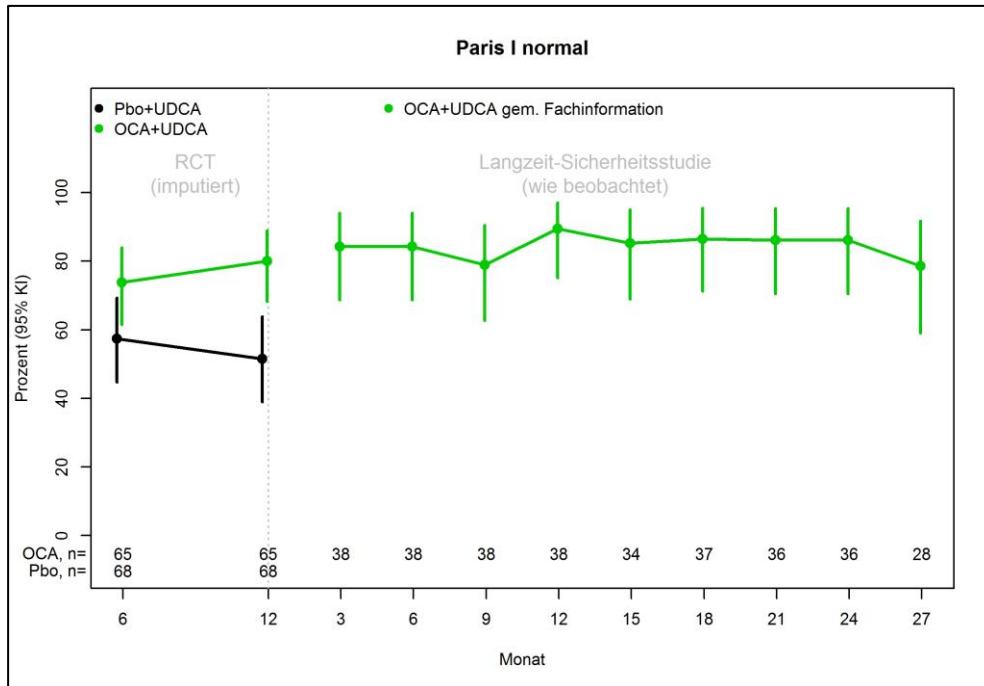


Abbildung 4-14: Ergänzende Ergebnisse für Anteil Patienten, die das Kriterium PARIS I erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

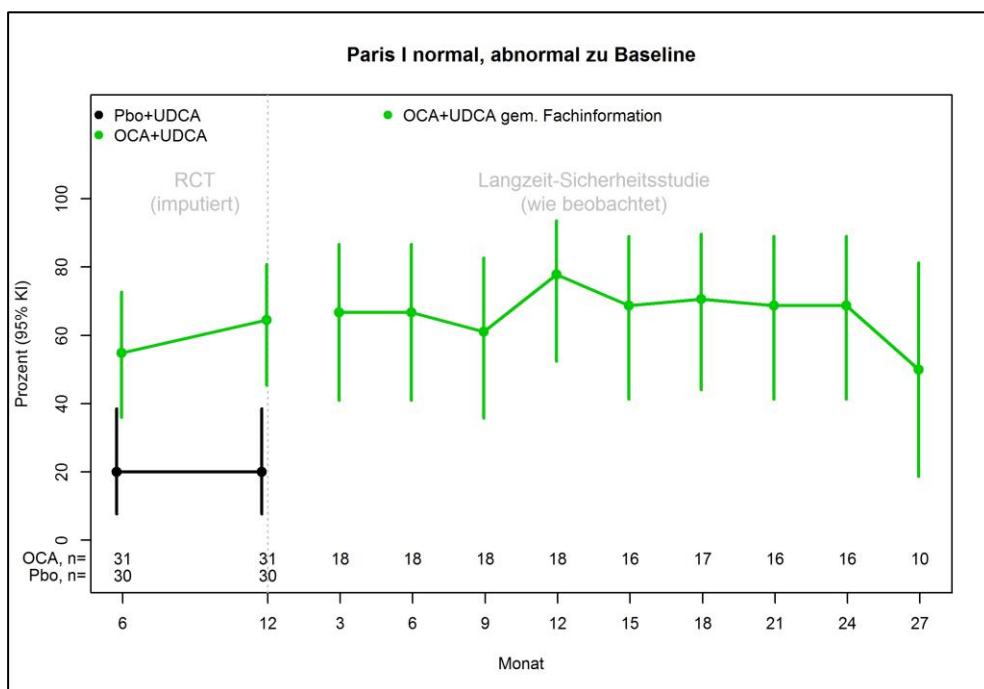


Abbildung 4-15: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium PARIS I erfüllt haben, Patienten aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie

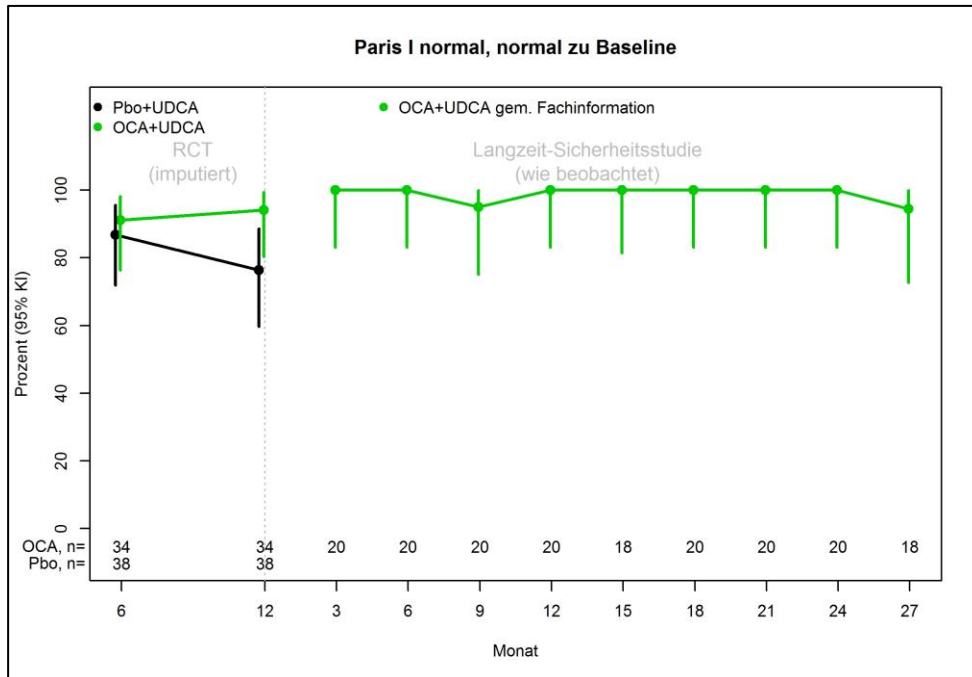


Abbildung 4-16: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium PARIS I erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline erfüllt haben – Patienten mit Kombinationstherapie

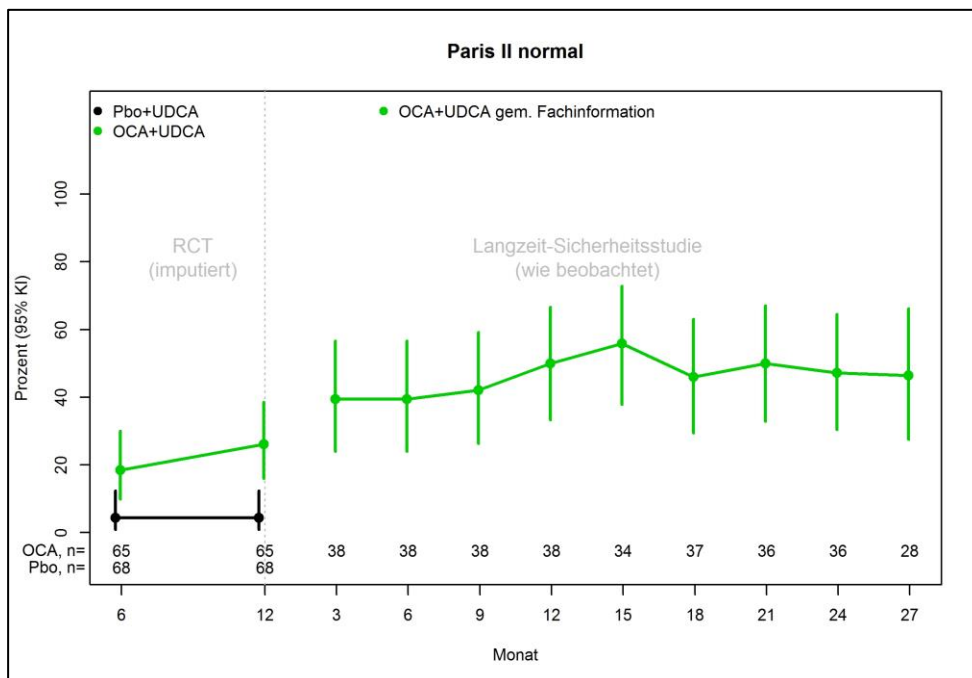


Abbildung 4-17: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium PARIS II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

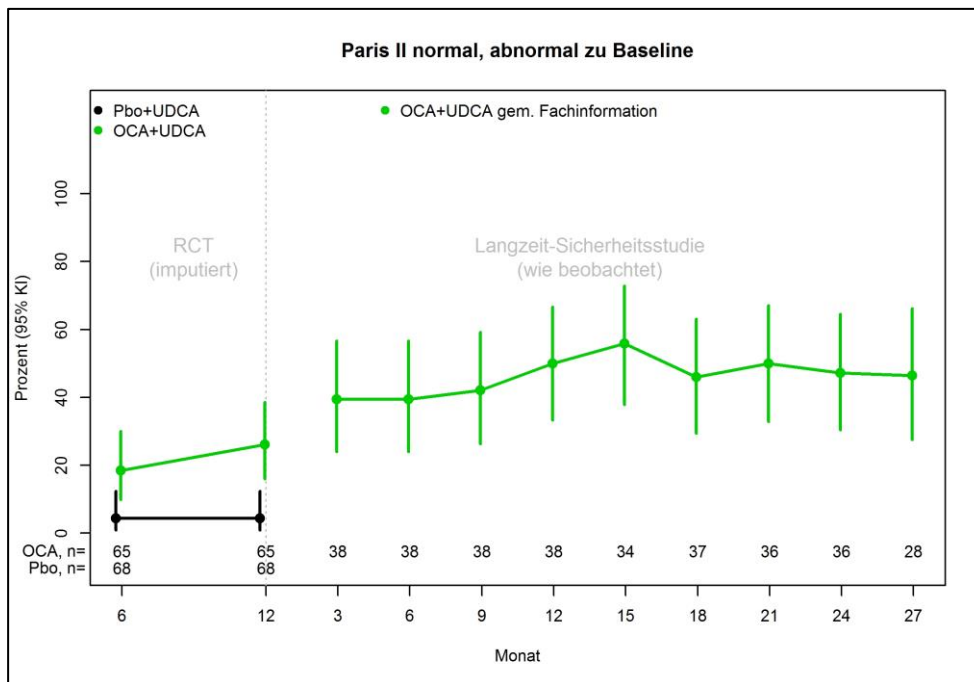


Abbildung 4-18: Ergänzende Ergebnisse den Anteil Patienten, die das Kriterium PARIS II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie

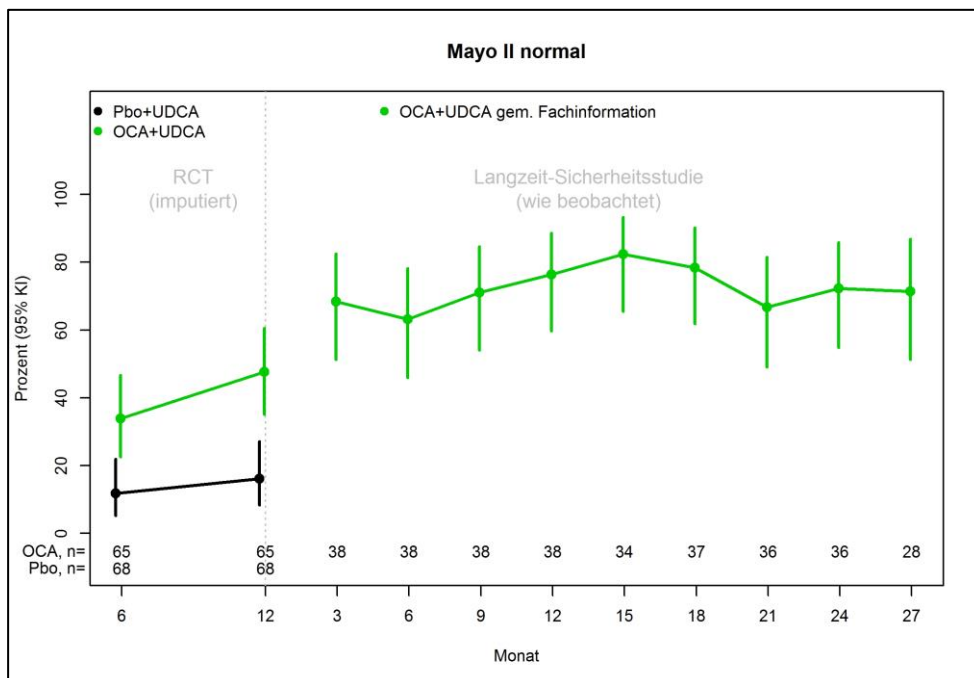


Abbildung 4-19: Ergänzende Ergebnisse den Anteil Patienten, die das Kriterium MAYO II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

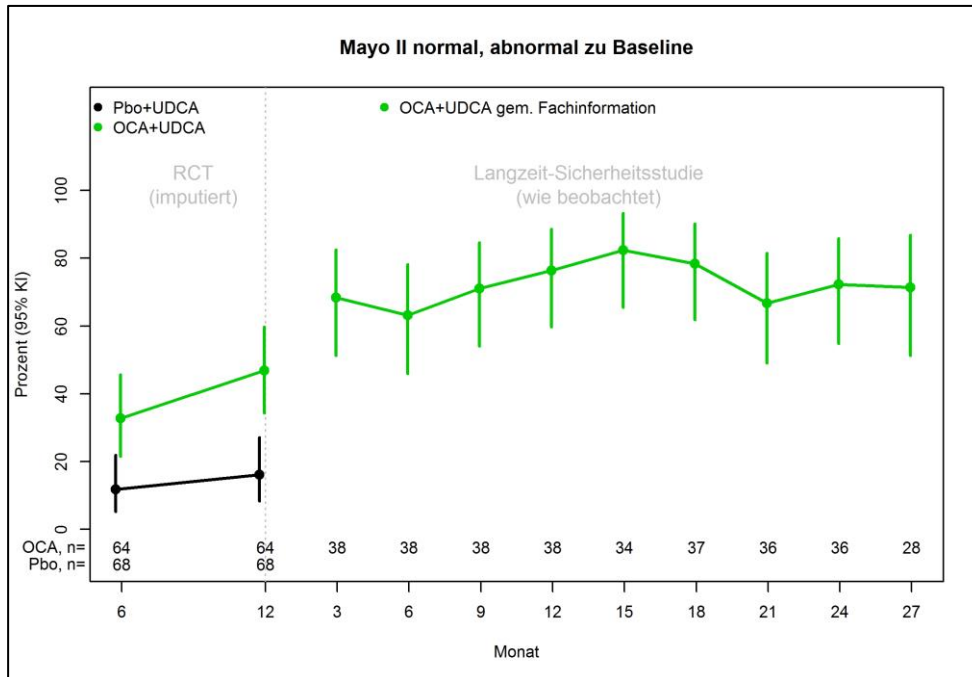


Abbildung 4-20: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium MAYO II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie

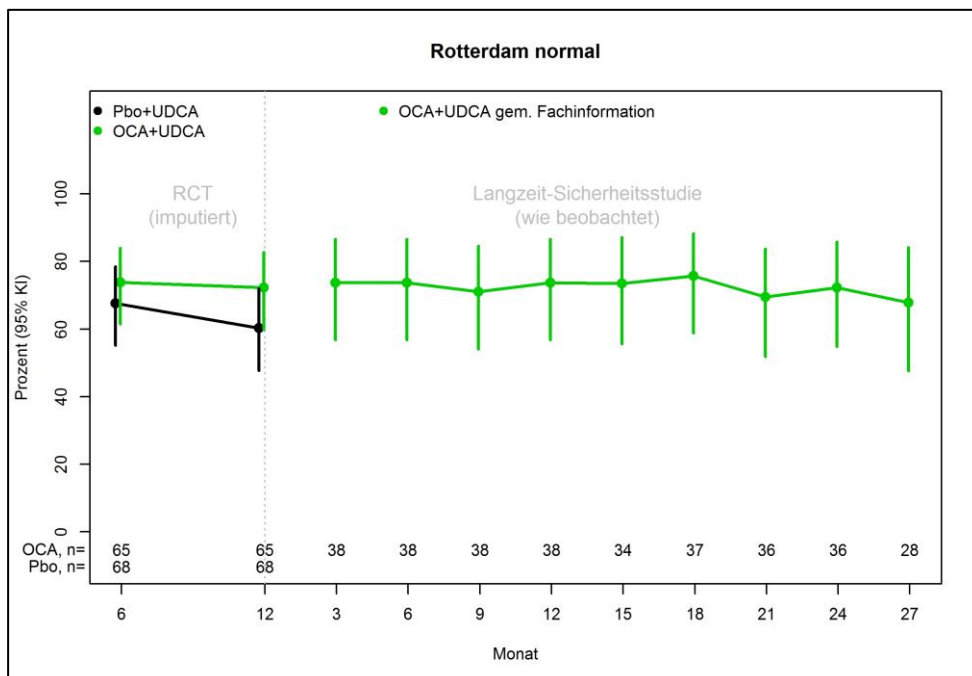


Abbildung 4-21: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium ROTTERDAM erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

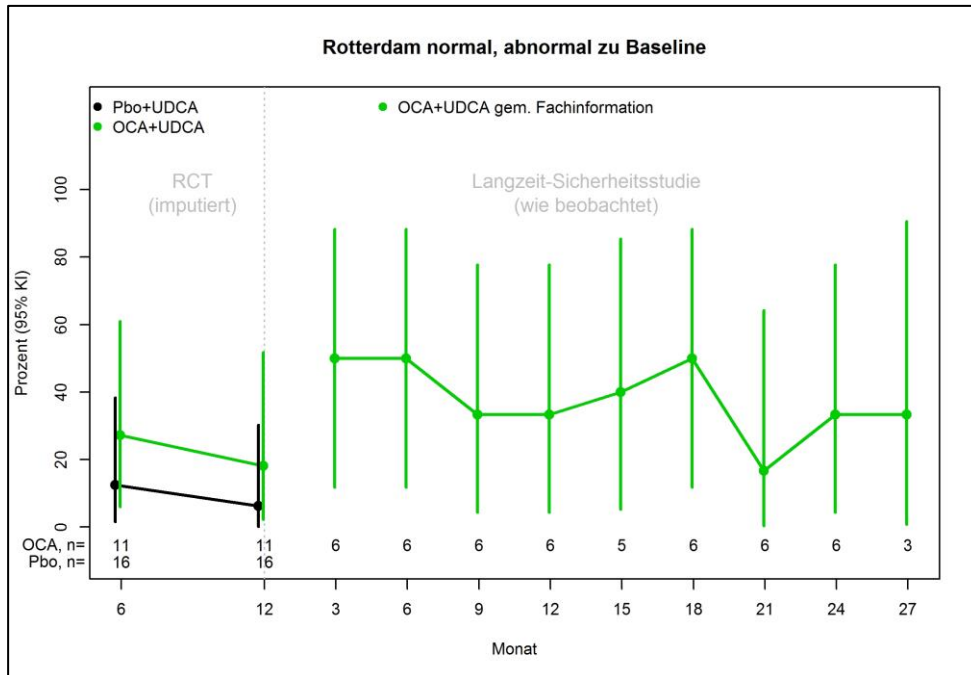


Abbildung 4-22: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium ROTTERDAM erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie

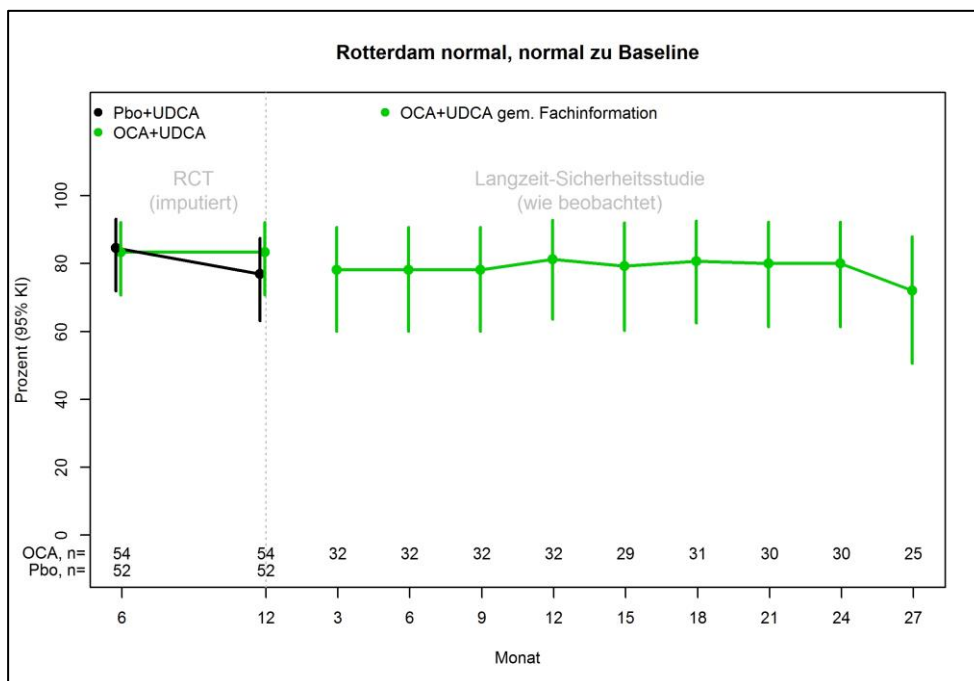


Abbildung 4-23: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium ROTTERDAM erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline erfüllt haben – Patienten mit Kombinationstherapie

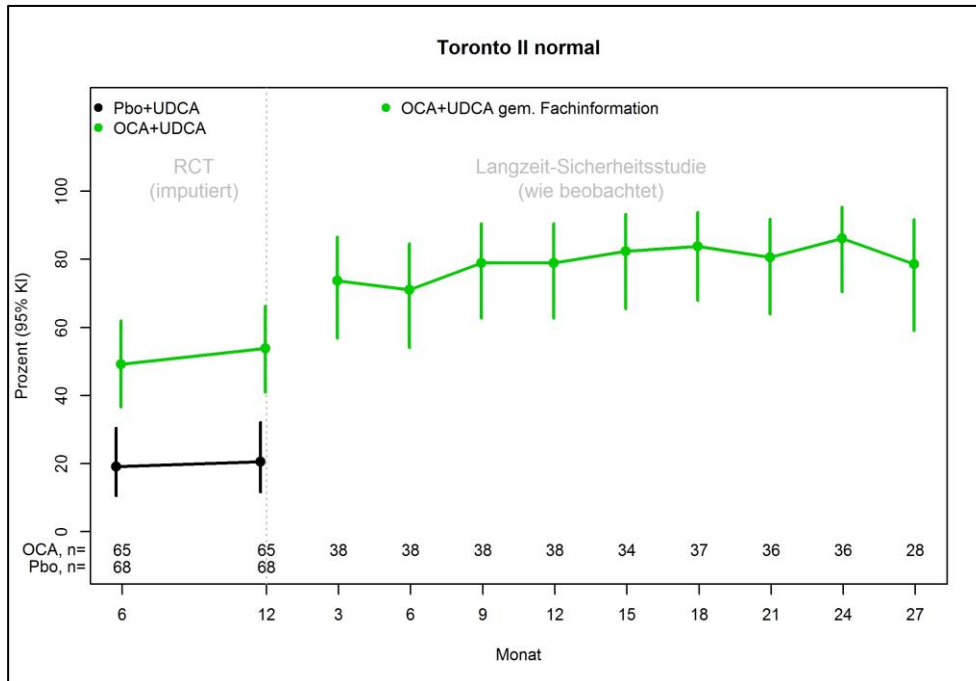


Abbildung 4-24: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium TORONTO II erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

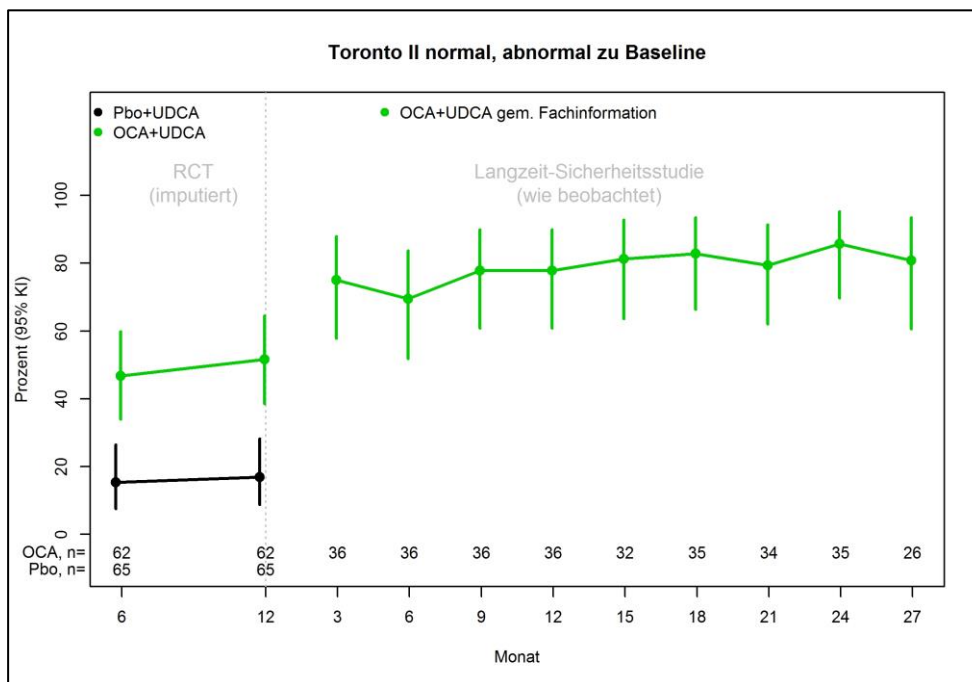


Abbildung 4-25: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium TORONTO II erfüllt haben, Patienten aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel für Patienten, die das Kriterium zu Baseline nicht erfüllt haben – Patienten mit Kombinationstherapie

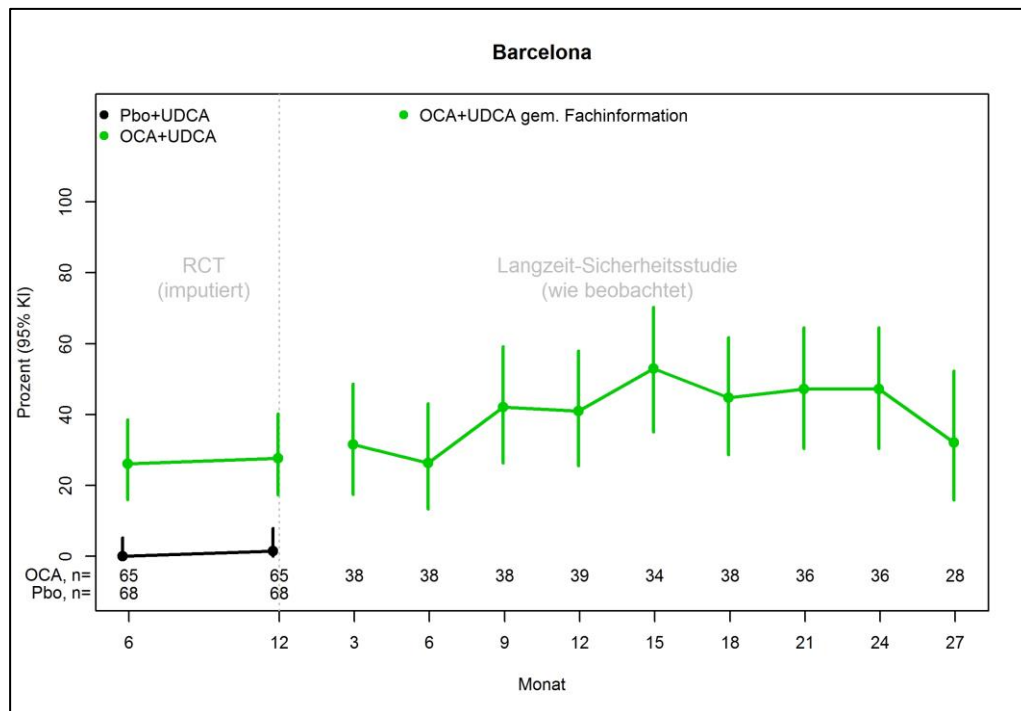


Abbildung 4-26: Ergänzende Ergebnisse für den Anteil Patienten, die das Kriterium BARCELONA erfüllt haben, aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.2) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Auch die ergänzenden Analysen für AP zeigten statistisch signifikante Vorteile für Obeticholsäure+UDCA gegenüber UDCA+Placebo in der Anzahl der Ansprecher, mit Ausnahme des Kriteriums „Rotterdam“. Da nach „Rotterdam“ lediglich die Endpunkte Normalisierung von Bilirubin und/oder Albumin in die Berechnung der Überlebenswahrscheinlichkeit einfließen, in der POISE-Studie allerdings über 90% der Patienten Gesamt-Bilirubinwerte im Normbereich und alle Patienten einen normalen Albuminwert aufwiesen, ist eine Signifikanz für Obeticholsäure+UDCA gegenüber UDCA+Placebo bei Betrachtung des Rotterdam Kriteriums nicht möglich. AP findet innerhalb des Rotterdam Kriteriums keine Berücksichtigung. Im Gegensatz zur AP lassen sich erhöhte Werte für Bilirubin und Albumin typischerweise erst in terminalen Stadien der PBC beobachten. Dies ist therapeutisch zu spät.

Die beobachteten Vorteile von Obeticholsäure+UDCA bleiben auch in der einarmigen Langzeit-Sicherheitsstudie bestehen.

In einer weiteren post hoc definierten Analyse wurde die Anzahl Patienten bestimmt, die insgesamt eine bestimmte Anzahl x der bestehenden Therapieverlaufs- Kriterien (Paris I, Paris II, Mayo II, Rotterdam, Toronto II, Barcelona) erfüllen.

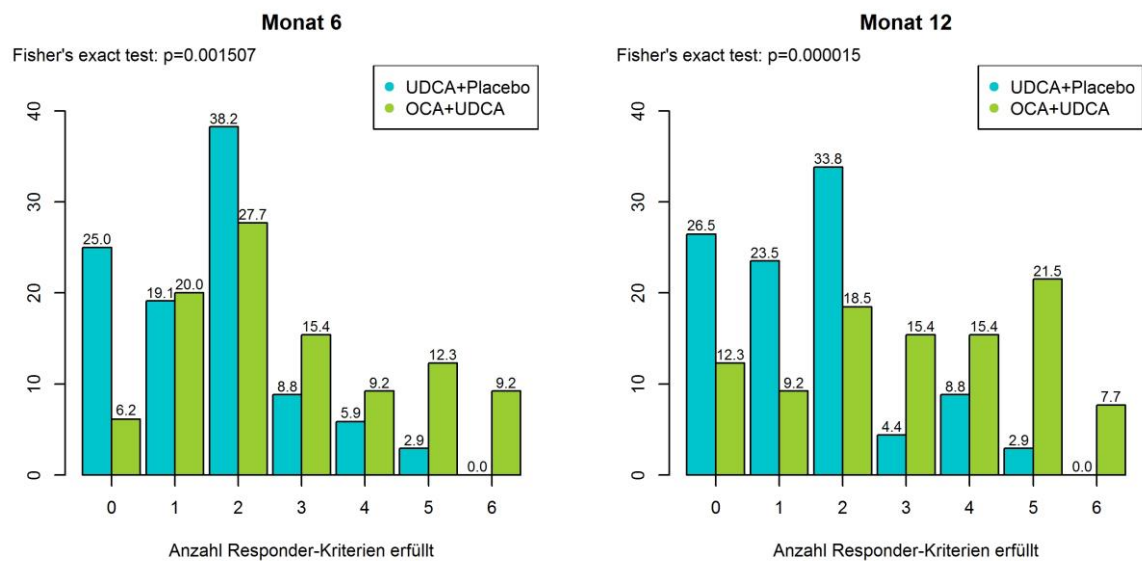


Abbildung 4-27: Ergänzende Analyse der Patienten, die x von sechs möglichen Kriterien für AP und Bilirubin (Paris I, Paris II, Mayo II, Rotterdam, Toronto II, Barcelona) erfüllt haben – Patienten mit Kombinationstherapie aus RCT; Obeticholsäure+UDCA versus UDCA+Placebo

Für Monat 12 zeigte sich ein signifikanter Vorteil für Obeticholsäure+UDCA gegenüber UDCA+Placebo. Insgesamt erfüllten unter Obeticholsäure+UDCA etwa 45% der Patienten vier oder mehr Kriterien, gegenüber der Kontrollgruppe mit 12% der Patienten. Fünf oder mehr Kriterien erfüllten 29% der Patienten mit Obeticholsäure+UDCA gegenüber 3% in der Kontrollgruppe. Weniger als zwei Kriterien erfüllten unter Obeticholsäure+UDCA 22% der Patienten, gegenüber 50% in der Kontrollgruppe. Die Analyse zum Monat 6 zeigte dazu konsistente Ergebnisse.

Damit ergab auch diese Analyse signifikante und relevante Vorteile für Obeticholsäure+UDCA gegenüber der ZVT.

Patienten mit Monotherapie

In der folgenden Tabelle wurden die Ergebnisse des primären Endpunkts der POISE-Studie dargestellt.

Tabelle 4-26: Ergebnisse für den primären Endpunkt (Anteil Patienten mit AP $<1,67 \times$ ULN, Gesamt-Bilirubin $\leq$ ULN und AP-Reduktion von $\geq 15\%$) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Monat	Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
6	2/5 (40)	0/5 (0)	5 [0,3; 83,69] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,0833
12	2/5 (40)	0/5 (0)	5 [0,3; 83,69] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,0833
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test							

In der Teilpopulation von Patienten mit Monotherapie ergaben sich Trends für einen Vorteil von Obeticholsäure gegenüber Placebo. Auf Grund des kleinen Stichprobenumfangs wurde hier keine Signifikanz erreicht.

Tabelle 4-27: Ergebnisse für die Einzelkomponenten des primären Endpunkts aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Monat	Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
AP $< 1,67 \times$ ULN							
6	2/5 (40)	0/5 (0)	5 [0,3; 83,69] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,0833
12	2/5 (40)	0/5 (0)	5 [0,3; 83,69] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,0833
Bilirubin $\leq$ ULN							
6	5/5 (100)	4/5 (80)	1,22 [0,73; 2,06] ^(a)	0,82 [0,49; 1,38] ^(a)	3,67 [0,12; 113,73] ^(a)	-0,2 [-0,55; 0,15] ^(a)	0,3173
12	5/5 (100)	3/5 (60)	1,57 [0,77; 3,22] ^(a)	0,64 [0,31; 1,3] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,1573
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test ULN = Upper limit of normal (Obere Grenze des Normalbereichs)							

Monat	Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
AP-Reduktion $\geq 15\%$ siehe Tabelle 4-28.							

Auch in den drei Einzelkomponenten zeigten sich Vorteile von Obeticholsäure als Trend gegenüber Placebo. Ein numerischer und im OR und RD signifikanter Vorteil zeigte sich dabei hinsichtlich der Komponente der prozentualen Reduktion der AP zu Monat 12 in drei bzw. allen vier Kategorien.

Tabelle 4-28: Ergebnisse für den Anteil Patienten mit AP-Reduktion zu Baseline aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Monat	Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
AP-Reduktion $\geq 10\%$							
6	3/5 (60)	1/5 (20)	3 [0,45; 19,93] ^(a)	0,33 [0,05; 2,21] ^(a)	6,44 [0,35; 119,17]	-0,38 [-0,95; 0,19]	0,2429
12	4/5 (80)	0/5 (0)	9 [0,61; 133,08] ^(a)	0,11 [0,01; 1,64] ^(a)	33 [1,06; 1023,56] ^(a)	-0,8 [-1,15; -0,45] ^(a)	0,0195
AP-Reduktion $\geq 15\%$							
6	3/5 (60)	0/5 (0)	7 [0,45; 108,26] ^(a)	0,14 [0,01; 2,21] ^(a)	15,4 [0,56; 425,53] ^(a)	-0,6 [-1,03; -0,17] ^(a)	0,0495
12	4/5 (80)	0/5 (0)	9 [0,61; 133,08] ^(a)	0,11 [0,01; 1,64] ^(a)	33 [1,06; 1023,56] ^(a)	-0,8 [-1,15; -0,45] ^(a)	0,0195
AP-Reduktion $\geq 20\%$							
6	3/5 (60)	0/5 (0)	7 [0,45; 108,26] ^(a)	0,14 [0,01; 2,21] ^(a)	15,4 [0,56; 425,53] ^(a)	-0,6 [-1,03; -0,17] ^(a)	0,0495
12	4/5 (80)	0/5 (0)	9 [0,61; 133,08] ^(a)	0,11 [0,01; 1,64] ^(a)	33 [1,06; 1023,56] ^(a)	-0,8 [-1,15; -0,45] ^(a)	0,0195
AP-Reduktion $\geq 40\%$							
6	2/5 (40)	0/5 (0)	5 [0,3; 83,69] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,1573
12	3/5 (60)	0/5 (0)	7 [0,45; 108,26] ^(a)	0,14 [0,01; 2,21] ^(a)	15,4 [0,56; 425,53] ^(a)	-0,6 [-1,03; -0,17] ^(a)	0,0495
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test							

Tabelle 4-29: Ergebnisse für den Anteil Patienten mit $AP \leq ULN$ aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Monat	Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
6	1/5 (20)	0/5 (0)	3 [0,15; 59,89] ^(a)	0,33 [0,02; 6,65] ^(a)	3,67 [0,12; 113,73] ^(a)	-0,2 [-0,55; 0,15] ^(a)	0,3173
12	0/5 (0)	0/5 (0)	1 [0,02; 42,65] ^(a)	1 [0,02; 42,65] ^(a)	1 [0,02; 59,99] ^(a)	0 [-0,01; 0,01] ^(a)	NA
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test ULN = Upper limit of normal (Obere Grenze des Normalbereichs) NA = nicht berechenbar							

Auch in der Teilpopulation der Patienten mit Monotherapie gab es nur einen Patienten, der das Kriterium von $AP \leq ULN$ erreicht hatte.

Tabelle 4-30: Ergebnisse für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
AP, abs. Veränderung							
5	485,7 (234)	-120,43 (20,81)	5	366,7 (125)	11,2 (30,57)	-131,63 [-207,81; -55,45] 0,0013	-2,01 [-3,53; -0,49]
Gesamt-Bilirubin, abs. Veränderung							
5	8,8 (3,0)	0,17 (0,66)	5	12,6 (3,5)	-0,36 (0,77)	0,53 [-1,65; 2,72] 0,6223	0,28 [-0,96; 1,53]
Direktes Bilirubin, abs. Veränderung							
5	2,8 (1,1)	0,63 (0,34)	5	5,8 (2,3)	-0,13 (0,41)	0,76 [-0,46; 1,97] 0,2138	0,72 [-0,56; 2,00]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

In den absoluten Veränderungen zum Vorwert zeigten sich keine signifikanten Änderungen für Bilirubin, wohl aber eine signifikante und klinisch relevante Reduktion von AP zu Gunsten von Obeticholsäure.

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.

Für beide Teilpopulationen lag exakt eine Studie mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.1.3 Morbidität: 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 4-31: Operationalisierung von Fibrose (Morbidität)

Studie	Operationalisierung
POISE	Die Fibrosierung der Leber wurde mittels ELF-Score (14) ermittelt. Dieser misst anhand von Reagenzien die Fibrosierung der Leber als Surrogat für die mittels Biopsie ermittelten Fibrose. Die Fibrosierung der Leber wurde zudem mittels transienter Elastografie (Fibroscan®) in den Zentren erhoben, in denen das Medizinprodukt (Echosens, Paris, Frankreich) verfügbar war. Der Fibroscan® ist ein validiertes, nicht-invasives Verfahren (15) im Rahmen der TE zur Bestimmung der Leberfibrose. Für den ELF-Score und dessen Komponenten sowie den Fibroscan® wurde die mittlere Änderung zu Baseline mittels eines gemischten Modells für Wiederholungsmessungen (MMRM) berechnet. Als Basis diente die ITT-Population. Dabei wurden die Faktoren "AP > 3x ULN und/oder AST > 2x ULN und/oder Bilirubin > ULN zu Baseline" und "Intoleranz gegenüber UDCA" als Strata sowie der Baselinewert als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde berechnet auf Basis der Kleinste-Quadrate (Least-Square)-Methode. Die LS-Mittelwerte pro Behandlungsarm und die Differenz der LS-Mittelwerte wurden dargestellt. Zudem wurde die standardisierte Mittelwertdifferenz (SMD) mit 95% Konfidenzintervall berechnet. Zusätzlich wurde die Präsenz von Fibrose (Grenze für die Lebersteifigkeit $\geq 17,3 \text{ kPa}$ (30) als Ansprecheranalyse auf Basis des Fibroscan® durchgeführt. Die Analyse erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline vorlagen. Für die Ansprecheranalyse wurden alle Patienten eingeschlossen, für die ein Fibroscan® zum Monat 12 vorlag.

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

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

Tabelle 4-32: Bewertung des Verzerrungspotenzials für Fibrose (Morbidität) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
ELF-Score						
POISE	niedrig	ja	nein	ja	nein	hoch
Fibroscan®						
POISE	niedrig	ja	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Der ELF-Score und dessen Komponenten wurden in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die mindestens der Baseline-Wert und ein Beobachtungswert nach Baseline vorlagen. Das ITT-Prinzip wurde nicht adäquat umgesetzt, da in der Teilpopulation der Patienten mit Kombinationstherapie in der Obeticholsäure Gruppe nur 59 von 65 (91%) der Patienten in die Analyse eingingen. In der Placebo-Gruppe waren es 64 von 68 (94%). In der Teilpopulation der Patienten mit Monotherapie wurden nur 4 von 5 Patienten der Obeticholsäure-Gruppe und 3 von 5 Placebo-Patienten in die Analyse aufgenommen. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da der ELF-Score ein validiertes Surrogat für die mittels Biopsie bestimmte Zirrhose ist, wird das Verzerrungspotenzial für diese Endpunkte als hoch bewertet (siehe Tabelle 4-32 und Anhang 4-F).

Der Fibroscan® wurde in einer Teilpopulation von Patienten erhoben und dokumentiert. Es lag keine informative Selektion vor, sondern eine zufällige Auswahl, da der Fibroscan® nur in einigen Zentren dokumentiert wurde. Sonstige Aspekte gab es nicht. Durch die Einschränkungen bezüglich der ITT-Population (64 von 133 Patienten mit Fibroscan®) wurde das Verzerrungspotenzial als hoch eingestuft (siehe Tabelle 4-32 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Tabelle 4-33: Ergebnisse für ELF-Score aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
ELF-Score, abs. Veränderung							
59	9,7 (1)	0,14 (0,06)	64	9,9 (1,2)	0,14 (0,05)	0 [-0,15; 0,15] 0,9843	0 [-0,36; 0,35]
HA (ng/mL), abs. Veränderung							
59	104,1 (132,5)	20,3 (8,69)	64	122,9 (126)	20,95 (8,19)	-0,64 [-23,61; 22,33] 0,9562	-0,01 [-0,36; 0,34]
P3NP (ng/mL), abs. Veränderung							
59	10,4 (5,1)	-0,34 (0,32)	64	11,5 (5,4)	0,14 (0,3)	-0,48 [-1,33; 0,37] 0,2669	-0,2 [-0,55; 0,16]
TIMP-1 (ng/mL), abs. Veränderung							
59	282,1 (101,9)	31,27 (84,43)	64	304,1 (109,7)	50,33 (84,39)	-19,06	-0,26

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
						[-44,61; 6,49] 0,1430	[-0,62; 0,09]

MW = Mittelwert
 SD = Standardabweichung
 SE = Standardfehler
 MWD = Mittelwertdifferenz
 SMD = Standardisierte Mittelwertdifferenz
 KI = Konfidenzintervall
 LS = Least-square
 Pbo = Placebo

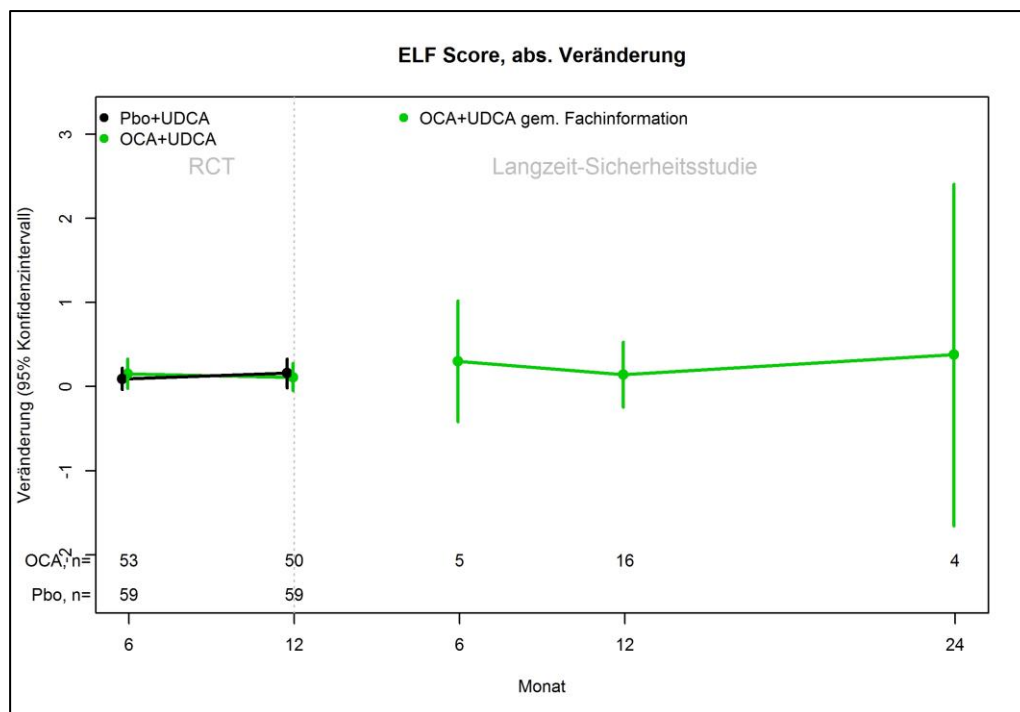


Abbildung 4-28: Ergebnisse für ELF-Score, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

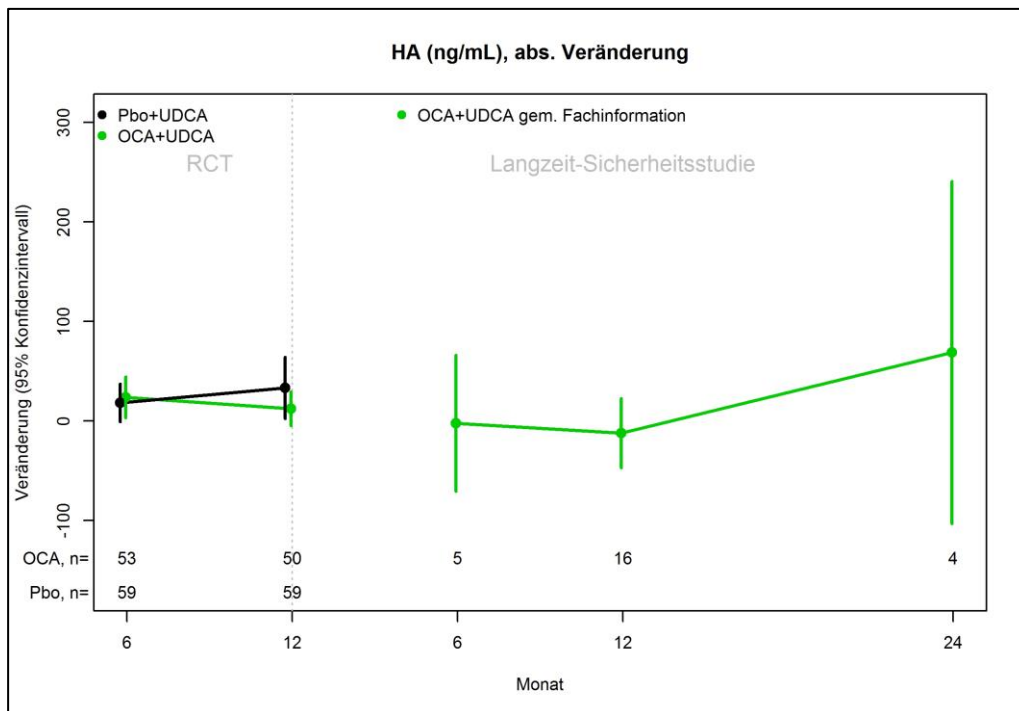


Abbildung 4-29: Ergebnisse für HA (ng/mL), absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

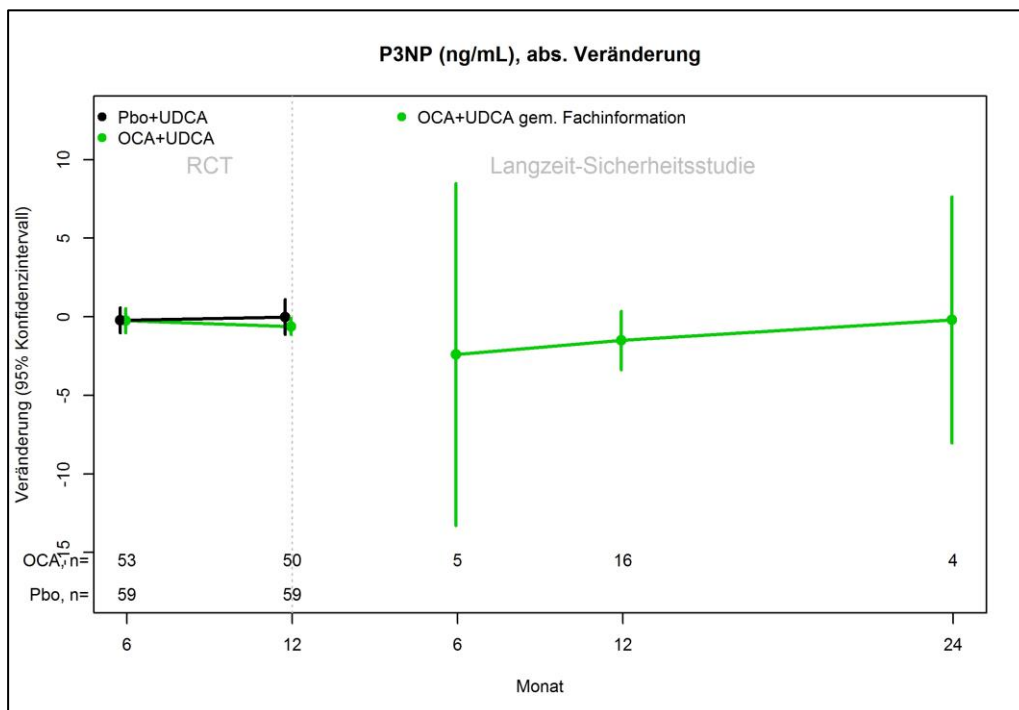


Abbildung 4-30: Ergebnisse für P3NP (ng/mL), absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

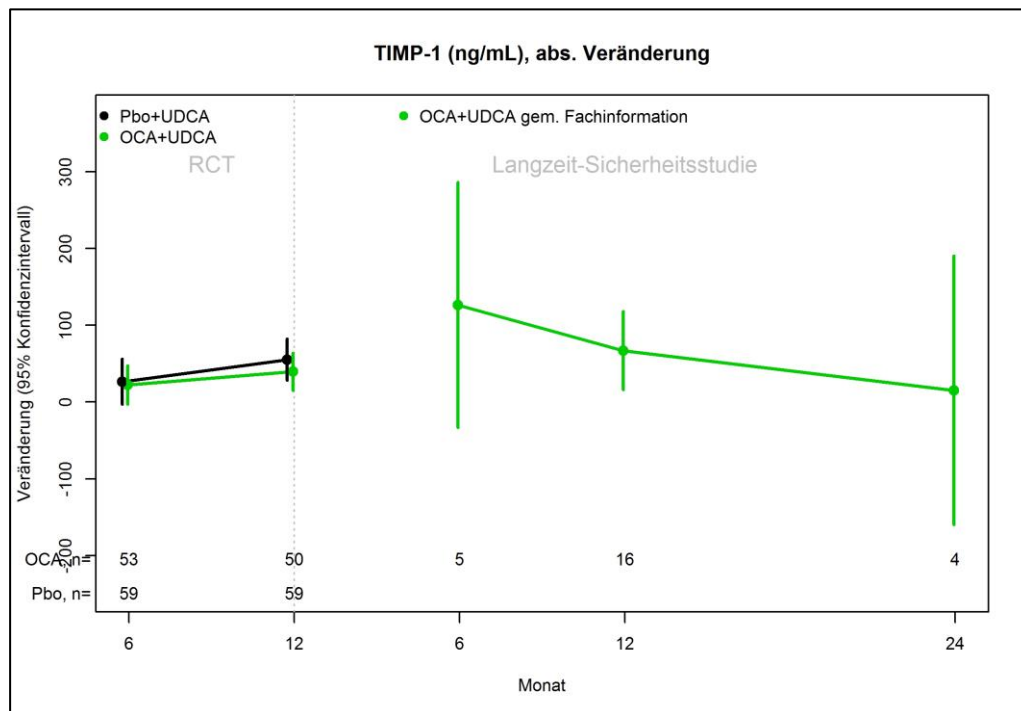


Abbildung 4-31: Ergebnisse für TIMP-1 (ng/mL), absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

In der Teilpopulation der Patienten mit Kombinationstherapie zeigten sich im ELF-Score und in dessen Einzelkomponenten keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen.

Tabelle 4-34: Ergebnisse für Fibrose anhand des Fibroscan® (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
Leber-Steifigkeit, abs. Veränderung							
30	10,2 (7,5)	0,47 (0,52)	32	12,6 (11,2)	0,73 (0,56)	-0,26 [-1,72; 1,2] 0,7230	-0,09 [-0,59; 0,41]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

Wegen der geringen Anzahl an vorliegenden Zeitpunkten (RCT Monat 12, Langzeit-Sicherheitsstudie Monat 12 und 24) wurde keine Abbildung für den Endpunkt Leber-Steifigkeit (absolute Veränderung zum Vorwert) generiert.

Im Fibroscan® lag der Punktschätzer für die Fibrose unterhalb von Null und zeigte einen numerischen Vorteil zu Gunsten von Obeticholsäure; allerdings fand sich kein relevanter oder signifikanter Unterschied.

Tabelle 4-35: Ergebnisse der Ansprecheranalysen für Fibrose anhand des Fibroscan® (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Monat	Ereignisse/N (%) OCA+ UDCA	Ereignisse/N (%) UDCA+ Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Leber-Steifigkeit $\geq 17,3$ kPa							
12	5/35 (14.29)	8/37 (21.62)	0,65 [0,23; 1,83]	1,53 [0,55; 4,26]	0,6 [0,18; 2,06]	0,07 [-0,1; 0,25]	0,4229
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test kPa = Kilopascal							

Die Analyse aller Patienten mit einem Fibroscan® zu Monat 12 zeigte einen numerischen Vorteil für Obeticholsäure+UDCA im Vergleich zur Kontrolle, der jedoch nicht statistisch signifikant war.

Patienten mit Monotherapie

Tabelle 4-36: Ergebnisse für ELF-Score aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
ELF-Score, abs. Veränderung							
4	10,1 (1,2)	-0,17 (0,19)	3	9,8 (0,4)	0,33 (0,23)	-0,5 [-1,2; 0,2] 0,1325	-1,12 [-2,73; 0,49]
HA (ng/mL), abs. Veränderung							
4	116,5 (96)	7,66 (128,64)	3	71,1 (7,2)	43,96 (129,96)	-36,31 [-139,11; 66,5] 0,4207	-0,56 [-2,08; 0,97]
P3NP (ng/mL), abs. Veränderung							
4	11,2 (7,7)	-1,14 (0,66)	3	10 (2,5)	-1,06 (0,78)	-0,08 [-2,49; 2,33] 0,9385	-0,05 [-1,55; 1,45]
TIMP-1 (ng/mL), abs. Veränderung							
4	389,3 (191)	-12,07 (71,9)	3	252 (124,4)	88,08 (97,18)	-100,15 [-401,71; 201,41] 0,4475	-0,52 [-2,04; 1,00]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

In der Teilpopulation der Patienten mit Monotherapie zeigten sich im ELF-Score und in dessen Einzelkomponenten keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen.

Tabelle 4-37: Ergebnisse für Fibrose anhand des Fibroscan® (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
Leber-Steifigkeit, abs. Veränderung							
2	4,6 (1,9)	-32,1 (0)	2	13,1 (1,8)	36,5 (0)	-68,6 [NA; NA] NA	NA [NA; NA]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz							

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo NA = nicht berechenbar							

Tabelle 4-38: Ergebnisse der Ansprecheranalysen für Fibrose anhand des Fibroscan® (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Monat	Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Leber-Steifigkeit $\geq 17,3$ kPa							
12	0/3 (0)	0/2 (0)	0,75 [0,02; 28,08] ^(a)	1,33 [0,04; 49,93] ^(a)	0,71 [0,01; 49,71] ^(a)	0 [-0,01; 0,01] ^(a)	NA
(a) Modell ohne Stratifikations-Faktoren							
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test kPa = Kilopascal							

Für den Fibroscan® lagen insgesamt nur Daten von 4 bzw. 5 Patienten vor, so dass keine Bewertung der Ergebnisse vorgenommen wurde.

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.

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.1.4 Morbidität: Osteoporose – RCT

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

Tabelle 4-39: Operationalisierung von Osteoporose (Morbidität)

Studie	Operationalisierung
POISE	Der DEXA-Scan ist eine nicht-invasive Methode und wurde zur Bestimmung der Osteoporose bei postmenopausalen Frauen verwendet. In einer Teilpopulation der POISE-Studie wurde der DEXA-Scan zur Messung der Dichte des femoralen Nackenknöchens und der Lendenwirbelsäule verwendet. Die Ergebnisse des DEXA-Scans wurden als T-Score und Z-Score zu Baseline und Monat 12 berichtet. Für T-Score, Z-Score und Knochendichte wurde jeweils die mittlere Änderung zu Baseline mittels eines gemischten Modells für Wiederholungsmessungen (MMRM) berechnet. Als Basis diente der Teil der Sicherheitspopulation, in dem der DEXA-Scan durchgeführt wurde. Dabei wurden die Faktoren "AP > 3x ULN und/oder AST > 2x ULN und/oder Bilirubin > ULN zu Baseline" und "Intoleranz gegenüber UDCA" als Strata sowie der Baselinewert als Co-Faktor in der Berechnung berücksichtigt. Für den Ausgangswert wurden der Mittelwert und die Standardabweichung berechnet. Die mittlere Änderung wurde berechnet auf Basis der Kleinst-Quadrat-(Least-Square-)Methode und die LS-Mittelwerte mit Standardfehler pro Behandlungsarm sowie die Differenz der LS-Mittelwerte dargestellt. Zudem wurde die standardisierte Mittelwertdifferenz (SMD) mit 95%-Konfidenzintervall berechnet.

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

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

Tabelle 4-40: Bewertung des Verzerrungspotenzials für Osteoporose (Morbidität) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
DEXA Scan						
POISE	niedrig	ja	nein	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig eingestuft wurde.

Die Studie wurde verblindet durchgeführt. Der DEXA-Scan wurde in einer Teilpopulation von Patienten erhoben und dokumentiert. Die Analyse erfolgte auf Basis der Sicherheitspopulation. Es lag keine informative Selektion vor, sondern eine zufällige Auswahl, da der DEXA-Scan nur in einigen Zentren dokumentiert wurde. Sonstige Aspekte liegen nicht vor. Durch die Einschränkungen bezüglich der ITT-Population wurde das Verzerrungspotenzial als hoch eingestuft (siehe Tabelle 4-40 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Tabelle 4-41: Ergebnisse für Osteoporose anhand des DEXA-Scan (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel in der Sicherheitspopulation – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
Schenkelhals-Scan: T-Score, abs. Veränderung							
37	-1,2 (0,9)	-0,02 (0,1)	33	-1,1 (1,3)	-0,31 (0,1)	0,29 [0,01; 0,57] 0,0396	0,5 [0,02; 0,97]
Schenkelhals-Scan: Z-Score, abs. Veränderung							
36	-0,2 (0,9)	0 (0,06)	32	-0,3 (1)	-0,11 (0,06)	0,11 [-0,05; 0,28] 0,1631	0,34 [-0,14; 0,82]
Schenkelhals-Scan: Knochendichte (g/cm²), abs. Veränderung							
36	0,8 (0,1)	0 (0,01)	31	0,8 (0,1)	-0,01 (0,01)	0,01 [-0,01; 0,04] 0,3556	0,23 [-0,26; 0,71]
Lendenwirbelsäulen-Scan: T-Score, abs. Veränderung							
39	-1,1 (1,6)	-0,07 (0,12)	40	-1,1 (1,6)	-0,31 (0,12)	0,24 [-0,08; 0,56] 0,1374	0,33 [-0,11; 0,78]
Lendenwirbelsäulen-Scan: Z-Score, abs. Veränderung							
38	-0,1 (1,8)	-0,01 (0,06)	36	-0,3 (1,2)	0,03 (0,07)	-0,04 [-0,21; 0,14] 0,6939	-0,09 [-0,55; 0,37]
Lendenwirbelsäulen-Scan: Knochendichte (g/cm²), abs. Veränderung							
38	1 (0,2)	-0,02 (0,01)	38	1 (0,2)	-0,01 (0,01)	-0,01 [-0,03; 0,01] 0,4994	-0,15 [-0,6; 0,3]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz							

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
KI = Konfidenzintervall LS = Least-square Pbo = Placebo g = Gramm cm ² = Quadratcentimeter							

Wegen der geringen Anzahl der vorliegenden Zeitpunkten (RCT Monat 12) wurde keine Abbildung für die Endpunkte Osteoporose anhand des DEXA-Scan vorgenommen.

In der Teilpopulation der Patienten mit Kombinationstherapie zeigten sich im DEXA-Scan keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen.

Patienten mit Monotherapie

Tabelle 4-42: Ergebnisse für Osteoporose anhand des DEXA-Scan (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel in der Sicherheitspopulation – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
Schenkelhals-Scan: T-Score, abs. Veränderung							
3	-1,1 (1,5)	-0,14 (0,02)	3	-1,4 (1)	-0,12 (0,02)	-0,03 [-0,12; 0,07] 0,3363	-0,82 [-2,49; 0,85]
Schenkelhals-Scan: Z-Score, abs. Veränderung							
3	0 (0,6)	-0,03 (0,06)	2	-0,6 (0,5)	-0,07 (0,07)	0,04 [-1,27; 1,36] 0,7442	0,28 [-1,52; 2,08]
Schenkelhals-Scan: Knochendichte (g/cm²), abs. Veränderung							
3	0,8 (0,1)	-0,01 (0)	2	0,7 (0,2)	-0,01 (0)	0 [-0,08; 0,08] 0,8675	0,14 [-1,65; 1,93]
Lendenwirbelsäulen-Scan: T-Score, abs. Veränderung							
3	-0,2 (1,7)	0,1 (0,1)	3	-1,9 (1,1)	-0,05 (0,1)	0,14 [-0,49; 0,78] 0,4339	0,63 [-1,01; 2,27]
Lendenwirbelsäulen-Scan: Z-Score, abs. Veränderung							
3	0,9 (2,1)	0,15 (0,22)	2	-0,9 (0,5)	-0,06 (0,25)	0,21 [-4,59; 5,02] 0,6734	0,37 [-1,43; 2,18]
Lendenwirbelsäulen-Scan: Knochendichte (g/cm²), abs. Veränderung							
3	1,1 (0,2)	0,01 (0,02)	2	0,9 (0,2)	-0,01 (0,02)	0,01 [-0,41; 0,44] 0,7416	0,29 [-1,51; 2,08]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo g = Gramm cm ² = Quadratzentimeter							

Für den DEXA-Scan lagen nur Daten von 5 Patienten vor, so dass keine Bewertung der Ergebnisse durchführbar war.

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.

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.1.5 Morbidität: 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 4-43: Operationalisierung von PBC-40 (Morbidität)

Studie	Operationalisierung
POISE	Morbidität wurde anhand des validierten Fragebogens PBC-40 bewertet (16). Der Fragebogen PBC-40 besteht aus 40 Items in sechs Subdomänen, von denen die Subdomänen Generelle Symptome, Jucken, Fatigue und kognitive Funktion als Ausprägungen der Morbidität dargestellt sind. Den Antworten zu den Fragen werden sogenannte Score-Punkte zugeordnet (1=überhaupt nicht, 2=selten, 3=manchmal, 4=häufig, 5=immer). Morbidität wurde jeweils bei Einschluss (an Tag 0), nach 2 Wochen, 3, 6, 9 und 12 Monaten, vor Verabreichung der Studienmedikation evaluiert. Der Unterschied zwischen den Behandlungsgruppen hinsichtlich Morbidität wurde anhand der Differenz der Score-Änderung zur Baseline (mit 95% KI) in einem gemischten Modell mit Wiederholungsmessungen (MMRM) berechnet. Dabei wurden die Faktoren "AP > 3x ULN und/oder AST > 2x ULN und/oder Bilirubin > ULN zu Baseline" und "Intoleranz gegenüber UDCA" als Strata sowie der Baselinewert des Scores als Co-Faktor in der Berechnung des Kleinste-Quadrate Mittelwerts und des zugehörigen 95%-Konfidenzintervalls berücksichtigt. Die Analyse erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die der Baseline-Wert sowie mindestens ein Beobachtungswert nach Baseline vorlagen.

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

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

Tabelle 4-44: Bewertung des Verzerrungspotenzials für PBC-40 (Morbidität) in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig eingestuft wurde.

Die Studie wurde verblindet durchgeführt. Die Analyse erfolgte auf Basis der ITT-Population, in die alle Patienten aufgenommen wurden, die der in der Fachinformation beschriebenen (zugelassenen) Population entsprachen. Für beide Teilpopulationen wurden mehr als 95% der Patienten in die ITT-Analyse eingeschlossen. Es lagen keine sonstigen Aspekte einer möglichen Verzerrung vor. Daher wurde das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (siehe Tabelle 4-44 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Tabelle 4-45: Rücklaufquoten für PBC-40 (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

	Generelle Symptome		Jucken		Fatigue		Kognitive Funktion	
	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo
Woche 2	56/65 (86.2%)	61/68 (89.7%)	55/65 (84.6%)	61/68 (89.7%)	56/65 (86.2%)	61/68 (89.7%)	56/65 (86.2%)	61/68 (89.7%)
Monat 3	58/65 (89.2%)	60/68 (88.2%)	57/65 (87.7%)	59/68 (86.8%)	58/65 (89.2%)	61/68 (89.7%)	58/65 (89.2%)	61/68 (89.7%)
Monat 6	58/65 (89.2%)	62/68 (91.2%)	57/65 (87.7%)	62/68 (91.2%)	58/65 (89.2%)	62/68 (91.2%)	58/65 (89.2%)	62/68 (91.2%)
Monat 9	58/65 (89.2%)	62/68 (91.2%)	57/65 (87.7%)	62/68 (91.2%)	58/65 (89.2%)	62/68 (91.2%)	58/65 (89.2%)	62/68 (91.2%)
Monat 12	56/65	61/68	56/65	61/68	56/65	61/68	56/65	61/68

	Generelle Symptome		Jucken		Fatigue		Kognitive Funktion	
	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo
	(86.2%)	(89.7%)	(86.2%)	(89.7%)	(86.2%)	(89.7%)	(86.2%)	(89.7%)
Pbo = Placebo								

Tabelle 4-46: Ergebnisse für PBC-40 (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
PBC-40 Generelle Symptome, abs. Veränderung							
64	13,6 (5,2)	-0,61 (0,17)	67	13,6 (5,3)	-0,3 (0,16)	-0,31 [-0,76; 0,13] 0,1699	-0,24 [-0,58; 0,11]
PBC-40 Jucken, abs. Veränderung							
63	4 (3,4)	0,55 (0,15)	67	4,3 (3,8)	0,27 (0,14)	0,29 [-0,11; 0,68] 0,1549	0,25 [-0,1; 0,59]
PBC-40 Fatigue, abs. Veränderung							
64	25,5 (10,9)	-0,83 (0,32)	67	26,3 (11,2)	-1,93 (0,31)	1,1 [0,25; 1,94] 0,0110	0,44 [0,10; 0,79]
PBC-40 Kognitive Funktion, abs. Veränderung							
64	11,1 (5,4)	0,16 (0,18)	67	11,5 (5,9)	-0,08 (0,17)	0,24 [-0,24; 0,72] 0,3226	0,17 [-0,17; 0,52]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

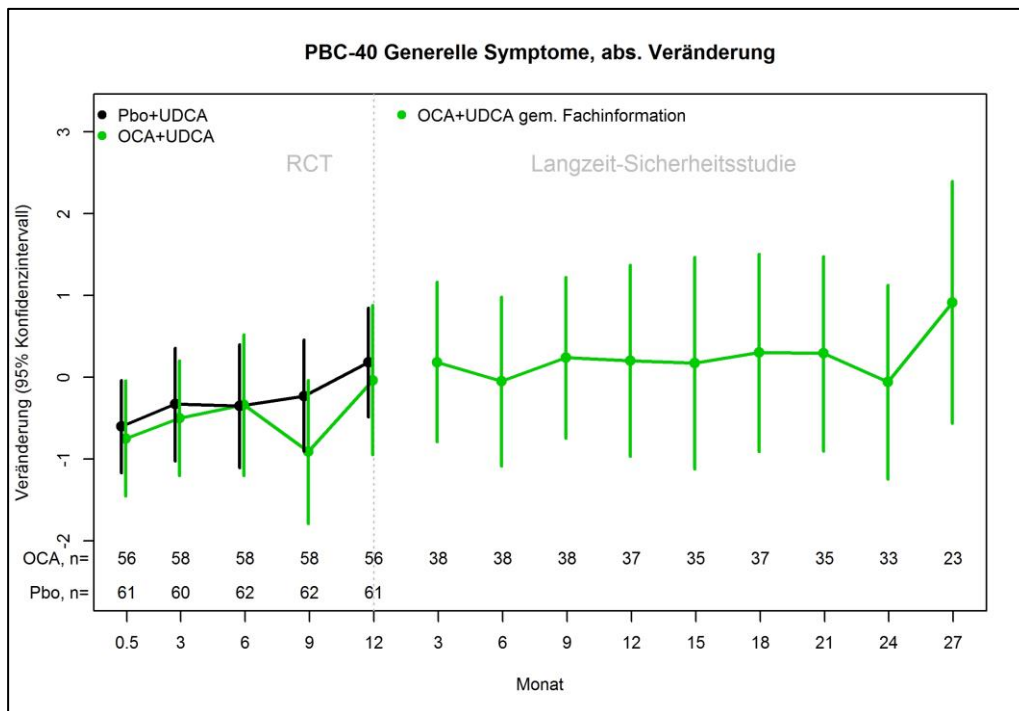


Abbildung 4-32: Ergebnisse für PBC-40 generelle Symptome, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

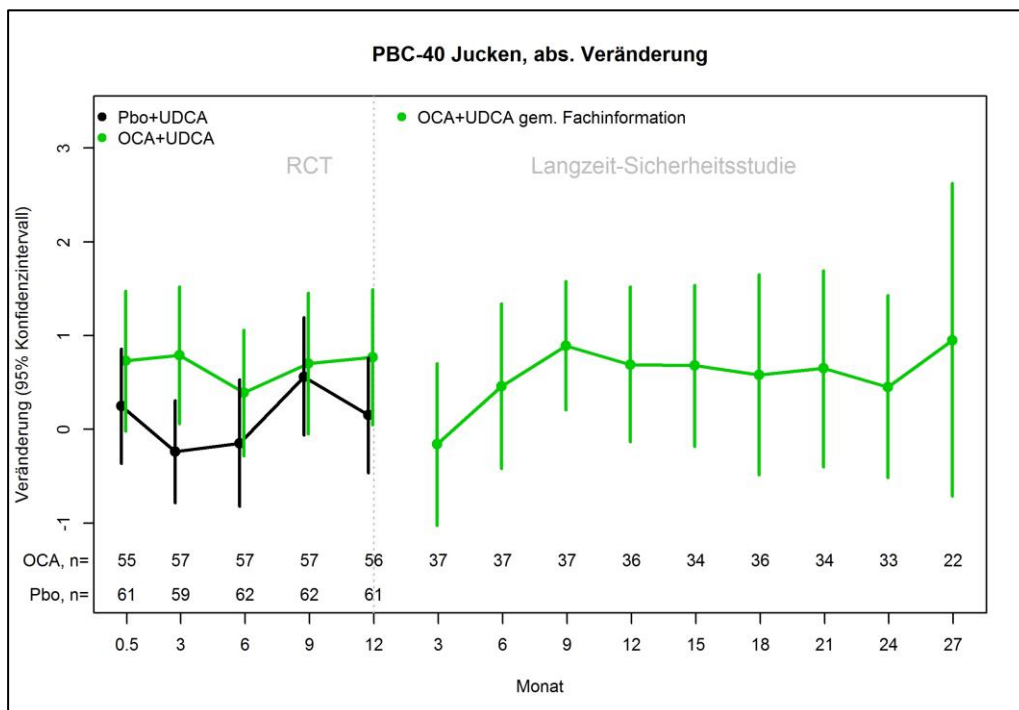


Abbildung 4-33: Ergebnisse für PBC-40 Jucken, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

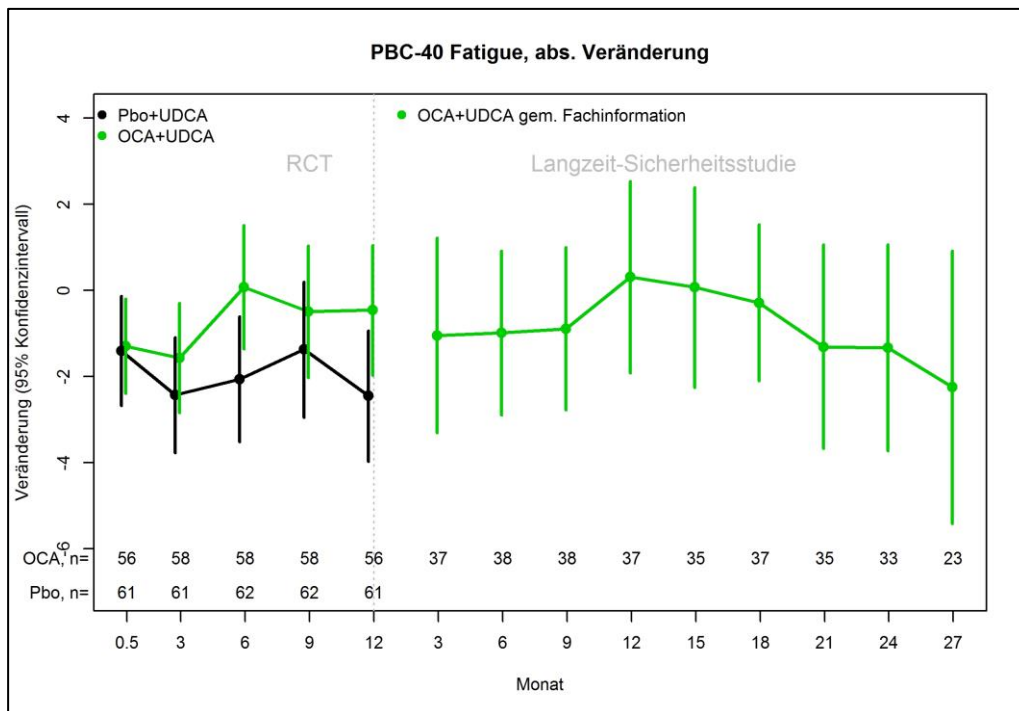


Abbildung 4-34: Ergebnisse für PBC-40 Fatigue, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

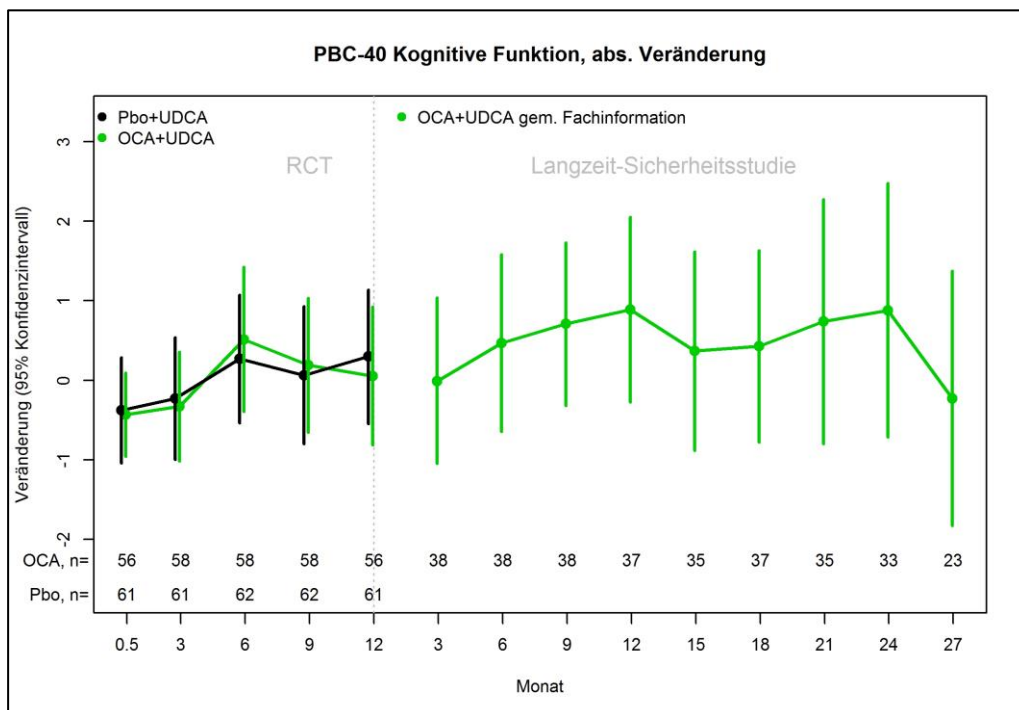


Abbildung 4-35: Ergebnisse für PBC-40 kognitive Funktion, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.5) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Für die Patienten mit Kombinationstherapie zeigte sich in keiner der vier Domänen des PBC-40 bzgl. Symptomen ein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen. Die 95%-Konfidenzintervalle der SMD überlappten mit dem Bereich von -0,2 bis 0,2. In der Langzeitstudie zeigen sich sehr ähnliche Ergebnisse wie in der RCT-Phase der POISE-Studie.

Patienten mit Monotherapie

Tabelle 4-47: Ergebnisse für PBC-40 (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
PBC-40 Generelle Symptome, abs. Veränderung							
5	14,6 (5,9)	0,6 (0,64)	5	14 (5,1)	1,54 (0,71)	-0,94 [-2,4; 0,53] 0,2028	-0,74 [-2,02; 0,54]
PBC-40 Jucken, abs. Veränderung							
5	4,6 (5,5)	1,98 (0,55)	5	4 (2,2)	-1,65 (0,68)	3,62 [1,87; 5,38] 0,0002	2,40 [0,77; 4,02]
PBC-40 Fatigue, abs. Veränderung							
5	28,4 (7,2)	-2,48 (1,18)	5	24,6 (11,8)	0,08 (1,47)	-2,57 [-6,52; 1,39] 0,1958	-0,75 [-2,04; 0,53]
PBC-40 Kognitive Funktion, abs. Veränderung							
5	13,2 (5,1)	-0,95 (0,74)	5	13,8 (6,6)	1,85 (0,86)	-2,8 [-4,8; -0,81] 0,0074	-1,63 [-3,06; -0,20]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

Für die Patienten mit Monotherapie zeigte sich in den Domänen "Generelle Symptome", "Fatigue" und "Kognitive Funktion" des PBC-40 bzgl. Symptomen kein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen. Die 95%-Konfidenzintervalle der SMD überlappten mit dem Bereich von -0,2 bis 0,2. Für das Symptom "Jucken" zeigte sich ein signifikanter Nachteil für Obeticholsäure. Diese Beobachtung deckt sich mit den Beobachtungen zum Pruritus-Fragebogen 5-D und der Pruritus-VAS. Zu berücksichtigen ist dabei die sehr kleine Stichprobengröße von 5 gegen 5 Patienten, die keine valide Aussage zum Zusatznutzen oder -schaden zulässt.

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.

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.1.6 Morbidität: 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 4-48: Operationalisierung von Pruritus (Morbidität)

Studie	Operationalisierung
POISE	Der Endpunkt Pruritus wurde anhand des unerwünschten Ereignisses Pruritus und dem Fragebogen 5-D sowie Pruritus-VAS erhoben. UE Pruritus Dichotomer Endpunkt hinsichtlich des Auftretens eines unerwünschten Ereignisses (UE, treatment-emergent adverse event) innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt dokumentiert wurde. Die Analyse basiert auf der Sicherheitspopulation, in die alle Patienten eingeschlossen sind, denen mindestens eine Dosis einer Studienbehandlung verabreicht wurde. Die Gruppenzuteilung erfolgte auf Basis der Studienmedikation (Obeticholsäure+UDCA vs. UDCA+Placebo bzw. Obeticholsäure vs. Placebo), nicht der Randomisierung. Pruritus-Fragebogen 5-D Zudem wurde Pruritus anhand des validierten Fragebogens 5-D (Descriptive Pruritus Scale) bewertet (12). Der Fragebogen 5-D besteht aus 27 Items in fünf Subdomänen, die als Ausprägungen der Morbidität dargestellt sind. Zudem wurde die Visuelle Analogskala erhoben und dargestellt. Pruritus-VAS Die Pruritus-Visuelle Analogskala (Pruritus-VAS) wurde zur weiteren Untersuchung des Symptoms Pruritus verwendet (13). Die Morbidität wurde jeweils bei Einschluss (an Tag 0), nach 2 Wochen, 3, 6, 9 und 12 Monaten vor Verabreichung der Studienmedikation evaluiert. Analysen Der Unterschied zwischen den Behandlungsgruppen hinsichtlich Morbidität wurde anhand der Differenz der Score-Änderung zur Baseline (mit 95% KI) in einem gemischten Modell mit Wiederholungsmessungen (MMRM) berechnet. Dabei wurden die Faktoren "AP > 3x ULN und/oder AST > 2x ULN und/oder Bilirubin > ULN zu Baseline" und "Intoleranz gegenüber UDCA" als Strata sowie der Baselinewert des Scores als Co-Faktor in der Berechnung des Kleinst-Quadrat Mittelwerts und des zugehörigen 95%-Konfidenzintervalls berücksichtigt. Die Analyse erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline vorlagen. Für das UE Pruritus wurden die Inzidenzraten mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen berechnet. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet. Zudem wurde die mediane Zeit bis zum ersten Ereignis mittels Kaplan-Meier Methode bestimmt und das Hazard Ratio mit 95% KI mittels Cox-Proportional Hazard Modell für den Vergleich der Behandlungsgruppen berechnet.

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 4-49: Bewertung des Verzerrungspotenzials für 5-D und Pruritus-VAS (Morbidität) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UE Pruritus						
POISE	niedrig	ja	ja	ja	ja	niedrig
5-D						
POISE	niedrig	ja	nein	ja	ja	hoch
Pruritus-VAS						
POISE	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig eingestuft wurde.

Die Studie wurde verblindet durchgeführt. Die Analyse erfolgte auf Basis der ITT-Population, in die alle Patienten aufgenommen wurden, die der in der Fachinformation beschriebenen Population entsprachen. In der Teilpopulation der Patienten mit Kombinationstherapie wurden für das UE Pruritus die Sicherheitspopulation herangezogen, die mehr als 95% der ITT abdeckt. Für den 5-D wurden in den Domänen Schweregrad und Verteilung jeweils mehr als 95% der Patienten der ITT-Population in die Analyse eingeschlossen. In den Domänen Dauer, Richtung und Behinderung wurde dieses Kriterium verfehlt. Es lagen keine sonstigen Aspekte einer möglichen Verzerrung vor. Für die Pruritus-VAS wurden mehr als 95% der Patienten mit Kombinationstherapie in die Analyse eingeschlossen. In der Teilpopulation der Patienten mit Monotherapie lagen nur für die 5-D Domäne Richtung für die Placebo-Gruppe Daten von weniger als 95% der ITT-Population vor. Insgesamt wurde das Verzerrungspotenzial für den Endpunkt 5-D daher als hoch und für das UE Pruritus und die Pruritus-VAS als niedrig eingestuft (siehe Tabelle 4-49 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Insgesamt trat in der POISE-Studie im Obeticholsäure+UDCA Arm in 35 von 65 (54%) auf, im Kontrollarm UDCA+Placebo in 28 von 68 (41%). Im Weiteren wird das UE Pruritus dargestellt.

Tabelle 4-50: Ergebnisse für das UE Pruritus (Inzidenzraten) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Patienten mit Pruritus						
35/65 (53,85)	28/68 (41,18)	1,32 [0,92; 1,9]	0,76 [0,53; 1,09]	1,67 [0,84; 3,32]	-0,13 [-0,29; 0,04]	0,1460
Patienten mit schwerem Pruritus						
11/65 (16,92)	5/68 (7,35)	2,31 [0,85; 6,26]	0,43 [0,16; 1,18]	2,58 [0,84; 7,92]	-0,09 [-0,2; 0,02]	0,0916
UE Pruritus, das zum Therapieabbruch führte						
1/65 (1,54)	0/68 (0)	3,14 [0,13; 75,63] ^(a)	0,32 [0,01; 7,69] ^(a)	3,19 [0,13; 79,63] ^(a)	-0,02 [-0,05; 0,01] ^(a)	0,3066
Intervention wegen Pruritus^(b)						
21/35 (60)	14/28 (50)	1,24 [0,8; 1,94]	0,8 [0,52; 1,26]	1,44 [0,52; 3,98]	-0,1 [-0,34; 0,15]	0,4856
(a) Modell ohne Stratifikations-Faktoren						
(b) Basierend nur auf Patienten mit Pruritus						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test UE = unerwünschtes Ereignis						

Tabelle 4-51: Ergebnisse für das UE Pruritus (Time-to-first-event) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Ereignisse/N (%) OCA+UDCA	Median [95% KI] OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	Median [95% KI] UDCA+Pbo	Hazard Ratio [95% KI] p-Wert
Zeit bis zum Beginn des Pruritus				
35/65 (53,8)	191 [56; NA]	28/68 (41,2)	NA [272; NA]	1,51 [0,92; 2,50] 0,1008

Ereignisse/N (%) OCA+UDCA	Median [95% KI] OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	Median [95% KI] UDCA+Pbo	Hazard Ratio [95% KI] p-Wert
Zeit bis zum Beginn des schweren Pruritus				
11/65 (16,9)	NA [NA; NA]	5/68 (7,4)	NA [NA; NA]	2,43 [0,84; 7,00] 0,0895
KI = Konfidenzintervall Pbo = Placebo NA = nicht berechenbar				

In den verschiedenen Operationalisierungen des Endpunktes Pruritus fand sich kein signifikant geringerer oder größerer Schaden von Obeticholsäure+UDCA gegenüber UDCA+Placebo. Im Folgenden sind die Ergebnisse des Fragebogens 5-D und der Pruritus-VAS dargestellt.

Tabelle 4-52: Rücklaufquoten für 5-D Gesamtscore und Pruritus-VAS (Morbidity) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

	5-D Gesamtscore		Pruritus-VAS	
	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo
Woche 2	48/65 (73.8%)	53/68 (77.9%)	54/65 (83.1%)	59/68 (86.8%)
Monat 3	49/65 (75.4%)	55/68 (80.9%)	56/65 (86.2%)	59/68 (86.8%)
Monat 6	51/65 (78.5%)	54/68 (79.4%)	56/65 (86.2%)	59/68 (86.8%)
Monat 9	51/65 (78.5%)	55/68 (80.9%)	56/65 (86.2%)	60/68 (88.2%)
Monat 12	49/65 (75.4%)	53/68 (77.9%)	55/65 (84.6%)	60/68 (88.2%)
Pbo = Placebo				

Tabelle 4-53: Rücklaufquoten für 5-D Einzelskalen aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

	5-D, Dauer		5-D, Schweregrad		5-D, Richtung		5-D, Behinderung		5-D, Verteilung	
	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo
Woche 2	50/65 (76.9%)	55/68 (80.9%)	53/65 (81.5%)	58/68 (85.3%)	50/65 (76.9%)	54/68 (79.4%)	53/65 (81.5%)	57/68 (83.8%)	53/65 (81.5%)	58/68 (85.3%)
Monat 3	51/65 (78.5%)	57/68 (83.8%)	54/65 (83.1%)	57/68 (83.8%)	50/65 (76.9%)	56/68 (82.4%)	53/65 (81.5%)	56/68 (82.4%)	54/65 (83.1%)	57/68 (83.8%)
Monat 6	53/65 (81.5%)	57/68 (83.8%)	56/65 (86.2%)	59/68 (86.8%)	52/65 (80%)	54/68 (79.4%)	55/65 (84.6%)	57/68 (83.8%)	56/65 (86.2%)	59/68 (86.8%)
Monat 9	52/65 (80%)	58/68 (85.3%)	55/65 (84.6%)	58/68 (85.3%)	52/65 (80%)	56/68 (82.4%)	54/65 (83.1%)	57/68 (83.8%)	55/65 (84.6%)	59/68 (86.8%)

	5-D, Dauer		5-D, Schweregrad		5-D, Richtung		5-D, Behinderung		5-D, Verteilung	
	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA + Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo
Monat 12	51/65 (78.5%)	58/68 (85.3%)	54/65 (83.1%)	58/68 (85.3%)	50/65 (76.9%)	54/68 (79.4%)	53/65 (81.5%)	58/68 (85.3%)	54/65 (83.1%)	59/68 (86.8%)
Pbo = Placebo										

Tabelle 4-54: Ergebnisse für 5-D und Pruritus-VAS (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
5-D, Gesamtscore, abs. Veränderung							
57	10,2 (3,6)	1,24 (0,29)	60	11 (5,1)	0,51 (0,29)	0,73 [0,12; 1,33] 0,0182	0,44 [0,07; 0,8]
5-D, Dauer, abs. Veränderung							
59	1,2 (0,6)	0,29 (0,07)	64	1,4 (1)	0,2 (0,06)	0,09 [-0,07; 0,25] 0,2908	0,19 [-0,16; 0,54]
5-D, Schweregrad, abs. Veränderung							
62	1,9 (1)	0,2 (0,05)	65	2 (1)	0,1 (0,05)	0,1 [-0,03; 0,22] 0,1369	0,26 [-0,09; 0,61]
5-D, Richtung, abs. Veränderung							
58	3,1 (1,2)	0,24 (0,15)	60	3,1 (1,5)	-0,06 (0,15)	0,3 [0,1; 0,5] 0,0037	0,53 [0,17; 0,9]
5-D, Behinderung, abs. Veränderung							
61	1,8 (1,1)	0,2 (0,06)	64	1,9 (1,4)	0,08 (0,06)	0,13 [-0,05; 0,3] 0,1523	0,25 [-0,1; 0,61]
5-D, Verteilung, abs. Veränderung							
62	2 (1,2)	0,31 (0,06)	65	2,3 (1,5)	0,26 (0,06)	0,05 [-0,11; 0,21] 0,5256	0,11 [-0,24; 0,46]
Pruritus-VAS, abs. Veränderung							
62	21,3 (25,5)	4,16 (1,41)	65	25,7 (27,1)	0,18 (1,36)	3,98 [0,75; 7,21] 0,0159	0,43 [0,07; 0,78]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo VAS = visuelle Analogskala							

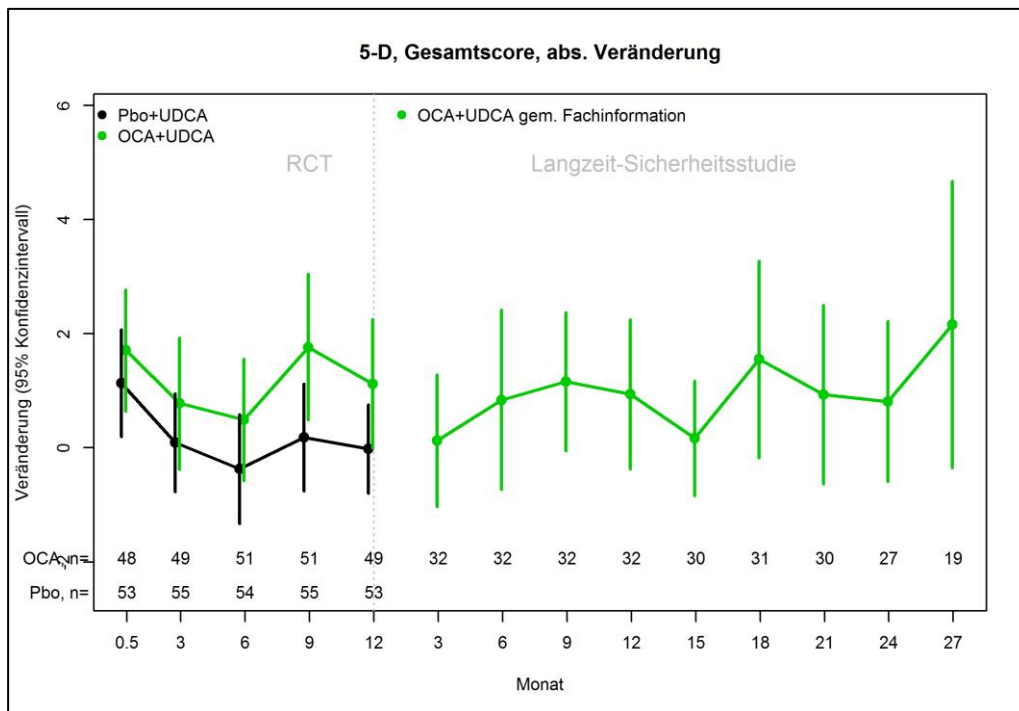


Abbildung 4-36: Ergebnisse für den 5-D Gesamtscore, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

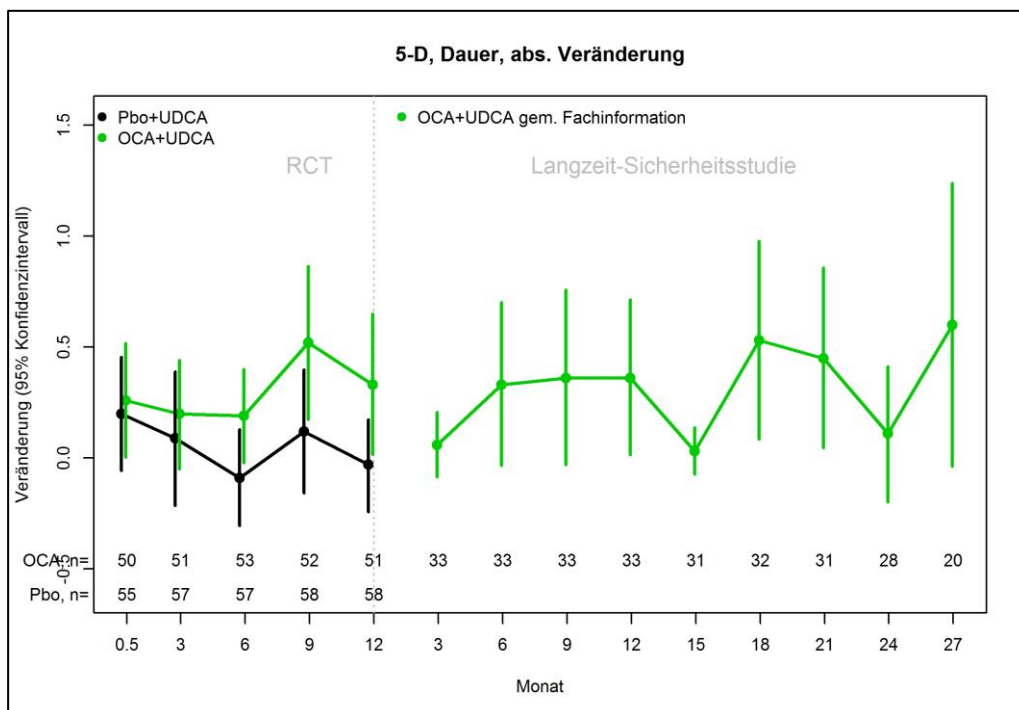


Abbildung 4-37: Ergebnisse für 5-D Dauer, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

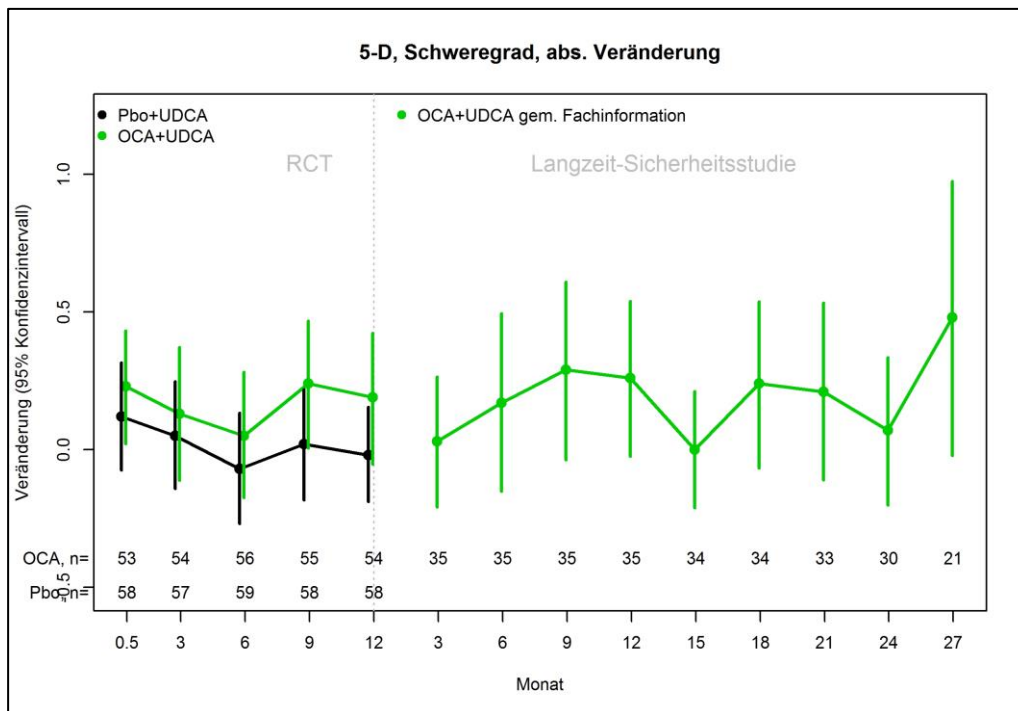


Abbildung 4-38: Ergebnisse für 5-D Schweregrad, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

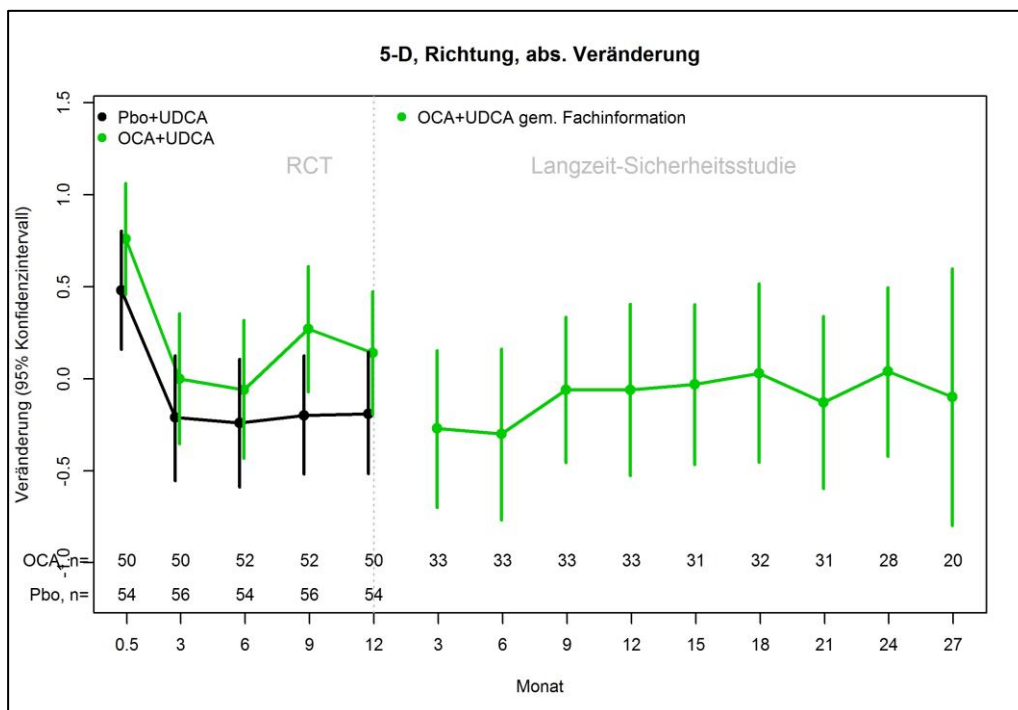


Abbildung 4-39: Ergebnisse für 5-D Richtung, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

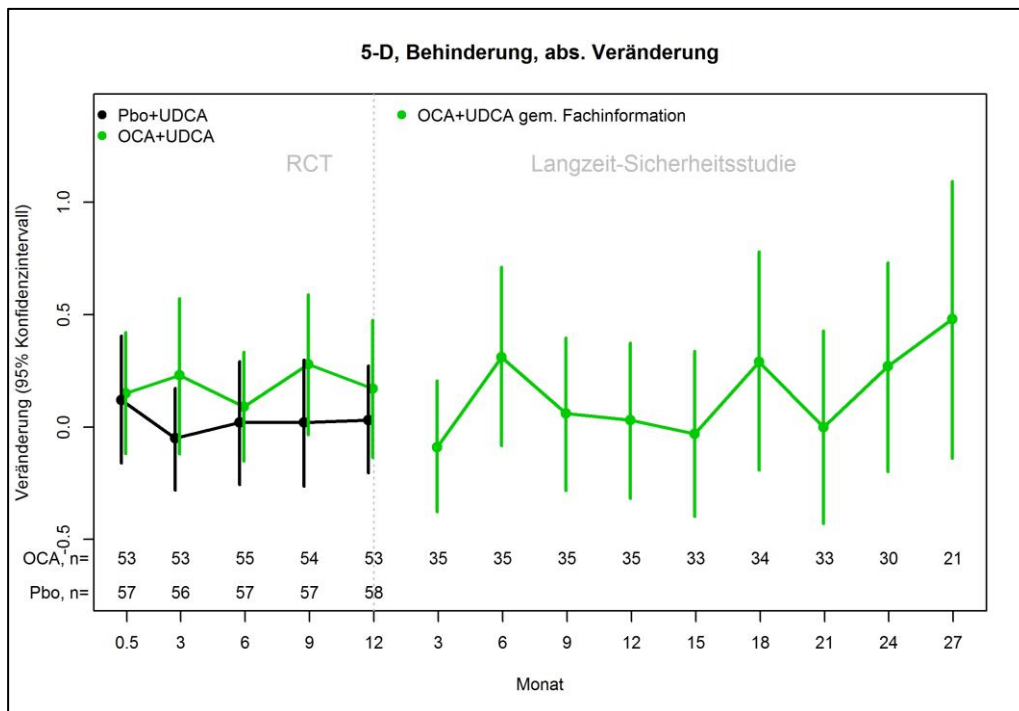


Abbildung 4-40: Ergebnisse für 5-D Behinderung, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

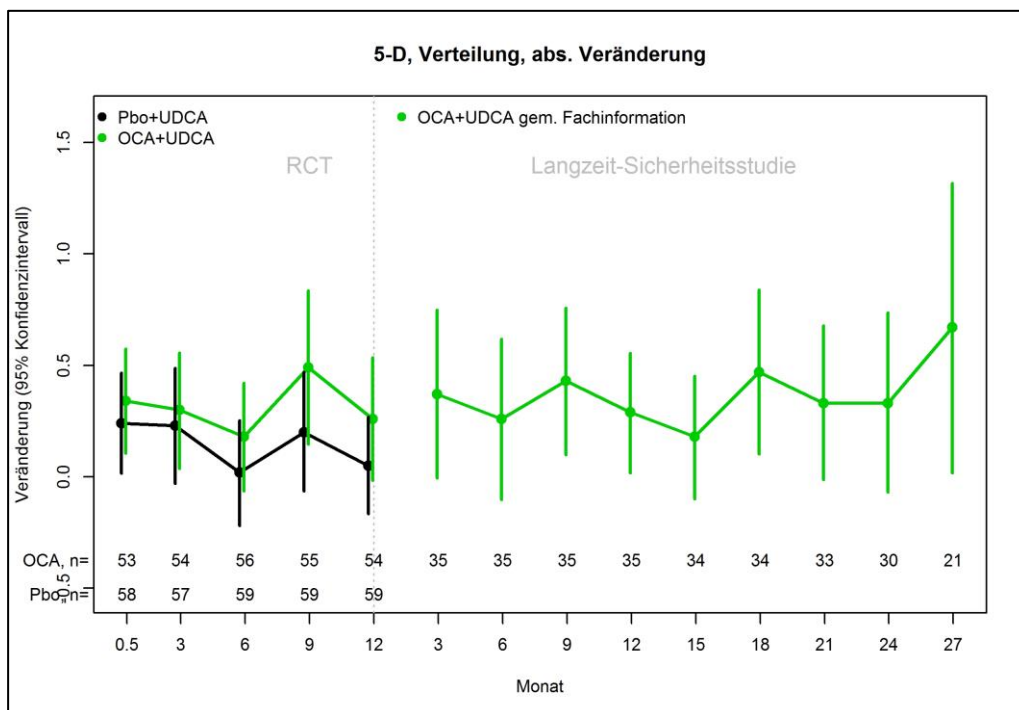


Abbildung 4-41: Ergebnisse für 5-D Verteilung, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

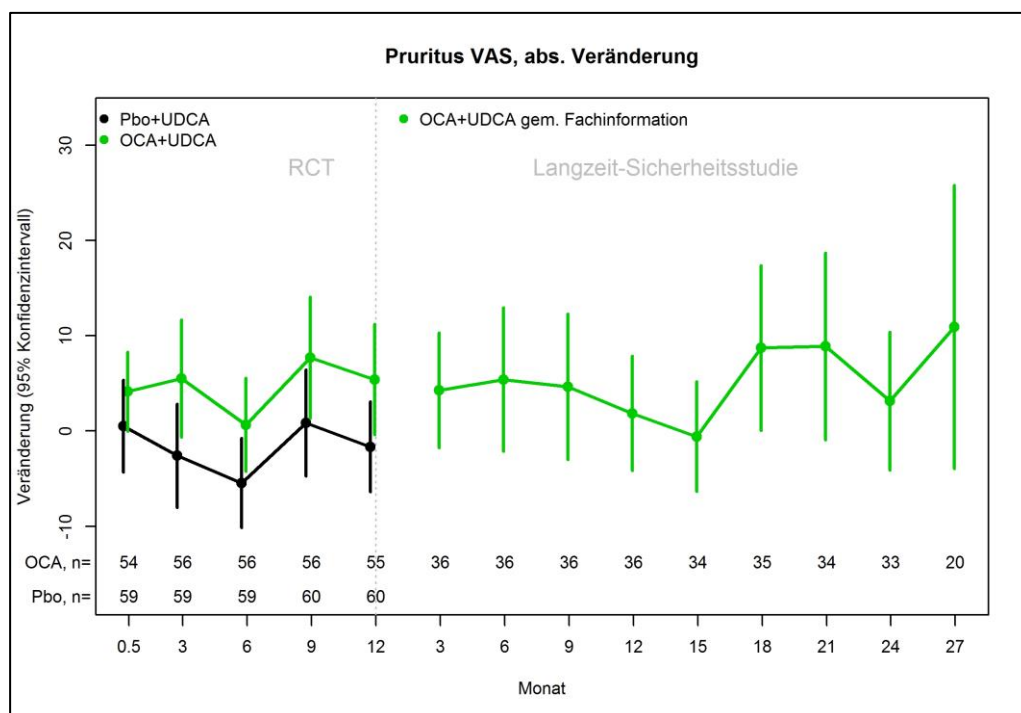


Abbildung 4-42: Ergebnisse für 5-D Pruritus-VAS, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

In den Patienten mit Kombinationstherapie zeigten sich keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen. Die 95%-Konfidenzintervalle der SMD überlappten für alle Domänen und die VAS mit dem Bereich von -0,2 bis 0,2. In der Langzeit-Sicherheitsstudie zeigen sich sehr ähnliche Effekte wie in der RCT.

Patienten mit Monotherapie

Insgesamt trat in der POISE-Studie im Arm Obeticholsäure+UDCA Pruritus als UE bei 4 von 5 Patienten auf; im Kontrollarm UDCA+Placebo bei keinem der 5 Patienten. Im Weiteren wird das UE Pruritus tabellarisch dargestellt.

Tabelle 4-55: Ergebnisse für Pruritus (Inzidenzraten) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Patienten mit Pruritus						
4/5 (80)	0/5 (0)	9 [0,61; 133,08] ^(a)	0,11 [0,01; 1,64] ^(a)	33 [1,06; 1023,56] ^(a)	-0,8 [-1,15; -0,45] ^(a)	0,0168
Patienten mit schwerem Pruritus						
2/5 (40)	0/5 (0)	5 [0,3; 83,69] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,1138

Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
UE Pruritus, das zum Therapieabbruch führte						
0/5 (0)	0/5 (0)	1 [0,02; 42,65] ^(a)	1 [0,02; 42,65] ^(a)	1 [0,02; 59,99] ^(a)	0 [-0,01; 0,01] ^(a)	NA
Intervention wegen Pruritus^(b)						
4/3 (75)	0/0	NA	NA	NA	NA	NA
(a) Modell ohne Stratifikations-Faktoren						
(b) Basierend nur auf Patienten mit Pruritus						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test UE = unerwünschtes Ereignis NA = nicht berechenbar						

Auf Grund der geringen Ereignisraten wurden keine Time-to-Event-Analysen durchgeführt.

In den verschiedenen Operationalisierungen des Endpunktes Pruritus zeigte sich ein Trend für einen größeren Schaden von Obeticholsäure gegenüber Placebo. Im Folgenden sind die Ergebnisse des Fragebogens 5-D und der Pruritus-VAS dargestellt.

Tabelle 4-56: Ergebnisse für 5-D und Pruritus-VAS (Morbidität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
5-D, Gesamtscore, abs. Veränderung							
5	10 (4,9)	2,79 (1,08)	4	10 (4,7)	-1,73 (1,28)	4,52 [1,08; 7,97] 0,0123	1,62 [0,11; 3,13]
5-D, Dauer, abs. Veränderung							
5	1,2 (0,4)	0,99 (0,24)	5	1,4 (0,9)	-0,16 (0,28)	1,15 [0,41; 1,88] 0,0034	1,83 [0,35; 3,30]
5-D, Schweregrad, abs. Veränderung							
5	2 (1,4)	0,69 (0,18)	5	1,6 (0,9)	-0,29 (0,23)	0,99 [0,39; 1,59] 0,0020	1,92 [0,42; 3,42]
5-D, Richtung, abs. Veränderung							
5	3 (1,4)	-0,04 (0,29)	4	3,5 (1)	-1,06 (0,38)	1,02 [0,03; 2,01] 0,0447	1,26 [-0,18; 2,70]
5-D, Behinderung, abs. Veränderung							

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
5	1,8 (1,3)	0,5 (0,2)	5	1,4 (0,9)	-0,24 (0,25)	0,74 [0,09; 1,39] 0,0262	1,33 [-0,04; 2,70]
5-D, Verteilung, abs. Veränderung							
5	2 (1)	0,5 (0,12)	5	1,6 (1,3)	0,13 (0,15)	0,36 [-0,05; 0,78] 0,0834	1,02 [-0,30; 2,34]
Pruritus-VAS, abs. Veränderung							
5	22,4 (36,1)	16,59 (4,79)	5	19,5 (35,8)	-6,4 (5,83)	22,99 [7,46; 38,52] 0,0050	1,72 [0,27; 3,17]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo VAS = visuelle Analogskala							

In den Patienten mit Monotherapie zeigten sich in den 5-D Domänen „Dauer“ und „Schweregrad“ sowie in der „Pruritus-VAS“ signifikante und klinisch relevante Unterschiede zum Nachteil von Obeticholsäure. Zu berücksichtigen ist dabei die sehr kleine Stichprobengröße von 5 gegen 5 Patienten, die keine valide Aussage zum Zusatznutzen oder -schaden zulassen.

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.

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.1.7 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 4-57: Operationalisierung von PBC-40 (Lebensqualität)

Studie	Operationalisierung
POISE	Gesundheitsbezogene Lebensqualität wurde anhand des validierten Fragebogens PBC-40 bewertet (16). Der Fragebogen PBC-40 besteht aus 40 Items in sechs Subdomänen, von denen die Subdomänen soziale Domäne und emotionale Domäne als Ausprägungen der gesundheitsbezogenen Lebensqualität dargestellt sind. Den Antworten zu den Fragen werden sogenannte Score-Punkte zugeordnet (1=überhaupt nicht, 2=selten, 3=manchmal, 4=häufig, 5=immer). Gesundheitsbezogene Lebensqualität wurde jeweils bei Einschluss (an Tag 0), nach 2 Wochen, 3, 6, 9 und 12 Monaten vor Verabreichung der Studienmedikation evaluiert. Der Unterschied zwischen den Behandlungsgruppen hinsichtlich gesundheitsbezogener Lebensqualität wurde anhand der Differenz der Score-Änderung zur Baseline (mit 95% KI) in einem gemischten Modell mit Wiederholungsmessungen (MMRM) berechnet. Dabei wurden die Faktoren "AP > 3x ULN und/oder AST > 2x ULN und/oder Bilirubin > ULN zu Baseline" und "Intoleranz gegenüber UDCA" als Strata sowie der Baselinewert des Scores als Co-Faktor in der Berechnung des Kleinst-Quadrat Mittelwerts und des zugehörigen 95% Konfidenzintervalls berücksichtigt. Die Analyse erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline vorlagen.

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

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

Tabelle 4-58: Bewertung des Verzerrungspotenzials für PBC-40 (Lebensqualität) in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig eingestuft wurde.

Die Studie wurde verblindet durchgeführt. Die Analyse erfolgte auf Basis der ITT-Population, in die alle Patienten aufgenommen wurden, die der in der Fachinformation beschriebenen Population entsprachen. Für beide Teilpopulationen wurden mehr als 95% der Patienten in die ITT-Analyse eingeschlossen. Es lagen keine sonstigen Aspekte einer möglichen Verzerrung vor. Daher wurde das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (siehe Tabelle 4-58 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Tabelle 4-59: Rücklaufquoten für PBC-40 (Lebensqualität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

	Soziale Domäne		Emotionale Domäne	
	N (%) OCA+ UDCA	N (%) UDCA+ Pbo	N (%) OCA+ UDCA	N (%) UDCA+ Pbo
Woche 2	56/65 (86.2%)	60/68 (88.2%)	56/65 (86.2%)	61/68 (89.7%)
Monat 3	58/65 (89.2%)	61/68 (89.7%)	58/65 (89.2%)	61/68 (89.7%)
Monat 6	58/65 (89.2%)	62/68 (91.2%)	58/65 (89.2%)	62/68 (91.2%)
Monat 9	58/65 (89.2%)	61/68 (89.7%)	58/65 (89.2%)	62/68 (91.2%)
Monat 12	56/65 (86.2%)	61/68 (89.7%)	56/65 (86.2%)	61/68 (89.7%)
Pbo = Placebo				

Tabelle 4-60: Ergebnisse für PBC-40 (Lebensqualität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
PBC-40 Soziale Domäne, abs. Veränderung							
64	19,1 (8,4)	0,02 (0,31)	67	21 (8,8)	-0,94 (0,3)	0,96 [0,13; 1,79] 0,0233	0,40 [0,05; 0,74]
PBC-40 Emotionale Domäne, abs. Veränderung							
64	6,3 (2,7)	-0,12 (0,11)	67	7 (3,3)	-0,46 (0,11)	0,34 [0,05; 0,63] 0,0233	0,40 [0,05; 0,74]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler							

N OCA+ UDCA	MW (SD) Baseline OCA+UDCA	LS-MW (SE) change OCA+UDCA	N UDCA+ Pbo	MW (SD) Baseline UDCA+Pbo	LS-MW (SE) change UDCA+Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
MWD = Mittelwertdifferenz SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

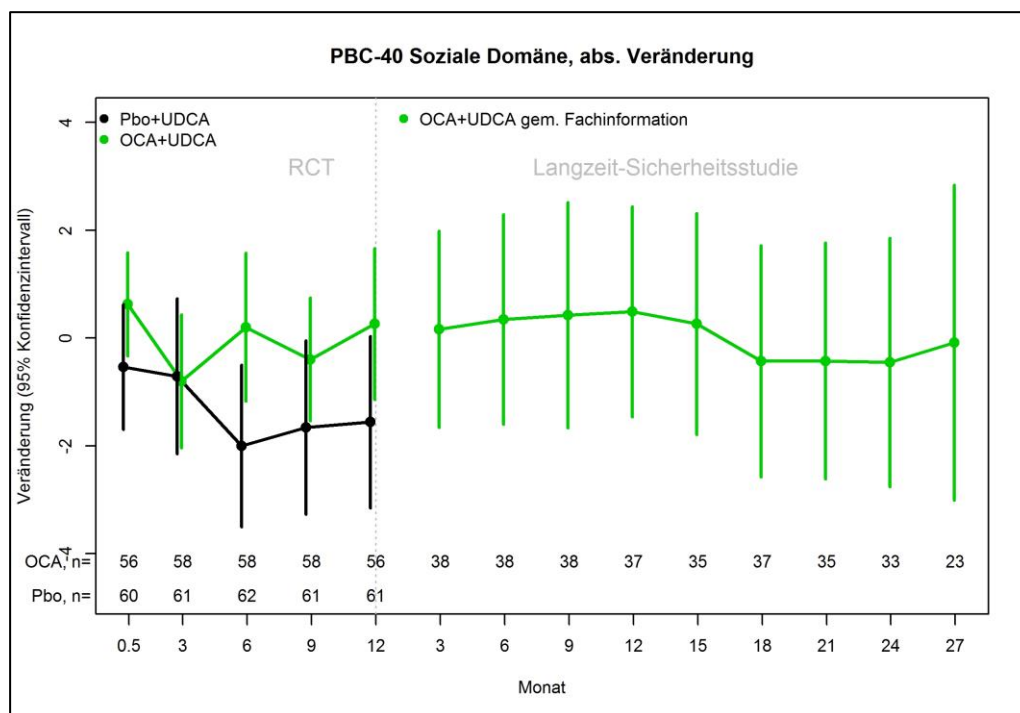


Abbildung 4-43: Ergebnisse für PBC-40, soziale Domäne, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

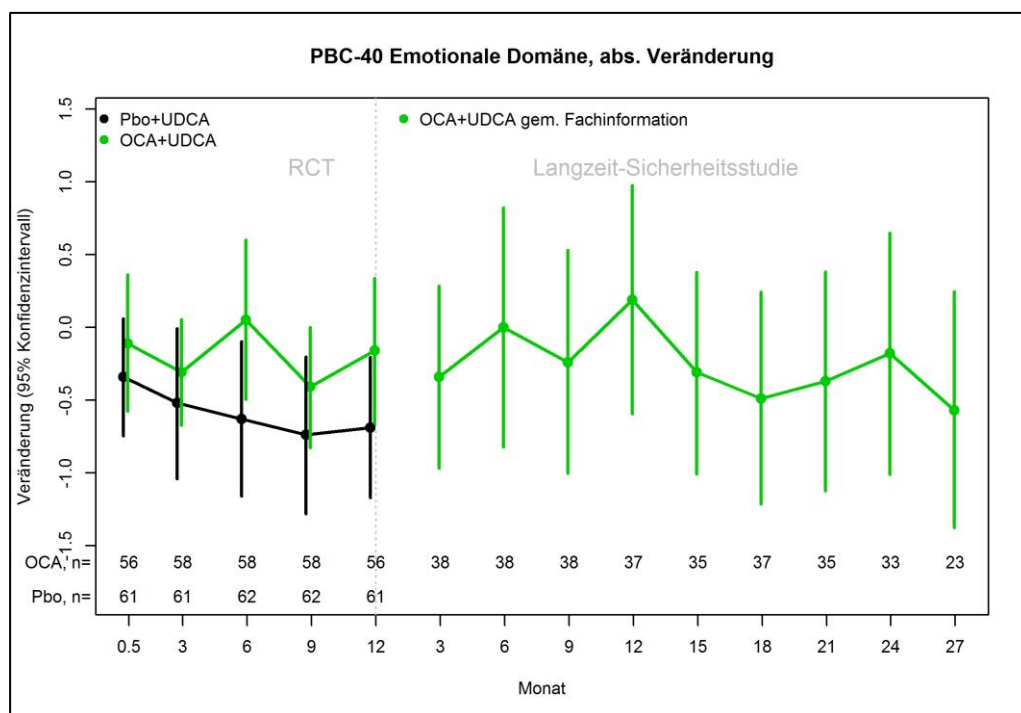


Abbildung 4-44: Ergebnisse für PBC-40, emotionale Domäne, absolute Veränderung zum Vorwert aus RCT und Langzeit-Sicherheitsstudie (Abschnitt 4.3.2.3.3.6) mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Für die Patienten mit Kombinationstherapie zeigte sich in beiden Domänen des PBC-40 zur gesundheitsbezogenen Lebensqualität kein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen. Die 95%-Konfidenzintervalle der SMD überlappten mit dem Bereich von -0,2 bis 0,2.

Patienten mit Monotherapie

Tabelle 4-61: Ergebnisse für PBC-40 (Lebensqualität) aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
PBC-40 Soziale Domäne, abs. Veränderung							
5	22,8 (5,5)	-2,32 (1,11)	5	21 (9,3)	1,63 (1,39)	-3,95 [-7,72; -0,19] 0,0403	-1,22 [-2,57; 0,13]
PBC-40 Emotionale Domäne, abs. Veränderung							
5	7 (3,8)	-0,54 (0,43)	5	5,2 (3,5)	1,19 (0,52)	-1,73 [-3,15; -0,31] 0,0183	-1,42 [-2,80; -0,03]
MW = Mittelwert SD = Standardabweichung SE = Standardfehler MWD = Mittelwertdifferenz							

N OCA	MW (SD) Baseline OCA	LS-MW (SE) change OCA	N Pbo	MW (SD) Baseline Pbo	LS-MW (SE) change Pbo	LS-MWD [95% KI] p-Wert	SMD [95% KI]
SMD = Standardisierte Mittelwertdifferenz KI = Konfidenzintervall LS = Least-square Pbo = Placebo							

Für die Patienten mit Monotherapie zeigte sich in beiden Domänen des PBC-40 zur gesundheitsbezogenen Lebensqualität kein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen. Die 95%-Konfidenzintervalle der SMD überlappten mit dem Bereich von -0,2 bis 0,2.

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.

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.1.8 Verträglichkeit – RCT

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

Tabelle 4-62: Operationalisierung von Verträglichkeit

Studie	Operationalisierung
Gesamtrate UE	
POISE	Dichotomer Endpunkt hinsichtlich des Auftretens eines unerwünschten Ereignisses (UE, treatment-emergent adverse event) innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt dokumentiert wurde. Die Analyse basiert auf der Sicherheitspopulation, in die alle Patienten eingeschlossen sind, denen mindestens eine Dosis einer Studienbehandlung verabreicht wurde. Die Gruppenzuteilung erfolgte auf Basis der Studienmedikation (Obeticholsäure+UDCA vs. UDCA+Placebo bzw. Obeticholsäure vs. Placebo), nicht der Randomisierung. Analysiert wurden die Inzidenzraten mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.
Schwerwiegende UE (SUE)	
POISE	Dichotomer Endpunkt hinsichtlich des Auftretens eines UE innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt dokumentiert und als schwerwiegend (serious) eingestuft wurde. Die Analyse basiert auf der Sicherheitspopulation, in die alle Patienten eingeschlossen sind, denen mindestens eine Dosis einer Studienbehandlung verabreicht wurde. Die Gruppenzuteilung erfolgte auf Basis der Studienmedikation (Obeticholsäure+UDCA vs. UDCA+Placebo bzw. Obeticholsäure vs. Placebo), nicht der Randomisierung. Analysiert wurden die Inzidenzraten mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.
Schwere UE ohne Pruritus	
POISE	Dichotomer Endpunkt hinsichtlich des Auftretens eines UE außer Pruritus innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt als schwer (englisch: severe) dokumentiert wurde. Pruritus wurde ausgeschlossen, da dieses Symptom bereits durch die Darstellung des UE Pruritus, des 5-D, der Pruritus-VAS und der Domäne Jucken des PBC-40 dargestellt ist und eine Doppelzählung vermieden werden soll. Die Analyse basiert auf der Sicherheitspopulation, in die alle Patienten eingeschlossen sind, denen mindestens eine Dosis einer Studienbehandlung verabreicht wurde. Die Gruppenzuteilung erfolgte auf Basis der Studienmedikation (Obeticholsäure+UDCA vs. UDCA+Placebo bzw. Obeticholsäure vs. Placebo), nicht der Randomisierung. Analysiert wurden die Inzidenzraten mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet. Als ergänzende Analyse wird das Auftreten eines schweren UE inklusive Pruritus innerhalb des Beobachtungszeitraums dargestellt.
UE, die zum Therapieabbruch führten	
POISE	Dichotomer Endpunkt hinsichtlich des Auftretens eines UE, das zum Therapieabbruch führte, der durch den jeweiligen verblindeten Prüfarzt dokumentiert wurde. Die Analyse basiert auf der Sicherheitspopulation, in die alle Patienten eingeschlossen sind, denen mindestens eine Dosis einer Studienbehandlung verabreicht wurde. Die Gruppenzuteilung erfolgte auf Basis der Studienmedikation (Obeticholsäure+UDCA vs. UDCA+Placebo bzw. Obeticholsäure vs. Placebo), nicht der Randomisierung.

Analysiert wurden die Inzidenzraten mittels Odds Ratio, relativem Risiko und der absoluten Risikoreduktion sowie den dazugehörigen 95%-Konfidenzintervallen. Der Unterschied (Obeticholsäure+UDCA bzw. Obeticholsäure minus Kontrolle) wurde mittels Cochran-Mantel-Haenszel-Test getestet.
Abkürzungen: OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; 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 4-63: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Gesamtrate UE						
POISE	niedrig	ja	ja	ja	ja	niedrig
Schwerwiegende UE						
POISE	niedrig	ja	ja	ja	ja	niedrig
Schwere UE ohne Pruritus						
POISE	niedrig	ja	ja	ja	ja	niedrig
UE, die zum Therapieabbruch führten						
POISE	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. In der POISE-Studie waren sowohl Patienten als auch Behandler verblindet. Die Analyse zum Endpunkt Verträglichkeit erfolgte auf Basis der Sicherheitspopulation, in die alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhielten. Eine ergebnisabhängige Berichterstattung ist ausgeschlossen, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es lagen auch keine sonstigen Aspekte einer möglichen Verzerrung vor.

Somit wurde das Verzerrungspotential der Endpunkte UE, SUE, schwere UE ohne Pruritus und UE, die zum Therapieabbruch führten, als niedrig eingestuft (siehe Tabelle 4-63 und Anhang 4-F).

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.

Patienten mit Kombinationstherapie

Tabelle 4-64: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Patienten mit UE						
60/65 (92,31)	61/68 (89,71)	1,03 [0,92; 1,14] ^(a)	0,97 [0,87; 1,08] ^(a)	1,38 [0,41; 4,59]	-0,07 [-0,21; 0,07]	0,6027
Patienten mit SUE						
8/65 (12,31)	3/68 (4,41)	2,79 [0,78; 10,07]	0,36 [0,1; 1,29]	3,05 [0,77; 12,05]	-0,08 [-0,17; 0,02]	0,1006
Patienten mit schweren UE ohne UE Pruritus						
9/65 (13,85)	6/68 (8,82)	1,58 [0,6; 4,17]	0,63 [0,24; 1,68]	1,66 [0,56; 4,98]	-0,05 [-0,15; 0,06]	0,3627
UE, die zum Therapieabbruch führten						
5/65 (7,69)	1/68 (1,47)	5,29 [0,64; 43,92]	0,19 [0,02; 1,57]	5,63 [0,64; 49,75]	-0,06 [-0,13; 0,01] ^(a)	0,0858
(a) Modell ohne Stratifikations-Faktoren						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test UE = unerwünschtes Ereignis SUE = schwerwiegendes unerwünschtes Ereignis						

Für die Patienten mit Kombinationstherapie zeigte sich in den Kategorien zu den unerwünschten Ereignissen kein signifikanter Unterschied zwischen den Behandlungsgruppen. Als Hauptanalyse der schweren UE wird dabei die Analyse unter Ausschluss des einzelnen UE Pruritus herangezogen, da Pruritus bereits durch die Darstellung im 5-D, Pruritus-VAS und der Domäne Jucken des PBC-40 ausführlich dargestellt ist und eine Doppelzählung vermieden werden soll.

Als ergänzende Analyse wurden in der folgenden Tabelle die schwere UE mit dem UE Pruritus dargestellt. Hier fand sich ein signifikanter Nachteil für Obeticholsäure+UDCA gegenüber UDCA+Placebo.

Tabelle 4-65: Ergänzende Analyse der schwere UE mit dem einzelnen UE Pruritus aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

Ereignisse/N (%) OCA+UDCA	Ereignisse/N (%) UDCA+Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Patienten mit schweren UE						
19/65 (29,23)	9/68 (13,24)	2,21 [1,08; 4,52]	0,45 [0,22; 0,93]	2,71 [1,12; 6,54]	-0,16 [-0,3; -0,03]	0,0247
(a) Modell ohne Stratifikations-Faktoren						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko OR = Odds Ratio RD = Risikodifferenz KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test UE = unerwünschtes Ereignis						

Darüberhinaus wurden im Folgenden die UE mit einer Inzidenz $\geq 5\%$ in mindestens einer Behandlungsgruppe dargestellt.

Tabelle 4-66: Ergänzende Ergebnisse für UE mit einer Inzidenz $\geq 5\%$ in mindestens einer Behandlungsgruppe aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Kombinationstherapie

	OCA+UDCA (N = 65)	UDCA+Placebo (N = 68)
Preferred Term	n (%)	n (%)
Gesamtrate UE	60 (92)	61 (90)
Pruritus	35 (54)	28 (41)
Ekzem	4 (6)	0
Fatigue	11 (17)	9 (13)
Nasopharyngitis	16 (25)	10 (15)
Infektionen der oberen Atemwege	4 (6)	8 (12)
Infektionen der Uringänge	3 (5)	8 (12)
Influenza	5 (8)	4 (6)
Bronchitis	4 (6)	0
Übelkeit	4 (6)	7 (10)
Erbrechen	2 (3)	4 (6)
Diarrhoe	2 (3)	7 (10)

Verstopfung	5 (8)	3 (4)
Schmerzen im oberen Abdomen	5 (8)	5 (7)
Schmerzen im Abdomen	2 (3)	4 (6)
Dyspepsie	4 (6)	6 (9)
Abdominales Unwohlsein	5 (8)	1 (1)
Abdominale Blähungen	3 (5)	6 (9)
Arthralgie	4 (6)	3 (4)
Rückenschmerzen	4 (6)	6 (9)
Kopfschmerzen	12 (18)	11 (16)
Husten	4 (6)	3 (4)
Anämie	2 (3)	4 (6)
Hypothyreoidismus	4 (6)	1 (1)
Insomnie	2 (3)	7 (10)
n = Anzahl Patienten mit Ereignis, UE = unerwünschtes Ereignis		

In den Einzel-UE mit einer Inzidenz $\geq 5\%$ zeigte Obeticholsäure+UDCA ein mit UDCA+Placebo vergleichbares Sicherheitsprofil. Das häufigste UE stellt Pruritus dar. Dieses ist in Abschnitt 4.3.1.3.1.6 als gesonderter Endpunkt beschrieben und wird daher hier nicht weiter diskutiert.

Patienten mit Monotherapie

Tabelle 4-67: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Patienten mit UE						
5/5 (100)	5/5 (100)	1 [0,71; 1,41] ^(a)	1 [0,71; 1,41] ^(a)	1 [0,02; 59,99] ^(a)	0 [-0,01; 0,01] ^(a)	NA
Patienten mit SUE						
3/5 (60)	0/5 (0)	7 [0,45; 108,26] ^(a)	0,14 [0,01; 2,21] ^(a)	15,4 [0,56; 425,53] ^(a)	-0,6 [-1,03; -0,17] ^(a)	0,0628
Patienten mit schweren UE ohne UE Pruritus						
2/5 (40)	0/5 (0)	5 [0,3; 83,68] ^(a)	0,2 [0,01; 3,35] ^(a)	7,86 [0,28; 217,11] ^(a)	-0,4 [-0,83; 0,03] ^(a)	0,1573
UE, die zum Therapieabbruch führten						
0/5 (0)	1/5 (20)	0,33 [0,02; 6,65] ^(a)	3 [0,15; 59,89] ^(a)	0,27 [0,01; 8,46] ^(a)	0,2 [-0,15; 0,55] ^(a)	0,3173
(a) Modell ohne Stratifikations-Faktoren						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio						

Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test UE = unerwünschtes Ereignis SUE = schwerwiegendes unerwünschtes Ereignis NA = nicht berechenbar						

Für die Patienten mit Monotherapie zeigte sich in den Kategorien zu den unerwünschten Ereignissen kein signifikanter Unterschied zwischen den Behandlungsgruppen. Als Hauptanalyse der schweren UE wird dabei die Analyse unter Ausschluss des einzelnen UE Pruritus herangezogen, da Pruritus bereits durch die Darstellung im 5-D, Pruritus-VAS und der Domäne Jucken des PBC-40 ausführlich dargestellt ist und eine Doppelzählung vermieden werden soll.

Als ergänzende Analyse wurden in der folgenden Tabelle die schwere UE mit dem UE Pruritus dargestellt. Auch hier fand sich kein signifikanter Nachteil für Obeticholsäure gegenüber Placebo.

Tabelle 4-68: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

Ereignisse/N (%) OCA	Ereignisse/N (%) Pbo	RR [95% KI]	RRR [95% KI]	OR [95% KI]	RD [95% KI]	p CMH
Patienten mit schweren UE						
3/5 (60)	0/5 (0)	7 [0,45; 108,26] ^(a)	0,14 [0,01; 2,21] ^(a)	15,4 [0,56; 425,53] ^(a)	-0,6 [-1,03; -0,17] ^(a)	0,0628
(a) Modell ohne Stratifikations-Faktoren						
RR = Relatives Risiko RRR = Invertiertes relatives Risiko RD = Risikodifferenz OR = Odds Ratio KI = Konfidenzintervall Pbo = Placebo CMH = Cochran-Mantel-Haenszel Test UE = unerwünschtes Ereignis						

Zusätzlich wurden im Folgenden die UE mit einer Inzidenz $\geq 40\%$ (also mindestens 2 der 5 Patienten) in mindestens einer Behandlungsgruppe dargestellt.

Tabelle 4-69: Ergänzende Ergebnisse für UE mit einer Inzidenz $\geq 40\%$ in mindestens einer Behandlungsgruppe aus RCT mit dem zu bewertenden Arzneimittel – Patienten mit Monotherapie

	OCA (N = 5)	Placebo (N = 5)
Preferred Term	n (%)	n (%)
Gesamtrate UE	5 (100)	5 (100)
Pruritus	4 (80)	0
Nasopharyngitis	1 (20)	3 (60)
Übelkeit	0	2 (40)
Schmerzen im Abdomen	1 (20)	2 (40)
Gastrooesophageale Reflux Krankheit	1 (20)	2 (40)
Dyspepsie	0	2 (40)
Rückenschmerzen	0	2 (40)
Kopfschmerzen	0	2 (40)
Husten	0	2 (40)
Oropharyngeale Schmerzen	2 (40)	0
Synkope	2 (40)	0
Nachtschweiß	2 (40)	0
Muskelspasmen	0	2 (40)
n = Anzahl Patienten mit Ereignis, UE = unerwünschtes Ereignis		

Als Einzel-UE trat Pruritus unter Obeticholsäure mit 4 von 5 Patienten häufiger auf als unter Placebo (0 von 5 Patienten). Pruritus ist in Abschnitt 4.3.1.3.1.6 als gesonderter Endpunkt beschrieben.

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.

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

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

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

Im vorliegenden Nutzendossier werden keine Subgruppenanalysen dargestellt und nicht zum Nachweis des Ausmaßes des Zusatznutzens herangezogen, da diese Angaben gemäß 1.3 der Anlage II.1 zum 5. Kapitel der VerfO nicht zu machen sind..

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Patienten mit Kombinationstherapie

In der Gesamtmortalität trat bei den Patienten mit Kombinationstherapie ein Todesfall auf. Es zeigten sich keine signifikanten Unterschiede zwischen den Behandlungsgruppen (RR=3,14; 95%-KI [0,13; 75,63]).

Die Morbidität wurde durch verschiedene Operationalisierungen der biochemischen Kenngrößen AP und Bilirubin untersucht. Für den primären Endpunkt der POISE-Studie, definiert als Ereignisrate der Patienten mit AP $<1.67 \times$ ULN, Gesamt-Bilirubin $\leq$ ULN bei AP-Reduktion von $\geq 15\%$, zeigte sich ein statistisch signifikanter Vorteil bezüglich der Ansprech-Kriterien für Obeticholsäure+UDCA gegenüber UDCA+Placebo (Monat 12, RR = 4,45; 95%-KI [2,12; 9,34]). Dieser Effekt wurde bestätigt durch die Einzelkomponenten des primären Endpunkts. Auch hier zeigten sich statistisch signifikante Vorteile von Obeticholsäure+UDCA gegenüber UDCA+Placebo. Diese Ergebnisse wurden unterstützt durch die weiteren im Studienprotokoll als sekundäre Analysen definierten Ansprech-Kriterien bzgl. prozentualer Reduktion von AP im Vergleich zu Baseline. In allen Analysen fand sich ein statistisch signifikanter Vorteil für Obeticholsäure+UDCA gegenüber UDCA+Placebo. Die ebenfalls a priori definierte Analyse für Patienten mit AP $\leq$ ULN nach 6 und 12 Monaten zeigte keine Unterschiede, da nur ein Patient der Obeticholsäure+UDCA Gruppe dieses Kriterium erreichte, aber kein Patient der UDCA+Placebo-Gruppe. Die zusätzlich durchgeführten und a priori geplanten Analysen bezüglich der Änderung zu Baseline zeigten in der SMD für die AP (SMD = -2,35; 95%-KI [-2,80; -1,91]) und dem Gesamt-Bilirubin (SMD = -0,73; 95%-KI [-1,08; -0,37]) sowie dem direkten Bilirubin (SMD = -0,99; 95%-KI [-1,35; -0,63]) statistisch signifikante und klinisch relevante Vorteile für Obeticholsäure+UDCA.

Die Fibrose wurde in der POISE-Studie anhand des ELF-Score und der transienten Elastographie (Fibroscan®) operationalisiert. In der Teilpopulation der Patienten mit Kombinationstherapie zeigten sich im ELF-Score und in dessen Einzelkomponenten keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen. Im Fibroscan® lag der Punktschätzer für die Fibrose unterhalb von Null und zeigte einen numerischen Vorteil zu Gunsten von Obeticholsäure; allerdings fand sich kein relevanter oder signifikanter Unterschied (SMD = -0,09; 95%-KI [-0,58; 0,41]).

Der Endpunkt Osteoporose wurde anhand des DEXA-Scan ermittelt, der nur in einer Teilpopulation erhoben wurde. Hierbei zeigten sich keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen.

Symptome der PBC wurden anhand des Fragebogens PBC-40 untersucht. In keiner der vier Domänen des PBC-40 zu den Symptomen "generelle Symptome", "Jucken", "Fatigue" und "kognitive Funktion" zeigte sich ein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen.

Zur weiteren Untersuchung des Symptoms Pruritus wurden zudem das UE Pruritus sowie der Fragebogen 5-D und die Pruritus-VAS verwendet. Hier zeigten sich im UE Pruritus und in den Instrumenten bezüglich der SMD keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen.

Die gesundheitsbezogene Lebensqualität wurde anhand der beiden Domänen des PBC-40, soziale und emotionale Domäne, untersucht. Hierbei zeigten sich keine signifikanten oder klinisch relevanten Unterschiede zwischen den Behandlungsgruppen.

In der Verträglichkeit zeigte sich in den Kategorien zu den unerwünschten Ereignissen kein signifikanter Unterschied zwischen den Behandlungsgruppen. In den Einzel-UE mit einer Inzidenz $\geq 5\%$ zeigte Obeticholsäure+UDCA ein mit UDCA+Placebo vergleichbares Sicherheitsprofil. Einzig Pruritus trat unter Obeticholsäure+UDCA signifikant häufiger auf als unter UDCA+Placebo. Die Ergebnisse zum Einzel-UE Pruritus wurden im Endpunkt Pruritus ausführlich dargestellt und sind hier nicht weiter diskutiert, um Doppelzählungen zu vermeiden.

Patienten mit Monotherapie

In der Teilpopulation der Patienten mit Monotherapie traten im Endpunkt Gesamtmortalität in beiden Behandlungsgruppen keine Todesfälle auf.

In der Teilpopulation der Patienten mit Monotherapie ergab sich für den primären Endpunkt aus AP $<1.67 \times \text{ULN}$, Gesamt-Bilirubin $\leq \text{ULN}$ bei AP-Reduktion von $\geq 15\%$ Trends für einen Vorteil von Obeticholsäure gegenüber Placebo. Auf Grund des kleinen Stichprobenumfangs wurde hier keine Signifikanz erreicht (RR = 5; 95%-KI [0,3; 83,69]). Auch in den drei Einzelkomponenten zeigen sich die Trends für einen Vorteil von Obeticholsäure gegenüber Placebo. Ein numerischer und im OR und RD signifikanter Vorteil zeigt sich dabei in der Komponente der prozentualen Reduktion der AP zu Monat 12 in drei bzw. allen vier Kategorien im OR und RD. Auch in der Teilpopulation der Patienten mit Monotherapie gab es nur einen Patienten, der das Kriterium von AP $\leq \text{ULN}$ erreichte. In den absoluten Veränderungen zum Vorwert zeigen sich keine signifikanten Änderungen für Bilirubin, wohl aber eine signifikante und klinisch relevante Reduktion von AP zu Gunsten von Obeticholsäure.

Für die Darstellung der Fibrose wurden die Ergebnisse des ELF-Score und Fibroscan® herangezogen. In der Teilpopulation der Patienten mit Monotherapie zeigten sich im ELF-Score und in dessen Einzelkomponenten keine klinisch relevanten Unterschiede zwischen den Behandlungsgruppen. Für den Fibroscan® lagen insgesamt nur Daten von 4 Patienten vor, so dass keine Bewertung der Ergebnisse durchgeführt wurde.

Für den DEXA-Scan zum Endpunkt Osteoporose lagen nur Daten von 5 Patienten vor, so dass keine Bewertung der Ergebnisse durchführbar war.

Für die Patienten mit Monotherapie zeigte sich in den Domänen "Generelle Symptome", "Fatigue" und "Kognitive Funktion" des PBC-40 bzgl. der Symptome kein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen. Für das Symptom "Jucken" zeigte sich ein signifikanter Nachteil für Obeticholsäure (SMD = 2,40; 95%-KI [0,77; 4,02]). Diese Beobachtung deckt sich mit den Beobachtungen zum Pruritus-Fragebogen 5-D und der Pruritus-VAS. Als UE trat Pruritus unter Obeticholsäure mit 4 von 5 Patienten häufiger auf als unter Placebo (0 von 5 Patienten). In den verschiedenen Operationalisierungen des UE Pruritus zeigt sich ein Trend für einen größeren Schaden von Obeticholsäure gegenüber Placebo. Dieses ist konform mit den Beobachtungen in den Fragebögen zur Morbidität. Zu berücksichtigen ist dabei die sehr kleine Stichprobengröße von 5 gegen 5 Patienten, die keine valide Aussage zum Zusatznutzen oder -schaden zulässt.

Die zur weiteren Untersuchung des Symptoms Pruritus verwendeten Instrumente zeigten im 5-D in den Domänen Dauer und Schweregrad sowie in der Pruritus-VAS signifikante und klinisch relevante Unterschiede zum Nachteil von Obeticholsäure (Dauer SMD = 1,83; 95%-KI [0,35; 3,30]; Schweregrad SMD = 1,92; 95%-KI [0,42; 3,42]; Pruritus-VAS SMD = 1,72; 95%-KI [0,27; 3,17]). Keine relevanten Unterschiede zeigten sich in den 5-D Domänen Gesamtscore, Richtung, Behinderung und Verteilung. Zu berücksichtigen ist bei der Bewertung die sehr kleine Stichprobengröße von 5 gegen 5 Patienten, die keine valide Aussage zum Zusatznutzen oder -schaden zulässt.

In den beiden Domänen des PBC-40 zur gesundheitsbezogenen Lebensqualität, der sozialen sowie emotionalen Domäne, zeigte sich kein signifikanter oder klinisch relevanter Unterschied zwischen den Behandlungsgruppen.

In der Verträglichkeit zeigte sich in den Kategorien zu den unerwünschten Ereignissen (jegliche UE, SUE, schwere UE und UE, die zum Studienabbruch führten) kein signifikanter Unterschied zwischen den Behandlungsgruppen.

Subgruppen und Metaanalysen

Für beide Teilpopulationen lag genau eine Studie für Obeticholsäure bei PBC mit dem zugelassenen Dosierungsschema vor. Daher wurde keine Metaanalyse durchgeführt.

Im vorliegenden Nutzendossier werden keine Subgruppenanalysen dargestellt und nicht zum Nachweis des Ausmaßes des Zusatznutzens herangezogen, da diese Angaben gemäß 1.3 der Anlage II.1 zum 5. Kapitel der Verfo nicht zu machen sind.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

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

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Es wurden keine indirekten Vergleiche durchgeführt.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

*Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Nicht relevant.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

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

Tabelle 4-70: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.**

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

Tabelle 4-71: 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 relevant.

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

Tabelle 4-72: 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 4-73: 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 relevant.

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

Tabelle 4-74: 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 relevant.

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

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

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

Nicht relevant.

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.2.

Nicht relevant.

4.3.2.2 Nicht randomisierte vergleichende Studien

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

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

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

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht relevant.

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 4-75: 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 relevant.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-76: 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 4-77: 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 relevant.

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

Nicht relevant.

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.2.5.5.

Nicht relevant.

4.3.2.3 Weitere Untersuchungen

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

Im Folgenden werden die Ergebnisse der einarmigen Langzeit-Sicherheitsstudie dargestellt, in die alle Patienten der POISE-Studie eingeschlossen werden konnten. Da die Patienten des Kontrollarms in der Langzeit-Sicherheitsstudie ebenfalls mit Obeticholsäure behandelt wurden, liegen keine vergleichenden Daten vor. Die Langzeitdaten erlauben allerdings eine Abschätzung der langfristigen Sicherheit und Wirksamkeit von Obeticholsäure in dieser chronischen und progressiven Erkrankung. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet, einmal die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, und als zweite Population alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die aber in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

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

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Trifft nicht zu.

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.

Die einarmige Langzeitsicherheitsstudie der POISE-Studie umfaßt alle Patienten, die in den RCT Teil der POISE eingeschlossen wurden. Patienten aus dem Placebo-Arm wurden in der Langzeitsicherheitsstudie mit Obeticholsäure behandelt. Im Dossier werden nur die Patienten dargestellt, die im RCT Teil der POISE-Studie in den Titrationsarm randomisiert wurden. Die Charakteristik der Patienten ist in Abschnitt 4.3.1.2. dargestellt.

Tabelle 4-78: Charakterisierung der eingeschlossenen Studien – weitere Unterlagen mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
Langzeitsicherheitsstudie der POISE (Studie 747-301)	einarmig, offen; Patienten im Kontrollarm ^a wurden weiter beobachtet und dort optional mit Obeticholsäure offen behandelt, Weiterbehandlung der Patienten der Obeticholsäure-Arme	Erwachsene mit primärer biliärer Cholangitis	Obeticholsäure Titration (n=193, davon 58 aus dem Titrationsarm der POISE) Therapiebeginn mit 5 mg täglich, Möglichkeit der Titration zu 10 mg nach 3 Monaten und 25 mg nach 6 Monaten einmal täglich	Behandlungsdauer offen Möglichkeit des Cross-over zur offenen Behandlung mit Obeticholsäure für Patienten des Placeboarms und Weiterbehandlung der Patienten der Obeticholsäure-Arme	Europa (Deutschland, Vereinigtes Königreich, Polen, Niederlande, Italien, Spanien, Österreich, Belgien, Frankreich, Schweden), Nordamerika (USA, Kanada), Australien ab 12/2013	<u>Endpunkte:</u> Kombinierter Endpunkt als Ansprech-Endpunkt mit den Komponenten AP < 1,67x Obere Normgrenze (ULN) bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt Bilirubin ≤ ULN zu den Nachbeobachtungszeitpunkten AP, Bilirubin, Biochemische Kriterien zum Behandlungserfolg, erkrankungsspezifische Symptome (PBC-40), Sicherheitsparameter
a) Der Kontrollarm umfaßte für Patienten mit Kombinationstherapie UDCA + Placebo, für Patienten mit Monotherapie Placebo						

Tabelle 4-79: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Therapie	Obeticholsäure Titration	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in- Phase etc.</i>
POISE (Studie 747-301)	Kombina- tions- therapie	UDCA Basistherapie plus Obeticholsäure 5 mg Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück Möglichkeit der Titration zu 10 mg nach 3 Monaten und 25 mg nach 6 Monaten einmal täglich	Patienten im Kontrollarm wurden weiterbeobachtet und dort optional mit Obeticholsäure offen behandelt, Weiterbehandlung der Patienten der Obeticholsäure-Arme
	Mono- therapie	Obeticholsäure 5 mg Tablette, einmal täglich ca. 30 Minuten vor dem Frühstück Möglichkeit der Titration zu 10 mg nach 3 Monaten und 25 mg nach 6 Monaten einmal täglich	Patienten im Kontrollarm wurden weiterbeobachtet und dort optional mit Obeticholsäure offen behandelt, Weiterbehandlung der Patienten der Obeticholsäure-Arme

Die Langzeitsicherheitsstudie deckt zwei Therapieschemata ab, die Kombination mit UDCA und als Monotherapie im Vergleich zum beobachtenden Abwarten.

Die Indikation für Obeticholsäure bei PBC sieht eine Titration der Behandlung vor. Somit werden Patienten, die im RCT im Behandlungsarm "Obeticholsäure 10 mg" und in der Kontrollgruppe waren, im Folgenden nicht weiter dargestellt. Auf Grund der wenigen Patienten unter Monotherapie wurde auf die Darstellung dieser Patienten verzichtet.

In den folgenden Tabellen sind die Demographie und Baseline Charakteristika der Patienten mit Kombinationstherapie für die Patienten des relevanten Behandlungsarm "Obeticholsäure Titration" aus der RCT dargestellt.

Tabelle 4-80: Charakterisierung der Studienpopulationen – Langzeitsicherheitsstudie mit dem zu bewertenden Arzneimittel

	Patienten, die gemäß Fach-information behandelt wurden (N=38)	Gesamtgruppe (N=58)
Geschlecht, n (%)		
Männlich	3 (8%)	4 (7%)
Weiblich	35 (92%)	54 (93%)
Alter (Jahre)		
n	38	58
MW (SD)	53,9 (8,42)	55,0 (8,43)
Median	53	54
Minimum; Maximum	29; 77	29; 77
Altersklasse, n (%)		
<65 Jahre	36 (95%)	53 (91%)
≥65 Jahre	2 (5%)	5 (9%)
Region, n (%)		
Europa	25 (66%)	39 (67%)
Nordamerika	10 (26%)	14 (24%)
Australien	3 (8%)	5 (9%)
Ethnische Gruppe, n (%)		
Kaukasier	36 (95%)	56 (97%)
Nicht-Kaukasier	2 (5%)	2 (3%)
Körpergewicht (kg)		
n	38	58
MW (SD)	67,05 (12,579)	67,90 (12,976)
Median	64,25	65,16
Minimum; Maximum	46,7; 95,0	46,7; 101,8
Body Mass Index (BMI) (kg/m²)		
n	38	58
MW (SD)	25,29 (4,504)	25,62 (4,944)
Median	24,2	24,37
Minimum; Maximum	17,7; 38,1	17,7; 40,7
BMI-Klasse, n (%)		
<30 kg/m ²	32 (84%)	48 (83%)
≥30 kg/m ²	6 (16%)	10 (17%)
Leber Biopsie vor Studienbeginn, n (%)		
Ja	5 (13%)	9 (16%)
Nein	33 (87%)	49 (84%)
UDCA Behandlung zu Baseline, n (%)		
Ja	38 (100%)	58 (100%)
Nein	0	0
Pruritus in der Anamnese, n (%)		
Ja	22 (58%)	36 (62%)
Nein	16 (42%)	22 (38%)
Intensität des letzten Pruritus-Ereignisses mit Pruritus in der Anamnese, n (%)		
Mild	14 (64%)	25 (69%)
Moderat	5 (23%)	8 (22%)
Schwer	2 (9%)	2 (6%)
unbekannt	1 (5%)	1 (3%)
Pruritus zu Baseline, n (%)		
Ja	18 (47%)	29 (50%)
Nein	20 (53%)	29 (50%)
Fatigue in der Anamnese, n (%)		
Ja	18 (47%)	27 (47%)
Nein	20 (53%)	31 (53%)
Gesamt-Intensität der PBC-bedingten Fatigue, n (%)		
Mild	9 (24%)	10 (17%)
Moderat	6 (16%)	12 (21%)
Schwer	3 (8%)	5 (9%)

	Patienten, die gemäß Fach-information behandelt wurden (N=38)	Gesamtgruppe (N=58)
Alter bei PBC Diagnose (Jahre)		
n	38	58
MW (SD)	46,1 (9,02)	46,4 (9,24)
Median	47,5	47,5
Minimum; Maximum	25; 67	25; 67
Altersklasse bei PBC Diagnose, n (%)		
<50 Jahre	20 (53%)	33 (57%)
≥50 Jahre	18 (47%)	25 (43%)
Dauer der PBC (Jahre)		
n	38	58
MW (SD)	7,9 (4,97)	8,6 (5,53)
Median	7	8,7
Minimum; Maximum	1,7; 18,7	0,8; 27,0
Dauer der PBC – Klassen, n (%)		
≤7,5 Jahre	20 (53%)	27 (47%)
>7,5 Jahre	18 (47%)	31 (53%)
AP zu Baseline (U/L)		
n	38	58
MW (SD)	296,51 (91,806)	310,42 (97,099)
Median	259,75	270,73
Minimum; Maximum	202,7; 540,1	186,8; 566,3
Totales Bilirubin zu Baseline (µmol/L)		
n	38	58
MW (SD)	10,457 (5,171)	10,571 (5,908)
Median	9,676	9,191
Minimum; Maximum	2,05; 22,23	2,05; 36,42
Konjugiertes Bilirubin zu Baseline (µmol/L)		
n	38	58
MW (SD)	4,437 (2,900)	4,650 (4,862)
Median	3,634	3,377
Minimum; Maximum	1,54; 12,14	1,54; 35,23
GGT zu Baseline (U/L)		
n	38	58
MW (SD)	252,76 (180,732)	242,38 (169,945)
Median	182,5	178,45
Minimum; Maximum	43,1; 727,8	43,1; 727,8
ALT zu Baseline (U/L)		
n	38	58
MW (SD)	63,69 (35,417)	60,64 (40,238)
Median	54,15	51,9
Minimum; Maximum	18,4; 172,9	18,4; 245,5
AST zu Baseline		
n	38	58
MW (SD)	52,57 (23,805)	51,28 (26,905)
Median	49,93	45,45
Minimum; Maximum	22,7; 133,3	20,1; 173,4
Albumin zu Baseline (g/L)		
n	38	58
MW (SD)	43,12 (3,019)	43,42 (2,956)
Median	43	43
Minimum; Maximum	34,0; 50,5	34,0; 50,5
INR zu Baseline		
n	37	57
MW (SD)	0,99 (0,152)	1,04 (0,359)
Median	1	1
Minimum; Maximum	0,9; 1,8	0,9; 3,5
kg = Kilogramm, U = Einheit, L = Liter, µmol = mikromol, g = Gramm, AP = Alkalische Phosphotase, INR = Internationales normalisiertes Verhältnis		

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Im folgenden ist die Langzeitsicherheitsstudie der POISE Studie dargestellt. Die graphische Aufbereitung der Ergebnisse findet sich in Abschnitt 4.3.1 zusammen mit dem RCT-Teil der POISE-Studie.

Die Übertragbarkeit der Langzeit-Sicherheitsstudie auf den deutschen Versorgungskontext gilt als Fortführung der POISE-Studie wie oben dargelegt (siehe Abschnitt 4.3.2.3).

4.3.2.3.3.1 Mortalität: Gesamtmortalität – 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 4-81: Operationalisierung von Gesamtmortalität - weitere Untersuchungen

Studie	Operationalisierung
POISE – Langzeit- sicherheits- studie	Gesamtrate der Todesfälle zwischen Studienbeginn und Studienende basierend auf der Sicherheitspopulation. Zusätzlich zur den Inzidenzraten wurden die zugehörigen 95%-Konfidenzintervalle berechnet. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet, einmal die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, und als zweite Population alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die aber in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse.

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Das Verzerrungspotenzial dieser einarmigen Langzeit-Sicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

Tabelle 4-82: Ergebnisse der Obeticholsäure Patienten für Gesamtmortalität aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Todesfälle/N	Prozent [95% KI]
0/38	0,00 [0,00; 9,25]
KI = Konfidenzintervall	

Gesamtgruppe

Tabelle 4-83: Ergebnisse der Obeticholsäure Patienten für Gesamtmortalität aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Todesfälle/N	Prozent [95% KI]
0/58	0,00 [0,00; 6,16]
KI = Konfidenzintervall	

In beiden Populationen wurden keine Todesfälle berichtet.

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

4.3.2.3.3.2 Morbidität: AP und Bilirubin – 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 4-84: Operationalisierung von AP und Bilirubin – weitere Untersuchungen

Studie	Operationalisierung
Primärer Endpunkt	
POISE – Langzeit-Sicherheitsstudie	Der primäre Endpunkt ist definiert als dichotomer kombinierter Ansprecher-Endpunkt mit aus AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% im Vergleich zum Ausgangswert (Baseline im RCT) und Gesamt-Bilirubin ≤ULN zu Monat 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 und 33 der Langzeit-Sicherheitsstudie. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum zweiten alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die dazugehörigen 95%-Konfidenzintervalle dargestellt.
AP-Reduktion zu Baseline	
POISE – Langzeit-Sicherheitsstudie	Der dichotome Endpunkt ist definiert als Reduktion der AP vom Ausgangswert (Baseline im RCT) zu Monat 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 und 33 der Langzeit-Sicherheitsstudie >10%, >15%, >20% und >40%. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die dazugehörigen 95%-Konfidenzintervalle dargestellt.
AP ≤ ULN	
POISE – Langzeit-Sicherheitsstudie	Der dichotome Endpunkt ist definiert als Reduktion der AP vom Ausgangswert (Baseline im RCT) zu Monat 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 und 33 der Langzeit-Sicherheitsstudie ≤ der oberen Grenze des Normbereichs (upper limit of normal). In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die dazugehörigen 95%-Konfidenzintervalle dargestellt.
Änderung zu Baseline in den biochemischen Kenngrößen AP und Bilirubin	
POISE – Langzeit-	Für die metrischen Endpunkte "Änderung zu Baseline in den biochemischen Kenngrößen" AP und Bilirubin wurde die mittlere Änderung der biochemischen Kenngrößen AP und

Sicherheitsstudie	Bilirubin (Gesamt und direktes) zu Baseline der RCT mittels gemischtem Modell für Wiederholungsmessungen (MMRM) berechnet. Als Basis diente die ITT-Population, alle vorliegenden Zeitpunkte (RCT + Langzeit-Sicherheitsstudie) wurden in die Berechnung einbezogen. Der Baselinewert wurde als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde auf Basis der Kleinste-Quadrate-(Least-Square-) Methode berechnet und die LS-Mittelwerte dargestellt. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Alle vorliegenden Zeitpunkte (RCT + Langzeit-Sicherheitsstudie) wurden in die Berechnung einbezogen.
-------------------	--

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Das Verzerrungspotenzial dieser einarmigen Langzeit-Sicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

In den folgenden Tabellen sind die Ergebnisse für die Patienten dargestellt, die gemäß Fachinformation behandelt wurden. Die graphische Aufbereitung und Bewertung findet sich in Abschnitt 4.3.1.3 als Ergänzung zu den Ergebnissen der RCT-Phase der POISE-Studie.

Tabelle 4-85: Ergebnisse der Obeticholsäure Patienten für den primären Endpunkt (AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin ≤ULN) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Monat	Ereignisse/N	Prozent [95%-KI]
3	26/38	68,42 [51,35; 82,50]
6	24/38	63,16 [45,99; 78,19]
9	27/38	71,05 [54,10; 84,58]

Monat	Ereignisse/N	Prozent [95%-KI]
12	28/38	73,68 [56,90; 86,60]
15	27/34	79,41 [62,10; 91,30]
18	29/37	78,38 [61,79; 90,17]
21	22/36	61,11 [43,46; 76,86]
24	25/36	69,44 [51,89; 83,65]
27	20/28	71,43 [51,33; 86,78]
30	7/10	70,00 [34,75; 93,33]
33	1/2	50,00 [1,26; 98,74]
KI = Konfidenzintervall		

Tabelle 4-86: Ergebnisse der Obeticholsäure Patienten für die Einzelkomponenten des primären Endpunkts aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Monat	Ereignisse/N	Prozent [95%-KI]
AP < 1.67xULN		
3	26/38	68,42 [51,35; 82,50]
6	25/38	65,79 [48,65; 80,37]
9	28/38	73,68 [56,90; 86,60]
12	31/39	79,49 [63,54; 90,70]
15	28/34	82,35 [65,47; 93,24]
18	31/38	81,58 [65,67; 92,26]
21	25/36	69,44 [51,89; 83,65]
24	27/36	75,00 [57,80; 87,88]
27	20/28	71,43 [51,33; 86,78]
30	7/10	70,00 [34,75; 93,33]
33	2/2	100,00 [15,81; 100,00]
Bilirubin ≤ ULN		
3	37/38	97,37 [86,19; 99,93]
6	37/38	97,37 [86,19; 99,93]
9	36/38	94,74 [82,25; 99,36]
12	38/39	97,44 [86,52; 99,94]
15	33/34	97,06 [84,67; 99,93]
18	36/38	94,74 [82,25; 99,36]
21	34/36	94,44 [81,34; 99,32]
24	34/36	94,44 [81,34; 99,32]
27	27/28	96,43 [81,65; 99,91]
30	10/10	100,00 [69,15; 100,00]
33	2/2	100,00 [15,81; 100,00]
KI = Konfidenzintervall ULN = Upper limit of normal (Obere Grenze des Normalbereichs) AP Reduktion ≥15%		

Tabelle 4-87: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP-Reduktion zu Baseline aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Monat	Ereignisse/N	Prozent [95%-KI]
AP Reduktion >= 10%		
3	34/38	89,47 [75,20; 97,06]
6	33/38	86,84 [71,91; 95,59]
9	32/38	84,21 [68,75; 93,98]
12	33/38	86,84 [71,91; 95,59]
15	31/34	91,18 [76,32; 98,14]
18	35/37	94,59 [81,81; 99,34]
21	32/36	88,89 [73,94; 96,89]
24	33/36	91,67 [77,53; 98,25]
27	25/28	89,29 [71,77; 97,73]
30	9/10	90,00 [55,50; 99,75]
33	2/2	100,00 [15,81; 100,00]
AP Reduktion >= 15%		
3	32/38	84,21 [68,75; 93,98]
6	32/38	84,21 [68,75; 93,98]
9	32/38	84,21 [68,75; 93,98]
12	32/38	84,21 [68,75; 93,98]
15	31/34	91,18 [76,32; 98,14]
18	34/37	91,89 [78,09; 98,30]
21	30/36	83,33 [67,19; 93,63]
24	32/36	88,89 [73,94; 96,89]
27	24/28	85,71 [67,33; 95,97]
30	8/10	80,00 [44,39; 97,48]
33	1/2	50,00 [1,26; 98,74]
AP Reduktion >= 20%		
3	30/38	78,95 [62,68; 90,45]
6	31/38	81,58 [65,67; 92,26]
9	31/38	81,58 [65,67; 92,26]
12	32/38	84,21 [68,75; 93,98]
15	30/34	88,24 [72,55; 96,70]
18	32/37	86,49 [71,23; 95,46]
21	27/36	75,00 [57,80; 87,88]
24	30/36	83,33 [67,19; 93,63]
27	20/28	71,43 [51,33; 86,78]
30	7/10	70,00 [34,75; 93,33]
33	1/2	50,00 [1,26; 98,74]
AP Reduktion >= 40%		
3	12/38	31,58 [17,50; 48,65]
6	10/38	26,32 [13,40; 43,10]
9	16/38	42,11 [26,31; 59,18]
12	16/38	42,11 [26,31; 59,18]
15	18/34	52,94 [35,13; 70,22]
18	17/37	45,95 [29,49; 63,08]

Monat	Ereignisse/N	Prozent [95%-KI]
21	17/36	47,22 [30,41; 64,51]
24	17/36	47,22 [30,41; 64,51]
27	9/28	32,14 [15,88; 52,35]
30	2/10	20,00 [2,52; 55,61]
33	0/2	0,00 [0,00; 84,19]
KI = Konfidenzintervall		

Tabelle 4-88: Ergebnisse der Obeticholsäure Patienten für den Endpunkt $AP \leq ULN$ aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Monat	Ereignisse/N	Prozent [95%-KI]
3	1/38	2,63 [0,07; 13,81]
6	1/38	2,63 [0,07; 13,81]
9	3/38	7,89 [1,66; 21,38]
12	3/38	7,89 [1,66; 21,38]
15	4/34	11,76 [3,30; 27,45]
18	3/37	8,11 [1,70; 21,91]
21	3/36	8,33 [1,75; 22,47]
24	4/36	11,11 [3,11; 26,06]
27	2/28	7,14 [0,88; 23,50]
30	1/10	10,00 [0,25; 44,50]
33	0/2	0,00 [0,00; 84,19]
KI = Konfidenzintervall		

Tabelle 4-89: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
AP, abs. Veränderung			
38	38	296,5 (91,8)	-99,46 (3,63)
Gesamt-Bilirubin, abs. Veränderung			
38	38	10,5 (5,2)	-0,66 (0,14)
Direktes Bilirubin, abs. Veränderung			
38	38	4,4 (2,9)	-0,46 (0,07)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Tabelle 4-90: Ergänzende Ergebnisse der Obeticholsäure Patienten für AP gemäß der Kriterien nach PARIS, MAYO, ROTTERDAM, TORONTO und BARCELONA aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Monat	Ereignisse/N	Prozent [95%-KI]
Paris I normal		
3	32/38	84,21 [68,75; 93,98]
6	32/38	84,21 [68,75; 93,98]
9	30/38	78,95 [62,68; 90,45]
12	34/38	89,47 [75,20; 97,06]
15	29/34	85,29 [68,94; 95,05]
18	32/37	86,49 [71,23; 95,46]
21	31/36	86,11 [70,50; 95,33]
24	31/36	86,11 [70,50; 95,33]
27	22/28	78,57 [59,05; 91,70]
30	10/10	100,00 [69,15; 100,00]
33	1/2	50,00 [1,26; 98,74]
Paris I normal, abnormal zu Baseline		
3	12/18	66,67 [40,99; 86,66]
6	12/18	66,67 [40,99; 86,66]
9	11/18	61,11 [35,75; 82,70]
12	14/18	77,78 [52,36; 93,59]
15	11/16	68,75 [41,34; 88,98]
18	12/17	70,59 [44,04; 89,69]
21	11/16	68,75 [41,34; 88,98]
24	11/16	68,75 [41,34; 88,98]
27	5/10	50,00 [18,71; 81,29]
30	2/2	100,00 [15,81; 100,00]
33	0/0	NA [NA; NA]
Paris I normal, normal zu Baseline		
3	20/20	100,00 [83,16; 100,00]
6	20/20	100,00 [83,16; 100,00]
9	19/20	95,00 [75,13; 99,87]
12	20/20	100,00 [83,16; 100,00]
15	18/18	100,00 [81,47; 100,00]
18	20/20	100,00 [83,16; 100,00]
21	20/20	100,00 [83,16; 100,00]
24	20/20	100,00 [83,16; 100,00]
27	17/18	94,44 [72,71; 99,86]
30	8/8	100,00 [63,06; 100,00]
33	1/2	50,00 [1,26; 98,74]
Paris II normal		
3	15/38	39,47 [24,04; 56,61]
6	15/38	39,47 [24,04; 56,61]
9	16/38	42,11 [26,31; 59,18]
12	19/38	50,00 [33,38; 66,62]
15	19/34	55,88 [37,89; 72,81]

Monat	Ereignisse/N	Prozent [95%-KI]
18	17/37	45,95 [29,49; 63,08]
21	18/36	50,00 [32,92; 67,08]
24	17/36	47,22 [30,41; 64,51]
27	13/28	46,43 [27,51; 66,13]
30	7/10	70,00 [34,75; 93,33]
33	0/2	0,00 [0,00; 84,19]
Paris II normal, abnormal zu Baseline		
3	15/38	39,47 [24,04; 56,61]
6	15/38	39,47 [24,04; 56,61]
9	16/38	42,11 [26,31; 59,18]
12	19/38	50,00 [33,38; 66,62]
15	19/34	55,88 [37,89; 72,81]
18	17/37	45,95 [29,49; 63,08]
21	18/36	50,00 [32,92; 67,08]
24	17/36	47,22 [30,41; 64,51]
27	13/28	46,43 [27,51; 66,13]
30	7/10	70,00 [34,75; 93,33]
33	0/2	0,00 [0,00; 84,19]
Mayo II normal		
3	26/38	68,42 [51,35; 82,50]
6	24/38	63,16 [45,99; 78,19]
9	27/38	71,05 [54,10; 84,58]
12	29/38	76,32 [59,76; 88,56]
15	28/34	82,35 [65,47; 93,24]
18	29/37	78,38 [61,79; 90,17]
21	24/36	66,67 [49,03; 81,44]
24	26/36	72,22 [54,81; 85,80]
27	20/28	71,43 [51,33; 86,78]
30	7/10	70,00 [34,75; 93,33]
33	2/2	100,00 [15,81; 100,00]
Mayo II normal, abnormal zu Baseline		
3	26/38	68,42 [51,35; 82,50]
6	24/38	63,16 [45,99; 78,19]
9	27/38	71,05 [54,10; 84,58]
12	29/38	76,32 [59,76; 88,56]
15	28/34	82,35 [65,47; 93,24]
18	29/37	78,38 [61,79; 90,17]
21	24/36	66,67 [49,03; 81,44]
24	26/36	72,22 [54,81; 85,80]
27	20/28	71,43 [51,33; 86,78]
30	7/10	70,00 [34,75; 93,33]
33	2/2	100,00 [15,81; 100,00]
Rotterdam normal		
3	28/38	73,68 [56,90; 86,60]
6	28/38	73,68 [56,90; 86,60]
9	27/38	71,05 [54,10; 84,58]
12	28/38	73,68 [56,90; 86,60]

Monat	Ereignisse/N	Prozent [95%-KI]
15	25/34	73,53 [55,64; 87,12]
18	28/37	75,68 [58,80; 88,23]
21	25/36	69,44 [51,89; 83,65]
24	26/36	72,22 [54,81; 85,80]
27	19/28	67,86 [47,65; 84,12]
30	6/10	60,00 [26,24; 87,84]
33	1/2	50,00 [1,26; 98,74]
Rotterdam normal, abnormal zu Baseline		
3	3/6	50,00 [11,81; 88,19]
6	3/6	50,00 [11,81; 88,19]
9	2/6	33,33 [4,33; 77,72]
12	2/6	33,33 [4,33; 77,72]
15	2/5	40,00 [5,27; 85,34]
18	3/6	50,00 [11,81; 88,19]
21	1/6	16,67 [0,42; 64,12]
24	2/6	33,33 [4,33; 77,72]
27	1/3	33,33 [0,84; 90,57]
30	0/0	NA [NA; NA]
33	0/0	NA [NA; NA]
Rotterdam normal, normal zu Baseline		
3	25/32	78,13 [60,03; 90,72]
6	25/32	78,13 [60,03; 90,72]
9	25/32	78,13 [60,03; 90,72]
12	26/32	81,25 [63,56; 92,79]
15	23/29	79,31 [60,28; 92,01]
18	25/31	80,65 [62,53; 92,55]
21	24/30	80,00 [61,43; 92,29]
24	24/30	80,00 [61,43; 92,29]
27	18/25	72,00 [50,61; 87,93]
30	6/10	60,00 [26,24; 87,84]
33	1/2	50,00 [1,26; 98,74]
Toronto II normal		
3	28/38	73,68 [56,90; 86,60]
6	27/38	71,05 [54,10; 84,58]
9	30/38	78,95 [62,68; 90,45]
12	30/38	78,95 [62,68; 90,45]
15	28/34	82,35 [65,47; 93,24]
18	31/37	83,78 [67,99; 93,81]
21	29/36	80,56 [63,98; 91,81]
24	31/36	86,11 [70,50; 95,33]
27	22/28	78,57 [59,05; 91,70]
30	7/10	70,00 [34,75; 93,33]
33	2/2	100,00 [15,81; 100,00]
Toronto II normal, abnormal zu Baseline		
3	27/36	75,00 [57,80; 87,88]
6	25/36	69,44 [51,89; 83,65]
9	28/36	77,78 [60,85; 89,88]

Monat	Ereignisse/N	Prozent [95%-KI]
12	28/36	77,78 [60,85; 89,88]
15	26/32	81,25 [63,56; 92,79]
18	29/35	82,86 [66,35; 93,44]
21	27/34	79,41 [62,10; 91,30]
24	30/35	85,71 [69,74; 95,19]
27	21/26	80,77 [60,65; 93,45]
30	7/10	70,00 [34,75; 93,33]
33	2/2	100,00 [15,81; 100,00]
Barcelona normal		
3	12/38	31,58 [17,50; 48,65]
6	10/38	26,32 [13,40; 43,10]
9	16/38	42,11 [26,31; 59,18]
12	16/39	41,03 [25,57; 57,90]
15	18/34	52,94 [35,13; 70,22]
18	17/38	44,74 [28,62; 61,70]
21	17/36	47,22 [30,41; 64,51]
24	17/36	47,22 [30,41; 64,51]
27	9/28	32,14 [15,88; 52,35]
30	2/10	20,00 [2,52; 55,61]
33	0/2	0,00 [0,00; 84,19]
KI = Konfidenzintervall		

Gesamtgruppe

Die Analyse der sogenannten Gesamtgruppe, also aller Patienten des Titrationsarms der RCT-Phase der POISE-Studie zeigte sehr ähnliche Ergebnisse wie die oben beschriebene Population.

Tabelle 4-91: Ergebnisse der Obeticholsäure Patienten für den primären Endpunkt (AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin ≤ULN) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Monat	Ereignisse/N	Prozent [95%-KI]
3	28/58	48,28 [34,95; 61,78]
6	28/58	48,28 [34,95; 61,78]
9	33/58	56,90 [43,23; 69,84]
12	32/57	56,14 [42,36; 69,26]
15	31/54	57,41 [43,21; 70,77]
18	34/55	61,82 [47,73; 74,59]
21	27/55	49,09 [35,35; 62,93]
24	33/55	60,00 [45,91; 72,98]
27	23/38	60,53 [43,39; 75,96]
30	9/15	60,00 [32,29; 83,66]
33	1/3	33,33 [0,84; 90,57]
KI = Konfidenzintervall		

Tabelle 4-92: Ergebnisse der Obeticholsäure Patienten für die Einzelkomponenten des primären Endpunkts aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Monat	Ereignisse/N	Prozent [95%-KI]
AP < 1.67xULN		
3	28/58	48,28 [34,95; 61,78]
6	30/58	51,72 [38,22; 65,05]
9	35/58	60,34 [46,64; 72,95]
12	35/58	60,34 [46,64; 72,95]
15	32/54	59,26 [45,03; 72,43]
18	36/56	64,29 [50,36; 76,64]
21	31/55	56,36 [42,32; 69,70]
24	35/55	63,64 [49,56; 76,19]
27	23/38	60,53 [43,39; 75,96]
30	9/15	60,00 [32,29; 83,66]
33	2/3	66,67 [9,43; 99,16]
Bilirubin ≤ ULN		
3	55/58	94,83 [85,62; 98,92]
6	55/58	94,83 [85,62; 98,92]
9	54/58	93,10 [83,27; 98,09]
12	56/58	96,55 [88,09; 99,58]
15	52/54	96,30 [87,25; 99,55]
18	54/56	96,43 [87,69; 99,56]
21	52/55	94,55 [84,88; 98,86]
24	53/55	96,36 [87,47; 99,56]
27	36/38	94,74 [82,25; 99,36]
30	15/15	100,00 [78,20; 100,00]
33	2/3	66,67 [9,43; 99,16]
KI = Konfidenzintervall ULN = Upper limit of normal (Obere Grenze des Normalbereichs) AP Reduktion ≥15% siehe Tabelle x-xx		

Tabelle 4-93: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP-Reduktion zu Baseline aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Monat	Ereignisse/N	Prozent [95%-KI]
AP Reduktion ≥ 10%		
3	49/58	84,48 [72,58; 92,65]
6	50/58	86,21 [74,62; 93,85]
9	48/58	82,76 [70,57; 91,41]
12	47/57	82,46 [70,09; 91,25]
15	46/54	85,19 [72,88; 93,38]
18	49/55	89,09 [77,75; 95,89]
21	48/55	87,27 [75,52; 94,73]

Monat	Ereignisse/N	Prozent [95%-KI]
24	49/55	89,09 [77,75; 95,89]
27	35/38	92,11 [78,62; 98,34]
30	14/15	93,33 [68,05; 99,83]
33	3/3	100,00 [29,24; 100,00]
AP Reduktion >= 15%		
3	47/58	81,03 [68,59; 90,13]
6	46/58	79,31 [66,65; 88,83]
9	48/58	82,76 [70,57; 91,41]
12	46/57	80,70 [68,09; 89,95]
15	44/54	81,48 [68,57; 90,75]
18	48/55	87,27 [75,52; 94,73]
21	46/55	83,64 [71,20; 92,23]
24	48/55	87,27 [75,52; 94,73]
27	34/38	89,47 [75,20; 97,06]
30	13/15	86,67 [59,54; 98,34]
33	2/3	66,67 [9,43; 99,16]
AP Reduktion >= 20%		
3	44/58	75,86 [62,83; 86,13]
6	44/58	75,86 [62,83; 86,13]
9	47/58	81,03 [68,59; 90,13]
12	44/57	77,19 [64,16; 87,26]
15	43/54	79,63 [66,47; 89,37]
18	45/55	81,82 [69,10; 90,92]
21	43/55	78,18 [64,99; 88,19]
24	43/55	78,18 [64,99; 88,19]
27	30/38	78,95 [62,68; 90,45]
30	12/15	80,00 [51,91; 95,67]
33	2/3	66,67 [9,43; 99,16]
AP Reduktion >= 40%		
3	13/58	22,41 [12,51; 35,27]
6	16/58	27,59 [16,66; 40,90]
9	21/58	36,21 [23,99; 49,88]
12	21/57	36,84 [24,45; 50,66]
15	24/54	44,44 [30,92; 58,60]
18	24/55	43,64 [30,30; 57,68]
21	25/55	45,45 [31,97; 59,45]
24	25/55	45,45 [31,97; 59,45]
27	12/38	31,58 [17,50; 48,65]
30	2/15	13,33 [1,66; 40,46]
33	0/3	0,00 [0,00; 70,76]
KI = Konfidenzintervall		

Tabelle 4-94: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP $\leq$ ULN aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Monat	Ereignisse/N	Prozent [95%-KI]
3	1/58	1,72 [0,04; 9,24]
6	1/58	1,72 [0,04; 9,24]
9	3/58	5,17 [1,08; 14,38]
12	3/57	5,26 [1,10; 14,62]
15	5/54	9,26 [3,08; 20,30]
18	3/55	5,45 [1,14; 15,12]
21	3/55	5,45 [1,14; 15,12]
24	4/55	7,27 [2,02; 17,59]
27	2/38	5,26 [0,64; 17,75]
30	1/15	6,67 [0,17; 31,95]
33	0/3	0,00 [0,00; 70,76]
KI = Konfidenzintervall		

Tabelle 4-95: Ergebnisse der Obeticholsäure Patienten für den Endpunkt AP und Bilirubin mit Änderung zum Vorwert aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
AP, abs. Veränderung			
58	58	310,4 (97,1)	-98,22 (3,03)
Gesamt-Bilirubin, abs. Veränderung			
58	58	10,6 (5,9)	-0,59 (0,12)
Direktes Bilirubin, abs. Veränderung			
58	58	4,7 (4,9)	-0,35 (0,07)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Tabelle 4-96: Ergänzende Ergebnisse der Obeticholsäure Patienten für AP gemäß der Kriterien nach PARIS, MAYO, ROTTERDAM, TORONTO und BARCELONA aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Monat	Ereignisse/N	Prozent [95%-KI]
Paris I normal		
3	49/58	84,48 [72,58; 92,65]
6	49/58	84,48 [72,58; 92,65]
9	47/58	81,03 [68,59; 90,13]
12	50/57	87,72 [76,32; 94,92]
15	46/54	85,19 [72,88; 93,38]

Monat	Ereignisse/N	Prozent [95%-KI]
18	49/55	89,09 [77,75; 95,89]
21	49/55	89,09 [77,75; 95,89]
24	50/55	90,91 [80,05; 96,98]
27	31/38	81,58 [65,67; 92,26]
30	15/15	100,00 [78,20; 100,00]
33	1/3	33,33 [0,84; 90,57]
Paris I normal, abnormal zu Baseline		
3	18/26	69,23 [48,21; 85,67]
6	17/26	65,38 [44,33; 82,79]
9	17/26	65,38 [44,33; 82,79]
12	19/25	76,00 [54,87; 90,64]
15	17/24	70,83 [48,91; 87,38]
18	18/24	75,00 [53,29; 90,23]
21	17/23	73,91 [51,59; 89,77]
24	18/23	78,26 [56,30; 92,54]
27	9/15	60,00 [32,29; 83,66]
30	3/3	100,00 [29,24; 100,00]
33	0/1	0,00 [0,00; 97,50]
Paris I normal, normal zu Baseline		
3	31/32	96,88 [83,78; 99,92]
6	32/32	100,00 [89,11; 100,00]
9	30/32	93,75 [79,19; 99,23]
12	31/32	96,88 [83,78; 99,92]
15	29/30	96,67 [82,78; 99,92]
18	31/31	100,00 [88,78; 100,00]
21	32/32	100,00 [89,11; 100,00]
24	32/32	100,00 [89,11; 100,00]
27	22/23	95,65 [78,05; 99,89]
30	12/12	100,00 [73,54; 100,00]
33	1/2	50,00 [1,26; 98,74]
Paris II normal		
3	16/58	27,59 [16,66; 40,90]
6	16/58	27,59 [16,66; 40,90]
9	18/58	31,03 [19,54; 44,54]
12	21/57	36,84 [24,45; 50,66]
15	21/54	38,89 [25,92; 53,12]
18	20/55	36,36 [23,81; 50,44]
21	22/55	40,00 [27,02; 54,09]
24	22/55	40,00 [27,02; 54,09]
27	14/38	36,84 [21,81; 54,01]
30	9/15	60,00 [32,29; 83,66]
33	0/3	0,00 [0,00; 70,76]
Paris II normal, abnormal zu Baseline		
3	16/58	27,59 [16,66; 40,90]
6	16/58	27,59 [16,66; 40,90]
9	18/58	31,03 [19,54; 44,54]
12	21/57	36,84 [24,45; 50,66]

Monat	Ereignisse/N	Prozent [95%-KI]
15	21/54	38,89 [25,92; 53,12]
18	20/55	36,36 [23,81; 50,44]
21	22/55	40,00 [27,02; 54,09]
24	22/55	40,00 [27,02; 54,09]
27	14/38	36,84 [21,81; 54,01]
30	9/15	60,00 [32,29; 83,66]
33	0/3	0,00 [0,00; 70,76]
Mayo II normal		
3	28/58	48,28 [34,95; 61,78]
6	29/58	50,00 [36,58; 63,42]
9	33/58	56,90 [43,23; 69,84]
12	33/57	57,89 [44,08; 70,86]
15	32/54	59,26 [45,03; 72,43]
18	34/55	61,82 [47,73; 74,59]
21	30/55	54,55 [40,55; 68,03]
24	34/55	61,82 [47,73; 74,59]
27	23/38	60,53 [43,39; 75,96]
30	9/15	60,00 [32,29; 83,66]
33	2/3	66,67 [9,43; 99,16]
Mayo II normal, abnormal zu Baseline		
3	27/57	47,37 [33,98; 61,03]
6	28/57	49,12 [35,63; 62,71]
9	33/57	57,89 [44,08; 70,86]
12	33/56	58,93 [44,98; 71,90]
15	32/53	60,38 [46,00; 73,55]
18	34/54	62,96 [48,74; 75,71]
21	29/54	53,70 [39,61; 67,38]
24	33/54	61,11 [46,88; 74,08]
27	22/37	59,46 [42,10; 75,25]
30	8/14	57,14 [28,86; 82,34]
33	2/3	66,67 [9,43; 99,16]
Rotterdam normal		
3	43/58	74,14 [60,96; 84,74]
6	43/58	74,14 [60,96; 84,74]
9	44/58	75,86 [62,83; 86,13]
12	45/57	78,95 [66,11; 88,62]
15	41/54	75,93 [62,36; 86,51]
18	44/55	80,00 [67,03; 89,57]
21	41/55	74,55 [61,00; 85,33]
24	43/55	78,18 [64,99; 88,19]
27	25/38	65,79 [48,65; 80,37]
30	10/15	66,67 [38,38; 88,18]
33	1/3	33,33 [0,84; 90,57]
Rotterdam normal, abnormal zu Baseline		
3	3/9	33,33 [7,49; 70,07]
6	3/9	33,33 [7,49; 70,07]
9	3/9	33,33 [7,49; 70,07]

Monat	Ereignisse/N	Prozent [95%-KI]
12	3/8	37,50 [8,52; 75,51]
15	2/8	25,00 [3,19; 65,09]
18	4/8	50,00 [15,70; 84,30]
21	2/8	25,00 [3,19; 65,09]
24	3/8	37,50 [8,52; 75,51]
27	1/4	25,00 [0,63; 80,59]
30	0/1	0,00 [0,00; 97,50]
33	0/1	0,00 [0,00; 97,50]
Rotterdam normal, normal zu Baseline		
3	40/49	81,63 [67,98; 91,24]
6	40/49	81,63 [67,98; 91,24]
9	41/49	83,67 [70,34; 92,68]
12	42/49	85,71 [72,76; 94,06]
15	39/46	84,78 [71,13; 93,66]
18	40/47	85,11 [71,69; 93,80]
21	39/47	82,98 [69,19; 92,35]
24	40/47	85,11 [71,69; 93,80]
27	24/34	70,59 [52,52; 84,90]
30	10/14	71,43 [41,90; 91,61]
33	1/2	50,00 [1,26; 98,74]
Toronto II normal		
3	32/58	55,17 [41,54; 68,26]
6	35/58	60,34 [46,64; 72,95]
9	40/58	68,97 [55,46; 80,46]
12	35/57	61,40 [47,57; 74,00]
15	37/54	68,52 [54,45; 80,48]
18	37/55	67,27 [53,29; 79,32]
21	37/55	67,27 [53,29; 79,32]
24	40/55	72,73 [59,04; 83,86]
27	27/38	71,05 [54,10; 84,58]
30	9/15	60,00 [32,29; 83,66]
33	2/3	66,67 [9,43; 99,16]
Toronto II normal, abnormal zu Baseline		
3	30/55	54,55 [40,55; 68,03]
6	32/55	58,18 [44,11; 71,35]
9	37/55	67,27 [53,29; 79,32]
12	33/54	61,11 [46,88; 74,08]
15	35/51	68,63 [54,11; 80,89]
18	35/52	67,31 [52,89; 79,67]
21	34/52	65,38 [50,91; 78,03]
24	38/53	71,70 [57,65; 83,21]
27	25/35	71,43 [53,70; 85,36]
30	8/14	57,14 [28,86; 82,34]
33	2/3	66,67 [9,43; 99,16]
Barcelona normal		
3	13/58	22,41 [12,51; 35,27]
6	16/58	27,59 [16,66; 40,90]

Monat	Ereignisse/N	Prozent [95%-KI]
9	21/58	36,21 [23,99; 49,88]
12	21/58	36,21 [23,99; 49,88]
15	24/54	44,44 [30,92; 58,60]
18	24/56	42,86 [29,71; 56,78]
21	25/55	45,45 [31,97; 59,45]
24	25/55	45,45 [31,97; 59,45]
27	12/38	31,58 [17,50; 48,65]
30	2/15	13,33 [1,66; 40,46]
33	0/3	0,00 [0,00; 70,76]
KI = Konfidenzintervall		

4.3.2.3.3 Morbidität: Fibrose – 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 4-97: Operationalisierung von Fibrose: Fibroscan® und ELF-Score - weitere Untersuchungen

Studie	Operationalisierung
POISE – Langzeit-sicherheitsstudie	Die Fibrosierung der Leber wurde mittels ELF-Score (14) ermittelt. Dieser misst anhand von Reagenzien die Fibrosierung der Leber als Surrogat für die mittels Biopsie ermittelten Fibrose. Die Fibrosierung der Leber wurde zudem mittels transientser Elastografie (Fibroscan®) in den Zentren erhoben, in denen das Medizinprodukt (Echosens, Paris, Frankreich) verfügbar war. Der Fibroscan ist ein validiertes nicht-invasives Verfahren (15) im Rahmen der TE zur Bestimmung der Leberfibrose. Für den ELF-Score und dessen Komponenten sowie den Fibroscan wurde die mittlere Änderung zu Baseline mittels eines gemischten Modells für Wiederholungsmessungen (MMRM) berechnet. Der Baselinewert wurde als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde auf Basis der Kleinst-Quadrate-(Least-Square-) Methode berechnet und die LS-Mittelwerte dargestellt. Zusätzlich wurde die Präsenz von Fibrose (Grenze für die Lebersteifigkeit $\geq 17,3$ kPa) als Ansprecheranalyse auf Basis des Fibroscan durchgeführt. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Die Analyse erfolgte auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlagen. Alle vorliegenden Zeitpunkte (RCT + Langzeit-Sicherheitsstudie) wurden in die Berechnung einbezogen.

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

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

Das Verzerrungspotenzial dieser einarmigen Langzeitsicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

In den folgenden Tabellen sind die Ergebnisse für die Patienten dargestellt, die gemäß Fachinformation behandelt wurden. Die graphische Aufbereitung und Bewertung findet sich in Abschnitt 4.3.1.3 als Ergänzung zu den Ergebnissen der RCT-Phase der POISE-Studie

Tabelle 4-98: Ergebnisse der Obeticholsäure Patienten für ELF-Score aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
ELF-Score, abs. Veränderung			
34	22	9,8 (1,1)	0,21 (0,07)
HA (ng/mL), abs. Veränderung			
34	22	106 (130,2)	18,93 (16,11)
P3NP (ng/mL), abs. Veränderung			
34	22	10,8 (5,5)	-0,68 (0,35)
TIMP-1 (ng/mL), abs. Veränderung			
34	22	273,6 (87,1)	40,81 (11,31)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Tabelle 4-99: Ergebnisse der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
Leber-Steifigkeit, abs. Veränderung			
18	18	11,1 (7,9)	-0,08 (0,69)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Tabelle 4-100: Ergebnisse der Ansprecheranalysen der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidity) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Monat	Ereignisse/N	Prozent [95%-KI]
Leber-Steifigkeit $\geq 17,3$ kPa		
12	3/17	17,65 [3,80; 43,43]
24	4/15	26,67 [7,79; 55,10]
KI = Konfidenzintervall		

Gesamtgruppe

Die Analyse der sogenannten Gesamtgruppe, also aller Patienten des Titrationsarms der RCT-Phase der POISE-Studie zeigte sehr ähnliche Ergebnisse wie die oben beschriebene Population.

Tabelle 4-101: Ergebnisse der Obeticholsäure Patienten für ELF-Score aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
ELF-Score, abs. Veränderung			
53	34	9,7 (1)	0,13 (0,05)
HA (ng/mL), abs. Veränderung			
53	34	97,7 (113,9)	14,44 (16,4)
P3NP (ng/mL), abs. Veränderung			
53	34	10,2 (5,1)	-0,49 (0,32)
TIMP-1 (ng/mL), abs. Veränderung			
53	34	274,3 (96,3)	37,6 (7,56)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten			

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Tabelle 4-102: Ergebnisse der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
Leber-Steifigkeit, abs. Veränderung			
30	30	10,2 (7,5)	1,08 (0,76)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Tabelle 4-103: Ergebnisse der Ansprecheranalysen der Obeticholsäure Patienten für Fibrose anhand des Fibroscan® (Morbidität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Monat	Ereignisse/N	Prozent [95%-KI]
Leber-Steifigkeit $\geq 17,3$ kPa		
12	4/33	12,12 [3,40; 28,20]
24	5/29	17,24 [5,85; 35,77]
KI = Konfidenzintervall		

4.3.2.3.4 Morbidität: Osteoporose – weitere Untersuchungen

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

Für den Endpunkt Osteoporose lagen für die weiteren Untersuchungen keine Daten vor.

4.3.2.3.5 Morbidität: PBC-40– 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 4-104: Operationalisierung von PBC-40 (Morbidität) - weitere Untersuchungen

Studie	Operationalisierung
POISE – Langzeit- sicherheits- studie	Morbidität wurde anhand des validierten Fragebogens PBC-40 bewertet (16). Der Fragebogen PBC-40 besteht aus 40 Items in sechs Subdomänen, von denen die Subdomänen Generelle Symptome, Jucken, Fatigue und kognitive Funktion als Ausprägungen der Morbidität dargestellt sind. Den Antworten zu den Fragen werden sogenannte Score-Punkte zugeordnet (1=überhaupt nicht, 2=selten, 3=manchmal, 4=häufig, 5=immer). Morbidität wurde jeweils bei Einschluss (an Tag 0), nach 2 Wochen, 3, 6, 9 und 12 Monaten im RCT und bei 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 und 33 Monaten in der Langzeit-Sicherheitsstudie, vor Verabreichung der Studienmedikation evaluiert. Für die Items des PBC-40 Fragebogens wurde die mittlere Änderung zu Baseline mittels eines gemischten Modells für Wiederholungsmessungen (MMRM) berechnet. Der Baselinewert wurde als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde auf Basis der Kleinste-Quadrate-(Least-Square-) Methode berechnet und die LS-Mittelwerte dargestellt. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet, einmal die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, und als zweite Population alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die aber in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlagen. Alle vorliegenden Zeitpunkte (RCT + Langzeit-Sicherheitsstudie) wurden in die Berechnung einbezogen.

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Das Verzerrungspotenzial dieser einarmigen Langzeitsicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

In den folgenden Tabellen sind die Ergebnisse für die Patienten dargestellt, die gemäß Fachinformation behandelt wurden. Die graphische Aufbereitung und Bewertung findet sich in Abschnitt 4.3.1.3 als Ergänzung zu den Ergebnissen der RCT-Phase der POISE-Studie

Tabelle 4-105: Ergebnisse der Obeticholsäure Patienten für PBC-40 (Morbidität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
PBC-40 Generelle Symptome, abs. Veränderung			
38	38	13,5 (4,6)	0,05 (0,14)
PBC-40 Jucken, abs. Veränderung			
37	37	3,5 (3,4)	0,64 (0,12)
PBC-40 Fatigue, abs. Veränderung			
38	38	25,8 (10,3)	-0,73 (0,26)
PBC-40 Kognitive Funktion, abs. Veränderung			
38	38	11 (5,2)	0,22 (0,15)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Gesamtgruppe

Die Analyse der sogenannten Gesamtgruppe, also aller Patienten des Titrationsarms der RCT-Phase der POISE-Studie zeigte sehr ähnliche Ergebnisse wie die oben beschriebene Population.

Tabelle 4-106: Ergebnisse der Obeticholsäure Patienten für PBC-40 (Morbidität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
PBC-40 Generelle Symptome, abs. Veränderung			
58	58	13,4 (4,9)	-0,32 (0,12)
PBC-40 Jucken, abs. Veränderung			
57	57	3,6 (3,1)	0,73 (0,1)
PBC-40 Fatigue, abs. Veränderung			
58	58	24,8 (10,6)	-0,69 (0,21)
PBC-40 Kognitive Funktion, abs. Veränderung			
58	58	11,1 (5,3)	0,18 (0,12)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

4.3.2.3.3.6 Morbidität: Pruritus – 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 4-107: Operationalisierung von Pruritus (Morbidität) - weitere Untersuchungen

Studie	Operationalisierung
POISE – Langzeit-sicherheitsstudie	Der Endpunkt Pruritus wurde anhand des unerwünschten Ereignisses Pruritus und des Fragebogens 5-D und der Pruritus-VAS erhoben. UE Pruritus Dichotomer Endpunkt hinsichtlich des Auftretens eines unerwünschten Ereignisses (UE, treatment-emergent adverse event) innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt dokumentiert wurde. Die Analyse basiert auf der Sicherheitspopulation, in die alle Patienten eingeschlossen sind, denen mindestens eine Dosis einer Studienbehandlung verabreicht wurde. Die Gruppenzuteilung erfolgte auf Basis der Studienmedikation (Obeticholsäure+UDCA vs. UDCA+Placebo bzw. Obeticholsäure vs. Placebo), nicht der Randomisierung. Pruritus-Fragebogen 5-D Zudem wurde Pruritus anhand des validierten Fragebogens 5-D (Descriptive Pruritus Scale) bewertet (12). Der Fragebogen 5-D besteht aus 27 Items in fünf Subdomänen, die als Ausprägungen der Morbidität dargestellt sind. Zudem wurde der Visual Analogue Scale erhoben und dargestellt. Pruritus-VAS Die Pruritus-Visuelle Analogskala (Pruritus-VAS) wurde zur weiteren Untersuchung des Symptoms Pruritus verwendet (13). Die Morbidität wurde jeweils bei Einschluss (an Tag 0), nach 2 Wochen, 3, 6, 9 und 12 Monaten im RCT und bei 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 und 33 Monaten in der Langzeit-Sicherheitsstudie, vor Verabreichung der Studienmedikation evaluiert. Analysen Für die Items des 5-D Fragebogens und für die Pruritus-VAS wurde die mittlere Änderung zu Baseline mittels eines gemischten Modells für Wiederholungsmessungen (MMRM) berechnet. Der Baselinewert wurde als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde auf Basis der Kleinste-Quadrate-(Least-Square-) Methode berechnet und die LS-Mittelwerte dargestellt. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlagen. Alle vorliegenden Zeitpunkte (RCT + Langzeit-Sicherheitsstudie) wurden in die Berechnung einbezogen. Für das UE Pruritus wurden die Inzidenzraten und die mediane Zeit bis zum Ereignis auf Basis der Kaplan-Meier-Methode dargestellt.

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Das Verzerrungspotenzial dieser einarmigen Langzeitsicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

In den folgenden Tabellen sind die Ergebnisse für die Patienten dargestellt, die gemäß Fachinformation behandelt wurden. Die graphische Aufbereitung und Bewertung findet sich in Abschnitt 4.3.1.3 als Ergänzung zu den Ergebnissen der RCT-Phase der POISE-Studie

Im Folgenden sind die Ergebnisse des UE Pruritus dargestellt.

Tabelle 4-108: Ergebnisse für Pruritus (Inzidenzraten) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Ereignisse/N	Prozent [95%-KI]
Patienten mit Pruritus	
20/38	52,63 [35,82; 69,02]
Patienten mit schwerem Pruritus	
7/38	18,42 [7,74; 34,33]
KI = Konfidenzintervall	

Tabelle 4-109: Ergebnisse für Pruritus (Time-to-first-event) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Ereignisse/N (%)	Median [95%-KI]
Zeit bis zum Beginn des Pruritus	
23/38 (60,53%)	475,50 [56,00; NA]
Zeit bis zum Beginn des schweren Pruritus	
10/38 (26,32%)	NA [NA; NA]
KI = Konfidenzintervall NA = nicht berechenbar	

Im Folgenden sind die Ergebnisse des Fragebogens 5-D und der Pruritus-VAS dargestellt.

Tabelle 4-110: Ergebnisse der Obeticholsäure Patienten für 5-D und Pruritus-VAS aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
5-D, Gesamtscore, abs. Veränderung			
32	32	10 (3,3)	1,11 (0,19)
5-D, Dauer, abs. Veränderung			
33	33	1,1 (0,2)	0,36 (0,06)
5-D, Schweregrad, abs. Veränderung			
36	36	1,7 (0,8)	0,24 (0,04)
5-D, Richtung, abs. Veränderung			
33	33	3,3 (1,2)	-0,06 (0,06)
5-D, Behinderung, abs. Veränderung			
35	35	1,8 (1,1)	0,19 (0,05)
5-D, Verteilung, abs. Veränderung			
36	36	1,8 (1,2)	0,36 (0,05)
Pruritus-VAS, abs. Veränderung			
36	36	16,6 (21,9)	5,38 (0,99)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Gesamtgruppe

Die Analyse der sogenannten Gesamtgruppe, also aller Patienten des Titrationsarms der RCT-Phase der POISE-Studie zeigt sehr ähnliche Ergebnisse wie die oben beschriebene Population.

Im Folgenden werden die Ergebnisse zum UE Pruritus dargestellt.

Tabelle 4-111: Ergebnisse für Pruritus (Inzidenzraten) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Ereignisse/N	Prozent [95%-KI]
Patienten mit Pruritus	
36/58	62,07 [48,37; 74,49]
Patienten mit schwerem Pruritus	

Ereignisse/N	Prozent [95%-KI]
11/58	18,97 [9,87; 31,41]
KI = Konfidenzintervall	

Tabelle 4-112: Ergebnisse für Pruritus (Time-to-first-event) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Ereignisse/N (%)	Median [95%-KI]
Zeit bis zum Beginn des Pruritus	
40/58 (68,97%)	254,50 [66,00; 738,00]
Zeit bis zum Beginn des schweren Pruritus	
14/58 (24,14%)	NA [NA; NA]
KI = Konfidenzintervall NA = nicht berechenbar	

Im Folgenden sind die Ergebnisse des Fragebogens 5-D und der Pruritus-VAS dargestellt.

Tabelle 4-113: Ergebnisse der Obeticholsäure Patienten für 5-D und Pruritus-VAS aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
5-D, Gesamtscore, abs. Veränderung			
51	51	9,8 (3,2)	1,09 (0,16)
5-D, Dauer, abs. Veränderung			
53	53	1,1 (0,2)	0,27 (0,04)
5-D, Schweregrad, abs. Veränderung			
56	56	1,9 (0,9)	0,14 (0,03)
5-D, Richtung, abs. Veränderung			
52	52	3,1 (1,2)	0,08 (0,06)
5-D, Behinderung, abs. Veränderung			
55	55	1,7 (1)	0,2 (0,04)
5-D, Verteilung, abs. Veränderung			
56	56	1,9 (1,2)	0,38 (0,04)
Pruritus-VAS, abs. Veränderung			
56	56	18,3 (22,6)	4,83 (0,82)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

4.3.2.3.3.7 Gesundheitsbezogene Lebensqualität: PBC-40 – 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 4-114: Operationalisierung von PBC-40 (Lebensqualität) - weitere Untersuchungen

Studie	Operationalisierung
POISE – Langzeit- sicherheits- studie	Gesundheitsbezogene Lebensqualität wurde anhand des validierten Fragebogens PBC-40 bewertet (16). Der Fragebogen PBC-40 besteht aus 40 Items in sechs Subdomänen, von denen die Subdomänen soziale Domäne und emotionale Domäne als Ausprägungen der gesundheitsbezogenen Lebensqualität dargestellt sind. Den Antworten zu den Fragen werden sogenannte Score-Punkte zugeordnet (1=überhaupt nicht, 2=selten, 3=manchmal, 4=häufig, 5=immer). Gesundheitsbezogene Lebensqualität wurde jeweils bei Einschluss (an Tag 0), nach 2 Wochen, 3, 6, 9 und 12 Monaten und bei 3, 6, 9, 12, 15, 18, 21,24, 27, 30 und 33 Monaten in der Langzeit-Sicherheitsstudie, vor Verabreichung der Studienmedikation evaluiert. Für die Items des PBC-40 wurde die mittlere Änderung zu Baseline mittels eines gemischten Modells für Wiederholungsmessungen (MMRM) berechnet. Der Baselinewert wurde als Co-Faktor in der Berechnung berücksichtigt. Die mittlere Änderung wurde auf Basis der Kleinst-Quadrat-(Least-Square-) Methode berechnet und die LS-Mittelwerte dargestellt. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der ITT-Population der beiden oben beschriebenen Populationen, für die der Baseline-Wert und mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlagen. Alle vorliegenden Zeitpunkte (RCT + Langzeit-Sicherheitsstudie) wurden in die Berechnung einbezogen.

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Das Verzerrungspotenzial dieser einarmigen Langzeitsicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

In den folgenden Tabellen sind die Ergebnisse für die Patienten dargestellt, die gemäß Fachinformation behandelt wurden. Die graphische Aufbereitung und Bewertung findet sich in Abschnitt 4.3.1.3 als Ergänzung zu den Ergebnissen der RCT-Phase der POISE-Studie

Tabelle 4-115: Ergebnisse für PBC-40 (Lebensqualität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
PBC-40 Soziale Domäne, abs. Veränderung			
38	38	18,9 (7,8)	0,08 (0,25)
PBC-40 Emotionale Domäne, abs. Veränderung			
38	38	6,4 (3)	-0,23 (0,09)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

Gesamtgruppe

Die Analyse der sogenannten Gesamtgruppe, also aller Patienten des Titrationsarms der RCT-Phase der POISE-Studie zeigte sehr ähnliche Ergebnisse wie die oben beschriebene Population.

Tabelle 4-116: Ergebnisse für PBC-40 (Lebensqualität) aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

N Pat. bei Baseline	N Pat. in LZ-SS	MW (SD) Baseline	LS-MW (SE) change
PBC-40 Soziale Domäne, abs. Veränderung			
58	58	18,5 (7,3)	-0,16 (0,19)
PBC-40 Emotionale Domäne, abs. Veränderung			
58	58	6,3 (2,8)	-0,27 (0,07)
LZ-SS = Langzeit-Sicherheitsstudie Pat. = Patienten MW = Mittelwert SD = Standardabweichung SE = Standardfehler LS = Least-square			

4.3.2.3.3.8 Verträglichkeit – 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 4-117: Operationalisierung von Verträglichkeit - weitere Untersuchungen

Studie	Operationalisierung
Gesamtrate UE	
POISE – Langzeit-sicherheitsstudie	Dichotomer Endpunkt bezüglich des Auftretens eines Unerwünschten Ereignisses (UE, treatment-emergent adverse event) innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt dokumentiert wurde. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der Sicherheitspopulation der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die dazugehörigen 95%-Konfidenzintervalle dargestellt.
Schwerwiegende UE (SUE)	
POISE – Langzeit-sicherheitsstudie	Dichotomer Endpunkt bezüglich des Auftretens eines UE innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt dokumentiert und als schwerwiegend (serious) eingestuft wurde. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der Sicherheitspopulation der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die dazugehörigen 95%-Konfidenzintervalle dargestellt.
Schwere UE ohne Pruritus	
POISE – Langzeit-sicherheitsstudie	Dichotomer Endpunkt hinsichtlich des Auftretens eines UE außer Pruritus innerhalb des Beobachtungszeitraums, der durch den jeweiligen verblindeten Prüfarzt als schwer (englisch: severe) dokumentiert wurde. Pruritus wurde ausgeschlossen, da dieses Symptom bereits durch die Darstellung des 5-D, der Pruritus-VAS und der Domäne Jucken des PBC-40 dargestellt ist und eine Doppelzählung vermieden werden soll. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der Sicherheitspopulation der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die

dazugehörigen 95%-Konfidenzintervalle dargestellt.	
UE, die zum Therapieabbruch führten	
POISE – Langzeit- sicherheits- studie	Dichotomer Endpunkt bezüglich des Auftretens eines UE, das zum Therapieabbruch führte, der durch den jeweiligen verblindeten Prüfarzt dokumentiert wurde. In die Analyse wurden dabei ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT gemäß deutscher Fachinformation behandelten Patienten, die in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse. Analysiert wurden die Inzidenzraten auf Basis aller Patienten der Sicherheitspopulation der beiden oben beschriebenen Populationen, für die mindestens ein Beobachtungswert nach Baseline in der Langzeit-Sicherheitsstudie vorlag. Zusätzlich zu den Inzidenzraten wurden die dazugehörigen 95%-Konfidenzintervalle dargestellt.
Abkürzungen: OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis.	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Das Verzerrungspotenzial dieser einarmigen Langzeitsicherheitsstudie wird als hoch eingestuft.

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.

Patienten, die gemäß Fachinformation behandelt wurden

In den folgenden Tabellen sind die Ergebnisse für die Patienten dargestellt, die gemäß Fachinformation behandelt wurden. Im Vergleich zur RCT-Phase der POISE-Studie war der Anteil Patienten, die die Studie wegen eines UE abbrachen, sehr ähnlich. Auch die Anteile mit schwerem Pruritus lagen auf ähnlichem Niveau.

Tabelle 4-118: Ergebnisse für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Ereignisse/N	Prozent [95%-KI]
Patienten mit UE	
36/38	94,74 [82,25; 99,36]
Patienten mit SUE	
10/38	26,32 [13,40; 43,10]
Patienten mit schweren UE ohne Pruritus	
6/38	15,79[6,02;31,25]

Ereignisse/N	Prozent [95%-KI]
UE, die zum Therapieabbruch führten	
2/38	5,26 [0,64; 17,75]
KI = Konfidenzintervall UE = unerwünschtes Ereignis SUE = schwerwiegendes unerwünschtes Ereignis	

Im Folgenden sind die ergänzenden Analysen zum Endpunkt schwere UE unter Einschluss des UE Pruritus dargestellt.

Tabelle 4-119: Ergebnisse der ergänzenden Analysen für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Ereignisse/N	Prozent [95%-KI]
Patienten mit schweren UE	
12/38	31,58 [17,50; 48,65]
KI = Konfidenzintervall UE = unerwünschtes Ereignis	

Gesamtgruppe

In der Gesamtgruppe, die auch Patienten einschloss, die in der Langzeit-Sicherheitsstudie Obeticholsäure Dosen bis zu 25 mg erhielten, war das Sicherheitsprofil sehr ähnlich zur Population der Patienten, die gemäß Fachinformation behandelt wurden. Dabei zeigte sich eine erwartete erhöhte Rate an Patienten mit Pruritus (62% vs. 53%), wobei die Rate an Patienten mit schwerem Pruritus in den beiden Populationen mit ca. 18% auf ähnlichem Niveau lag.

Tabelle 4-120: Ergebnisse für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Gesamtgruppe

Ereignisse/N	Prozent [95%-KI]
Patienten mit UE	
54/58	93,10 [83,27; 98,09]
Patienten mit SUE	
13/58	22,41 [12,51; 35,27]
Patienten mit schweren UE ohne Pruritus	
11/58	18,97 [9,87; 31,41]
UE, die zum Therapieabbruch führten	
3/58	5,17 [1,08; 14,38]
KI = Konfidenzintervall UE = unerwünschtes Ereignis SUE = schwerwiegendes unerwünschtes Ereignis	

Im Folgenden sind die ergänzenden Analysen zum Endpunkt schwere UE unter Einschluss des UE Pruritus dargestellt.

Tabelle 4-121: Ergebnisse der ergänzenden Analysen für Verträglichkeit aus Langzeit-Sicherheitsstudie mit dem zu bewertenden Arzneimittel – Patienten, die gemäß Fachinformation behandelt wurden

Ereignisse/N	Prozent [95%-KI]
Patienten mit schweren UE	
20/58	34,48 [22,49; 48,12]
KI = Konfidenzintervall UE = unerwünschtes Ereignis	

4.3.2.3.9 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.2.5.5.

Im vorliegenden Nutzendossier werden keine Subgruppenanalysen dargestellt und nicht zum Nachweis des Ausmaßes des Zusatznutzens herangezogen, da diese Angaben gemäß 1.3 der Anlage II.1 zum 5. Kapitel der Verfo nicht zu machen sind.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

In die Analyse der Langzeit-Sicherheitsstudie wurden ausschließlich Patienten eingeschlossen, die in der RCT-Phase der POISE-Studie im Titrationsarm, d.h. gemäß der deutschen Fachinformation, behandelt wurde. Es wurden dabei zwei Populationen betrachtet: Zum einen die Patienten, die auch in der Langzeit-Sicherheitsstudie gemäß deutscher Fachinformation behandelt wurden, zum anderen alle in der RCT-Phase gemäß deutscher Fachinformation behandelten Patienten, die aber in der Langzeit-Sicherheitsstudie auch Dosen über 10 mg erhalten konnten. Diese Analyse dient als Sensitivitätsanalyse.

In den Endpunkten Mortalität, Morbidität und gesundheitsbezogene Lebensqualität zeigten sich dabei Änderungen zum Ausgangswert und Ansprecherraten, die auf dem Niveau des Monats 12 der RCT-Phase der POISE-Studie lagen. Insbesondere bei AP und Bilirubin zeigt sich damit eine langfristige Wirksamkeit, die über den gesamten Beobachtungszeitraum (POISE, 12 Monate, und LTSE, 27 Monate mit ausreichender Evidenz) bis zu 39 Monaten bestehen bleibt.

Im Endpunkt Verträglichkeit wurden die Inzidenzraten der oben beschriebenen Populationen dargestellt. Wie erwartet lag der Anteil der Patienten in den Kategorien schwere UE und SUE höher als in der RCT-Phase der POISE-Studie, da die Patienten bis zu 27 Monate in der Sicherheitsstudie nachverfolgt wurden. Im Vergleich zur RCT-Phase der POISE-Studie war der Anteil der Patienten, die die Studie wegen eines UE abbrachen, ähnlich niedrig (7% vs. 5%). Auch die Anteile der Patienten mit schwerem Pruritus lagen auf ähnlichem Niveau (ca. 18%).

In der Gesamtgruppe, die auch Patienten einschloss, die in der Langzeit-Sicherheitsstudie Obeticholsäure Dosen bis zu 25 mg erhielten, war das Sicherheitsprofil sehr ähnlich zur Population der Patienten, die gemäß Fachinformation behandelt wurden. Dabei zeigte sich eine erwartete erhöhte Rate an Patienten mit Pruritus (62% vs. 53%), wobei die Rate an Patienten mit schwerem Pruritus in den beiden Populationen mit ca. 18% auf ähnlichem Niveau lag.

Im vorliegenden Nutzendossier werden keine Subgruppenanalysen dargestellt und nicht zum Nachweis des Ausmaßes des Zusatznutzens herangezogen, da diese Angaben gemäß 1.3 der Anlage II.1 zum 5. Kapitel der VerfO nicht zu machen sind.

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.

Die bibliographische, systematische Literaturrecherche (Stand 31.10.2016), eine umfassende Recherche in Studienregistern (Stand 17.10.2016) und die ausführliche Sichtung der Studien des pharmazeutischen Unternehmers (Stand 31.10.2016) stellen sicher, dass in die Bestimmung des Ausmaßes des Zusatznutzens alle verfügbare Evidenz aus randomisierten kontrollierten Studien eingegangen ist. Es wurde eine doppelblinde, randomisierte kontrollierte Phase-III-Zulassungsstudie („POISE“, 747-301, NCT 01473524) identifiziert.

Die POISE-Studie entspricht nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie der Evidenzstufe Ib und ist damit für die Herleitung des Ausmaßes des Zusatznutzens geeignet.

In der POISE-Studie wurde die Randomisierung adäquat durchgeführt und die Gruppenzuteilung verdeckt über ein Sprachdialogsystem (IxRS) durchgeführt. Patienten und behandelte Personen waren während der randomisierten Phase verblindet. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren oder – wie bei den Sicherheitsendpunkten – gemäß Verfahrensordnung berichtet werden mussten (z.B. die Berechnung der Effektmaße für den Vergleich der Behandlungsgruppen). Es lagen keine sonstigen Aspekte für ein erhöhtes Verzerrungspotenzial vor, so dass das Verzerrungspotenzial insgesamt als niedrig eingestuft wird.

Zusätzlich wird die Langzeit-Sicherheitsstudie als einarmige Anschlußstudie zur POISE-Studie herangezogen. Diese entspricht als einarmige Studie dem Evidenzgrad IV. Die Langzeitdaten erlauben allerdings eine Abschätzung der langfristigen Sicherheit und Wirksamkeit von Obeticholsäure in dieser chronischen und progressiven Erkrankung.

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.

Im Folgenden wird die Zusammenfassung der Ergebnisdarstellung aus Abschnitt 4.3.1 für alle patientenrelevanten Endpunkte unter Angabe des Ausmaßes des nachgewiesenen Zusatznutzens von Obeticholsäure für die beiden Teilpopulationen der Patienten mit Kombinationstherapie mit UDCA und mit Monotherapie dargestellt.

Patienten mit Kombinationstherapie

Die randomisierte, kontrollierte POISE-Studie („747-301“, NCT 01473524) stellt nach G-BA-Verfahrensordnung eine Evidenzhierarchie der Stufe Ib dar und ist damit zur Herleitung des Ausmaßes des Zusatznutzens sowohl für Patienten mit Kombinationstherapie als auch für Patienten mit Monotherapie geeignet.

Tabelle 4-122: Ausmaß des Zusatznutzens von Obeticholsäure auf Endpunktebene - Patienten mit Kombinationstherapie

	Effektschätzer (95% KI) / Ereignisanteil OCA vs. Placebo / p-Wert	Ableitung des Ausmaßes
Mortalität		
Gesamtüberleben	RRR=0,32 (95% KI: 0,01; 7,69) 2% vs. 0% p=0,3063	Kein Zusatznutzen nachweisbar
Morbidität		
Primärer Endpunkt ^a Monat 12	RRR=0,22 (95% KI: 0,11; 0,47) 46% vs. 10% p<0,0001	Zusatznutzen mit Ausmaß beträchtlich
Fibrose	Keine relevanten Unterschiede in der SMD für ELF-Score (alle Komponenten) und den Fibroscan	Kein Zusatznutzen nachweisbar
Osteoporose	Keine relevanten Unterschiede in der SMD im DEXA-Scan (alle Komponenten)	Kein Zusatznutzen nachweisbar
PBC-40	Keine relevanten Unterschiede in der SMD in den Symptomen generelle Symptome, Jucken, Fatigue, kognitive Funktion	Kein Zusatznutzen nachweisbar
Pruritus	Keine relevanten Unterschiede in der SMD für den 5-D (alle Komponenten), Pruritus-VAS und im UE Pruritus in den Kategorien alle UE Pruritus, schwere UE Pruritus, Pruritus, der zum Therapieabbruch führte, Interventionen wegen Pruritus, Zeit bis zum Beginn des Pruritus, Zeit bis zum Beginn des schweren Pruritus	Kein Zusatznutzen nachweisbar
Gesundheitsbezogene Lebensqualität		
PBC-40	Keine relevanten Unterschiede in der SMD in beiden Domänen (soziale und emotionale Domäne)	Kein Zusatznutzen nachweisbar
Verträglichkeit		
UE	OCA/Placebo: RR=1,03 (95% KI: 0,92; 1,14) 92% vs. 90%	Kein geringerer oder größerer Schaden nachweisbar
SUE	OCA/Placebo: RR=2,79 (95% KI: 0,78; 10,07) 12% vs. 4%	Kein geringerer oder größerer Schaden nachweisbar
Schwere UE ohne UE Pruritus	OCA/Placebo:	Kein geringerer oder größerer

	RR=1,58 (95% KI: 0,60; 4,17) 14% vs. 9%	Schaden nachweisbar
UE, die zum Therapieabbruch führten	OCA/Placebo: RR=5,29 (95% KI: 0,64; 43,92) 8% vs. 1%	Kein geringerer oder größerer Schaden nachweisbar
^a Primärer Endpunkt: Anteil Patienten mit AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin ≤ULN Abkürzungen: KI: Konfidenzintervall; n. b.: nicht berechenbar; RR: relatives Risiko; RRR: reversiertes relatives Risiko; SUE: schwerwiegende unerwünschte Ereignisse; UE: unerwünschte Ereignisse; vs. versus.		

Für den Endpunkt Mortalität ließ sich kein Zusatznutzen für Obeticholsäure nachweisen.

In der Endpunktdimension Morbidität wurde der Zusatznutzen anhand der Endpunktkomplexe AP und Bilirubin, Fibrose, Osteoporose und PBC-40 hergeleitet. Für den Endpunktkomplex AP und Bilirubin wurde der primäre Endpunkt der POISE-Studie als Hauptoperationalisierung herangezogen (Anteil Patienten mit AP <1.67x ULN bei gleichzeitiger Reduktion des AP-Wertes um mindestens 15% und Gesamt-Bilirubin ≤ULN). Der primäre Endpunkt wurde in 46% der Patienten der OCA+UDCA-Gruppe erreicht im Vergleich zu 10% in der Kontrollgruppe ($p < 0,0001$) erreicht. Dies entspricht einem relativen Risiko von 0,22 (95% KI: 0,11; 0,47). Da die biochemischen Kenngrößen AP und Bilirubin mit dem Grad der Leberschädigung korrelieren, handelt es sich um einen validen Surrogatendpunkt für eine patientenrelevante Morbiditäts- und Mortalitätsveränderung. Wie in Abschnitt 4.5.4 ausführlich beschrieben, ist die Verbesserung der Surrogatparameter AP und Bilirubin gemäß der Analysen des UK-PBC- und des GLOBE-Registers prädiktiv für die Krankheitsbelastung und insbesondere die Überlebensdauer und die Zeit bis zu einer Lebertransplantation.

Mit dem durchgängigen, signifikanten Erreichen des primären Endpunkts in der POISE-Studie ergibt sich eine patientenrelevante, spürbare Linderung der Erkrankung, wodurch ein beträchtlicher Zusatznutzen gegenüber der Kontrollgruppe belegt ist.

Die Einzelkomponenten und die übrigen Operationalisierungen des Endpunktes AP und Bilirubin unterstützen den beträchtlichen Zusatznutzen.

Bzgl. der übrigen Morbiditätsendpunkte zeigten sich keine Vor- oder Nachteile für Obeticholsäure gegenüber der Kontrollgruppe UDCA+Placebo. Dies ist unter anderem der sehr langsam fortschreitenden Erkrankung und dem Zeitpunkt der Behandlung geschuldet, da die klinischen Ereignisse Leberversagen oder Leber-bedingte Mortalität sowie andere Krankheitssymptome erst im Verlauf der Erkrankung entstehen können. Die frühzeitige Behandlung der PBC hat das Ziel, diese schwerwiegenden Ereignisse zu vermeiden. Eine Erhebung solcher Ereignisse in einer klinischen Studie erfordert eine Nachbeobachtungszeit von mehreren Jahren / Jahrzehnten und kann zum Zeitpunkt der Zulassung von Obeticholsäure daher noch nicht vorliegen.

Bei den Verträglichkeitsendpunkten bestand kein Unterschied in den UE, den SUE, den schweren UE ohne Pruritus und den UE, die zum Therapieabbruch führten. Pruritus ist ein krankheitsbedingtes Grundsymptom der PBC. In den verschiedenen Operationalisierungen der POISE-Studie zeigten sich keine signifikanten Unterschiede gegenüber der Kontrollgruppe. Die Ergebnisse der Verträglichkeitsendpunkte mindern somit nicht den beträchtlichen Zusatznutzen von Obeticholsäure.

In der einarmigen Langzeit-Sicherheitsstudie (LTSE) bestätigte sich der in der randomisierten Therapiephase nachgewiesene Zusatznutzen von Obeticholsäure. Die Langzeitdaten erlauben dabei eine Abschätzung der langfristigen Sicherheit und Wirksamkeit von Obeticholsäure in dieser chronischen und progressiven Erkrankung. Bei den Endpunkten Mortalität, Morbidität und gesundheitsbezogene Lebensqualität zeigten sich Ergebnisse, die auf dem Niveau des Monats 12 der RCT-Phase der POISE-Studie oder darüber lagen. Insbesondere bei AP und Bilirubin bestätigt sich damit eine langfristige Wirksamkeit, die über den gesamten Beobachtungszeitraum (POISE, 12 Monate, und LTSE, 27 Monate mit ausreichender Evidenz) bis zu 39 Monaten bestehen bleibt. Hinsichtlich der Sicherheits-Endpunkte liegen keine Auffälligkeiten vor. Insbesondere die Rate an Patienten, die die Langzeit-Sicherheitsstudie wegen UE abbrechen, lag dabei mit 5% sehr niedrig. Die in der LTSE beobachteten Effekte sind daher zu den Ergebnissen der POISE-Studie gleichgerichtet und bestätigen den dort gefundenen beträchtlichen Zusatznutzen.

In der Gesamtschau ergibt sich damit für Obeticholsäure bei Patienten mit Kombinationstherapie ein beträchtlicher Zusatznutzen.

Patienten mit Monotherapie

Für die Patienten mit Monotherapie lagen insgesamt nur Daten von 5 Patienten aus der Obeticholsäure-Behandlungsgruppe und 5 Patienten aus der Placebo-Behandlungsgruppe vor.

Auf Grund der geringen, nicht ausreichend belastbaren Evidenzlage wurde auf die Herleitung eines Zusatznutzens auf Endpunktebene verzichtet. Gemäß § 35a Abs. 1 Satz 10 1. HS SGB V gilt der medizinische Zusatznutzen in dieser Teilpopulation durch die Zulassung als belegt.

In der Gesamtschau ergibt sich damit für Obeticholsäure bei Patienten mit Monotherapie ein nicht quantifizierbarer Zusatznutzen.

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 4-123: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten mit Kombinationstherapie	Zusatznutzen mit Ausmaß beträchtlich
Patienten mit Monotherapie	Zusatznutzen mit Ausmaß nicht quantifizierbar

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.

Trifft nicht zu.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte 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.

Trifft nicht zu.

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.

Trifft nicht zu.

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 2005¹³,

¹³ Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

Molenberghs 2010¹⁴). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁵) 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 2006¹⁶) 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 klinischen Studien zur PBC werden Surrogatparameter verwendet, um die Wirksamkeit bezüglich der klinischen Endpunkte, insbesondere Leber-bedingte Mortalität und Lebertransplantation abzubilden. Dabei schlagen Silveira et al. (4) für RCTs bei PBC als

¹⁴ 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.

¹⁵ Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat* 2006; 5(3): 173-186.

¹⁶ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006; 25(2): 183-203.

primären Endpunkt eine Kombination aus AP und Bilirubin vor, um eine bestmögliche Abschätzung der Wirksamkeit zu erzielen.

Eine formale Validierung der Surrogatendpunkte AP und Bilirubin bezüglich der klinischen Endpunkte Leber-bedingte Mortalität und Lebertransplantation in RCTs ist durch die lange zeitliche Verzögerung von eventuell mehreren Jahrzehnten zwischen der Schädigung der Leber und den klinischen Endpunkten nicht möglich. Der Verlauf der Erkrankung wird dabei mit Hilfe der biochemischen Kenngrößen AP und Bilirubin abgeschätzt, mit denen man auf den Grad der Leberschädigung rückschließt. Für die PBC-Therapie mittels UDCA wurden diverse Ansprech-Kriterien auf Basis verschiedener Kombinationen aus AP, AST und Bilirubin entwickelt und validiert. Diese umfassen unter anderem die Kriterien Paris I und II, Rotterdam, Toronto usw. Eine Liste der verfügbaren Kriterien und eine Diskussion in Abschnitt 4.2.5.2 (siehe insbesondere Tabelle 4-3).

Die Datenbasis für die Validierung der Surrogate AP und Bilirubin kann daher nur aus Registern stammen, die Langzeitbeobachtungen von Patienten über viele Jahrzehnte beinhalten.

Wie in Modul 3 dargestellt, sind bei Diagnosestellung die Mehrzahl der Patienten asymptomatisch, es können aber auch Grundsymptome der PBC wie Fatigue, Pruritus und/oder Ikterus bestehen. Meist wird die Diagnose gestellt, nachdem andere Ursachen einer Lebererkrankung ausgeschlossen wurden (37), (38). Die meisten Patienten haben abnormale Leberwerte bezüglich Alanin-Aminotransferase (ALT), Asparat-Aminotransferase (AST) und alkalische Phosphatase (AP). Zudem sind bei etwa 95% der Patienten anti-mitochondriale Antikörper (AMA) nachweisbar (37), (38). Für die Diagnose einer PBC sollten mindestens zwei der folgenden drei Kriterien erfüllt sein (37), (38):

1. Biochemischer Nachweis von Cholestase anhand von Serum-Lebertests, vor allem erhöhte alkalische Phosphatase [AP]
2. Nachweis von anti-mitochondrialen Antikörpern (AMA) ($\geq 1:40$)
3. Histologischer Nachweis einer nicht-eitrigen (sog. nicht-suppurativen) destruktiven Cholangitis und Zerstörung intralobulärer Gallengänge

Erste Hinweise auf die Validität von Bilirubin für das Gesamtüberleben wurden von Shapiro et al (1979) (39) berichtet. In den vorgestellten Analysen, getrennt nach lebenden und verstorbenen Patienten, ergibt sich ein deutlicher Hinweis, dass mit weniger steilem Anstieg des Bilirubins eine geringere Todesrate einhergeht:

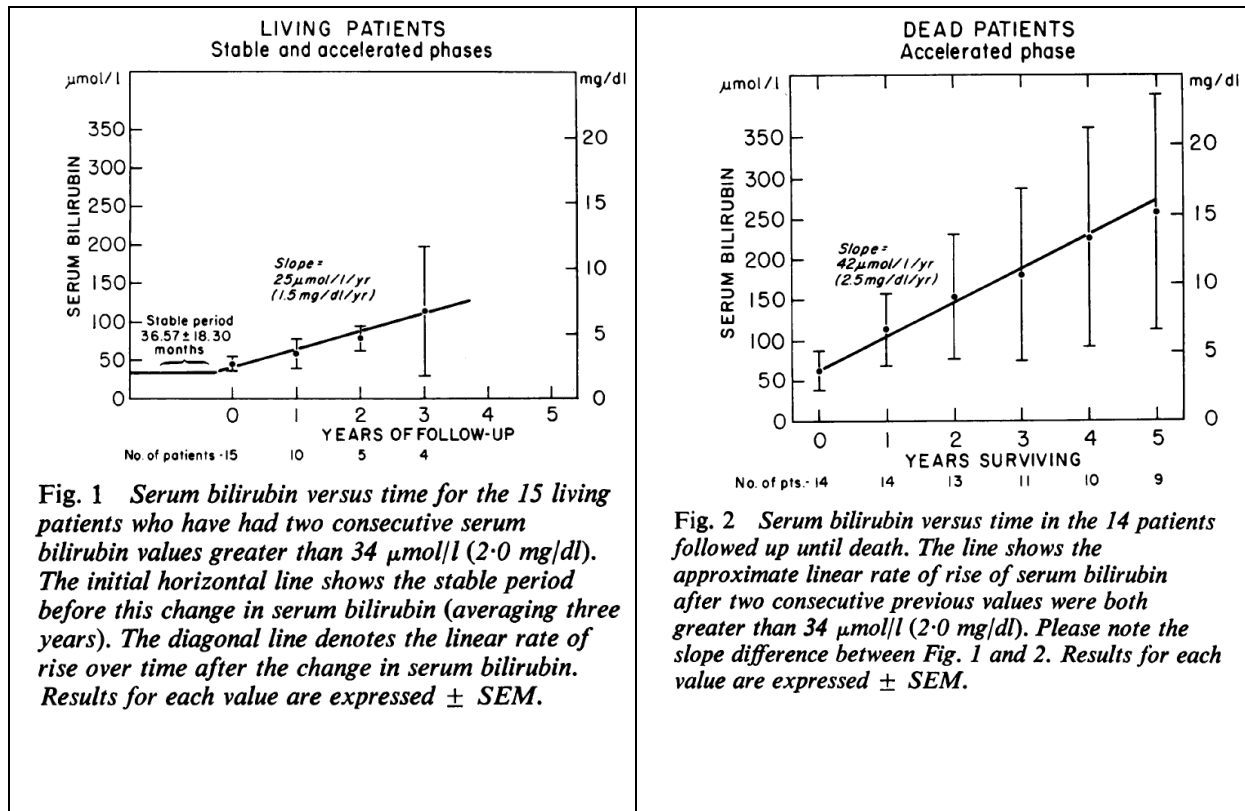


Abbildung 4-45: Serum-Bilirubin gegen Zeit der Nachbeobachtung (39)

Dieser erste Trend suggeriert einen gewissen Zusammenhang zwischen Bilirubin und Mortalität, wobei Confounding durch unbekannte Faktoren nicht ausgeschlossen werden kann.

Pares et al 2006 (23) berichten ebenfalls für Ansprecher bezüglich AP Reduktion > 40% oder Normalisierung des AP innerhalb eines Jahres nach Behandlung mit UDCA Unterschiede beim Überleben zwischen Ansprechern und Nicht-Ansprechern:

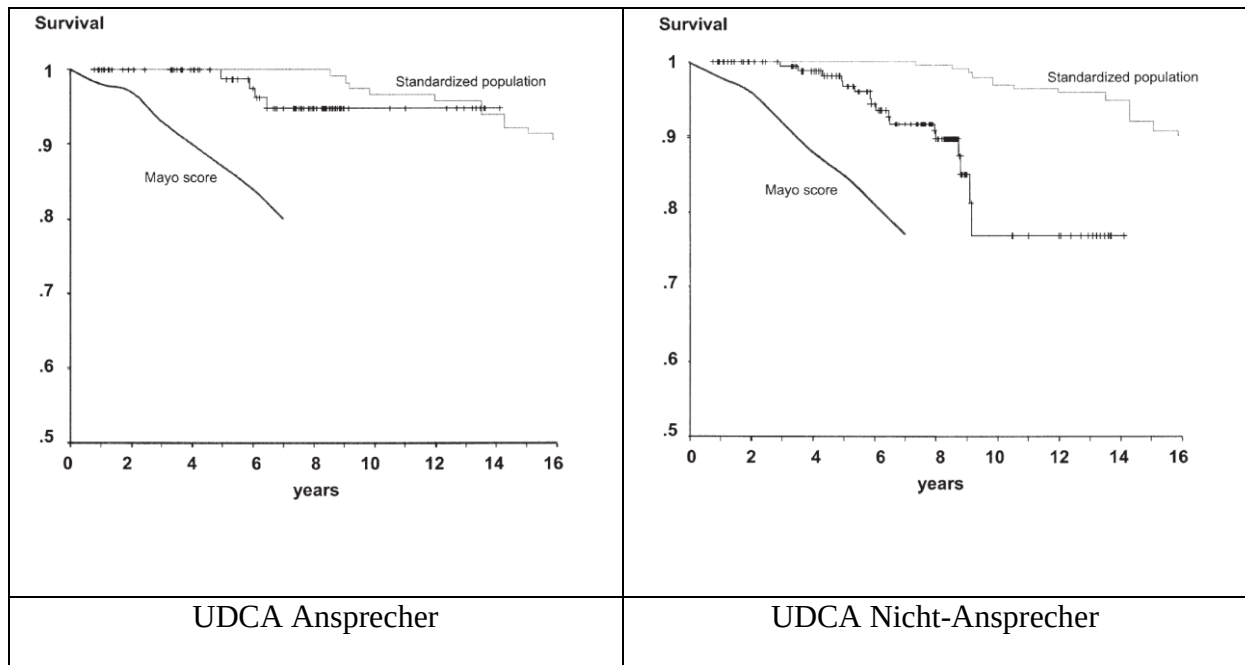


Abbildung 4-46: Mayo Score gegen die Zeit der Nachbeobachtung, getrennt nach UDCA Ansprechern und Nicht-Ansprechern (23)

Im Folgenden werden zwei große Register herangezogen, um die Validität der Surrogate AP und Bilirubin zu zeigen.

Validierung der Surrogate AP und Bilirubin anhand UK-PBC Register

Wie schon im Modul 3 (Seite 21 ff.) dargestellt, wurde aus dem Register der UK-PBC Study Group auf der Basis der biochemischen Variablen AP, Transaminasen (ALT wenn verfügbar, sonst AST), Bilirubin gesamt, Albumin sowie Thrombozytenzahl nach 1 Jahr Therapie mit UDCA ein Vorhersagemodell für den klinisch relevanten Endpunkt Tod auf Grund Leberveränderungen oder Lebertransplantation entwickelt (5). Dieses – an einer europäischen Population entwickelte – Modell zeigt, wie sich die Prognose bei Patienten mit einer guten versus einer schlechten Antwort auf UDCA verändert.

Eine ausführliche Beschreibung des UK-PBC Registers liegt nicht in publizierter Form vor.

Neben der Vorhersage von Tod oder Lebertransplantation wurde anhand der UK-PBC Kohorte auch nachgewiesen, dass ein Nicht-Ansprechen nach 12 Monaten – unabhängig vom

verwendeten Ansprech-Kriterium – ein erhöhtes Risiko für ein hepatozelluläres Karzinom als Morbiditätsendpunkt bedeutet. (6)

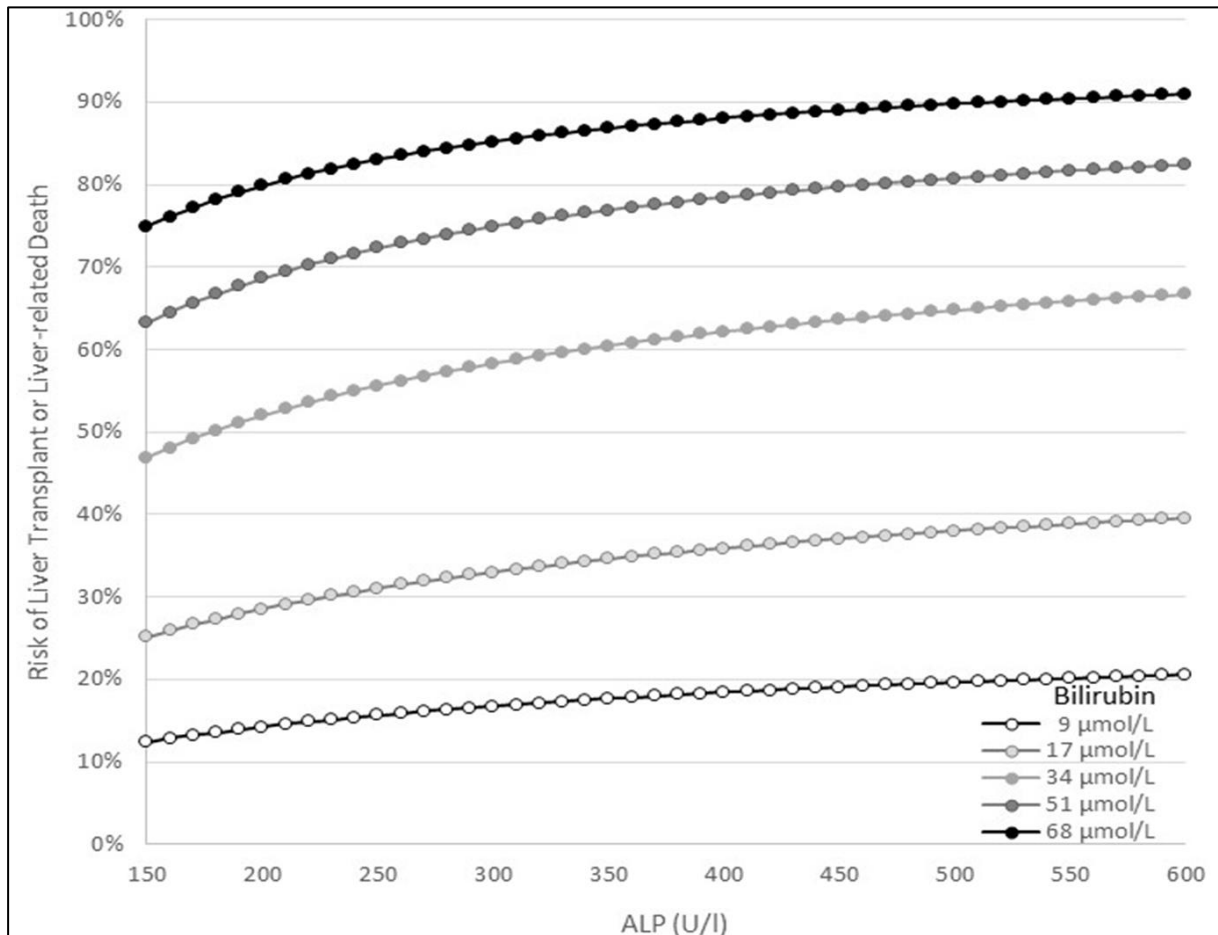


Abbildung 4-47: UK-PBC Vorhersagemodell für Tod auf Grund Leberveränderungen oder Lebertransplantation in 15 Jahren bei unterschiedlicher AP nach 1 Jahre Therapie mit UDCA. Jeweils dargestellt bei Annahme von ALT=84 U/l, Albumin 40 g/l (1,143 des unteren Normwerts 35 g/l) sowie einer Plättchenzahl von $120 \times 10^9/l$. Es wird je nach Bilirubin in den Kurven für unterschiedliche Werte der AP jeweils das 15-Jahres Risiko für Tod auf Grund Leberveränderungen oder Lebertransplantation dargestellt. Der Normalwert für Gesamt-Bilirubin wird dabei $<18 \mu\text{mol/L}$ angegeben. (5)

Validierung der Surrogate AP und Bilirubin anhand der GLOBE-Kohorte

In der internationalen GLOBE Kohorte wurden ebenfalls Untersuchungen zur Validität der Surrogate AP und Bilirubin durchgeführt.

Die Patienten der GLOBE Kohorte umfaßte 6.191 Patienten. Davon wurden 23 wegen unsicherer Diagnose ausgeschlossen, 286 wegen fehlender Laborwerte, 268 wegen unbekanntem Start der UDCA Therapie oder unbekanntem letzten Nachbeobachtungsdatum, 679 wegen zu kurzer Nachbeobachtungszeit (< sechs Monate) und 90 Patienten wegen AIH Overlap Erkrankung. Damit verblieben 4.845 Patienten für die Analyse (7).

Die Population umfaßte 4.119 Patienten mit UDCA-Therapie und 640 Patienten ohne berichtete UDCA-Therapie; von diesen war bei 83 Patienten unbekannt, ob sie eine UDCA Therapie erhalten haben oder nicht. Das mittlere Alter lag in der betrachteten Gesamtkohorte bei 53 Jahren (SD=12) mit einem Anteil Frauen von 4.348/4.845 (89,7%). Das mediane Jahr des Eintritts in das Register lag bei 1998 (Interquartilsrange von 1991 bis 2004), die mediane Nachbeobachtungszeit bei 7,3 Jahren (Interquartilsrange von 3,6 bis 11,5 Jahre). Von 4.845 Patienten lebten noch 3.727 (76,9%) und hatten keine Lebertransplantation; 1.118 (23,1%) waren verstorben oder hatten eine Lebertransplantation. Bei 729 (15,0%) lag Tod jeglicher Ursache vor, davon erlitten 358 (7,4%) einen leberbedingten Tod. Lebertransplantationen lagen bei 389 (8,0%) Patienten vor.

Aus dieser Kohorte ergeben sich bestätigende Ergebnisse (8): das GLOBE-Vorhersagemodell (GLOBE-SCORE) bezieht sich auf den kombinierten Endpunkt Tod unabhängig von der Ursache sowie Lebertransplantation. Es bezieht sich ebenfalls auf den Zeitpunkt nach einem Jahr Therapie mit UDCA, zu dem der Therapieeffekt bzw. Ansprechen auf UDCA gut bestimmbar ist. Der GLOBE-Score beinhaltet das Alter bei Therapiestart, sowie die biochemischen Variablen Bilirubin, AP, Albumin und Thrombozytenzahl (8).

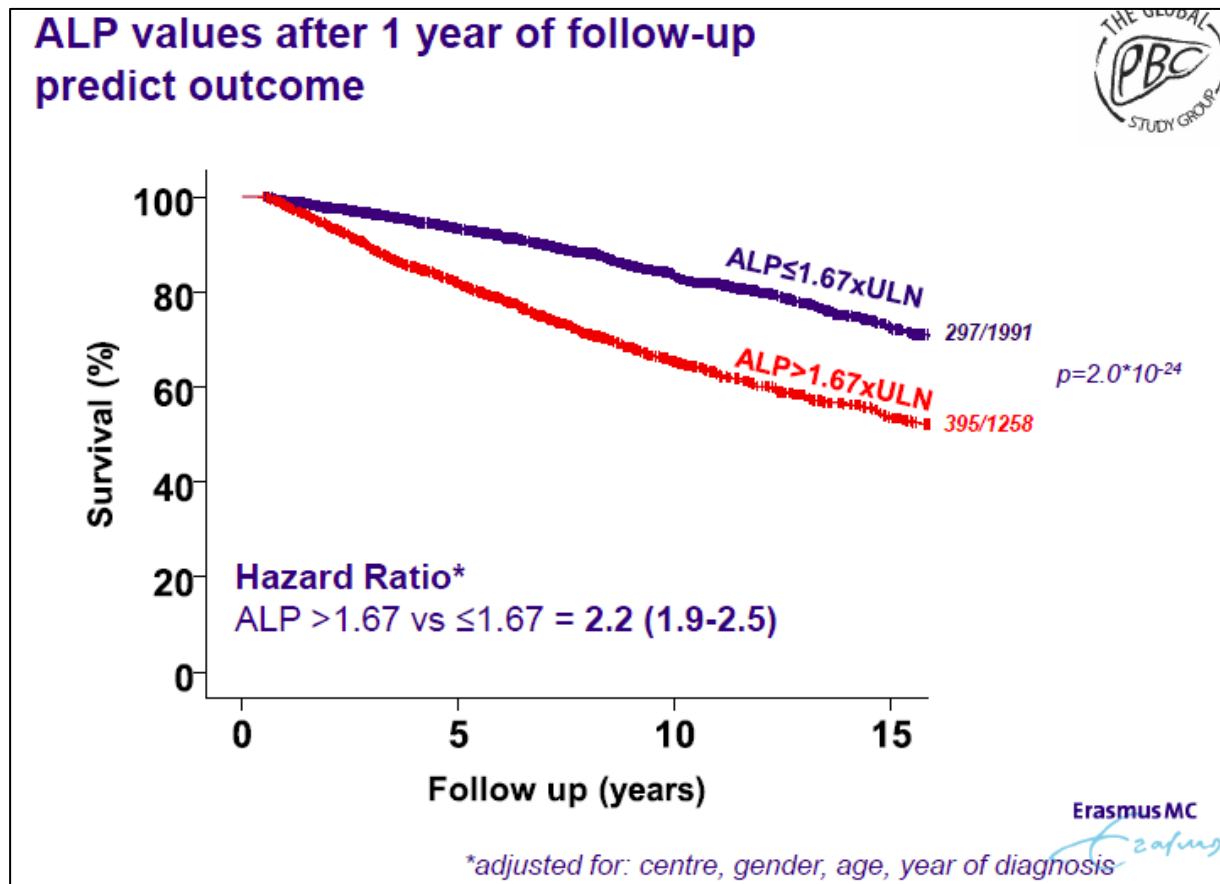


Abbildung 4-48: Hansen et al (2015): Kaplan-Meier-Kurve für den Vergleich der AP-Werte (7)

Die Hazard Ratios lagen für den Vergleich des AP bei 2,2 (95%-KI [1,9; 2,5]) zu Ungunsten der Patienten, die ein AP $> 1,67 \times \text{ULN}$ nach einem Jahr aufwiesen. Für die UDCA-Behandelten lag das HR im gleichen Rahmen (HR=2,2 (95%-KI [1,9; 2,6]), für Patienten ohne Bericht einer UDCA-Behandlung bei 1,8 (95%-KI [1,1; 2,9]). Das Ausmaß des HR zeigt, dass das Risiko der Mortalität bei AP-Werten $> 1,67 \times \text{ULN}$ um mehr als das Doppelte erhöht ist. Dieser Effekt zeigt sich durchgängig auch in allen Kategorien der Subgruppen nach Alter, Geschlecht, Jahr der Diagnose, Alter bei Diagnose und histologischem Status, und war somit über alle Subgruppen konsistent.

Ein gleiches Bild zeigt sich für das Bilirubin mit einem HR von 5,1 (95%-KI [4,3; 5,9]).

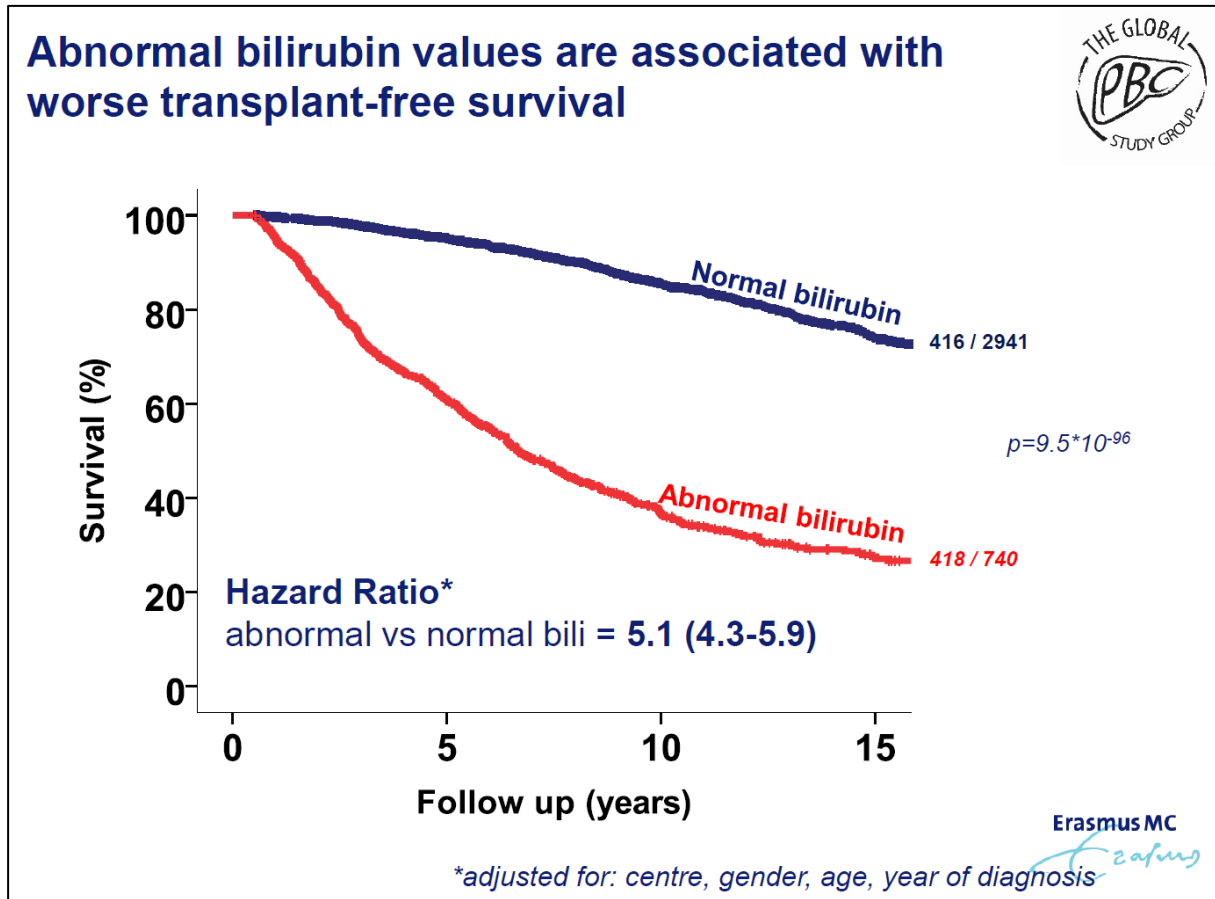


Abbildung 4-49: Hansen et al (2015): Kaplan-Meier-Kurve für den Vergleich der Bilirubin-Werte (7)

Auch eine Kombination der beiden Parameter AP und Bilirubin, wie sie in der POISE-Studie Verwendung fand, zeigte deutliche positive Effekte der Surrogate auf das Gesamtüberleben.

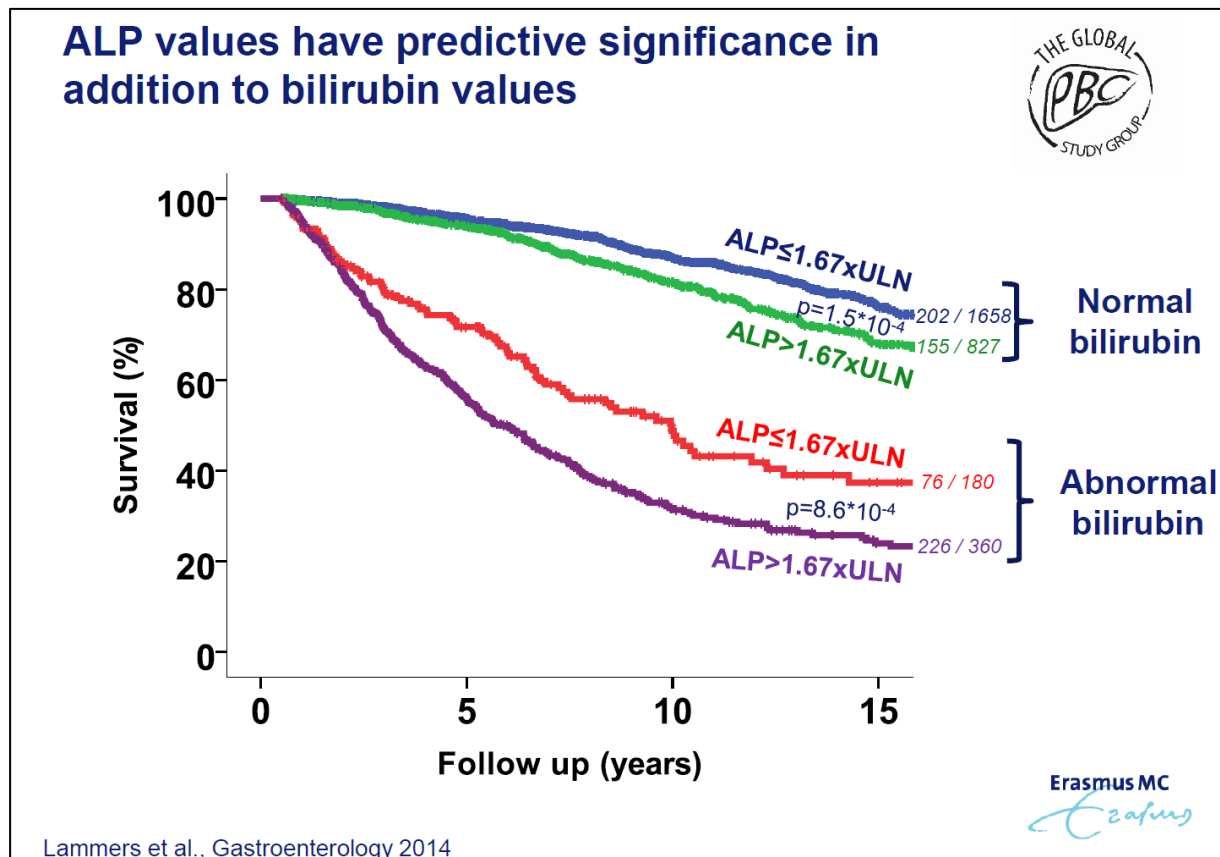


Abbildung 4-50: Lammers et al (2014): Kaplan-Meier-Kurve für den Vergleich der Bilirubin-Werte (40)

In einer weiteren Analyse wurden die UDCA-Ansprecher gegen die Nicht-Ansprecher verglichen. Mit einem Grenzwert von 0,30 hatten Ansprecher auf UDCA nach einem Jahr eine Lebenserwartung, die der Normalbevölkerung entspricht: das transplantationsfreie 5-,10- bzw. 15-Jahresüberleben von Ansprechern auf UDCA lag bei 98%, 92% bzw. 82% (8). Hingegen hatten Nicht-Ansprecher auf UDCA (rund 40% aller Patienten in GLOBE) eine deutlich reduzierte Prognose: in der Gruppe mit >90tem Perzentil des GLOBE-Scores reduzierte sich bereits die transplantatfreie 5 Jahres Überlebensrate auf nur rund 46% (8). Die folgende Abbildung 4-51 zeigt das Lebertransplantat-freie Überleben in Abhängigkeit vom GLOBE-Score.

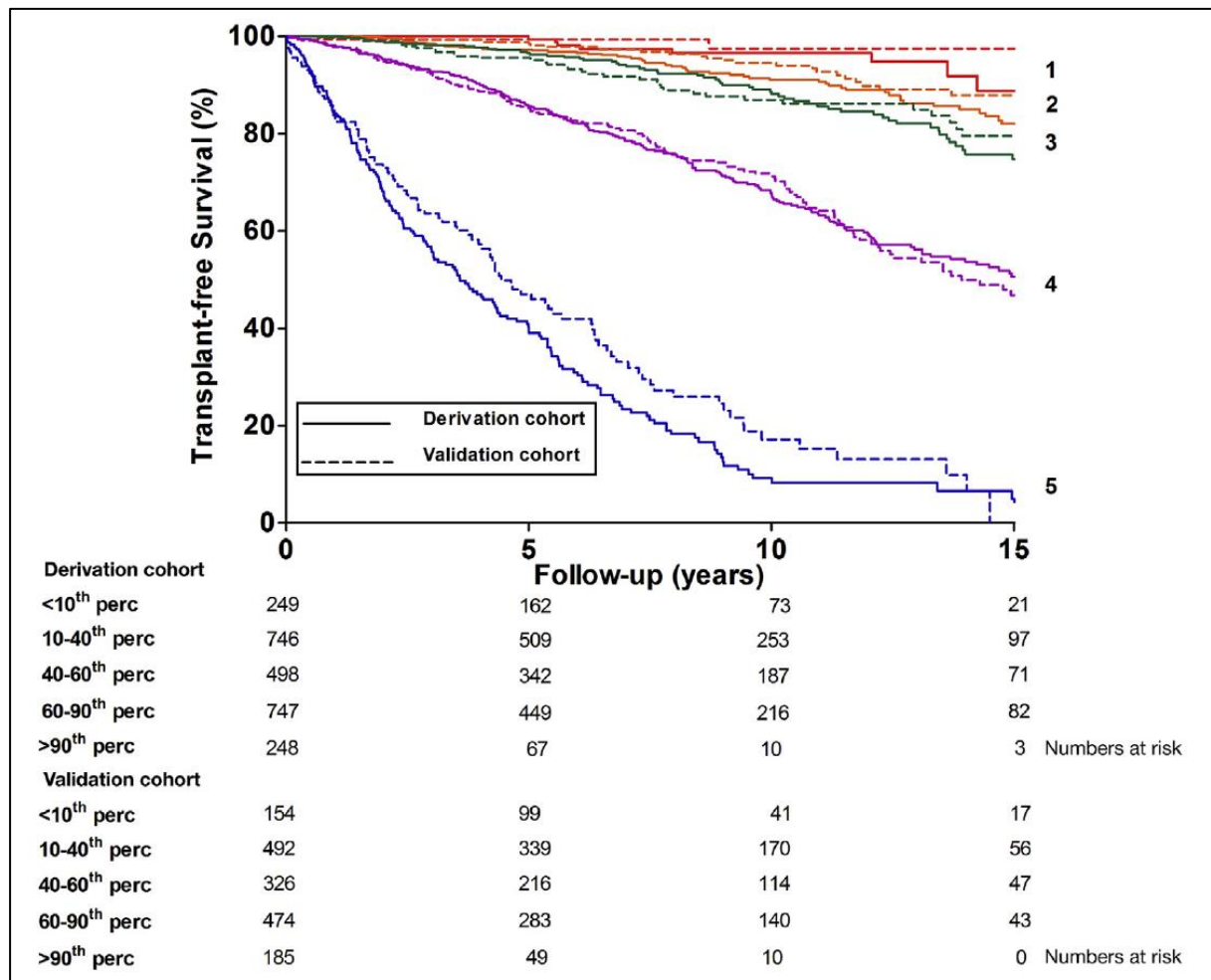


Abbildung 4-51: GLOBE Vorhersagemodell für Lebertransplantat-freies Überleben

Entnommen aus (8). Dargestellt sind 5 Risikogruppen entsprechend der Perzentile des GLOBE Scores: 1: <10te Perzentile, 2: 10te-40te Perzentile, 3: 40te-60te Perzentile, 4: 60te-90te Perzentile, 5: >90te Perzentile. Die gestrichelte Kurve gibt dabei jeweils die Ergebnisse der Validierungskohorte, die durchgezogene Kurve die Derivationskohorte an.

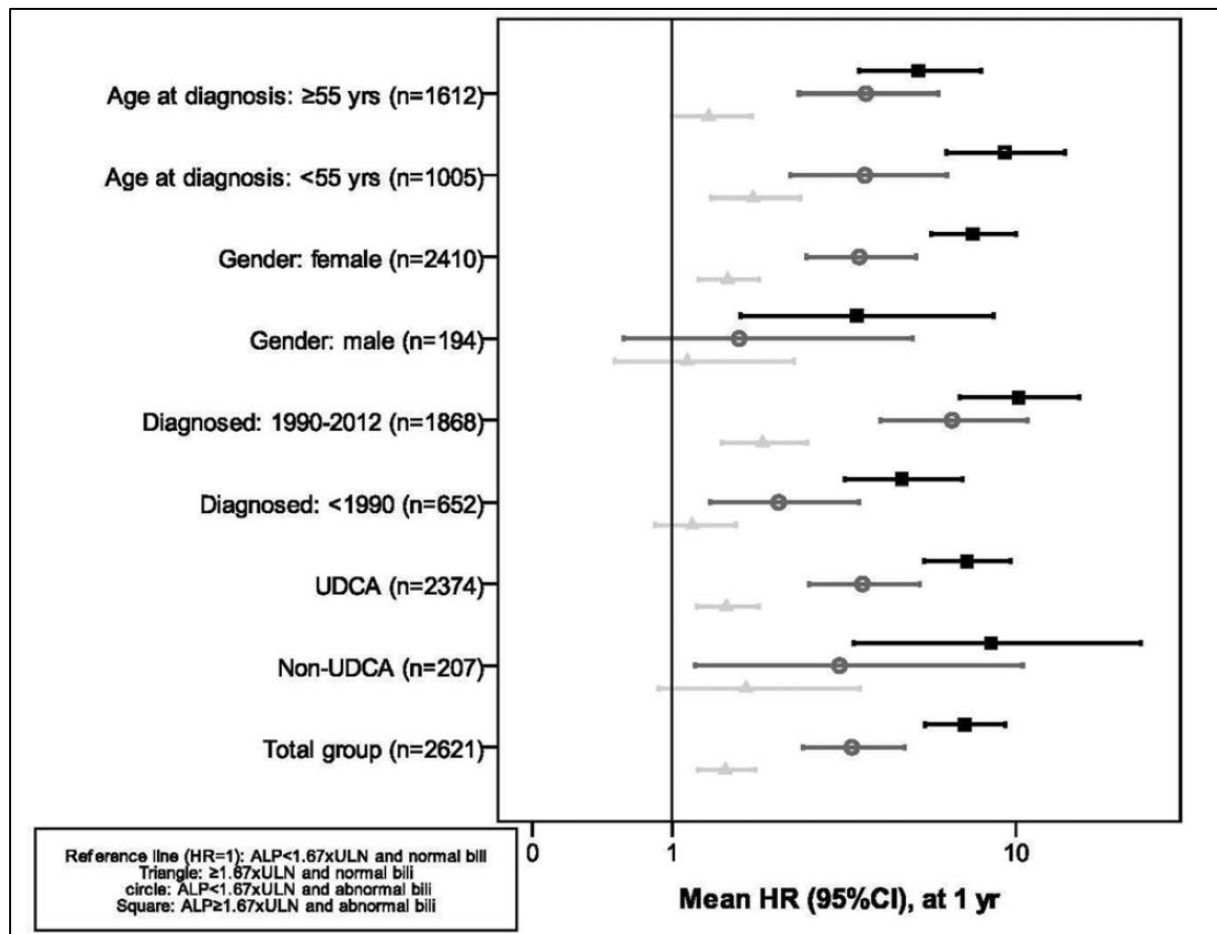


Abbildung 4-52: Lammers et al (2013): Forestplot der Hazard Ratios im Vergleich verschiedener Szenarien für den kombinierten Endpunkt aus AP und Bilirubin - Subgruppenanalysen (9)

Lammers et al (2013) schliessen aus den Ergebnissen, dass diese Analysen, der größten verfügbaren Datenbank, von PBC-Patienten eine Korrelation von AP und Bilirubin mit der Gesamtmortalität bei UDCA behandelten und nicht-behandelten Patienten zeigt und damit eine starke Evidenz ist, dass diese biochemischen Kenngrößen hochgradig valide Surrogatendpunkte für die PBC darstellen (9).

Insgesamt ergibt sich daraus insbesondere wegen der nicht verfügbaren Langzeitdaten aus RCT mit den beiden Registern die best-verfügbare Evidenz für die Validierung der Surrogate AP und Bilirubin für die klinischen Endpunkte Tod und Lebertransplantation. Aus den vorliegenden Ergebnissen lassen sich folgende Schlussfolgerungen ableiten:

Die Surrogate AP und Bilirubin und die aus ihnen zusammengesetzten Endpunkte weisen einen starken Zusammenhang mit den klinischen Endpunkten Tod und Lebertransplantation auf. Die Hazard Ratios für den Vergleich von Patienten mit einer Nachbeobachtung der

biochemischen Kenngrößen nach einem Jahr unter UDCA Therapie und ohne diese Therapie weisen deutlich darauf hin, dass eine Senkung des AP und Bilirubin mit einer Überlebenszeitverlängerung bzw. einer Verlängerung der Zeit bis zur Lebertransplantation oder deren Vermeidung zusammenhängt. Somit wirkt sich die Therapie zur Absenkung dieser biochemischen Kenngrößen vorteilhaft für die Patienten aus. Die Verwendung der biochemischen Kenngrößen AP und Bilirubin als kombinierter Wirksamkeitsendpunkt ist daher als adäquat anzusehen. Dem schliesst sich an, dass auch die Herleitung des Zusatznutzens auf Basis dieser Endpunkte möglich ist und eine Abwägung des Nutzens und möglichen geringeren oder größeren Schadens erlaubt.

4.6 Liste der eingeschlossenen Studien

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

POISE-Studie (Studie 747-301)

Im Dossier verwendete Studienbezeichnung: POISE-Studie

Publikation: Nevens F, Andreaone G, Strasser C, Bowlus C, Invernizzi P, Drenth JPH, et al. A Placebo-controlled trial of obeticholic acid in primary biliary cholangitis. *The New England Journal of Medicine*. 2016;375(7):631-43. (36)

Studienbericht: Intercept Inc. Clinical Study Report. 747-301. The ‘POISE‘ Trial. A Phase 3, double blind, placebo controlled trial and long term safety extension of obeticholic acid in patients with primary biliary cirrhosis. 2015. (35)

Studienbezeichnung laut Studienregistereintrag: Phase 3 Study of Obeticholic Acid in Patients With Primary Biliary Cirrhosis (POISE); Identifier: NCT01473524. (31), (33)

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Amtsblatt der Europäischen Gemeinschaften. VERORDNUNG (EG) Nr. 141/2000 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. 2000.

2. European Medicines Agency. Public summary of opinion on orphan drug designation, EMA/COMP/653816/2016 [online]. 2016. [Aufgerufen am 14.11.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Minutes/2016/11/WC500216203.pdf.
3. Intercept Pharmaceuticals. Fachinformation OCALIVA (OCA) Tabletten [online]. 12.2016. [Aufgerufen am 23.12.2016]. URL: www.fachinfo.de.
4. 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;52.
5. Carbone M, Sharp SJ, Flack S, Paximadas D, Spiess K, Adgey C, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cirrhosis. Hepatology. 2015 Jul 29.
6. Trivedi PJ, Lammers WJ, van Buuren HR, Pares A, Floreani A, Janssen HL, et al. Stratification of hepatocellular carcinoma risk in primary biliary cirrhosis: a multicentre international study. Gut. 2016 Feb;65:321-9.
7. Hansen BE. The Global PBC Study Group EASL, Vienna 2015[online]. 2015. [Aufgerufen am 25.11.2016]. URL: http://www.hivforum.org/storage/documents/2015/Liver_Forum/hansen_presentation_liverforum2.pdf.
8. 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;149:1804-12 e4.
9. Lammers WJ, van Buuren HR, Janssen HL, Invernizzi P, Battezzati PM, Floreani A, et al. Validation of Alkaline Phosphatase and Bilirubin Values as a Surrogate Endpoint in Primary Biliary Cirrhosis - an International, Collaborative Study. Hepatology. 2013;58:250A-1A.
10. Food and Drug Administration - Center for Drug Evaluation Research. Summary Review - OCALIVA (obeticholic acid) [online]. 2016. [Aufgerufen am 25.11.2016]. URL: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/207999Orig1s000SumR.pdf.
11. European Medicines Agency. New medicine for rare, chronic liver disease - Ocaliva recommended for conditional approval, EMA/CHMP/666310/2016 [online]. 2016. [Aufgerufen am 25.10.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2016/10/WC500214166.pdf.

12. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. *Br J Dermatol.* 2010;162:587-93.
13. 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;92:502-7.
14. Guéchet J, Trocmé C, Renversez JC, Sturm N, Zarski JP. ANRS HC EP 23 Fibrostar Study Group. 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;50:693-9.
15. Friedrich-Rust M, Zeuzem S. [Transient elastography (FibroScan) for the non-invasive assessment of liver fibrosis: current status and perspectives]. *Z Gastroenterol.* 2007;45:387-94.
16. Jacoby A, Rannard A, Buck D, M B, Newton JL, James OFW, 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;54:1622-9.
17. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden Version 4.2 vom 22.04.2015 [online]. 2015. [Aufgerufen am 23.12.2016]. URL: https://www.iqwig.de/de/methoden/methodenpapiere/allgemeine_methoden.3020.html.
18. European Medicines Agency. Public summary of opinion on orphan designation (6alpha-ethyl-chenodeoxycholic acid for treatment of primary biliary cirrhosis, EMA/COMP/235947/2010) [online]. 2010. [Aufgerufen am 17.02.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Orphan_designation/2010/08/WC500095691.pdf.
19. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses. In der Fassung vom 18. Dezember 2008, veröffentlicht im Bundesanzeiger Nr. 84a (Beilage) vom 10. Juni 2009, in Kraft getreten am 1. April 2009. Zuletzt geändert am 21. Juli 2016 veröffentlicht im Bundesanzeiger BAnz AT 28.10.2016 B5 in Kraft getreten am 29. Oktober 2016 [online]. 2016. [Aufgerufen am 31.10.2016]. URL: https://www.g-ba.de/downloads/62-492-1282/VerfO_2016-07-21_iK-2016-10-29.pdf.
20. Wong SSL, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE*. *J Med Libr Assoc.* 2006 Jan;94:41-7.
21. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c332.

22. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*. 2010;63:e1-e37.
23. Pares A, Caballeria L, Rodes J. Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic Acid. *Gastroenterology*. 2006 Mar;130(3):715-20.
24. Angulo P, Lindor KD, Therneau TM, Jorgensen RA, Malinchoc M, Kamath PS, et al. Utilization of the Mayo risk score in patients with primary biliary cirrhosis receiving ursodeoxycholic acid. *Liver*. 1999 Apr;19(2):115-21.
25. Momah N, Silveira MG, Jorgensen R, Sinakos E, Lindor KD. Optimizing biochemical markers as endpoints for clinical trials in primary biliary cirrhosis. *Liver Int*. 2012 May;32(5):790-5.
26. Corpechot C, Abenavoli L, Rabahi N, Chretien Y, Andreani T, Johanet C, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. *Hepatology*. 2008 Sep;48(3):871-7.
27. 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.
28. Kuiper EM, Hansen BE, de Vries RA, den Ouden-Muller JW, van Ditzhuijsen TJ, Haagsma EB, et al. Improved prognosis of patients with primary biliary cirrhosis that have a biochemical response to ursodeoxycholic acid. *Gastroenterology*. 2009 Apr;136(4):1281-7.
29. Kumagi T, Guindi M, Fischer SE, Arenovich T, Abdalian R, Coltescu C, et al. Baseline ductopenia and treatment response predict long-term histological progression in primary biliary cirrhosis. *Am J Gastroenterol*. 2010 Oct;105(10):2186-94.
30. Corpechot C, El Naggar A, Poujol-Robert A, Ziol M, Wendum D, Chazouillères O, et al. Assessment of Biliary Fibrosis by Transient Elastography in Patients With PBC and PSC. *Hepatology*. 2006;43:1118-24.
31. clinicaltrials.gov. Phase 3 Study of Obeticholic Acid in Patients With Primary Biliary Cirrhosis (POISE) [online]. 2016. [Aufgerufen am 17.10.2016]. URL: <http://ClinicalTrials.gov/show/NCT01473524>.
32. EUCTR. A Phase 3, Double Blind, Placebo Controlled Trial and Long Term Safety Extension of Obeticholic Acid in Patients with Primary Biliary Cirrhosis [online]. 2011. [Aufgerufen am 17.10.2016]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-004728-36.

33. WHO ICTRP. Phase 3 Study of Obeticholic Acid in Patients With Primary Biliary Cirrhosis POISE [online]. 2016. [Aufgerufen am 17.10.2016]. URL: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01473524>.
34. PharmNet.Bund. A Phase 3, Double Blind, Placebo Controlled Trial and Long Term SafetyExtension of Obeticholic Acid in Patients with Primary Biliary Cirrhosis [online]. 2011. [Aufgerufen am 17.10.2016]. URL: <http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm> (Direktlink nicht möglich).
35. Intercept Pharmaceuticals. A Phase 3, Double Blind, Placebo Controlled Trial and Long Term Safety Extension of Obeticholic Acid in Patients with Primary Biliary Cirrhosis. PBC OCA International Study of Efficacy (POISE), Clinical Study Report Protocol 747-3012015.
36. Nevens F, Andreone P, Mazzella G, Strasser SI, Bowlus C, Invernizzi P, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. The New England journal of medicine. 2016;375:631-43.
37. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. J Hepatol. 2009 Aug;51:237-67.
38. Lindor KD, Gershwin ME, Poupon R, Kaplan M, Bergasa NV, Heathcote EJ, et al. Primary biliary cirrhosis. Hepatology. 2009 Jul;50:291-308.
39. Shapiro JM, Smith H, Schaffner F. Serum bilirubin: a prognostic factor in primary biliary cirrhosis. Gut. 1979;20:137-40.
40. 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;147:1338-49.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁷ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-124: Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in Cochrane Register of Controlled Trials

Datenbankname		EBM Reviews - Cochrane Register of Controlled Trials (Clinical Trials)
Suchoberfläche		Ovid
Datum der Suche		31.10.2016
Zeitsegment		Keine Einschränkung
Suchfilter		Filter für randomisierte kontrollierte Studien nach Wong 2006 Strategy minimizing difference between sensitivity and specificity (Wong 2006)
#	Suchbegriffe	Ergebnis
1	SBAS CCTR93; CDSR93	972642
Indikation		
2	CT=PRIMARY BILIARY CIRRHOSIS	302
3	PRIMARY BILIARY CIRRHOSIS	528
4	PBC	300
5	PRIMARY BILIARY CHOLANGITIS	3
6	PRIMAERE# BILIAERE# ZIRRHOSE	2
7	PRIMAERE# BILIAERE# LEBERZIRRHOSE	0
8	CHRONISCH## NICHT EITRIGE# DESTRUIERENDE# CHOLANGITIS	0
9	CHRONIC, NONPURULENT DESTRUCTIVE CHOLANGITIS	0
10	or/2-9	625
Therapie / Intervention		
11	CT=OBETICHOLIC ACID	30
12	OBETICHOLIC ACID	38
13	OCA	84
14	or/11-13	92
Zusammenfassung		
15	and/10, 14	28
16	RANDOM?/(TI;AB) OR CLINICAL TRIAL?/(TI;AB;TE;CT) OR CT D HEALTH CARE QUALITY	665249
17	RANDOM?/(TI;AB) OR CLINICAL TRIAL?/(TI;AB;TE;CT) OR CT D TREATMENT OUTCOME	587065
18	DOUBLE BLIND?/(TI;AB;TE;CT) OR PLACEBO?/(TI;AB) OR BLIND?/(TI;AB)	271423
19	RANDOM?/(TI;AB) OR PLACEBO?/(TI;AB;TE;CT) OR DOUBLE-BLIND?/(TI;AB)	574474

20	DT="RANDOMIZED CONTROLLED TRIAL" OR RANDOMIZED?/(TI;AB;CT;TE) OR CT=RANDOMIZED? OR PLACEBO?/(TI;AB;CT;TE) OR CT=PLACEBO? OR DOUBLE- BLIND?/(TI;AB;CT;TE)	648101
Zusammenfassung		
21	and/15, 16	22
22	and/15, 17	22
23	and/15, 18	25
24	and/15, 19	25
25	and/15, 20	25
26	or/21, 22, 23, 24, 25	25
27	26	25

Tabelle 4-125: Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in MEDLINE

Datenbankname		MEDLINE (Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R))
Suchoberfläche		Ovid
Datum der Suche		31.10.2016
Zeitsegment		Keine Einschränkung
Suchfilter		Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity (Wong 2006)
#	Suchbegriffe	Ergebnis
1	SBAS ME60	26547461
Indikation		
2	CT=PRIMARY BILIARY CIRRHOSIS	7480
3	PRIMARY BILIARY CIRRHOSIS	7361
4	PBC	4246
5	PRIMARY BILIARY CHOLANGITIS	83
6	PRIMAERE# BILIAERE# ZIRRHOSE	29
7	PRIMAERE# BILIAERE# LEBERZIRRHOSE	5
8	CHRONISCH## NICHT EITRIGE# DESTRUIERENDE# CHOLANGITIS	1
9	CHRONIC, NONPURULENT DESTRUCTIVE CHOLANGITIS	1
10	or/2-9	10780
Therapie / Intervention		
11	CT=OBETICHOLIC ACID	0
12	OBETICHOLIC ACID	92
13	OCA	802
14	or/11-13	873
Zusammenfassung		
15	and/10, 14	31
16	RANDOM?/(TI;AB) OR CLINICAL TRIAL?/(TI;AB;TE;CT) OR CT D HEALTH CARE QUALITY	6067906
17	RANDOM?/(TI;AB) OR CLINICAL TRIAL?/(TI;AB;TE;CT) OR CT D TREATMENT OUTCOME	1791690
18	DOUBLE BLIND?/(TI;AB;TE;CT) OR PLACEBO?/(TI;AB) OR BLIND?/(TI;AB)	357442
19	RANDOM?/(TI;AB) OR PLACEBO?/(TI;AB;TE;CT) OR DOUBLE-BLIND?/(TI;AB)	961982
20	DT="RANDOMIZED CONTROLLED TRIAL" OR RANDOMIZED?/(TI;AB;CT;TE) OR CT=RANDOMIZED? OR	1096790

	PLACEBO?/(TI;AB;CT;TE) OR CT=PLACEBO? OR DOUBLE-BLIND?/(TI;AB;CT;TE)	
Zusammenfassung		
21	and/15, 16	14
22	and/15, 17	14
23	and/15, 18	4
24	and/15, 19	4
25	and/15, 20	5
26	or/21, 22, 23, 24, 25	14
27	26	14

Tabelle 4-126: Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in EMBASE

Datenbankname		EMBASE
Suchoberfläche		Ovid
Datum der Suche		31.10.2016
Zeitsegment		Keine Einschränkung
Suchfilter		Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity (Wong 2006)
#	Suchbegriffe	Ergebnis
1	SBAS EM47	33570687
Indikation		
2	CT=PRIMARY BILIARY CIRRHOSIS	8882
3	PRIMARY BILIARY CIRRHOSIS	12144
4	PBC	6377
5	PRIMARY BILIARY CHOLANGITIS	124
6	PRIMAERE# BILIAERE# ZIRRHOSE	2
7	PRIMAERE# BILIAERE# LEBERZIRRHOSE	1
8	CHRONISCH## NICHT EITRIGE# DESTRUIERENDE# CHOLANGITIS	1
9	CHRONIC, NONPURULENT DESTRUCTIVE CHOLANGITIS	2
10	or/2-9	14206
Therapie / Intervention		
11	CT=OBETICHOLIC ACID	441
12	OBETICHOLIC ACID	437
13	OCA	1225
14	or/11-13	1582
Zusammenfassung		
15	and/10, 14	162
16	RANDOM?/(TI;AB) OR CLINICAL TRIAL?/(TI;AB;TE;CT) OR CT D HEALTH CARE QUALITY	4066269
17	RANDOM?/(TI;AB) OR CLINICAL TRIAL?/(TI;AB;TE;CT) OR CT D TREATMENT OUTCOME	3003755
18	DOUBLE BLIND?/(TI;AB;TE;CT) OR PLACEBO?/(TI;AB) OR BLIND?/(TI;AB)	494675
19	RANDOM?/(TI;AB) OR PLACEBO?/(TI;AB;TE;CT) OR DOUBLE-BLIND?/(TI;AB)	1371861
20	DT="RANDOMIZED CONTROLLED TRIAL" OR RANDOMIZED?/(TI;AB;CT;TE) OR CT=RANDOMIZED? OR PLACEBO?/(TI;AB;CT;TE) OR CT=PLACEBO? OR DOUBLE-BLIND?/(TI;AB;CT;TE)	1086884
Zusammenfassung		
21	and/15, 16	126

22	and/15, 17	124
23	and/15, 18	50
24	and/15, 19	82
25	and/15, 20	102
26	or/21, 22, 23, 24, 25	133
27	26	133
Löschen der Dokumente, die in Medline vorhanden sind		
28	27 NOT SU=MEDLINE	128

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht relevant.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht relevant.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

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

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

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

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	17.10.2016
Suchstrategie	("INT-747" OR "Obeticholic acid" OR "OCA") AND ("PBC" OR "Primary biliary cirrhosis" OR "Primary biliary cholangitis")
Treffer	5

Studienregister	EU-RCT
Internetadresse	http://www.clinicaltrialsregister.eu
Datum der Suche	17.10.2016
Suchstrategie	("INT-747" OR "Obeticholic acid" OR "OCA") AND ("PBC" OR "Primary biliary cirrhosis" OR "Primary biliary cholangitis")
Treffer	4

Studienregister	International Clinical Trials Registry Platform (ICTRP)
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	17.10.2016
Suchstrategie	("INT-747" OR "Obeticholic acid" OR "OCA") AND ("PBC" OR "Primary biliary cirrhosis" OR "Primary biliary cholangitis")
Treffer	6

Studienregister	PharmNet.Bund
Internetadresse	http://www.pharmnet-bund.de
Datum der Suche	17.10.2016
Suchstrategie	("INT-747" OR "Obeticholic acid" OR "OCA") AND ("PBC" OR "Primary biliary cirrhosis" OR "Primary biliary cholangitis")
Treffer	4

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht relevant.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht relevant.

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

Bei der bibliographischen Literaturrecherche wurde lediglich eine Publikation im Volltext gesichtet und im Anschluss eingeschlossen (36). Alle übrigen Literaturstellen konnten bereits anhand von Titel und Zusammenfassung ausgeschlossen werden.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht relevant.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht relevant.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

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

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-127 (Anhang): Aus der Bewertung ausgeschlossene RCT aus der Suche in Studienregistern mit Ausschlussgrund

Nr.	Studien-nummer	Identifikationsort	Titel (Offizieller Titel)	Ausschlussgrund	URL
1	NCT00550862	Clinicaltrials.gov	Study of INT 747 in Combination With URSO in Patients With Primary Biliary Cirrhosis (PBC)	A5	https://clinicaltrials.gov/ct2/show/NCT00550862?term=NCT00550862&rank=1 abgerufen am 17.10.2016
2	NCT00570765	Clinicaltrials.gov	Study of INT-747 as Monotherapy in Patients With PBC	A5	https://clinicaltrials.gov/ct2/show/NCT00570765?term=NCT00570765&rank=1 abgerufen am 17.10.2016
3	NCT01865812	Clinicaltrials.gov	Phase 2 Study on Effects of Obeticholic Acid (OCA) on Lipoprotein Metabolism in Subjects With Primary Biliary Cirrhosis	A5	https://clinicaltrials.gov/ct2/show/NCT01865812?term=NCT01865812&rank=1 abgerufen am 17.10.2016
4	NCT02308111	Clinicaltrials.gov	Phase 3 Study of Obeticholic Acid Evaluating Clinical Outcomes in Patients With Primary Biliary Cirrhosis (COBALT)	A13	https://clinicaltrials.gov/ct2/show/NCT02308111?term=NCT02308111&rank=1 abgerufen am 17.10.2016
5	EUCTR2007-001424-12-GB	ICTRP	A Study of INT 747 Monotherapy in Patients with Primary Biliary Cirrhosis	A5	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2007-

					<u>001424-12-GB</u> abgerufen am 17.10.2016
6	EUCTR2007-001425-10-GB	ICTRP	A Study of INT-747 (6-ECDC) in Combination with Ursodeoxycholic Acid (URSO®, UDCA) in Patients with Primary Biliary Cirrhosis	A5	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2007-001425-10-GB abgerufen am 17.10.2016
7	NCT01865812	ICTRP	Phase 2 Study on Effects of Obeticholic Acid (OCA) on Lipoprotein Metabolism in Subjects With Primary Biliary Cirrhosis	A5	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01865812 abgerufen am 17.10.2016
8	NCT02308111	ICTRP	A Phase 3b, Double Blind, Randomized, Placebo Controlled, Multicenter Study Evaluating the Effect of Obeticholic Acid on Clinical Outcomes in Subjects with Primary Biliary Cirrhosis	A13	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02308111 abgerufen am 17.10.2016
9	EUCTR2007-001424-12	EUCTR	A Study of INT 747 Monotherapy in Patients with Primary Biliary Cirrhosis	A5	https://www.clinicaltrialsregister.eu abgerufen am 17.10.2016
10	EUCTR2007-001425-10	EUCTR	A Study of INT-747 (6-ECDC) in Combination with Ursodeoxycholic Acid (URSO®, UDCA) in Patients with Primary Biliary Cirrhosis	A5	https://www.clinicaltrialsregister.eu abgerufen am 17.10.2016
11	EUCTR2014-005012-42	EUCTR	A Phase 3b, Double Blind, Randomized, Placebo Controlled, Multicenter Study Evaluating the Effect of Obeticholic Acid on Clinical Outcomes in Subjects with Primary Biliary Cirrhosis	A13	https://www.clinicaltrialsregister.eu abgerufen am 17.10.2016
12	EUCTR2007-001424-12	Pharmnet	A Study of INT 747 Monotherapy in Patients with Primary Biliary Cirrhosis	A5	www.pharmnet-bund.de abgerufen am 17.10.2016
13	EUCTR2007-001425-10	Pharmnet	A Study of INT-747 (6-ECDC) in Combination with Ursodeoxycholic Acid (URSO®, UDCA) in Patients with Primary Biliary Cirrhosis	A5	www.pharmnet-bund.de abgerufen am 17.10.2016
14	EUCTR2014-005012-42	Pharmnet	A Phase 3b, Double Blind, Randomized, Placebo Controlled, Multicenter Study Evaluating the Effect of Obeticholic Acid on Clinical Outcomes in Subjects with Primary Biliary Cirrhosis	A13	www.pharmnet-bund.de abgerufen am 17.10.2016

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht relevant.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht relevant.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-128 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-128 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-128 (Anhang): Studiendesign und -methodik für POISE-Studie (747-301)

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primärziele der Studie waren die Bewertung der Effektivität von Obeticholsäure bei Probanden mit PBC in Bezug auf alkalische Phosphatase (AP) und Gesamt-Bilirubin im Serum als kombinierter Endpunkt und auf Sicherheit. Sekundäre Ziele waren die Bewertung der Effekte von Obeticholsäure bei Probanden mit PBC auf die folgenden Endpunkte: • hepatozelluläre Schäden und Leberfunktion, inklusive Histologie (entzündliche, strukturelle [portal, parenchymal] und fibrotische Bewertung) • Krankheits-spezifische Symptome • Biomarker und nicht-invasive Bewertung von Leberfibrose • Gallensäuren • Andere exploratorische Auswertungen. Vergleichstherapie war die Behandlung mit UDCA+Placebo bei Patienten mit Kombinationstherapie und Placebo bei Patienten mit Monotherapie.
Methoden		
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	• Prospektiv randomisierte, Placebo-kontrollierte Phase-III-Studie • Multinational, multizentrisch • Doppelblind • Parallel mit Zuteilungsverhältnis 1:1:1 Behandlungsarme: (a) Kontrollgruppe, (b) 10 mg OCA, (c) 5 mg titriert auf 10 mg OCA.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Informationen aus Amendments und Addenden: <u>Änderung 07. Feb. 2012:</u> Das Protokoll wurde landesspezifisch für Österreich geändert, um eine Anforderung der österreichischen Gesundheitsbehörde aufzunehmen: • Schwangerschaftstests werden bei weiblichen Patienten im gebärfähigen Alter nicht wie geplant nur bei allen geplanten klinischen Visits, sondern zusätzlich mit heimischen Schwangerschaftstests durchgeführt.

Item ^a	Charakteristikum	Studieninformation
		<u>Änderung 04. April 2012:</u> Präzisierung einiger Passagen als Reaktion von Kommentaren der Prüfarzte: • Ausschlusskriterium #2: Portale Hypertonie mit Komplikationen. „Zugehörige Interventionen“ erweitert auf „zugehörige therapeutische oder prophylaktische Interventionen“; „Beta-Blocker“ in der Liste der Beispiele ergänzt. • Ausschlusskriterium #3: Präzisierung des Zeitraums, der zur Pruritus-Vorgeschichte gezählt wird auf 2 Monate vor Tag 0. • Ausschlusskriterium #6: Patienten mit kardialen Arrhythmien werden ausgeschlossen, wenn diese „wahrscheinlich Auswirkungen auf das Überleben während der klinischen Prüfung“ haben. • Kontrollmessungen für AP-Werte im Bereich von $1.67 \times \text{ULN} \leq \text{AP} < 2 \times \text{ULN}$ bei der Screening-Visit wurde auf alle Patienten mit $\text{AP} < 2 \times \text{ULN}$ ausgedehnt. • Unter „weitere Gründe für Studien- oder Therapie-Abbruch“ wurde der Zeitraum für Behandlungsunterbrechungen von bisher >14 Tage auf >28 Tage erhöht, so dass nun ein Studienabbruch erst ab 28 Tagen empfohlen wurde. • Die Anweisung, dass Patienten vor dem Screening-Beobachtungszeitpunkt über Nacht fasten sollten (mindestens 8 Stunden) wurde entfernt. • Die Erhebung einer transienten Elastographie nach 6 Monaten entfiel. • Erweiterung der jährlichen Nachbeobachtungszeitpunkte im Rahmen der Langzeit-Sicherheits-Extension um eine Messung mit Fibroscan® im Rahmen der TE. • Die Erhebung von 6 Ethyl UDCA (6E-UDCA) wurde als Sekundärendpunkt ergänzt. • Definition von „unwahrscheinlich“ als Ausprägung des Zusammenhangs von unerwünschten Ereignissen mit der Studienmedikation wurde korrigiert von: „Eine Reaktion, die in angemessener zeitlicher Nähe auf die Administration der Studienmedikation folgt; die einem bekannten oder vermuteten Reaktionsmuster für die vermutete

Item ^a	Charakteristikum	Studieninformation
		Studienmedikation folgt; aber die angemessen mit bekannten Charakteristika des klinischen Zustands des Probanden erklärt werden könnte.“ zu „Eine Reaktion, die nicht in angemessener zeitlicher Nähe auf die Administration der Studienmedikation folgt; die einem bekannten oder nicht vermuteten Reaktionsmuster für die vermutete Studienmedikation folgt; und die angemessen mit bekannten Charakteristika des klinischen Zustands des Probanden erklärt werden könnte. • Zur Liste der geplanten Laborwerte wurden hinzugefügt: Serum-Chemie: Natrium, Kalium, Kalzium, Chlorid, Magnesium, Phosphate, Bicarbonat. Gallensäuren: konjugiertes UDCA ohne 6E-UDCA und C4. <u>Änderung 24. Sept. 2012:</u> • Ausschlußkriterium #1: Patienten mit Hepatitis B und Serokonversion (Hbs Ag und Hbe Ag negativ) können nach Rücksprache mit dem Medical Monitor in die Studie aufgenommen werden. • Patienten, bei denen innerhalb von 6 Monaten vor Tag 0 bereits ein DEXA-Scan durchgeführt wurde, müssen diesen nicht für die Baseline-Messung wiederholen. • Die Studienmedikation darf den Patienten auch per Kurier zugestellt werden, wenn angemessene Vorsichtsmaßnahmen getroffen werden und der Versand dokumentiert wird. • Die in der vorangegangenen Änderung fälschlicherweise entfernte Anweisung, dass Patienten vor der Screening-Beobachtungszeitpunkt über Nacht fasten sollten (mindestens 8 Stunden) wurde wieder eingefügt. • Die Anweisung, dass Patienten, bei denen ein DEXA-Scan durchgeführt werden soll, 48 Stunden vor dem Scan keine Calcium-Supplemente einnehmen sollten, wurde entfernt. • Der fälschlicherweise in der vorangegangenen Änderung entfernte

Item ^a	Charakteristikum	Studieninformation
		Laborwert C4 wurde der Liste der Gallensäure-Analyse wieder hinzugefügt.
4	Probanden / Patienten	Erwachsene Patienten mit primärer biliärer Cholangitis (PBC)
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Aus dem Studienprotokoll: <u>Einschlußkriterien:</u> Patienten müssen alle folgenden Kriterium erfüllen: - Definitive oder wahrscheinliche PBC Diagnose (konsistent mit AASLD und EASL Leitlinien; [(38),(37)](belegt durch Zutreffen von ≥ 2 der folgenden 3 diagnostischen Faktoren): - Vorgeschichte von erhöhten AP Spiegeln seit mindestens 6 Monaten - Positive AMA Titer oder PBC-spezifische Antikörper - Leber-Biopsie konsistent mit PBC - Mindestens eins der folgenden qualifizierenden biochemischen Werte: $AP \geq 1.67 \times ULN$ - Gesamt-Bilirubin $> ULN$ aber $< 2 \times ULN$ - Alter ≥ 18 Jahre - Einnahme von UDCA seit mindestens 12 Monaten (stabile Dosis seit ≥ 3 Monaten) vor Tag 0, oder UDCA-Unverträglichkeit (keine UDCA seit ≥ 3 Monaten) vor Tag 0. - Kontrazeption: Weibliche Patienten im gebärfähigen Alter müssen während der Studie und für 30 Tage nach der Visite zum Behandlungsende mindestens eine Kontrazeptionsmethode nutzen ($\leq 1\%$ Ausfallrate). - Vorliegen der schriftlichen Einverständniserklärung und Einverständnis, sich nach dem Studienprotokoll zu richten. <u>Ausschlußkriterien:</u> Patienten werden von der Studie ausgeschlossen, wenn sie mindestens eins der folgenden Kriterien erfüllen: - Vorgeschichte oder Anwesenheit anderer gleichzeitig bestehender Lebererkrankungen, einschließlich: - Infektion mit dem Hepatitis C Virus (HCV); Patienten mit aktiver

Item ^a	Charakteristikum	Studieninformation
		Hepatitis B (HBV) werden ausgeschlossen, allerdings können Patienten mit Hepatitis B und Serokonversion (Hbs Ag und Hbe Ag negativ) nach Rücksprache mit dem Medical Monitor in die Studie aufgenommen werden. ○ Primäre sklerosierende Cholangitis (PSC) ○ Alkoholische Leberkrankheit ○ Definitive Autoimmunkrankheit der Leber oder overlap Hepatitis ○ Nicht-alkoholische Steatohepatitis (NASH) ○ Gilberts Syndrom (Ausschluß durch Interpretierbarkeit von Bilirubin-Spiegeln) 2. Vorliegen von klinischen Komplikationen von PBC oder klinische signifikanter Leberdekompensation, einschließlich: ○ Vorgeschichte einer Lebertransplantation, gegenwärtige Platzierung auf einer Lebertransplantationsliste oder gegenwärtiger MELD Score ≥ 15 ○ Portale Hypertonie mit Komplikationen, einschließlich: bekannte gastrische oder ösophageale Varizen, schlecht kontrollierte oder Diuretika-resistente Aszites, Vorgeschichte von Blutungen der Varizen oder zugehörigen therapeutischen oder prophylaktischen Interventionen (z.B. Beta-Blocker, Einsetzen von Varizen-Bändern oder transjugulären intrahepatischen portosystemischen Shunts [TIPS]) oder Leberenzephalopathie ○ Zirrhose mit Komplikationen, einschließlich Vorgeschichte oder Vorliegen von: spontaner bakterieller Peritonitis, hepatozellulärem Karzinom, Bilirubin $> 2 \times \text{ULN}$ ○ Hepatorenales Syndrom (Typ I oder II) oder Serum Creatinin bei Screening $> 2 \text{ mg/dL}$ ($178 \text{ } \mu\text{mol/L}$) 3. Patienten mit schwerem Pruritus oder solche, die systemische Behandlung für Pruritus (z.B. mit Gallensäure-Komplexbildnern [BAS] oder

Item ^a	Charakteristikum	Studieninformation
		Rifampicin) innerhalb von 2 Monaten von Tag 0 benötigten, werden ausgeschlossen. 4. Die Gabe der folgenden Arzneimittel ist unzulässig: ○ Unzulässig 6 Monate vor Tag 0 and durchgehend während der Studie (d.h. bis zur letzten Dosis und/oder Behandlungsende): Azathioprine, Colchicine, Cyclosporine, Methotrexat, Mycophenolate Mofetil, Pentoxifylline; Fenofibrate oder andere Fibrate; Budesonide und andere systemische Corticosteroide; potentiell hepatotoxische Arzneimittel (einschließlich a-Methyl-Dopa, Sodium Valproic Acid, Isoniazide oder Nitrofurantoin) ○ Unzulässig 12 Monate vor Tag 0 und durchgehend während der Studie (d.h. bis zur letzten Dosis und/oder Behandlungsende): Antikörper oder Immuntherapien, die sich gegen Interleukine oder andere Zytokine oder Chemokine richten. 5. Patienten, die schon vorher an einer klinischen Studie von Obeticholsäure teilgenommen haben, wird die Teilnahme nicht gestattet. 6. Vorgeschichte oder Vorliegen einer klinisch relevanten Herzrhythmusstörung, die sich wahrscheinlich auf das Überleben während der Studie auswirken wird, oder Verlängerung des QT -oder QTc-Intervals von >500 Millisekunden (msec) bei Screening (vor Behandlung). 7. Weibliche Patienten: Bekannte Schwangerschaft oder positiver Urin-Schwangerschaftstest (bestätigt durch einen positiven Serum Schwangerschaftstest) oder in der Stillzeit.
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrisch: Länder: Australien, Belgien, Deutschland, Frankreich, Großbritannien, Italien, Kanada, Niederlande, Österreich, Polen, Schweden, Spanien, USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	In der doppelblinden Phase der Studie (12 Monate) werden drei Behandlungsgruppen betrachtet:

Item ^a	Charakteristikum	Studieninformation
		○ Kontrolle • UDCA+Placebo bei Patienten mit Kombinationstherapie • Placebo bei Patienten mit Monotherapie als Proxi für ein beobachtendes Abwarten ○ Obeticholsäure 10 mg • OCA+UDCA bei Patienten mit Kombinationstherapie • OCA bei Patienten mit Monotherapie ○ Obeticholsäure Titration • OCA+UDCA bei Patienten mit Kombinationstherapie • OCA bei Patienten mit Monotherapie 5 mg titriert bis 10 mg Obeticholsäure. Patienten in diesem Behandlungsarm wurden bei der Monat 6 Visite von 5 mg zu 10 mg titriert, wenn dies angebracht war. Um die Verblindung zu erhalten wurden alle Patienten (d.h., alle Behandlungsgruppen) zu Monat 6 basierend auf AP und/oder Bilirubin-Ansprechen und Verträglichkeit der Studienmedikation für die Titration bewertet. Allerdings konnten nur Patienten im Behandlungsarm mit 5 mg, titriert bis 10 mg Obeticholsäure, tatsächlich eine Steigerung ihrer Dosis erhalten, was durch das interaktive Web-basierte Dialogsystem (IWRS) kontrolliert wurde. Eine Anfrage nach „Hoch-Titration“ konnten die Prüfarzte über das IWRS für Patienten stellen, die mindestens eins der folgenden Kriterien bei der Monat-6-Bewertung erfüllten: • $AP \geq 1.67 \times ULN$ und/oder • Gesamt-Bilirubin > ULN oder • <15% AP Reduktion bei Monat 6 im Vergleich zu den mittleren Vorbehandlungswerten in der doppelblinden Phase und • vorausgesetzt, dass unerwünschte Ereignisse (z.B. schwerer Pruritus) die Gabe einer höheren Dosis nicht beschränken. ○ In allen Gruppen begann die Dosierung am Tag nach Tag 0. Es wurde 1 Tablette (oder 2 Tabletten während der Langzeit-

Item ^a	Charakteristikum	Studieninformation
		Sicherheits-Extension) weisungsgemäß täglich mit Wasser eingenommen, etwa 30 Minuten vor dem Frühstück. Die Patienten wurden angewiesen, die Tabelle als Ganzes zu schlucken, sie durften die Tabelle nicht kauen, teilen oder zerstoßen.
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	<u>Primärer Wirksamkeitsendpunkt:</u> Kombinierter Endpunkt: Anteil von Patienten mit AP<1,67xULN und Gesamt-Bilirubin≤ULN und AP-Reduktion ≥15% im Vergleich zu Baseline bei Monat 12. <u>Sekundäre Wirksamkeitsendpunkte:</u> ○ Anteil von Patienten mit AP< 1,67xULN und Gesamt-Bilirubin≤ULN und AP-Reduktion ≥15% im Vergleich zu Baseline bei Monat 6. ○ Absolute und prozentuale Veränderung zu Baseline in AP bei Monat 6. ○ Absolute und prozentuale Veränderung zu Baseline in AP bei Monat 12. ○ Hepatozelluläre Verletzung und Leberfunktion: • Anteil von Patienten mit AP-Reduktion von ≥10%, ≥15%, ≥20% und ≥40% im Vergleich zu Baseline. • Anteil von Patienten mit AP≤ULN • Prognostisches Krankheitsrisiko: ○ AP≤3xULN und AST≤2xULN und Bilirubin≤ULN (Paris I) ○ AP≤1,5xULN und AST≤1,5xULN und Bilirubin≤ULN (Paris II) ○ AP≤1,67xULN und Bilirubin≤ULN (Mayo II) ○ Bilirubin≤ULN und Albumin≥LLN ○ AP≤1,76xULN (Toronto II) ○ Bilirubin≤ULN und /oder Albumin≥LLN (Rotterdam) ○ Klinisches Labor: AP, GGT, ALT, AST, Gesamt- und konjugiertes Bilirubin, Albumin, PT, INR; jeweils absolute und prozentuale Veränderung zu Baseline. ○ Biomarker und nicht-invasive Bewertung von Leberfibrose: • ELF-Score und seine Komponenten

Item ^a	Charakteristikum	Studieninformation
		(HA, P3NP, TIMP-1), absolute Veränderung zu Baseline. • Leber-Steifigkeits-Messungen, absolute Veränderung zu Baseline (nur in ausgewählten Prüfzentren) • Weitere Marker: CRP, TNF-α, TGF-β, FGF-19, IL-6, CK-18, jeweils absolute Veränderung zu Baseline. ○ Gallensäuren: • Gesamt-Gallensäuren, Gesamt endogene Gallensäuren und Summen für die individuellen Gallensäuren: UDCA, CDCA, DCA, CA und LCA. • Anteil jeder der individuellen Gallensäuren an den Gesamt-Gallensäuren. <u>Weitere patientenrelevante Endpunkte:</u> ○ Krankheitsspezifische Symptome: • <u>PBC-40 (Lebensqualitäts-Messung für PBC): Domänen-Scores (Symptome, Jucken, Fatigue, kognitive Funktion, soziale Domäne, emotionale Domäne), absolute Veränderung zu Baseline.</u> • <u>Patientenbefragungsbogen</u> <u>Exploratorische Endpunkte:</u> ○ Eine genetische Studie für Single-nucleotid Polymorphismen (SNPs), die an PBC beteiligt sein könnten, wurde in denjenigen Prüfzentren durchgeführt, die bereit waren, Proben zur Verfügung zu stellen. ○ Zusätzliche Analyse wie PBC-Autoimmunantikörper (IgA, IgG, IgM) und andere Zytokine und Interleukine (IL-12 (p40), IL-23) wurden im Rahmen explorativer Forschung geprüft. <u>Sicherheitsendpunkte:</u> ○ Unerwünschte Ereignisse ○ Schwerwiegende unerwünschte Ereignisse ○ Schwere unerwünschte Ereignisse ○ Zugehörige unerwünschte Ereignisse ○ Unerwünschte Ereignisse, die zum Therapieabbruch führten ○ Gesamtmortalität ○ Pruritus

Item ^a	Charakteristikum	Studieninformation
		○ Schwerer Pruritus ○ Unerwünschtes Ereignis Pruritus, das zum Therapieabbruch führte ○ Intervention wegen Pruritus ○ Zeit bis zum ersten Auftreten von Pruritus ○ Zeit bis zum Auftreten des ersten moderaten oder schweren Pruritus ○ Zeit bis zum Auftreten des ersten schweren Pruritus ○ 5-D Pruritus-Fragebogen ○ Pruritus-VAS ○ Klinisches Labor: Hematologie, Serumchemie, Urinchemie, Lipoprotein-Analyse, Urinanalyse ○ Mayo-Risiko-Score und MELD- (Model for End Stage Liver Disease, “Modell für Leberkrankheit im Endstadium”) Score ○ DEXA-Scan zur Messung der Knochendichte ○ Vitalparameter ○ Körpergewicht und BMI ○ Ärztliche Untersuchung ○ Elektrokardiogramm <u>Bewertung:</u> ○ Histologische Untersuchung: Jeweils im Studienzentrum, wenn dort histologische Bewertungen durchgeführt werden. ○ Fibroscan® TE Messungen wurden nur in solchen Studienzentren durchgeführt, in denen Mitarbeiter für die Nutzung der Geräte und die Interpretation der Daten geschult sind. ○ Elektrokardiogramme wurden vom Prüfarzt oder seinen Beauftragten durchgeführt und darüber hinaus zentral von einem spezialisierten Kardiologen oder einem zentralen EKG-Labor gelesen.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Trifft nicht zu.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlen wurden auf Basis der Ergebnisse der Phase-II geplant. Dort zeigten 9% der Probanden, die mit Placebo behandelt wurden, und 40% der Probanden, die mit Obeticholsäure 10 mg behandelt wurden, ein

Item ^a	Charakteristikum	Studieninformation
		Ansprechen im Sinne des primären Endpunkts. Der Stichprobenumfang wurde dann mit etwas konservativeren Annahmen berechnet: Unter der Annahme von 14% (Placebo) und 40% (Obeticholsäure 10 mg) ansprechenden Patienten wurde basierend auf einem 2-seitigen Bionomialtest auf gleiche Anteile bei einem Signifikanzniveau von 5% errechnet, dass 60 Patienten pro Gruppe 90% Power zur Detektion eines Unterschieds zwischen den Gruppen gewährleisten würden. Wegen eines Screening-Zeitfensters von bis zu 8 Wochen hatten einige Probanden bereits erfolgreich die Screening-Prozedur durchlaufen, alle Einschlusskriterien erfüllt und waren für die Randomisierung vorgesehen, obwohl die geplante Fallzahl von 180 Probanden bereits erreicht war. Diese Probanden wurden nach Absprache mit dem Sponsor randomisiert. Insgesamt wurden 217 Patienten randomisiert.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Zwischenanalysen durchgeführt, die Studie wurde nicht vorzeitig abgebrochen.
8	Randomisierung, Erzeugung der Behandlungsfolge	Für die Randomisierung wurde ein interaktives, Web-basiertes Dialogsystem (IWRS) verwendet. Der Prüfarzt oder sein Beauftragter musste jeden Probanden im IWRS registrieren und wurde aufgefordert, die notwendigen Daten für die Stratifizierung einzugeben. Eine eindeutige Randomisierungsnummer wurde entsprechend dem Hauptrandomisierungsplan durch das IWRS zugewiesen. Die Zuteilung der Behandlungsgruppen erfolgte im Verhältnis 1:1:1 innerhalb jedes Stratum.
8a	Methode zur Generierung der zufälligen Zuteilung	
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung erfolgte nach den folgenden Kriterien: 1. Vorbehandlungswert von $AP > 3 \times ULN$ und/oder $AST > 2 \times ULN$ und/oder Bilirubin $> ULN$, UDCA-Unverträglichkeit 2. Vorbehandlungswert von $AP \leq 3 \times ULN$ und $AST \leq 2 \times ULN$ und Bilirubin $\leq ULN$, UDCA-Unverträglichkeit 3. Vorbehandlungswert von $AP > 3 \times ULN$ und/oder $AST > 2 \times ULN$ und/oder Bilirubin $> ULN$, gegenwärtige Einnahme von UDCA 4. Vorbehandlungswert von $AP \leq 3 \times ULN$

Item ^a	Charakteristikum	Studieninformation
		und AST ≤ 2xULN und Bilirubin ≤ ULN, gegenwärtige Einnahme von UDCA
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Behandlungsfolge wurde geheim gehalten. Die Zuteilung der Medikation erfolgte durch ein interaktives, Web-basiertes Dialogsystem (IWRS). Die Medikamente wurden in Behältern zu je 30 Tabletten ausgeteilt. Die Behälter enthielten keine Information, die auf die Behandlung/Dosis hinwies. Um die Randomisierung während der Titrations-Phase zu gewährleisten, wurden alle Probanden für die Titrierung bei Monat 6 bewertet und für Probanden, die die Titrierungs-Kriterien erfüllten, wurde die Anfrage nach Titration durch das IWRS gestellt. Allerdings wurde mit Hilfe des IWRS nur bei den Probanden in der Titrations-Gruppe tatsächlich eine Titrierung durchgeführt. Eine Notfallentblindung konnte ebenfalls durch das IWRS durchgeführt werden.
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 Randomisierungsliste wurde von Sharp Clinical Services (ehemals Bilcare) erstellt. Die Probanden wurden vom Prüfarzt eingeschlossen. Der Prüfarzt kontaktierte das IWRS, um für den eingeschlossenen Patienten eine Randomisierungsnummer zu erhalten.
11	Verblindung	Doppelverblindung. Alle Endpunkterheber waren verblindet.
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?	○ Verblindung der Patienten: ja ○ Verblindung der Personen, die die Zielgrößen beurteilten: ja ○ Vorgehensweise der Verblindung: Identische aussehende Arzneimittelbehälter, doppelblind.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht relevant.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Zur Analyse des (binären) primären Endpunktes wurde ein Cochran-Mantel-Haenszel (CMH) Test stratifiziert nach den Stratifikationskriterien durchgeführt. Die Primäranalyse verglich die Obeticholsäure 10 mg-Gruppe mit der Kontrollgruppe zum Signifikanzniveau von 5%. Binäre Endpunkte wurden mit Hilfe des CMH Tests unter Berücksichtigung des Stratifikationsfaktors analysiert. Metrische Endpunkte wurden mit Hilfe eines Kovarianzanalyse-Modells (ANCOVA)

Item ^a	Charakteristikum	Studieninformation
		durchgeführt. Dabei wurden die Behandlungsgruppe und der Stratifizierungsfaktor als feste Effekte und der jeweilige Baseline-Wert als Covariate berücksichtigt. Parameter, die zu mehreren Zeitpunkten erhoben wurden, wurden ergänzend durch ein gemischtes Modell mit Messwiederholungen ausgewertet. Die Behandlungsgruppe, der Zeitpunkt, die Interaktion zwischen Behandlungsgruppe und Zeit, der Baseline-Wert und der Stratifikationsfaktor wurden als feste Effekte berücksichtigt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppenanalysen: Es wurden a priori im statistischen Analyseplan definierte Subgruppenanalysen durchgeführt, um den Einfluß potentieller Effektmodifikatoren auf den primären Endpunkt zu untersuchen. Untersucht wurden (jeweils in Kategorien) Alter, Alter bei PBC Diagnose, Geschlecht, ethnische Herkunft, Baseline BMI, Baseline AP-Spiegel, Baseline Bilirubin-Spiegel, Einnahme von UDCA, Jahre seit Diagnose von PBC, Region. Adjustierte Analysen: Die Analysen für den primären Endpunkt wurden für den Stratifizierungsfaktor adjustiert wie in 12a beschrieben.
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Anzahl der randomisierten Patienten: ○ Placebo: 73 ○ Obeticholsäure Titration: 71 ○ Obeticholsäure 10 mg: 73 Anzahl der Patienten, die tatsächlich die Behandlung erhalten haben: ○ Placebo: 73 ○ Obeticholsäure Titration: 70 ○ Obeticholsäure 10 mg: 73 Anzahl der Patienten, die in der Analyse des primären Zielkriteriums in den vordefinierten Analysepopulationen berücksichtigt wurden: ○ Placebo: • ITT: 73 • Completer Population: 70

Item ^a	Charakteristikum	Studieninformation
		• Wirksamkeits-bewertbare (EE, efficacy evaluable) Population: 67 ○ Obeticholsäure Titration: • ITT: 70 • Completer Population: 64 • Wirksamkeits-bewertbare (EE, efficacy evaluable) Population: 63 ○ Obeticholsäure 10 mg: • ITT: 73 • Completer Population: 64 • Wirksamkeits-bewertbare (EE, efficacy evaluable) Population: 62
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Ausgeschlossene Patienten nach Randomisierung: ○ Placebo: • ITT: 0 • Completer Population: 3 • Wirksamkeits-bewertbare (EE, efficacy evaluable) Population: 6 ○ Obeticholsäure Titration: • ITT: 1 • Completer Population: 7 • Wirksamkeits-bewertbare (EE, efficacy evaluable) Population: 8 ○ Obeticholsäure 10 mg: • ITT: 0 • Completer Population: 9 • Wirksamkeits-bewertbare (EE, efficacy evaluable) Population: 11 Ausschlußgründe: Ein Patient zog in der Obeticholsäure Titrationsgruppe seine Einverständniserklärung nach Randomisierung aber vor Gabe der ersten Dosis der Studienmedikation zurück (daher nicht in ITT). Die „Completer“ Population enthält nur diejenigen Patienten, die mindestens eine Dosis der Studienmedikation bekommen haben und bis zum Ende der doppelblinden Phase der Studie (Monat 12) teilgenommen haben. 1 (<1%) wegen Tod, 8 (4%) wegen Pruritus, 6 (3%) wegen anderen unerwünschten Nebenwirkungen (nicht Pruritus), 4 zogen ihre Einverständniserklärung zurück. Die Wirksamkeits-bewertbare Population

Item ^a	Charakteristikum	Studieninformation
		enthält nur diejenigen Patienten, die keine bedeutenden Protokollabweichungen, die potentiell die Wirksamkeit des Medikaments beeinträchtigen könnten, hatten.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Einschluss des ersten Patienten bis zur letzten Nachbeobachtung des letzten Patienten zwischen März 2012 und Dezember 2013. ○ Screening: Im Intervall von ≤ 1 bis zu 8 Wochen vor Tag 0, wenn nötig musste eine zweite Screening-Visit möglich sein (um biochemische Tests zu wiederholen) und es musste gewährleistet werden, dass genügend Zeit zum Erhalt der Laborresultate zur Verfügung stand. ○ Tag 0: Tag der Randomisierung; die erste Dosis der Studienmedikation sollte am folgenden Tag (Tag 1) genommen werden. ○ Bis Monat 6 wurde die Studienmedikation täglich eingenommen. ○ Monat 6: Bewertung der Patienten im Hinblick auf eine mögliche Titration. Nur Patienten in der Titrationsgruppe wurden tatsächlich titriert (gesteuert vom IWRS). ○ Bis Monat 12 wurde die Studienmedikation täglich eingenommen. ○ Ab Monat 12 / Langzeitsicherheitsbeobachtung Tag 1 wurden die Patienten auf Wunsch auf eine Open-Label Behandlung mit 5 mg Obeticholsäure umgestellt.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde regulär beendet.
a: nach CONSORT 2010.		

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

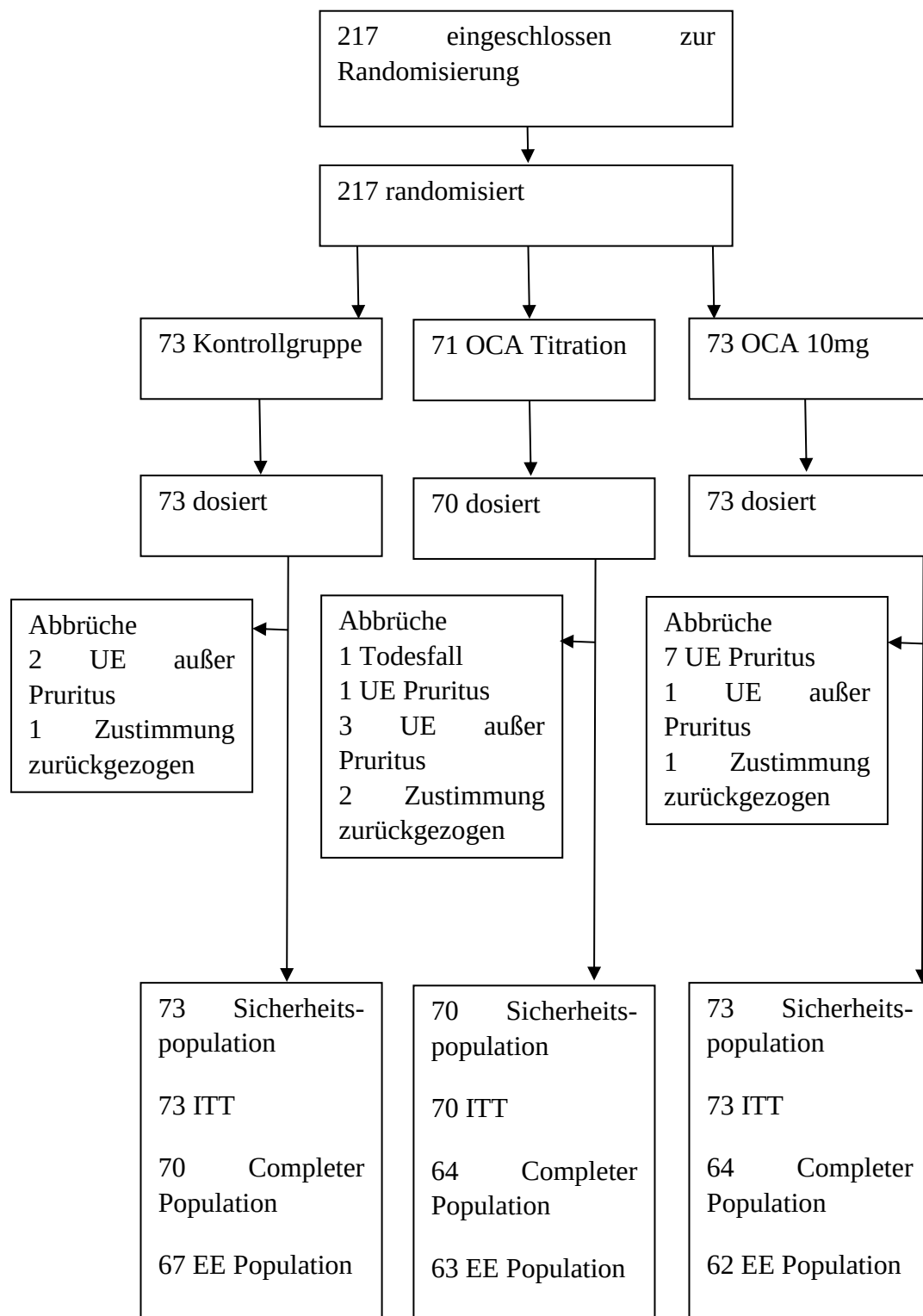


Abbildung 4-53: Flow-Chart der Patienten der POISE-Studie

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 4-129 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie <Studienbezeichnung>

Studie: _____ **POISE** _____

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
CSR Data on File: Clinical Study Report, Protocol 747-301	POISE

A: Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

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

Angaben zum Kriterium:

Die Studie wurde randomisiert kontrolliert durchgeführt.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

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

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

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

Die Zuteilung der Randomisierungsnummern wurde durch ein zentrales unabhängiges interaktives Web-basiertes Dialogsystem (IWRS) durchgeführt. Die Randomisierung erfolgte stratifiziert nach (1) $AP \leq 3 \times ULN$ und $AST \leq 2 \times ULN$ und Gesamt-Bilirubin $\leq ULN$ (ja/nein) und

(2) Toleranz von UDCA Therapie (ja/nein).

Das Randomisierungsschema wurde von einer unabhängigen Firma (Sharp Clinical Services [vorher Bilcare]) zur Verfügung gestellt und überprüft.

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:

Die Patienten wurden durch ein interaktives Web-basiertes Dialogsystem (IWRS) randomisiert. Während der doppelblinden Phase der Studie (12 Monate) waren alle Behälter für die Studienmedikation für beide Obeticholsäure-Behandlungsarme und für die Kontroll-Gruppe visuell identisch. Die UDCA-Medikation war nicht Teil der Studienmedikation. Die Behälter mit der Studienmedikation enthielten keine Informationen über die Stärke der Tabletten.

Um die Verblindung zum Zeitpunkt der Titrierung (Monat 6) zu erhalten, wurden alle Patienten (d.h., alle Behandlungsgruppen) zu Monat 6 basierend auf AP und/oder Bilirubin-Ansprechen und Verträglichkeit der Studienmedikation für die Titration bewertet. Allerdings konnten nur Patienten im Obeticholsäure-Behandlungsarm 5 mg, titriert auf 10 mg, tatsächlich eine Steigerung ihrer Dosis erhalten, was durch das interaktive Web-basierte Dialogsystem (IWRS) kontrolliert wurde. Eine Anfrage auf Titration konnten die Prüfarzte über das IWRS für Patienten stellen, die mindestens eins der folgenden Kriterien bei der Monat-6-Bewertung erfüllten:

- $AP \geq 1.67 \times \text{ULN}$ und/oder
- Gesamt-Bilirubin $> \text{ULN}$ oder
- $<15\%$ AP Reduktion bei Monat 6 im Vergleich zu den mittleren Vorbehandlungswerten in der doppelblinden Phase und
- vorausgesetzt, dass unerwünschte Ereignisse (z.B. schwerer Pruritus) die Gabe einer höheren Dosis nicht beschränken.

Bei Übertritt in die Open-Label Langzeit-Sicherheits-Extension der Studie nach Abschluß der doppelblinden Studienphase von 12 Monaten wurde die Verblindung weiterhin beibehalten. Informationen über die Behandlung während der doppelblinden Phase wurden nicht preisgegeben, alle Patienten starteten die Open-Label-Verlängerung der Studie mit der gleichen (Start-)Dosis von 5 mg Obeticholsäure.

behandelnde bzw. weiterbehandelnde Personen:

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten wurden durch ein interaktives Web-basiertes Dialogsystem (IWRS) randomisiert. Während der doppelblinden Phase der Studie (12 Monate) waren alle Behälter für die Studienmedikation für beide Obeticholsäure-Behandlungsarme und für die Kontroll-Gruppe visuell identisch. Die UDCA-Medikation war nicht Teil der Studienmedikation. Die Behälter mit der Studienmedikation enthielten keine Informationen über die Stärke der Tabletten.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

In der POISE-Studie wurde die Randomisierung adäquat durchgeführt und die Gruppeneinteilung verdeckt über ein Sprachdialogsystem (IxRS) durchgeführt. Patienten und behandelnde Personen waren während der randomisierten Phase verblindet. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren oder - wie bei den Sicherheitsendpunkten - gemäß Verfahrensordnung berichtet werden musste (z.B. die Berechnung der Effektmaße für den Vergleich der Behandlungsgruppen). Es lagen keine sonstigen Aspekte für ein erhöhtes Verzerrungspotenzial vor, so dass das Verzerrungspotenzial insgesamt als niedrig eingestuft wird.

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt**Endpunkt: _Gesamtmortalität_****1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Der Endpunkt „Gesamtmortalität“ wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte nach dem ITT-Prinzip auf Basis der Sicherheitspopulation. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Das Verzerrungspotenzial für den Endpunkt „Gesamtmortalität“ wird daher mit niedrig bewertet.

Endpunkt: _ Primärer Endpunkt (AP <1.67x ULN, Gesamt-Bilirubin ≤ULN und AP-Reduktion von ≥15% vom Ausgangswert)_____

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkte AP und Bilirubin und damit auch der hieraus abgeleitete primäre kombinierte Wirksamkeitsendpunkt AP<1.67xULN und Gesamt-Bilirubin≤ULN und AP-Reduktion zu Baseline ≥15% wurden verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☐ ja ☒ nein

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

Ein sonstiger Aspekt liegt vor, da der Endpunkt als Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist.

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Der Endpunkt wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis der ITT-Population. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen, für die Einstufung wesentlichen, Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da der Endpunkt als ein Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist, wird das Verzerrungspotenzial für diese Endpunkte daher mit hoch bewertet.

Endpunkt: _ AP-Reduktion zu Baseline _

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☐ ja ☒ nein

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

Ein sonstiger Aspekt liegt vor, da der Endpunkt als Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist.

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Der Endpunkt wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis der ITT-Population. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da der Endpunkt als ein Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist, wird das Verzerrungspotenzial für diese Endpunkte daher mit hoch bewertet.

Endpunkt: _ AP ≤ ULN _____**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☐ ja ☒ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Ein sonstiger Aspekt liegt vor, da der Endpunkt als Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist.

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

Begründung für die Einstufung:

Der Endpunkt wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis der ITT-Population. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da der Endpunkt als ein Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist, wird das Verzerrungspotenzial für diese Endpunkte daher mit hoch bewertet.

Endpunkt: _ Änderung zu Baseline in den biochemischen Kenngrößen AP und Bilirubin**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☐ ja ☒ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Ein sonstiger Aspekt liegt vor, da der Endpunkt als Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist.

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

Begründung für die Einstufung:

Der Endpunkt wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis der ITT-Population. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da der Endpunkt als ein Surrogat für die klinischen Endpunkte Mortalität und Lebertransplantation dient, aber nicht formal validiert ist, wird das Verzerrungspotenzial für diese Endpunkte daher mit hoch bewertet.

Endpunkt: _Fibrose (ELF-Score)_**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der ELF-Score und dessen Komponenten wurden in der POISE-Studie verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Zur Analyse des Endpunkts ELF-Score wurden nur beobachtete Daten genutzt, fehlende Werte wurden nicht ersetzt. Durch Einsatz einer Kovarianz-Analyse (ANCOVA) wurden nur Patienten in die Analyse einbezogen, für die mindestens die Baseline-Messung und ein Beobachtungswert nach Baseline vorlagen. Dies waren bei den Patienten mit Kombinationstherapie 59 von 65 (91%) der Patienten in der Obeticholsäure+UDCA-Gruppe und 64 von 68 (94%) der Patienten in der UDCA+Placebo-Gruppe. Bei den Patienten mit Monotherapie wurden 4 von 5 (80%) der Patienten in der Obeticholsäure-Gruppe und 3 von 5 (60%) der Patienten in der Placebo-Gruppe in die Analyse aufgenommen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☐ ja ☒ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der ELF-Score ist ein Surrogat für die mittels Biopsie bestimmte Zirrhose.

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

Begründung für die Einstufung:

Der ELF-Score und dessen Komponenten wurden in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte auf Basis einer modifizierten ITT-Population, die alle Patienten enthielt, für die mindestens der Baseline-Wert und ein Beobachtungswert nach Baseline vorlagen. Das ITT-Prinzip wurde nicht adäquat umgesetzt, da in die Teilpopulation der Patienten mit Kombinationstherapie in der Obeticholsäure+UDCA-Gruppe nur 59 von 65 (91%) Patienten in die Analyse eingingen. In der UDCA+Placebo-Gruppe waren es 64 von 68 (94%). In die Teilpopulation der Patienten mit Monotherapie wurden nur 4 von 5 Patienten der Obeticholsäure-Gruppe und 3 von 5 Placebo-Patienten in die Analyse aufgenommen. Die sonstigen für die

Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Da der ELF-Score ein Surrogat für die mittels Biopsie bestimmte Zirrhose ist, wird das Verzerrungspotenzial für diese Endpunkte mit hoch bewertet.

Endpunkt: _ Fibrose (Fibroscan®) _____

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde in der POISE-Studie verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips

☐ ja ☐ unklar ☒ nein

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

Die Messung der Leber-Steifigkeit wurde mit Hilfe des Fibroscan® TE-Geräts vorgenommen, allerdings nur in den Studienzentren, in denen die entsprechende Ausrüstung verfügbar war. Diese Studienzentren hatten ausgebildete Mitarbeiter bezüglich der Nutzung und Dateninterpretation des Geräts.
Der Endpunkt „Fibrose“ wurde bei den Patienten mit Kombinationstherapie bei 62 von 133 Patienten (47%; Obeticholsäure+UDCA: 46%, UDCA+Placebo: 47%) erhoben. Bei den Patienten mit Monotherapie standen bei 4 von 10 Patienten (40%; Obeticholsäure: 40%, Placebo: 40%) Daten für diesen Endpunkt zur Verfügung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☐ ja ☒ nein

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

Der Fibroscan® wird als Surrogat für die Messung fortgeschrittener Fibrose und Zirrhose genutzt.

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Der Fibroscan® wurde in einer Teilpopulation von Patienten erhoben und dokumentiert. Es lag keine informative Selektion vor, sondern eine zufällige Auswahl, da der Fibroscan® nur in einigen Zentren dokumentiert wurde. Sonstige Aspekte lagen nicht vor. Durch die Einschränkungen bezüglich der ITT-Population (64 von 133 Patienten mit Fibroscan®) wird das Verzerrungspotenzial als hoch eingestuft.

Endpunkt: _Osteoporose (DEXA-Scan)_**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde in der POISE-Studie verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt „Osteoporose“ wurde mittels DEXA-Scan auf Basis der Sicherheitspopulation nur in ausgewählten Studienzentren erhoben. Hierdurch wurde die Analysepopulation deutlich eingeschränkt:

Bei Patienten mit Kombinationstherapie wurden für den Endpunkt „Schenkelhals-Scan: T-Score“ 70 von 133 Patienten in die Analyse aufgenommen (53%; Obeticholsäure+UDCA: 57%, UDCA+Placebo: 49%), für den Endpunkt „Schenkelhals-Scan: Z-Score“ 68 von 133 Patienten (51%; Obeticholsäure+UDCA: 55%, UDCA+Placebo: 47%), für den Endpunkt „Schenkelhals-Scan: Knochendichte“ 67 von 133 Patienten (50%; Obeticholsäure+UDCA: 55%, UDCA+Placebo: 46%), für den Endpunkt „Lendenwirbelsäulen-Scan: T-Score“ 79 von 133 Patienten (59%; Obeticholsäure+UDCA: 60%, UDCA+Placebo: 59%), für den Endpunkt „Lendenwirbelsäulen-Scan: Z-Score“ 74 von 133 Patienten (56%; Obeticholsäure+UDCA: 58%, UDCA+Placebo: 53%), für den Endpunkt „Lendenwirbelsäulen-Scan: Knochendichte“ 76 von 133 Patienten (57%; Obeticholsäure+UDCA: 58%, UDCA+Placebo: 56%).

Bei den Patienten mit Monotherapie wurden für die Endpunkte „Schenkelhals-Scan: T-Score“ und „Lendenwirbelsäulen-Scan: T-Score“ 6 von 10 Patienten in die Analyse aufgenommen (60%; Obeticholsäure: 60%, Placebo: 60%) und für die Endpunkte „Schenkelhals-Scan: Z-Score“, „Schenkelhals-Scan: Knochendichte“, „Lendenwirbelsäulen-Scan: Z-Score“ und „Lendenwirbelsäulen-Scan: Knochendichte“ 5 von 10 Patienten (50%; Obeticholsäure: 60%, Placebo: 40%).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☐ ja ☒ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der Endpunkt Osteoporose wird anhand des Surrogats DEXA-Scan erhoben.

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

☐ **niedrig**☒ **hoch**

Begründung für die Einstufung:

Der DEXA-Scan wurde in einer Teilpopulation von Patienten erhoben und dokumentiert. Die Analyse erfolgte auf Basis der Sicherheitspopulation. Es lag keine informative Selektion vor, sondern eine zufällige Auswahl, da der DEXA-Scan nur in einigen Zentren dokumentiert wurde. Als sonstiger Aspekt ist zu nennen, dass der DEXA-Scan ein Surrogat für die Bewertung der Osteoporose gilt. Durch die Einschränkungen bezüglich der ITT-Population und die Verwendung eines Surrogats wird das Verzerrungspotenzial als hoch eingestuft.

Endpunkt: PBC-40**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse erfolgte auf Basis der ITT-Population, fehlende Werte wurden nicht ersetzt. Es wurden jedoch jeweils mehr als 95% der Patienten der Population in die Analysen aufgenommen.
Patienten mit Kombinationstherapie:
Generelle Symptome, Fatigue, kognitive Funktion, soziale und emotionale Domäne: 131 von 133 Patienten (98%; Obeticholsäure+UDCA: 98%, UDCA+Placebo: 99%)
Jucken: 130 von 133 Patienten (98%; Obeticholsäure+UDCA: 97%, UDCA+Placebo: 99%)
Patienten mit Monotherapie:
10 von 10 Patienten (100%) bei allen Komponenten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Der Endpunkt „PBC-40“ wurde in der POISE-Studie verblindet erhoben. Die Auswertung erfolgte nach dem ITT-Prinzip auf Basis der Sicherheitspopulation. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Die sonstigen für die Einstufung wesentlichen Aspekte (z. B. Verblindung und Transparenz der berichteten Ergebnisse) wiesen auf eine hohe methodische Qualität des Endpunktes hin. Das Verzerrungspotenzial für den Endpunkt „Gesamt mortalität“ wird daher mit niedrig bewertet.

Endpunkt: _UE Pruritus_____**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. In der POISE-Studie waren sowohl Patienten als auch Behandler verblindet. Die Analyse zum Endpunkt Verträglichkeit erfolgte auf Basis der Sicherheitspopulation, in die alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhielten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es lagen auch keine sonstigen Aspekte einer möglichen Verzerrung vor. Somit wird das Verzerrungspotential dieses Endpunkts als niedrig eingestuft.

Endpunkt: _5-D_**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse erfolgte auf Basis der ITT-Population, allerdings wurden nur beobachtete Daten genutzt, fehlende Werte wurden nicht ersetzt. Durch Einsatz einer Kovarianz-Analyse (ANCOVA) wurden nur Patienten in die Analyse einbezogen, für die mindestens die Baseline-Messung und ein Beobachtungswert nach Baseline vorlagen.

Dies waren bei den Patienten mit Kombinationstherapie für den Gesamtscore 117 von 133 Patienten (88%; Obeticholsäure+UDCA: 88%, UDCA+Placebo: 88%), für die Dauer 123 von 133 Patienten (92%; Obeticholsäure+UDCA: 91%, UDCA+Placebo: 94%), für den Schweregrad 127 von 133 Patienten (95%; Obeticholsäure+UDCA: 95%, UDCA+Placebo: 96%), für die Richtung 118 von 133 Patienten (89%; Obeticholsäure+UDCA: 89%, UDCA+Placebo: 88%), für die Behinderung 125 von 133 (94%; Obeticholsäure+UDCA: 94%, UDCA+Placebo: 94%) und für die Verteilung 127 von 133 Patienten (95%; Obeticholsäure+UDCA: 95%, UDCA+Placebo: 96%).

Bei den Patienten mit Monotherapie wurden für den Gesamtscore und die Richtung 9 von 10 Patienten (90%; Obeticholsäure: 100%, Placebo: 80%), für die Dauer, den Schweregrad, die Behinderung und die Verteilung 10 von 10 Patienten (100%) in die Analyse einbezogen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig

eingestuft wurde. Die Studie wurde verblindet durchgeführt. Die Analyse erfolgte auf Basis der ITT-Population, in die alle Patienten aufgenommen wurden, die der in der Fachinformation beschriebenen Population entsprachen. In der Teilpopulation der Patienten mit Kombinationstherapie wurden für den 5-D in den Domänen Schweregrad und Verteilung jeweils mehr als 95% der Patienten der ITT-Population in die Analyse eingeschlossen. In den Domänen Dauer, Richtung und Behinderung wurde dieses Kriterium verfehlt. Es lagen keine sonstigen Aspekte einer möglichen Verzerrung vor. In der Teilpopulation der Patienten mit Monotherapie lagen nur für die 5-D Domäne Richtung für die Kontrollgruppe Daten von weniger als 95% der ITT-Population vor. Insgesamt wird das Verzerrungspotenzial für diese Endpunkte daher als niedrig eingestuft.

Endpunkt: _Pruritus-VAS_____**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse erfolgte auf Basis der ITT-Population, fehlende Werte wurden nicht ersetzt. Durch Einsatz einer Kovarianz-Analyse (ANCOVA) wurden nur Patienten in die Analyse einbezogen, für die mindestens die Baseline-Messung und ein Beobachtungswert nach Baseline vorlagen.
Dies waren bei den Patienten mit Kombinationstherapie 127 von 133 Patienten (95%; Obeticholsäure+UDCA-Gruppe: 95%, UDCA+Placebo-Gruppe: 96%) und bei den Patienten mit Monotherapie 10 von 10 Patienten (100%).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig eingestuft wurde.

Die Studie wurde verblindet durchgeführt. Die Analyse erfolgte auf Basis der ITT-Population, in die alle Patienten aufgenommen wurden, die der in der Fachinformation beschriebenen Population entsprachen. Für die Pruritus-VAS wurden mehr als 95% der Patienten mit Kombinationstherapie in die Analyse eingeschlossen. Insgesamt wird das Verzerrungspotenzial für den Endpunkt Pruritus-VAS als niedrig eingestuft.

Endpunkt: Gesamtrate UE

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. In der POISE-Studie waren sowohl Patienten als auch Behandler verblindet. Die Analyse zum Endpunkt Verträglichkeit erfolgte auf Basis der Sicherheitspopulation, in die alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhielten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es lagen auch keine sonstigen Aspekte einer möglichen Verzerrung vor. Somit wird das Verzerrungspotential dieses Endpunkts als niedrig eingestuft.

Endpunkt: _Schwerwiegende UE_

5. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

6. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. In der POISE-Studie waren sowohl Patienten als auch Behandler verblindet. Die Analyse zum Endpunkt Verträglichkeit erfolgte auf Basis der Sicherheitspopulation, in die alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhielten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es lagen auch keine sonstigen Aspekte einer möglichen Verzerrung vor. Somit wird das Verzerrungspotential dieses Endpunkts als niedrig eingestuft.

Endpunkt: _Schwere UE ohne Pruritus_____**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. In der POISE-Studie waren sowohl Patienten als auch Behandler verblindet. Die Analyse zum Endpunkt Verträglichkeit erfolgte auf Basis der Sicherheitspopulation, in die alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhielten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es lagen auch keine sonstigen Aspekte einer möglichen Verzerrung vor. Somit wird das Verzerrungspotential dieses Endpunkts als niedrig eingestuft.

Endpunkt: _UE, die zum Therapieabbruch führten_____

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Studie wurde doppelblind durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die POISE-Studie wurde als RCT durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. In der POISE-Studie waren sowohl Patienten als auch Behandler verblindet. Die Analyse zum Endpunkt Verträglichkeit erfolgte auf Basis der Sicherheitspopulation, in die alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhielten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan definiert waren. Es lagen auch keine sonstigen Aspekte einer möglichen Verzerrung vor. Somit wird das Verzerrungspotential dieses Endpunkts als niedrig eingestuft.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Verzerrungsaspekte auf Studienebene:**Einstufung als randomisierte Studie**

☐ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein:** Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

☐ **ja:** Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

☐ **unklar:** Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

☐ **nein:** Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

☐ **ja:** Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.☐ **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**für nicht randomisierte Studien:****Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

☐ **unklar:** Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.☐ **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Verblindung von Patienten und behandelnden Personen****Patient**☐ **ja:** Die Patienten waren verblindet.☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.
- Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.
- Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.
- Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- Subgruppen
- Zeitpunkte/-räume
- Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)
- Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)
- Cut-off-points bei Dichotomisierung
- statistischer Verfahren

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).
- Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.

Zulässige Gründe sind:

- erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung
- Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe
- geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben
- Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.
- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

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

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

☐ **ja**

☐ **nein**

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

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

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- ☐ **hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ☐ **ja:** Der Endpunkt wurde verblindet erhoben.
- ☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.
- ☐ **nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ☐ **ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante 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 zu Gunsten 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:
