

von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V

**Bewertung von Arzneimitteln für seltene Leiden nach § 35a
Absatz 1 Satz 11 i. V. m. Kapitel 5 § 12 Nr. 1 Satz 2 Verfo**

Wirkstoff: Olezarsen

**Neues Anwendungsgebiet: Familiäres Chylomikronämie-
Syndrom**

Datum der Veröffentlichung: 2. März 2026

Version 1.0

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abkürzungsverzeichnis	4
Hintergrund.....	5
1 Fragestellung.....	6
2 Auswahl und Methodik der relevanten Studien	7
2.1 Studienbasis für die Nutzenbewertung	7
2.2 Beschreibung der eingeschlossenen Studie	7
2.3 Endpunkte	12
2.3.1 Mortalität.....	13
2.3.2 Morbidität.....	13
2.3.3 Lebensqualität	22
2.3.4 Sicherheit	27
2.3.5 Erhebungszeitpunkte	29
2.4 Statistische Methoden.....	30
2.5 Verzerrungspotential auf Studien- und Endpunktebene	32
3 Ergebnisse der eingeschlossenen Studie	35
3.1 Studiencharakteristika und Studienmedikation	35
3.2 Mortalität	38
3.3 Morbidität	39
3.4 Lebensqualität.....	43
3.5 Sicherheit	46
4 Diskussion der Methodik und Ergebnisse.....	49
4.1 Design und Methodik der Studie.....	49
4.2 Zulassungsstatus und Zulassungspopulation.....	51
4.3 Zusammenfassende Beurteilung zum Verzerrungspotential und zur Ergebnissicherheit	52
4.4 Mortalität	52
4.5 Morbidität	52
4.6 Lebensqualität.....	53
4.7 Sicherheit	54
5 Anforderung an eine qualitätsgesicherte Anwendung	56
6 Zusammenfassung der Nutzenbewertung	57
Referenzen	60
Anhang	62

Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis.....	7
Tabelle 2:	Charakterisierung der Studie Balance.....	7
Tabelle 3:	Für die Nutzenbewertung relevante Protokolländerungen der Studie Balance.....	11
Tabelle 4:	Charakterisierung der Intervention der Studie Balance (Kohorte B)	11
Tabelle 5:	Zusammenfassung der Endpunktbewertung der Studie Balance.....	12
Tabelle 6:	Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie Balance	29
Tabelle 7:	Verzerrungspotential der Studie Balance	32
Tabelle 8:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie Balance.....	33
Tabelle 9:	Allgemeine Angaben; Studie Balance	35
Tabelle 10:	Charakterisierung der Studienpopulation; Studie Balance.....	35
Tabelle 11:	Begleitmedikation; Studie Balance	38
Tabelle 12:	Nüchtern-Triglyzeridwert, Prozentuale Veränderung zu Monat 12 (ANCOVA); Studie Balance, FAS (DS: 14.07.2023)	39
Tabelle 13:	Bestätigte akute Pankreatitis bis Monat 12; Studie Balance; FAS (DS: 30.11.2023)	40
Tabelle 14:	Anteil der Personen mit einer Veränderung im PGIS zu Monat 12; Studie Balance, FAS (DS: 14.07.2023)	41
Tabelle 15:	Anteil der Personen mit einer Veränderung in der FCS-SIS (Symptom) zu Monat 12; Studie Balance, FAS (DS: 14.07.2023)	42
Tabelle 16:	Anteil der Personen mit einer Veränderung in der FCS-SIS (Impact) zu Monat 12; Studie Balance, FAS (DS: 14.07.2023)	43
Tabelle 17:	Veränderung des PROMIS-29+2 zu Monat 12 (ANCOVA); Studie Balance, FAS (DS: 14.07.2023)	44
Tabelle 18:	Veränderung des PROMIS Short Form – Pain Interference 8a zu Monat 12 (ANCOVA); Studie Balance, FAS (DS: 14.07.2023)	46
Tabelle 19:	Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der verblindeten Behandlungsphase; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)	46
Tabelle 20:	UE mit Inzidenz $\geq 10\%$; Studie Balance, Sicherheitspopulation (DS: 14.07.2023) ..	47
Tabelle 21:	Schwere UE (CTCAE-Grad ≥ 3), die bei $\geq 5\%$ der Personen in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)	48
Tabelle 22:	SUE, die bei $\geq 5\%$ der Personen in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)	48
Tabelle 23:	Zusammenfassende Darstellung der Ergebnisse der Studie Balance	57
Tabelle 24:	Hospitalisierungen aus jeglichen Gründen bis Monat 12; Studie Balance, FAS (DS: 14.07.2023)	62
Tabelle 25:	Anteil der Personen mit einer Verbesserung im PGIC zu Monat 6; Studie Balance, FAS (DS: 14.07.2023)	63

Abkürzungsverzeichnis

AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Analysis of Covariance
AP	Akute Pankreatitis
CTCAE	Common Terminology Criteria for Adverse Events
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FAS	Full Analysis Set
FCS	Familiäres Chylomikronämie-Syndrom
FCS-SIS	Familial Chylomicronemia Syndrome – Symptom and Impact Scale
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
GSRS	Gastrointestinal Symptom Rating Scale
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Item Response Theory
ITT	Intention to Treat
J2R	Jump to Reference
KI	Konfidenzintervall
LS	Least Squares
MACE	Major Adverse Cardiovascular Event
MAR	Missing at Random
MedDRA	Medical Dictionary for Regulatory Activities
N	Anzahl
PAC	Acute Pancreatitis Adjudication Committee
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PRO	Patient-Reported Outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
PROPr	PROMIS-Preference
PT	Preferred Term
pU	pharmazeutischer Unternehmer
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SGB	Sozialgesetzbuch
SOC	Systemorganklasse
SUE	Schwerwiegendes unerwünschtes Ereignis
TG	Triglyzerid
UE	Unerwünschtes Ereignis
ULN	Upper Limit of Normal
WHOQOL	World Health Organization's Quality of Life

Hintergrund

Olezarsen ist zugelassen als Arzneimittel zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt.

Gemäß § 35a Absatz 2 SGB V entscheidet der Gemeinsame Bundesausschuss (G-BA), ob er die Nutzenbewertung selbst durchführt oder das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beauftragt. Ausgehend von der gesetzlichen Vorgabe in § 35a Absatz 1 Satz 11 1. Halbs. SGB V, dass der Zusatznutzen eines Orphan Drug durch die Zulassung als belegt gilt, hat der G-BA in seiner Sitzung vom 15. März 2012 das Verfahren der Nutzenbewertung von Orphan Drugs dahingehend modifiziert, dass bei Orphan Drugs zunächst keine eigenständige Festlegung einer zweckmäßigen Vergleichstherapie mehr durch den G-BA als Grundlage der insoweit allein rechtlich zulässigen Bewertung des Ausmaßes eines gesetzlich zu unterstellenden Zusatznutzens erfolgt. Vielmehr wird ausschließlich auf der Grundlage der Zulassungsstudien das Ausmaß des Zusatznutzens durch den G-BA bewertet.

Bei Orphan Drugs erfolgt eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann, wenn der Umsatz des betreffenden Arzneimittels die Umsatzschwelle gemäß § 35a Absatz 1 Satz 12 SGB V überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt.

Der G-BA bestimmt gemäß Kapitel 5 § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Olezarsen zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (pU) heran und bewertet die darin enthaltenen Daten nach Maßgabe der in Kapitel 5 § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Olezarsen in seiner Sitzung am 24. Februar 2026 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 28. November 2025 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung (GKV)
- Anzahl der GKV-Patientinnen und -Patienten in der Zielpopulation

Die Nutzenbewertung wird am 2. März 2026 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Olezarsen (Tryngolza®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 30 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. Die Bewertung bezieht sich auf folgende Indikation [23]:

Bei erwachsenen Patienten ergänzend zu einer Diät zur Behandlung des genetisch bestätigten familiären Chylomikronämie-Syndroms (FCS).

Gemäß Fachinformation beträgt die zugelassene Dosierung 80 mg, verabreicht als subkutane Injektion einmal monatlich.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzen- bewertung	Ausschlussgrund
Studien zum Wirkstoff				
Balance (ISIS-678354- CS3) ¹⁾²⁾	Ja	Ja	Ja	

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

²⁾ Patientinnen und Patienten konnten aus der Studie Balance in die offene einarmige Extensionsstudie ISIS-678354-CS13 übergehen. Diese wird aufgrund des einarmigen Studiendesigns nicht zur Nutzenbewertung herangezogen.

Abkürzungen: EPAR: European Public Assessment Report; pU: pharmazeutischer Unternehmer.

Die für die Nutzenbewertung herangezogene Studie entspricht derjenigen Studie, auf der die Einschätzungen des pU zum Zusatznutzen basieren.

Zur Nutzenbewertung für Olezarsen herangezogene Unterlagen

- Dossier des pU zu Olezarsen [22]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [13]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie Balance (ISIS-678354-CS3) [14, 15, 16, 17, 18, 19]
- Fachinformation zu Olezarsen [23]

2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Olezarsen im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie Balance (ISIS-678354-CS3). Die Studie und die Intervention werden in Tabelle 2 und Tabelle 4 charakterisiert.

Tabelle 2: Charakterisierung der Studie Balance

Charakteris- tikum	Beschreibung
Design und Studienablauf	Design Multizentrische, doppelblinde, randomisiert kontrollierte Phase-III-Studie zur Untersuchung der Wirksamkeit und Sicherheit von Olezarsen gegenüber Placebo bei erwachsenen Patientinnen und Patienten mit genetisch bestätigtem familiärem Chylomikronämie-Syndrom (FCS) ergänzend zu einer Diät (≤ 20 g Fett pro Tag). Randomisierung Randomisierung im Verhältnis 1 : 1 in die Kohorten A und B: • Kohorte A: Behandlung mit 50 mg Olezarsen oder Placebo 0,5 ml • Kohorte B: Behandlung mit 80 mg Olezarsen oder Placebo 0,8 ml Anschließende Randomisierung in der jeweiligen Kohorte im Verhältnis 2 : 1 Olezarsen vs. Placebo.

Charakteristikum	Beschreibung
	<u>Stratifizierung nach</u> • Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein) • Vorbehandlung mit Volanesorsen (ja vs. nein) Die Behandlung mit Olezarsen in Kohorte A entspricht laut Fachinformation nicht der zugelassenen Dosierung von 80 mg [23]. Entsprechend wird ausschließlich der Behandlungsarm mit Olezarsen 80 mg aus Kohorte B gegenüber Placebo (gepoolt) der Studie Balance zur Nutzenbewertung herangezogen.¹⁾ Studienablauf (siehe Abbildung 1) • Screening-Phase (4–8 Wochen) ◦ Diätstabilisierungs-/Run-in-Phase (mind. 2 Wochen) ◦ Qualifizierungsphase²⁾ (2 Wochen) • Behandlungsphase (53 Wochen) • Nachbeobachtung (13 Wochen) oder Teilnahme an der offenen, einarmigen Extensionsstudie (ISIS-678354-CS13)
Population	Wesentliche Einschlusskriterien • Alter mind. 18 Jahre. • FCS-Diagnose mit bestätigter homozygoter, compound-heterozygoter oder doppelheterozygoter Mutation und Funktionsverlust in Typ-1-verursachenden Genen (wie LPL, GPIHBP1, ApoA5, ApoC2, GPD1 oder LMF1). • Nüchtern-TG-Wert ≥ 880 mg/dl (10 mmol/l) zum Screening³⁾. • Mit Ausnahmen Vorgeschichte einer bestätigten Pankreatitis⁴⁾⁵⁾ innerhalb von 10 Jahren vor dem Screening. • Bereitschaft Diät mit ≤ 20 g Fett pro Tag einzuhalten. Wesentliche Ausschlusskriterien • Klinisch bedeutsame Auffälligkeiten in der Anamnese (z. B. vorheriges akutes Koronarsyndrom innerhalb von 6 Monaten vor dem Screening, größere chirurgische Eingriffe innerhalb von 3 Monaten vor dem Screening) oder bei der körperlichen Untersuchung. • Aktive Pankreatitis innerhalb von 4 Wochen vor dem Screening. • Klinisch bedeutsame Abnormalitäten bei den Screening-Laborwerten: ◦ Thrombozytenzahl $< 100.000/\text{mm}^3$⁶⁾ ◦ ALT oder AST $> 3,0 \times \text{ULN}$ ◦ Gesamtbilirubin $> \text{ULN}$, es sei denn, es liegt ein Gilbert-Syndrom vor ◦ eGFR < 45 ml/min/1,73m² (gemäß CKD-EPI-Formel) ◦ UPCR ≥ 500 mg/g oder UACR ≥ 300 mg/g • Unkontrollierte arterielle Hypertonie ($> 160/100$ mmHg). • Blutungsdiathese oder Koagulopathie in der Vorgeschichte oder klinisch signifikante Anomalien der Gerinnungsparameter beim Screening. • Aktive Infektion mit COVID-19, HIV, Hepatitis C oder Hepatitis B und/oder aktive Infektion, die systemisch antivirale oder antimikrobielle Therapie erfordert und nicht vor ersten Studientag abgeschlossen ist. • Bösartige Erkrankungen innerhalb von 5 Jahren mit Ausnahme von Basal- oder Plattenepithelkarzinomen der Haut oder Carcinoma in situ des Gebärmutterhalses, die erfolgreich behandelt worden sind. • Blutspende von 50 bis 499 ml innerhalb von 30 Tagen vor dem Screening oder von > 499 ml innerhalb von 60 Tagen vor dem Screening.

Charakteristikum	Beschreibung
	Vorherige Behandlung mit einem Oligonukleotid (einschließlich siRNA) innerhalb von 4 Monaten vor dem Screening oder 5 Halbwertszeiten, je nachdem, welcher Zeitraum länger war; gilt nicht für Impfstoffe.
Intervention und Zahl der Patientinnen und Patienten	Gescreent gesamt: N = 144 Randomisiert/Eingeschlossen: N = 66 (Kohorten A und B) Intervention: N = 22 (Kohorte B) Kontrolle: N = 23 (gepoolt: Kohorte A und B)
Ort und Zeitraum der Durchführung; Datenschnitte	Studienzentren Insgesamt 29 Studienzentren in den Vereinigten Staaten (9), Kanada (3), Europa (16) sowie dem Vereinigten Königreich (1). Studienzeitraum - Erste Person, erste Dosis: 17.12.2020 - Letzter Person, letzte Dosis: 20.06.2023 Datenschnitt - Datenschnitt zum Ende der Behandlung: 14.07.2023 - Korrigierter Datenschnitt: 30.11.2023⁷⁾ Zum Zeitpunkt des Datenschnitts am 14.07.2023 befanden sich gemäß den Angaben des pU noch Patientinnen und Patienten in der Nachbeobachtung. Der Datenschnitt vom 30.11.2023 enthält keine neu erhobenen Daten, sondern lediglich eine Korrektur hinsichtlich eines Wirksamkeitsendpunkts⁷⁾. Ein weiterer bzw. finaler Datenschnitt wurde nicht vorgelegt. Gemäß Angaben des pU wurde die Studie planmäßig beendet.
Endpunkte gemäß Studienprotokoll	Primärer Endpunkt Prozentuale Veränderung des Nüchtern-TG-Werts gegenüber Baseline. Sekundäre Endpunkte - Prozentuale Veränderung des Nüchtern-ApoC3, -ApoB-48, -Non-HDL-C-Werts. - Anteil der Personen mit Reduktion des Nüchtern-TG-Werts $\geq 40\%$ bzw. $\geq 70\%$. - Anteil der Personen mit Nüchtern-TG-Wert ≤ 880 mg/dl (10 mmol/l) bzw. ≤ 500 mg/dl (5,7 mmol/l). - Ereignisrate bestätigter AP. Explorative Endpunkte - Krankenhausaufenthalte (Hospitalisierungen, Besuche in der Notaufnahme, Anzahl der stationären Tage) - PGIS - PGIC - FCS-SIS (Symptom) - FCS-SIS (Impact) - PROMIS-29+2 - PROMIS Short Form – Pain Interference 8a - PROMIS Short Form – Cognitive Function 4a⁸⁾ Sicherheit Unerwünschte Ereignisse

¹⁾ Entspricht dem Vorgehen des pU in Modul 4.

²⁾ Zur erneuten Prüfung bestimmter Ein- und Ausschlusskriterien (bspw. TG-Werte).

³⁾ War der Nüchtern-TG-Wert < 880 mg/dl, wurden bis zu zwei zusätzliche Tests durchgeführt, wobei jeder einzelne Test zur Qualifizierung verwendet wurde.

- ⁴⁾ Definiert als dokumentierte Diagnose einer AP oder Hospitalisierung oder einem Besuch in der Notaufnahme wegen starker Bauchschmerzen, die mit einer AP übereinstimmen und für die keine alternative Diagnose gestellt wurde.
- ⁵⁾ Patientinnen und Patienten ohne bestätigte Pankreatitis in der Anamnese oder ohne dokumentierte Anamnese innerhalb von 10 Jahren vor dem Screening waren ebenfalls teilnahmeberechtigt, ihre Aufnahme war jedoch ab Protokollamendment 6 (25.10.2021) auf 35 % begrenzt (d. h. ≤ 21 der 60 geplanten Patientinnen und Patienten).
- ⁶⁾ War die Thrombozytenzahl $< 100.000/\text{mm}^3$, können bis zu zwei zusätzliche Tests durchgeführt worden sein, um sich zu qualifizieren.
- ⁷⁾ Gemäß Angaben des pU wurde der Datenschnitt vom 14.07.2023 aufgrund eines Fehlers bei der Anzahl der gemeldeten bestätigten AP-Ereignisse im Olezarsen-80mg-Arm um ein Ereignis korrigiert. Nur die Ergebnisse für den Endpunkt „Bestätigte akute Pankreatitis“ stammen daher vom Datenschnitt am 30.11.2023.
- ⁸⁾ Der Endpunkt wird vom pU im vorliegenden Dossier nicht berücksichtigt. Dieser wurde erst mit Protokollamendment 2 (20.10.2020) eingeführt. Da die Rücklaufquoten bereits zu Baseline in mindestens einem Behandlungsarm unter 70 % lagen, wird auf eine weitere Beschreibung des Endpunkts in der Nutzenbewertung verzichtet.

Abkürzungen: ALT: Alanin-Aminotransferase; AP: Akute Pankreatitis; Apo: Apolipoprotein; AST: Aspartat-Aminotransferase; CKD-EPI: Chronic Kidney Disease Epidemiology Collaboration; eGFR: geschätzte glomeruläre Filtrationsrate; FCS: Familiäres Chylomikronämie-Syndrom; FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; GPD1: Glycerin-3-Phosphat-Dehydrogenase 1; GPIHBP1: Glycosylphosphatidylinositol-anchored High-Density Lipoprotein-Binding Protein 1; LMF1: Lipase-Maturation-Faktor 1; LPL: Lipoproteinlipase; Non-HDL-C: Non-High Density Lipoprotein Cholesterol; PGIC: Patient Global Impression of Change; PGIS: Patient Global Impression of Severity; PROMIS: Patient-Reported Outcomes Measurement Information System; pU: pharmazeutischer Unternehmer; siRNA: small interfering RNA; TG: Triglyzerid; UACR: Urine Albumin Creatinin Ratio; ULN: Upper Limit of Normal; UPCR: Urine Protein Creatinin Ratio.

Abbildung zum Studienverlauf

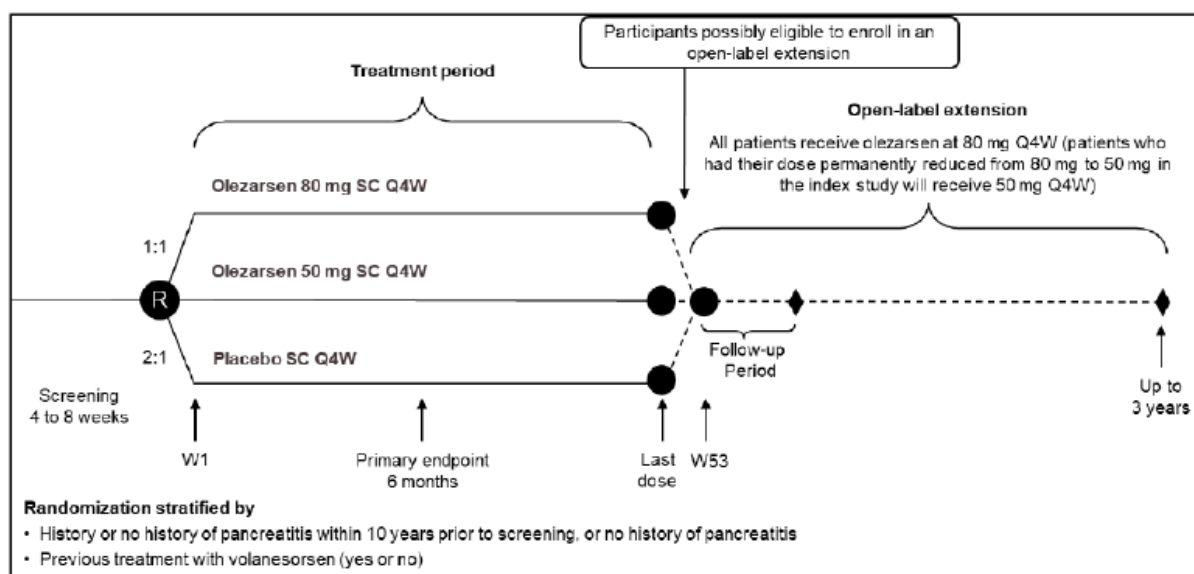


Abbildung 1: Schematische Darstellung des Studienablaufs der Studie Balance [22]

Protokolländerungen

Es wurden insgesamt neun Änderungen des Originalprotokolls vom 21.05.2020 vorgenommen. Amendment 1 vom 07.08.2020 und Amendment 2 vom 30.10.2020 fanden zeitlich vor Studienbeginn statt (Screening der ersten Person am 18.11.2020). Die wesentlichsten Änderungen nach Studienbeginn sind der nachfolgenden Tabelle 3 zu entnehmen.

Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie Balance

Amendment	Wesentliche Änderungen
Version 6 vom 25.10.2021 (Anzahl bis dahin eingeschlossene Personen: k. A.)	Erhöhung der Obergrenze für die Aufnahme von Patientinnen und Patienten ohne dokumentierte Pankreatitis in der Anamnese von 20 % auf 35 %.
Version 7 vom 26.10.2022 (Anzahl bis dahin eingeschlossene Personen: k. A.)	Für die SUE sieht die Definition einer Hospitalisierung nicht länger einen Mindestzeitraum in einer medizinischen Einrichtung von 24 Stunden vor.

Abkürzungen: k. A.: keine Angabe; SUE: Schwerwiegendes unerwünschtes Ereignis.

Charakterisierung der Intervention

Tabelle 4: Charakterisierung der Intervention der Studie Balance (Kohorte B)

Intervention	Kontrolle
Screeningphase (4 bis max.8 Wochen)	
Mind. 2-wöchige Diätstabilisierung auf ≤ 20 mg Fett pro Tag, welche während des gesamten Studienverlaufs beibehalten werden sollte.	
Behandlungsphase (53 Wochen)	
Olezarsen Dosierung: 80 mg (0,8 ml) subkutan einmal monatlich von Woche 1 bis 49.	Placebo (gepoolt) Dosierung: 0,5 ml oder 0,8 ml subkutan ¹⁾ einmal monatlich von Woche 1 bis 49.
Nicht erlaubte Begleitmedikation • Plasmapherese. • Experimentelle Wirkstoffe bzw. Dossierungen existierender Wirkstoffe zur Behandlung von Hypertriglyceridämie. • Initiierung oder Dosisanpassung der erlaubten Begleitmedikation (siehe unten). Erlaubte Begleitmedikation • Beibehaltung der Dossierung von GLP-1-Agonisten oder systemischen Kortikosteroiden. • Beibehaltung der Dossierung²⁾ von Nahrungsergänzungsmitteln bzw. Medikamenten mit lipidverändernder Wirkung (z. B. Statine, Omega-3-Fettsäuren), Antidiabetika, Antihypertonika, orale Antikoagulanzen (z. B. Warfarin), Tamoxifen, Östrogenen oder Progestine.	

¹⁾ Riboflavin wurde zugesetzt zur farblichen Übereinstimmung mit dem Verum.

²⁾ Wenn Dosierung ≥ 4 Wochen vor dem Screening stabil und keine Anpassung während des Studienverlaufs zu erwarten war.

Abkürzungen: GLP-1: Glucagon-like Peptide-1 (Glucagon-like Peptide 1).

2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Alle Endpunkte der eingeschlossenen Studie (siehe Tabelle 2) wurden einer Prüfung unterzogen. Endpunkte, die in der nachfolgenden Tabelle nicht gelistet sind, wurden weder vom pU noch im Rahmen der Nutzenbewertung als patientenrelevant bzw. bewertungsrelevant eingestuft. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie Balance

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Todesfälle ¹⁾	Mortalität	Nein	Ja
Nüchtern-Triglyzeridwert ²⁾	Morbidität	Ja	Ergänzend
Bestätigte akute Pankreatitis		Ja	Ja
Krankenhausaufenthalte • Hospitalisierungen • Besuche der Notaufnahme • Anzahl der stationären Tage		Ja Ja Ja	Nein ³⁾ Nein Nein
PGIS		Ja	Ergänzend ⁴⁾
PGIC		Ja	Nein ³⁾
FCS-SIS (Symptom)		Ja	Ja
FCS-SIS (Impact)		Ja	Ja
PROMIS-29+2	Gesundheitsbezogene Lebensqualität	Ja	Ja
PROMIS Short Form – Pain Interference 8a ⁵⁾		Nein ⁶⁾	Ergänzend
Unerwünschte Ereignisse ⁷⁾	Sicherheit	Ja	Ja

¹⁾Todesfälle wurden im Rahmen der Sicherheit erhoben.

²⁾ Primärer Endpunkt.

³⁾ Die Ergebnisse werden aus Transparenzgründen im Anhang dargestellt (siehe Kapitel 2.3.2).

⁴⁾ Es liegt lediglich eine Responderanalyse zur Verbesserung vor. Bevorzugte Responderanalysen zur Verschlechterung konnten nicht identifiziert werden. Daher erfolgt eine ergänzende Darstellung des Endpunkts und der Responderanalysen zur Verbesserung (siehe Kapitel 2.3.2 und 2.4).

⁵⁾ Der PROMIS Short Form – Pain Interference 8a wird in der vorliegenden Nutzenbewertung in der Endpunktkategorie „Gesundheitsbezogene Lebensqualität“ ergänzend dargestellt. Die gesundheitsbezogene Lebensqualität wird durch das Instrument nicht vollumfänglich erfasst. Eine Zuordnung zur Endpunktkategorie „Morbidität“ ist daher auch denkbar, wird in der vorliegenden Endpunktkonstellation mit dem PROMIS-29+2 jedoch nicht als sinnvoll erachtet (siehe Kapitel 2.3.3).

⁶⁾ Aufgrund einer teilweisen Mehrfacherfassung (PROMIS-29+2) wurde der Endpunkt vom pU nicht im Dossier berücksichtigt.

⁷⁾ Für den Endpunkt „MACE“ ist die Zuordnung zur Endpunktkategorie „Sicherheit“ unklar. Aufgrund der Erhebung in der Studie Balance im Rahmen der Sicherheit wird diese für die Nutzenbewertung beibehalten (siehe Kapitel 2.3.4).

Abkürzungen: FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; MACE: Major Adverse Cardiovascular Event; PGIC: Patient Global Impression of Change; PGIS: Patient Global Impression of Severity; pU: pharmazeutischer Unternehmer; PROMIS: Patient-Reported Outcomes Measurement Information System.

2.3.1 Mortalität

Todesfälle

Der Endpunkt „Todesfälle“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die während der Studie Balance aufgetretenen Todesfälle wurden im Rahmen der Sicherheit erfasst.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz und Validität

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen. Der Endpunkt wird als valide erachtet.

2.3.2 Morbidität

Nüchtern-Triglyzeridwert

Bei dem Endpunkt „Nüchtern-Triglyzeridwert“ handelt es sich um den primären Endpunkt. Dieser wird als nicht unmittelbar patientenrelevant eingeschätzt, ist jedoch ein klinisch relevanter Parameter, der unter anderem der Therapiesteuerung dienen kann. Der Endpunkt wird ergänzend dargestellt.

Operationalisierung

Beschreibung

Die Bestimmung der Triglyzerid (TG)-Werte im Blut fand in der Studie Balance im Rahmen der Laboruntersuchung statt. Die Blutchemie- und Lipidproben wurden nach einer Nüchternheit von mindestens 10 Stunden und vorzugsweise nicht mehr als 12 Stunden entnommen und von einem Zentrallabor analysiert. Patientinnen und Patienten war es erlaubt, während ihrer Nüchternheit Wasser zu trinken.

Präspezifiziert war die Erfassung der prozentualen Veränderung der Nüchtern-TG-Werte nach Monat 6 bzw. Monat 12 im Vergleich zu Baseline. Baseline war definiert als Mittelwert aus der letzten nicht fehlenden Messung während der Screening-Periode und aus der Messung vor Verabreichung der ersten Studienmedikation an Tag 1. Für die Auswertung zum jeweiligen Analysezeitpunkt wird ein Mittelwert über zwei bzw. drei Erhebungszeitpunkte gebildet (bspw. Woche 51 und 53 für Monat 12). Zusätzlich wurden deskriptive Verlaufsdaten dargestellt.

Bewertung

Die Operationalisierung ist nachvollziehbar.

In der vorliegenden Nutzenbewertung werden die Auswertungen zu Monat 12 dargestellt.

Patientenrelevanz

Bei dem TG-Wert handelt es sich um einen Laborparameter, der als nicht unmittelbar patientenrelevant eingeschätzt wird. Geeignete Studien zur Surrogatvalidierung für einen patientenrelevanten Endpunkt legt der pU nicht vor. Der TG-Wert stellt im vorliegenden Anwendungsgebiet jedoch ein zentrales diagnostisches Charakteristikum (Grenzwert > 885 mg/dl bzw. 10 mmol/l) und therapeutisches Zielkriterium dar [20, 21]. Der TG-Wert kann zudem für die Therapiesteuerung relevant sein [2]. Die Erkrankung wird häufig erst bei dauerhaft erhöhten TG-Werten von > 1.000 mg/dl klinisch symptomatisch, wobei eine akute Pankreatitis die schwerwiegendste

Komplikation erhöhter TG-Werte darstellt [20]. Dieses Risiko steigt weiter an, wenn die TG-Werte $> 2.000 \text{ mg/dl}$ liegen [6]. Entsprechend schreibt die „European Medicines Agency“ (EMA), dass das vorrangige Therapieziel bei FCS darin bestehe, die TG-Werte im Blut ausreichend zu senken, um das Risiko einer akuten Pankreatitis zu reduzieren [13]. Die akute Pankreatitis wird in der vorliegenden Nutzenbewertung bereits separat als Morbiditätsendpunkt erfasst (siehe unten).

Zusammenfassend wird der Endpunkt „Nüchtern-Triglyzeridwert“ in der vorliegenden Operationalisierung nicht als unmittelbar patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen. Da der primäre Endpunkt in der Studie Balance jedoch ein klinisch relevanter Parameter für die Erkrankung FCS darstellt, wird der Endpunkt ergänzend dargestellt.

Validität

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

Bestätigte akute Pankreatitis

Der Endpunkt „Bestätigte akute Pankreatitis“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Alle während der Studie auftretenden UE und SUE, die mit einer akuten Pankreatitis (AP) übereinstimmten und für die eine Notaufnahme aufgesucht wurde bzw. eine Hospitalisierung stattgefunden hat, sollten von einem verblindeten, unabhängigen Komitee attestiert werden.

Alle Patientinnen und Patienten erhielten eine Informationskarte, die bei Verdacht auf Pankreatitis der behandelnden Ärztin / dem behandelnden Arzt vorgelegt werden musste. Die Karte weist auf die Teilnahme an einer klinischen Studie hin und empfiehlt die Bestimmung von Serumlipase und/oder -amylase zum nachfolgenden Bestätigen des Ereignisses gemäß der modifizierten Atlanta-Klassifizierung (siehe unten). Lag mindestens einer der Werte unter dem 3-fachen des oberen Normwerts (Upper Limit of Normal (ULN)), sollte zur Bestätigung der Diagnose gemäß modifizierter Atlanta-Klassifizierung eine bildgebende Untersuchung in Erwägung gezogen werden.

Die Beurteilung erfolgte gemäß der modifizierten Atlanta-Klassifizierung [5] für AP und gemäß zusätzlicher Kriterien, die in einer studienspezifischen Charta des „Acute Pancreatitis Adjudication Committee“ (PAC) präspezifiziert wurden. Das PAC sah folgende Unterteilung vor:

- Dokumentierte AP: Eine Pankreatitis, die den modifizierten Atlanta-Diagnosekriterien für AP entspricht; diese erfordert mindestens zwei der folgenden Kriterien:
 - Bauchschmerzen, die stark auf eine AP hindeuten (akutes Auftreten eines anhaltenden, starken Oberbauchschmerzes, der häufig in den Rücken ausstrahlt);
 - Aktivität der Serumlipase (oder Amylase) $\geq 3 \times \text{ULN}$;
 - charakteristische Befunde einer AP bei einer kontrastverstärkten Computertomographie, einer Magnetresonanztomographie oder eines transabdominalen Ultraschalls.
- Wahrscheinliche AP: Klinische Pankreatitis mit Dokumentation typischer klinischer Merkmale und eines abnormalen Werts für Amylase/Lipase $< 3 \times \text{ULN}$; alle folgenden Merkmale müssen erfüllt sein:
 - Abdominaler Schmerz, der vereinbar mit einer AP ist (akuter Beginn eines persistierenden, schwerwiegenden, epigastrischen Schmerzes, der oft in den Rücken ausstrahlt);
 - typische unterstützende Behandlung einer Pankreatitis (z. B. keine orale Nahrungs- oder Flüssigkeitsaufnahme, intravenöse Verabreichung von Flüssigkeit, Verschreibung von Analgetika (stärker als Paracetamol allein));
 - Werte für Amylase oder Lipase sind erhöht, aber $\leq 3 \times \text{ULN}$;
 - radiologische Untersuchungen sind nicht durchgeführt oder nicht diagnostisch verwertbar;
 - Diagnose einer AP bei der Entlassung durch die Ärztin / den Arzt, ohne Vorschlag alternativer Diagnose.

- Mögliche AP: Dokumentation typischer klinischer Merkmale einer Pankreatitis bei einer Person mit einer Krankengeschichte von Pankreatitiden; alle folgenden Merkmale müssen erfüllt sein:
 - Vorhergehende medizinische Diagnose einer AP;
 - abdominaler Schmerz, der vereinbar mit einer AP ist (akuter Beginn eines persistierenden, schwerwiegenden, epigastrischen Schmerzes, der oft in den Rücken ausstrahlt);
 - typische unterstützende Behandlung einer Pankreatitis (z. B. keine orale Nahrungs- oder Flüssigkeitsaufnahme, intravenöse Verabreichung von Flüssigkeit, Verschreibung von Analgetika (stärker als Paracetamol allein));
 - Werte für die Serumamylase oder -lipase entweder normal, nicht bestimmt oder fehlend;
 - radiologische Untersuchungen sind nicht durchgeführt oder nicht diagnostisch verwertbar;
 - Diagnose einer AP bei der Entlassung durch die Ärztin / den Arzt, ohne Vorschlag alternativer Diagnose.
- Nicht bestätigte AP: Unzureichende Dokumentation, dass eine AP vorlag.
- Keine Diagnose einer AP: Schlussfolgerung auf Grundlage zureichender Dokumentation, dass keine AP vorlag.

Die Beilegung von Unstimmigkeiten zwischen den Gutachtern des Komitees sollte gemäß eines im PAC präspezifizierten Verfahrens erfolgen. Erzielten die Gutachter keinen Konsens, traf der Vorsitzende des PAC die endgültige Entscheidung. Für jedes Ereignis wurden die Entscheidungen der Bewerter und des Komitees festgehalten und berichtet, in wie vielen Fällen die Einschätzungen der einzelnen Bewerter von der endgültigen Entscheidung abwich.

Für die Auswertung galten dokumentierte, wahrscheinliche und mögliche AP gemäß der Einschätzung des Komitees als auswertungsrelevante bzw. bestätigte AP.

Ausgewertet wurde

- Ereignisrate (präspezifiziert) und Inzidenz (post hoc) bis Monat 12 (definiert als Tag 379 (± 14) oder als der letzte Kontakt mit den Patienten innerhalb der Studie, je nachdem, was früher eintritt) sowie
- Zeit bis zur ersten bestätigten AP (post hoc).

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Zwar wurden den Patientinnen und Patienten Informationskarten mitgegeben, welche der Verbesserung der Erhebungswahrscheinlichkeit von Serumlipase- und/oder -amylase dienen kann. Es geht jedoch nicht eindeutig aus den Unterlagen hervor, wie sichergestellt wurde, dass dem Komitee ausreichend Informationen zur Bewertung der AP anhand aller spezifizierten Kriterien vorlagen.

Es werden die Ereignisrate und Inzidenz bis Monat 12 in der Nutzenbewertung dargestellt.

Patientenrelevanz

AP stellt eine der wichtigsten Komplikationen des FCS dar und kann mit starken und anhaltenden Schmerzen sowie mit weiteren Symptomen wie Übelkeit und Erbrechen einhergehen [24]. Der Endpunkt stellt einen patientenrelevanten Endpunkt entsprechend § 2 Satz 3 AM-NutzenV dar.

Validität

Die Bestätigung des Endpunkts erfolgte auf Basis eines verblindeten, unabhängigen Expertenkomitees und wird als ausreichend standardisiert eingeschätzt. Es ist unklar, inwieweit die Klassifikation nach den modifizierten Atlanta-Kriterien und den zusätzlichen Kriterien gemäß PAC im deutschen Versorgungskontext als etabliert gilt. Es muss berücksichtigt werden, dass es sich vorliegend um eine multizentrische Studie handelt und dass regionale Unterschiede hinsichtlich der Notwendigkeit und Möglichkeit eines Besuchs in der Notaufnahme bzw. einer Hospitalisierung bestehen können. Eine stratifizierte Randomisierung auf Ebene der Studienzentren oder Länder war in der Studie nicht vorgesehen.

Krankenhausaufenthalte

Die Endpunkte zu Krankenhausaufenthalten werden aufgrund von wesentlichen Unklarheiten zur Operationalisierung nicht zur Nutzenbewertung herangezogen. Der Endpunkt „Hospitalisierung“ wird aus Transparenzgründen im Anhang dargestellt.

Operationalisierung

Beschreibung

Krankenhausaufenthalte wurden während der Studie Balance über folgende präspezifizierte Endpunkte erfasst:

- Hospitalisierungen (aus jeglichen Gründen)
- Besuche in der Notaufnahme
- Anzahl der stationären Tage

Post hoc wurden außerdem AP-assoziierte Hospitalisierungen ausgewertet.

In Modul 4 des Dossiers wird eine Hospitalisierung definiert als eine Einweisung in eine medizinische Einrichtung von mehr als 24 Stunden aus jeglichen Gründen und dabei auf die Studienprotokolle verwiesen. In den Studienprotokollen konnte eine solche Angabe ausschließlich für SUE im Rahmen der Sicherheit und nur bis Inkrafttreten von Amendment 7 des Studienprotokolls (26.10.2022) identifiziert werden. Ab Amendment 7 wurde die Mindestdauer von 24 Stunden aus dem Protokoll entfernt. Für den Morbiditätsendpunkt zu Hospitalisierungen finden sich über die Studienprotokolle hinweg keine Angaben zu einer Mindestdauer.

Die Analysen erfolgen präspezifiziert anhand der Ereignisrate bis Monat 12 (definiert als Tag 379 (± 14) oder als der letzte Kontakt mit den Patienten innerhalb der Studie, je nachdem, was früher eintrat) bzw. der jährlichen Rate der:

- Hospitalisierungen (aus jeglichen Gründen) und AP-assoziierte Hospitalisierungen (post hoc)
- Besuche der Notaufnahme aus jeglichen Gründen
- Stationäre Tage und AP-assoziierte stationäre Tage (post hoc)

Zudem liegen mit Modul 4 des Dossiers Post-hoc-Auswertungen bezüglich der Inzidenz (Personen mit ≥ 1 Ereignis) der Hospitalisierung (jegliche und AP-assoziiert) und der Besuche der Notaufnahme bis Monat 12 vor.

Bewertung

Die Operationalisierung der Endpunkte kann nicht nachvollzogen werden. Eingehendere Angaben zur Operationalisierung der Endpunkte zu Krankenhausaufenthalten konnten in den Studienunterlagen nicht identifiziert werden. Ein definiertes Zeitkriterium für den Morbiditätsendpunkt „Hospitalisierungen“ fehlt und es liegen im Dossier dazu widersprüchliche Angaben vor. Auch lassen sich keine Angaben dazu ableiten, inwiefern geplante Eingriffe durch die Endpunkte „Hospitalisierungen“ und „Anzahl der stationären Tage“ miterhoben wurden. Ob „Hospitalisierungen“ stationäre, teil-stationäre und/oder ambulante Krankenhausaufenthalte umfassen konnte, kann auf Grundlage der vorliegenden Informationen nicht abschließend bewertet werden. Für den Endpunkt „Anzahl der stationären Tage“ liegen ebenfalls keine Informationen vor, ob zeitliche Mindestkriterien erfüllt sein mussten.

Für die post hoc durchgeführten Auswertungen „AP-assoziierte Hospitalisierungen“ und „AP-assoziierte stationäre Tage“ liegen ebenfalls keine Kriterien vor. Es ist somit unklar, welche Hospitalisierungen als AP-assoziiert gewertet wurden und welche nicht. Es ist zudem unklar, ob die jeweilige AP ursächlich für die Einweisung in die medizinische Einrichtung sein musste.

Die Endpunkte zu den Krankenhausaufenthalten werden aufgrund unklarer Operationalisierung nicht zur Nutzenbewertung herangezogen. Ergebnisse für den Endpunkt „Hospitalisierungen“ werden aus Transparenzgründen im Anhang dargestellt.

Patientenrelevanz und Validität

Stationäre Hospitalisierungen jeglicher Ursache können grundsätzlich als patientenrelevant eingeschätzt werden. Aufgrund fehlender Informationen und damit einhergehend wesentlicher Unklarheiten zur Operationalisierung, erscheint die abschließende Bewertung der Validität und Patientenrelevanz der Endpunkte zu Krankenhausaufenthalten jedoch nicht möglich.

Grundsätzlich kann darauf hingewiesen werden, dass es sich vorliegend um eine multizentrische Studie handelt und dass regionale Unterschiede hinsichtlich der Notwendigkeit und Möglichkeit eines Besuchs in der Notaufnahme bzw. einer Hospitalisierung bestehen können. Eine stratifizierte Randomisierung auf Ebene der Studienzentren oder Länder war in der Studie nicht vorgesehen. Durch Erhebungen im Rahmen der Sicherheit (Hospitalisierungen aufgrund eines UE) kann zudem eine Mehrfacherfassung vorliegen.

PGIS

Für den Endpunkt „PGIS“ liegen lediglich Responderanalysen zur Verbesserung vor. Eine für die betrachtete Studienpopulation als relevanter eingeschätzte Responderanalyse zur Verschlechterung konnte in den Studienunterlagen nicht identifiziert werden. Der Endpunkt und die Responderanalyse zur Verbesserung werden daher ergänzend dargestellt.

Operationalisierung

Beschreibung

Die vorliegende Version des „Patient Global Impression of Severity“ (PGIS) soll in der Studie Balance die Schwere der patientenberichteten FCS-Symptomatik erheben.

Der PGIS bestand aus einem Item („Please choose the response below that best describes the severity of your Familial Chylomicronemias Syndrome over the past week.“), mit dem der Gesamtschweregrad der eigenen Symptome in der vergangenen Woche bewertet werden sollte. Die Schweregradeinteilung erfolgte anhand einer 5-Punkte-Likert-Skala von 0 („keine Symptome“) bis 4 („sehr schwere Symptome“).

Baseline war definiert als letzter nicht fehlender Wert vor der ersten Gabe der Studienmedikation. Für die Auswertung zum jeweiligen Analysezeitpunkt wird ein Mittelwert über drei Erhebungszeitpunkte zu Monat 6 bzw. zwei Erhebungszeitpunkte (Woche 51 und 53) zu Monat 12 gebildet.

Die Auswertung des PGIS sollte präspezifiziert mittels Responderanalyse zum Anteil der Personen mit einer Verbesserung (definiert als Veränderung im Vergleich zu Baseline < 0) bzw. keiner Verbesserung (definiert als Veränderung im Vergleich zu Baseline ≥ 0) zu Monat 6 bzw. 12 gegenüber Baseline erfolgen. Responderanalysen zur Verschlechterung liegen nicht vor.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar, es ergeben sich jedoch Einschränkungen hinsichtlich des Auswertungsvorgehens (siehe Patientenrelevanz und Validität). Es fehlen zudem Angaben, in welcher Reihenfolge die Instrumente der „Patient-Reported Outcomes“ (PRO) administriert wurden. Es ist daher unklar, ob und welchen Einfluss die möglicherweise zuvor erhobenen PRO-Instrumente auf das Ausfüllen des PGIS hatten.

Für die Nutzenbewertung werden die Responderanalysen zur Verbesserung zu Monat 12 ergänzend dargestellt (siehe Patientenrelevanz und Validität).

Patientenrelevanz und Validität

Die Einschätzung der Patientinnen und Patienten zum Schweregrad der jeweils eigenen erkrankungsspezifischen Symptomatik wird als patientenrelevant bewertet und eine Erhebung per PGIS wird vorliegend als geeignet eingeschätzt.

In der Studie Balance wird die FCS-Symptomatik über mehrere Erhebungsinstrumente (PGIS, PGIC und FCS-SIS) erfasst. Es liegen in Teilen inhaltliche Überschneidungen vor, sodass eine Mehrfacherfassung bei der Interpretation der Ergebnisse zu berücksichtigen ist. Während die FCS-SIS spezifische Symptome erhebt, erfragt der PGIS mit nur einem Item die allgemeine Syndrom-Ausprägung. In welchem Ausmaß der PGIS dadurch im Vergleich zum FCS-SIS wesentliche zusätzliche Symptome erfasst, ist unklar. Da zum FCS-SIS jedoch u. a. Bodeneffekte beschrieben werden und es sich bei FCS um ein heterogenes Erkrankungsbild handelt, kann der PGIS durch eine breite Fragestellung im Vergleich zur FCS-SIS mögliche zusätzliche Informationen generieren.

Einschränkungen ergeben sich jedoch hinsichtlich des Auswertungsvorgehens:

- Responderanalysen zur Verschlechterung, welche in der vorliegenden Studienpopulation bevorzugt werden (siehe Kapitel 2.4), liegen nicht vor (siehe auch Spiegelstrich 3).
- Die Responderanalyse zu keiner Verbesserung (definiert als Veränderung im Vergleich zu Baseline ≥ 0) entspricht nicht einem Responsekriterium von mindestens 15 % der Skalenspannweite und wird nicht zur Nutzenbewertung herangezogen.
- Zur Responderanalyse zur Verbesserung (definiert als Veränderung im Vergleich zu Baseline < 0), die in der vorliegenden Studienpopulation als nachrangig bewertet wird (siehe unten), liegen Einschränkungen vor: Zwar übersteigt eine mindestens ganzzahlige Änderung auf der Skala des PGIS das geforderte Responsekriterium von mindestens 15 %, jedoch sieht die vorliegende Operationalisierung eine Mittelwertbildung zweier Erhebungszeitpunkte vor. Daraus resultiert die Möglichkeit zu halbstufigen Verbesserungswerten bzw. einmaligen Verbesserungen (zu Woche 51 oder zu Woche 53). Dies ermöglicht punktuelle Verbesserungen zu nur einem Zeitpunkt gegenüber Baseline bzw. eine Verbesserung um 0,5 Punkte. Eine Änderung in dieser Höhe liegt unter einer akzeptablen Responseschwelle von mindestens 15 % der Skalenspannweite. Auswertungen zur bestätigten Response (zu Woche 51 und Woche 53) oder zur (einmaligen) Response zu Woche 53, jeweils mit einer Response von mindestens 15 % der Skalenspannweite, werden für die Nutzenbewertung bevorzugt. Dies gilt auch für Responderanalysen zur Verschlechterung. In der vorliegenden Datensituation mit deskriptiven Angaben zu Woche 51 und 53 stimmen die Responder zu Woche 53 mit jenen zu Woche 51 oder Woche 53 überein. Damit können die Auswertungen als Responder zu Woche 53 interpretiert werden und werden in der Nutzenbewertung ergänzend dargestellt.

In der Studienpopulation mit FCS der Studie Balance mit heterogenem Erkrankungsbild kann für einige Personen die Symptomverbesserung ein Therapieziel darstellen. Etwa ein Drittel der Studienpopulation berichtet zu Baseline jedoch keinerlei Symptome im PGIS, sodass nicht für alle Patientinnen und Patienten eine Verbesserung ein mögliches Therapieziel darstellt (siehe auch Kapitel 2.4). Daher werden für diese Studienpopulation Responderanalysen zur Verschlechterung bevorzugt gegenüber Analysen zur Verschlechterung dargestellt. Für den PGIS wurden jedoch keine adäquaten Responderanalysen zur Verschlechterung eingereicht.

In der Gesamtschau liegt lediglich eine Responderanalyse zur Verbesserung zu Monat 12 (bzw. zu Woche 53) vor, welche jedoch für die Nutzenbewertung gegenüber einer Responderanalyse zur Verschlechterung als nachrangig bewertet wird. Demnach erfolgt eine ergänzende Darstellung (siehe Kapitel 2.4).

Zudem ist bei der Ergebnisinterpretation die Mehrfacherfassung mit dem FCI-SIS zu berücksichtigen.

PGIC

Der Endpunkt „PGIC“ wird in der Nutzenbewertung aufgrund einer vorliegenden Mehrfacherfassung und Einschränkungen in der Aussagekraft der Erhebungen durch den PGIC in der Studie Balance nicht berücksichtigt. Aus Transparenzgründen erfolgt eine Darstellung im Anhang.

Operationalisierung

Beschreibung

Die vorliegende Version des „Patient Global Impression of Change“ (PGIC) sollte in der Studie Balance die patientenberichtete Veränderung der FCS-Symptomatik seit Behandlungsbeginn erheben. Der PGIC bestand aus einem Item („Please choose the response below that best describes the overall change in your Familial Chylomicronemias Syndrome since you started taking the study medication.“), mit dem die Gesamtveränderung der eigenen Symptome im Vergleich zum Start der Studie bewertet werden sollte. Die Selbsteinschätzung erfolgte auf einer 5-Punkte-Likert-Skala von 1 („viel besser“) bis 5 („viel schlechter“).

Die Auswertung des PGIC erfolgte präspezifiziert mittels Responderanalyse zum Anteil der Personen mit

- einer Verbesserung (definiert als Wert $< 2,5$) und
- keiner Verbesserung (definiert als Wert $\geq 2,5$).

Responderanalysen zur Verschlechterung liegen nicht vor. Für die Auswertung zum jeweiligen Analysezeitpunkt wird ein Mittelwert über drei Erhebungszeitpunkte zu Monat 6 bzw. zwei Erhebungszeitpunkte (Woche 51 und 53) zu Monat 12 gebildet.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Zu berücksichtigen sind die Unklarheiten zur Reihenfolge der PRO-Instrumente und möglicher Auswirkungen auf die Erhebung (siehe Abschnitt „PGIS“).

Aufgrund stark differierender Rückläufe zu Monat 12 wird im Anhang die Responderanalyse zur Verbesserung zu Monat 6 dargestellt.

Patientenrelevanz und Validität

Einschätzungen der Patientinnen und Patienten zur Veränderung der jeweils eigenen erkrankungsspezifischen Symptomatik werden grundsätzlich als patientenrelevant eingeschätzt.

Methodisch werden dabei i. d. R. Erhebungsinstrumente bevorzugt, die sowohl zu Baseline- als auch zu Post-Baseline-Studienvisiten erhoben werden. Veränderungseinschätzungen gegenüber Baseline ohne eigenständige Baseline-Erhebung können demgegenüber unter Umständen weniger aussagekräftige Ergebnisse liefern. Mögliche Einflussfaktoren auf die Aussagekraft der Erhebungen von Instrumenten ohne Baseline-Erhebungen (wie dem PGIC) sind u. a. Art und Schwere der Symptomatik, Länge der Erinnerungsintervalle und systematische Berücksichtigung geeigneter anderweitiger Baseline-Erhebungen.

In der Studie Balance werden die erkrankungsspezifische Symptomatik und diesbezügliche Veränderungen bis Woche 53 bereits mittels FCS-SIS und PGIS erhoben. Es wird davon ausgegangen, dass der PGIC darüber hinaus keine aussagekräftigen zusätzlichen Informationen erhebt.

In der Gesamtschau wird der PGIC daher nicht zur Nutzenbewertung herangezogen. Es erfolgt aus Transparenzgründen jedoch eine Darstellung der Ergebnisse zur Verbesserung im Anhang. Responderanalysen zur Verschlechterung, welche in der vorliegenden Studienpopulation grundsätzlich bevorzugt werden würden (siehe Kapitel 2.4), fehlen.

FCS-SIS (Symptom)

Der Endpunkt „FCS-SIS (Symptom)“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung*Beschreibung*

Die „Familial Chylomicronemia Syndrome – Symptom and Impact Scale“ (FCS-SIS) soll in der Studie Balance die Schwere der patientenberichteten FCS-Symptomatik erheben. Der Fragebogen wird aufgeteilt in die Domänen „Symptom“ und „Impact“, wobei die Symptom-Domäne der Endpunktkategorie „Morbidität“ zuzuordnen ist und die Impact-Domäne der „Lebensqualität“.

Die Symptom-Domäne enthält die folgenden Items:

- Abdominale Schmerzen
- Körperliche Ermüdung
- Probleme beim Denken
- Diarrhö

Die Schwere jedes Symptoms wurde mit einem Bezugszeitraum der letzten 24 Stunden täglich auf einer 11-stufigen Skala von 0 („kein Symptom“) bis 10 („schlimmstmögliches Symptom“) erfasst.

Der (tägliche) Symptom-Gesamtwert wird anhand des Mittelwerts der 4 Symptom-Items berechnet. Wenn 2 oder mehr Items fehlten, wurde der Gesamtwert für den jeweiligen Tag als fehlend betrachtet. Der ermittelte Symptom-Gesamtwert wurde anschließend für die entsprechende Studienwoche gemittelt und zur Auswertung verwendet. Wenn Patientinnen und Patienten in einer Woche weniger als 4 tägliche Bewertungen des FCS-SIS vornahmen, wurde der Fragebogen als fehlend für die Woche betrachtet.

Wie laut SAP präspezifiziert, wurde die Veränderung des Symptom-Gesamtwerts nach 6 bzw. 12 Monaten im Vergleich zu Baseline untersucht. Baseline wird definiert als der Durchschnitt während der Screening-Periode bis einschließlich Tag 1 der Studienmedikation. Für die Auswertung zum jeweiligen Erhebungszeitpunkt wird ein Mittelwert über je 4 Wochen gebildet (bspw. Woche 49 bis 53 für Monat 12).

Darüber hinaus liegen mit Modul 4 des Dossiers post hoc durchgeführte Responderanalysen hinsichtlich der Verbesserung bzw. Verschlechterung des Symptom-Gesamtwerts zu Monat 6 bzw. 12 mit einem Responsekriterium von 15 % der Skalenspannweite vor.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Zu berücksichtigen sind die Unklarheiten zur Reihenfolge der PRO-Instrumente und möglicher Auswirkungen auf die Erhebung (siehe Abschnitt „PGIS“).

Es wird die Responderanalyse zur Verschlechterung zu Monat 12 in der Nutzenbewertung dargestellt. Die Responderanalyse zur Verbesserung zu Monat 12 wird ergänzend dargestellt. Für Hintergründe siehe Abschnitt „PGIS“ und Kapitel 2.4.

Patientenrelevanz

Der Endpunkt wird unter Berücksichtigung der vorliegenden Operationalisierung als patienten-relevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen.

Validität

Die Entwicklung der FCS-SIS erfolgte im Rahmen einer qualitativen Studie mittels Literaturrecherche sowie zwei qualitativer Interviews mit 10 an FCS erkrankten Personen [11]. Die durchgeführten qualitativen Konzeptfindungs- und kognitiven Nachbesprechungsinterviews resultierten in einem Fragebogen mit 17 Items: 4 Items zur Bewertung der Symptome (Symptom-Domäne) und 13 Items zur Bewertung der Auswirkungen (Impact-Domäne) des FCS. Darüber hinaus

stimmten die in den qualitativen Interviews berichteten Symptome und Auswirkungen des FCS größtenteils mit den Ergebnissen einer zuvor durchgeführten Querschnittsstudie zur Krankheitslast des FCS überein [12]. Auch die Verständlichkeit und einfache Ausfüllbarkeit des Instruments konnte durch die Interviews mit Betroffenen bestätigt werden. Basierend auf dem Entwicklungsprozess wird die Inhaltsvalidität des Fragebogens angenommen.

In einer Validierungsstudie wurden die psychometrischen Gütekriterien der FCS-SIS anhand von 33 Patientinnen und Patienten mit FCS im Alter von 24 bis 72 Jahren überprüft (Brown et al. (2024) [7]). Für die Validierung der Symptom-Domäne wurde sowohl eine Version der FCS-SIS mit einem Referenzzeitraum von 24 Stunden als auch von 7 Tagen berücksichtigt. Laut der Publikation beträgt die Korrelation zwischen den Items der Symptom-Domäne (24 Stunden) $r = 0,32$ („Probleme beim Denken“ und „Diarrhö“) bis $r = 0,78$ („Körperliche Ermüdung“ und „Probleme beim Denken“). Es wird beschrieben, dass manche Items der Symptom-Domäne teils höher mit den Items der Impact-Domäne korrelieren als mit Items der Symptom-Domäne. Dies erscheint jedoch nachvollziehbar, da die in der Symptom-Domäne erfassten Symptome als eher distinkt zu bewerten sind als die Items der Impact-Domäne, welche teils einen inhaltlich nachvollziehbaren Bezug zu einzelnen Symptomen aufweisen können (bspw. „Bauchschmerz“ und „Sorge um Pankreatitisattacke“). Die Inter-Item-Korrelation der Impact-Domäne beträgt zumeist über $r = 0,80$. Drei von vier Items in der Symptom-Domäne (24 Stunden) weisen laut der Validierungsstudie Bodeneffekte und ein Item der Impact-Domäne einen Deckeneffekt auf. Laut der Autoren nehmen die Bodeneffekte der Symptom-Domäne unter Betrachtung des 7-tägigen Bezugszeitraums ab. Die Bodeneffekte führen Bown et al. vor allem auf die kleine Stichprobe der Querschnittsstudie und den episodischen Charakter der FCS-Symptome zurück. Im vorliegenden Fall wird aufgrund der Mittelung der täglich erhobenen Symptom-Domäne auf jeweils 7 Tage angenommen, dass mögliche Bodeneffekte entsprechend der Ergebnisse der Validierungsstudie reduziert werden.

Die Symptom-Domäne der FCS-SIS weist einen guten (Cronbach's $\alpha = 0,84$) bzw. hinsichtlich der Impact-Domäne einen exzellenten (Cronbach's $\alpha = 0,95$) Grad an interner Konsistenz auf. Des Weiteren wird in der Validierungsstudie die Konstruktvalidität anhand weiterer PRO-Instrumente („Gastrointestinal Symptom Rating Scale“ (GSRS), PROMIS-29 v2.0 und PGIS) sowie der AP der vorangegangenen 12 Monate überprüft. Dabei korrelieren für den 7-tägigen Bezugszeitraum sowohl die einzelnen Symptom-Items der FCS-SIS stark mit den Symptom-Items der PRO-Instrumente, die dieselben Konstrukte messen ($r = 0,87$ – $0,93$), als auch der Symptom-Gesamtwert mit dem jeweiligen Gesamtwert der PRO-Instrumente (PGIS: $r = 0,81$; GSRS: $r = 0,73$; PROMIS-29 v2.0 Global Health-Physical: $r = -0,80$ bzw. Global Health-Mental: $r = -0,73$). Ähnlich starke Korrelationen werden für den Impact-Gesamtwert und dem PROMIS-29 v2.0 hinsichtlich „Ängste / Sorgen“ ($r = 0,82$), „Depression“ ($r = 0,70$) sowie „Fähigkeit zur Teilnahme an sozialen Aktivitäten“ ($r = -0,91$) berichtet. Es konnte außerdem gezeigt werden, dass die Anzahl der Patientinnen und Patienten mit einer AP im vorangegangenen Jahr stark mit dem berichteten Schweregrad der abdominalen Schmerzen (7-tägiger Bezugszeitraum) korrelierte ($p = 0,77$). Limitierend sind die kleine Stichprobenzahl der Validierungsstudie ($n = 33$), welche auf die Seltenheit der Erkrankung zurückzuführen ist (eingeschränkte externe Validität), die in der Validierungsstudie berichteten Bodeneffekte hinsichtlich der Symptom-Items (eingeschränkte Sensitivität, insbesondere für den Bezugszeitraum von 24 Stunden) sowie eine fehlende Faktorenanalyse hinsichtlich der Strukturvalidität.

Zusammenfassend liegen zwar Einschränkungen zu einzelnen Kriterien der Validität der FCS-SIS vor. In der Gesamtschau wird der Fragebogen in der vorliegenden Nutzenbewertung jedoch als weitestgehend valide eingeschätzt und zur Nutzenbewertung herangezogen.

2.3.3 Lebensqualität

FCS-SIS (Impact)

Der Endpunkt „FCS-SIS (Impact)“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die von den Patientinnen und Patienten berichtete FCS-bezogenen Auswirkungen auf die gesundheitsbezogene Lebensqualität sollten in der Studie Balance monatlich anhand der Impact-Domäne der FCS-SIS erhoben werden. Die Impact-Domäne umfasst 13 Items, welche die gegenwärtige Lebensqualität auf einer 5-stufigen Skala von 0 („nie“) bis 4 („immer“) erfasst. Der Impact-Gesamtwert wurde als Durchschnitt der nicht fehlenden 13 Itemscores berechnet. Wenn mehr als 50 % der Items nicht beantwortet wurden, wurde der Impact-Gesamtwert für die Visite als fehlend gewertet. Wie präspezifiziert, wurde die Veränderung des Impact-Gesamtwerts nach 6 bzw. 12 Monaten im Vergleich zu Baseline untersucht. Baseline war definiert als letzter nicht fehlender Wert vor der ersten Gabe der Studienmedikation. Für die Auswertung zum jeweiligen Erhebungszeitpunkt wird ein Mittelwert über je 4 Wochen gebildet (bspw. Woche 49 bis 53 für Monat 12). Darüber hinaus liegen mit Modul 4 des Dossiers post hoc durchgeführte Responderanalysen hinsichtlich der Verbesserung bzw. Verschlechterung des Impact-Gesamtwerts zu Monat 6 bzw. 12 mit einem Responsekriterium von 15 % der Skalenspannweite vor.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Zu berücksichtigen sind die Unklarheiten zur Reihenfolge der PRO-Instrumente und möglicher Auswirkungen auf die Erhebung (siehe Abschnitt „PGIS“ in Kapitel 2.3.2).

Es wird die Responderanalyse zur Verschlechterung zu Monat 12 in der Nutzenbewertung dargestellt. Die Responderanalyse zur Verbesserung zu Monat 12 wird ergänzend dargestellt. Für Hintergründe siehe Abschnitt „PGIS“ (Kapitel 2.3.2) und Kapitel 2.4.

Patientenrelevanz und Validität

Der Endpunkt wird unter Berücksichtigung der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen.

Der Fragebogen enthält Items zur physischen, psychischen und sozialen Dimension der Lebensqualität. Die Bewertung der Validität der Impact-Domäne der FCS-SIS kann größtenteils der Bewertung zum Endpunkt „FCS-SIS (Symptom)“ in Kapitel 2.3.2 entnommen werden; der Fragebogen wird demnach als weitestgehend valide eingeschätzt.

Der Endpunkt wird zur Nutzenbewertung herangezogen. Es ist jedoch zu berücksichtigen, dass der Impact-Gesamtwert teils stark mit den Domänen „Ängste / Sorgen“, „Depression“ und „Fähigkeit zur Teilnahme an sozialen Aktivitäten“ des PROMIS-29 v2.0 korreliert [7] und in Teilen inhaltliche Mehrfacherfassungen vorliegen können.

PROMIS-29+2

Der Endpunkt „PROMIS-29+2“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die selbstberichtete allgemeine (d. h. indikationsübergreifende) gesundheitsbezogene Lebensqualität sollte in der Studie Balance mittels PROMIS-29+2 (Version 2.1) digital (YPrime-System) erhoben werden. Der Fragebogen umfasst insgesamt 29 + 2 Items aus dem Fragebogensystem des „Patient-Reported Outcomes Measurement Information System“ (PROMIS) und setzt sich aus den folgenden Domänen zusammen:

- Körperliche Funktionen (4 Items)
- Ängste / Sorgen (4 Items)
- Depression (4 Items)
- Müdigkeit (4 Items)
- Schmerzbeeinträchtigung (4 Items)
- Schlafstörungen (4 Items)
- Fähigkeit zur Teilnahme an sozialen Aktivitäten (4 Items)
- Kognitive Funktion (2 Items)

Die Items werden auf einer 5-stufigen Skala mit Bezugszeitraum der vergangenen Woche erfasst, wobei die Skalierung zwischen den Domänen teils unterschiedlich ist. Höhere Werte deuten bei der Mehrzahl der Domänen auf größere Einschränkungen hin. Mit Ausnahme der Domänen „Körperliche Funktionen“, „Fähigkeit zur Teilnahme an sozialen Aktivitäten“ und „Kognitive Funktion“, bei denen niedrigere Werte größere Einschränkungen bedeuten. Darüber hinaus wird durch den Fragebogen eine numerische Bewertungsskala von 0 bis 10 für die Schmerzintensität erhoben, welche die gleichnamige Domäne bildet.

Anhand der beschriebenen Domänen lassen sich Domänenscores, ein präferenzbasierter (PROMIS-Preference (PROPr)) Gesamtscore und zwei Summenscores (körperlicher und psychischer Summenscore) bilden, wobei die Domäne „Kognitive Funktion“ nicht in die jeweiligen Summenscores eingeht [1]. Es liegen Auswertungen hinsichtlich der einzelnen Domänenscores und des PROPr-Scores vor. Für die Domäne „Schmerzintensität“ wurden keine Auswertungen vorgelegt.

Der PROPr-Score wurde laut Studienunterlagen basierend auf den Domänenscores berechnet:

- Domänenscores
 - o Je Domäne wurde ein domänenspezifischer Summenscore berechnet. Wenn > 50 % der Items in einer Domäne fehlten, wurde der Wert für die jeweilige Visite als fehlend betrachtet. Wenn ≤ 50 % der Fragen in einer Domäne fehlten, wurden Werte für die fehlenden Items anhand des Durchschnitts der nicht fehlenden Items in derselben Domäne imputiert und auf die nächste ganze Zahl gerundet.
 - o Die domänenspezifischen Summenscores wurden anschließend mithilfe der Tabellen des PROMIS-Bewertungshandbuchs t-standardisiert. Ein T-Score von 50 entspricht dabei dem Durchschnitt der allgemeinen US-amerikanischen Bevölkerung. Bei den Domänen: „Ängste / Sorgen“, „Depression“, „Müdigkeit“, „Schmerzbeeinträchtigung“ und „Schlafstörungen“ bedeuten Werte höher als 50 eine höhere Einschränkung, bei den Domänen „Körperliche Funktionen“, „Fähigkeit zur Teilnahme an sozialen Aktivitäten“ und „Kognitive Funktion“ bedeuten Werte niedriger als 50 eine höhere Einschränkung.
- PROPr-Score:
 - o Die ermittelten T-Scores wurden in Theta-Werte umgewandelt: $\text{Theta} = (T - 50) / 10$
 - o Schließlich wurde der SAS-Code („Multi-Attribute Utility Theory“) unter Verwendung der Theta-Werte zur Berechnung des PROPr-Scores angewandt.

Der PROPr-Score reicht von 0 („sehr schlecht, dem Tode nahe“) bis 1 („perfekt, ideale Gesundheit“) und stellt einen gewichteten Nutzwert dar.

Wie präspezifiziert, wurde die Veränderung des PROPr-Scores bzw. der jeweiligen Domänenscores nach 6 bzw. 12 Monaten im Vergleich zu Baseline untersucht. Baseline wurde als die letzte nicht fehlende Bewertung vor der ersten Dosis des Studienmedikaments definiert. Für die Auswertung zum jeweiligen Analysezeitpunkt wird ein Mittelwert über zwei bzw. drei Erhebungszeitpunkte gebildet (bspw. Woche 51 und 53 für Monat 12). Für die Domänenscores und den PROPr-Score wurden kontinuierliche Auswertungen durchgeführt. Post hoc wurden zudem Responderanalysen hinsichtlich des PROPr-Scores eingereicht.

Bewertung

Die Operationalisierung für den Fragebogen inkl. der Domänenscores ist weitestgehend nachvollziehbar. Wesentliche Unklarheiten ergeben sich für die Berechnung des PROPr-Scores (siehe Patientenrelevanz).

Zu berücksichtigen sind die Unklarheiten zur Reihenfolge der PRO-Instrumente und möglicher Auswirkungen auf die Erhebung (siehe Abschnitt „PGIS“ in Kapitel 2.3.2).

Für die Nutzenbewertung werden die kontinuierlichen Auswertungen der Domänenscores zu Monat 12 dargestellt (siehe Patientenrelevanz). Responderanalysen lagen nicht vor.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen.

Bei dem vorliegenden PROPr-Score (auch „Health Utility Score“) handelt es sich um einen präferenzbasierten, d. h. gewichteten, Nutzwert. Die präsentierten Auswertungen zum Nutzwert sind für die Nutzenbewertung nicht verwertbar. Der pU liefert zudem keine Informationen aus denen hervorgeht, dass die Gewichtung auf einer Bewertung der Gesundheitszustände durch Patientinnen und Patienten mit FCS beruht. Der Nutzwert des PROMIS-29+2 wird somit nicht als patientenrelevant angesehen und nicht zur Nutzenbewertung herangezogen.

Auswertungen hinsichtlich des körperlichen und psychischen Summenscores legt der pU nicht vor. Es werden entsprechend die berichteten Domänenscores zur Nutzenbewertung herangezogen.

Validität

Die PROMIS-Profile wurden auf Basis umfangreicher Itembanken entwickelt, die aus bestehenden PRO-Datensätzen (z. B. WHOQOL) extrahiert, qualitativ geprüft und mithilfe eines modifizierten Delphi-Verfahrens durch ein Expertengremium sowie durch Fokusgruppen für verschiedene Erkrankungen inhaltlich abgestimmt wurden [10]. Die Itembanken zeigen moderate bis stark positive Korrelationen mit etablierten Messinstrumenten, die das gleiche Konstrukt messen [9]. Ausgehend von den PROMIS-Itembanken wurden Kurzformen für 7 zentrale Gesundheitsdomänen entwickelt, aus denen wiederum PROMIS-Profile gebildet werden können [8]. Für alle Kurzformen (4, 6 und 8 Items) zeigte sich, dass die Korrelationen der Kurzformen mit den Itembanken in 6 der 7 Domänen über 0,9 lag. Die Korrelationen für den Bereich der körperlichen Funktion reichen von 0,83 für die 4-Item-Kurzform bis 0,89 für die 8-Item-Kurzform.

Das PROMIS-29-Profil wurde in einem vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als weitestgehend valide eingeschätzt. Das vorliegende PROMIS-29+2-Profil enthält im Vergleich zum PROMIS-29-Profil zwei zusätzliche Items zur kognitiven Funktionsfähigkeit. Die durch 2 Items erhobene Domäne „Kognitive Funktion“ scheint keiner etablierten PROMIS-Kurzform zu entsprechen. Der pU legt zudem keine geeigneten Studien hinsichtlich der Validität des PROMIS-29+2-Profils vor. Entsprechend ist die Validität für die gegenüber dem PROMIS-29-Profil zusätzlich eingehende Domäne „Kognitive Funktion“ als unklar zu bewerten.

Für die Nutzenbewertung werden folglich lediglich die validierten Domänen (mit jeweils 4 Items) herangezogen, die dem PROMIS-29-Profil entsprechen.

Weiterhin sollte gemäß dem Manual das „Response Scoring Pattern“ für die Auswertung des Fragebogens bevorzugt werden [1]. Dieses berücksichtigt die jeweiligen Item-Level-Parameter für jedes Item und jede Antwort, wodurch ein besserer Umgang mit fehlenden Werten gewährleistet wird. Im vorliegenden Fall wurde jedoch die manuelle Transformierungsmethode mittels der Tabellen des Bewertungshandbuchs genutzt. Da das Ausfüllen des PROMIS-29+2 in der Studie Balance über ein digitales System erfolgte, welches laut pU sicherstellt, dass bei ausgefülltem Fragebogen keine fehlenden Werte auf Level der Einzelitems vorliegen, wird im vorliegenden Fall davon ausgegangen, dass sich die manuelle Konvertierung der Rohwerte in T-Scores anstelle der Nutzung des „Response Pattern Scoring“ nicht negativ auswirkte.

PROMIS Short Form – Pain Interference 8a

Der Endpunkt „PROMIS Short Form – Pain Interference 8a“ wird in der Nutzenbewertung aufgrund einer inhaltlichen Überschneidung (Hälfte der Fragebogen-Items) und somit einhergehenden Doppelerfassung mit dem PROMIS-29+2 ergänzend berücksichtigt.

Operationalisierung

Beschreibung

Die selbstberichtete Beeinträchtigung durch Schmerzen sollte in der Studie Balance mittels PROMIS Short Form – Pain Interference 8a (Version 1.1) erhoben werden.

Dieser erfasst mittels 8 Items auf einer 5-stufigen Skala von 1 („gar nicht“) bis 5 („sehr stark“), wie stark der Schmerz die Patientinnen und Patienten in den letzten 7 Tagen beeinträchtigt hat. Für die Berechnung des Gesamtwerts wurden zunächst die Antworten auf die einzelnen Fragen addiert, um einen Rohwert zu erhalten. Der Rohwert wurde anschließend mithilfe des PROMIS-Bewertungshandbuchs in einen T-Score umgewandelt. Ein T-Score von 50 entspricht dabei dem Durchschnitt der allgemeinen US-amerikanischen Bevölkerung. Höhere Werte des PROMIS Short Form – Pain Interference 8a weisen auf eine erhöhte Beeinträchtigung durch Schmerzen hin.

Wenn > 50 % der Items fehlten, wurde der Fragebogen für die Visite als fehlend betrachtet; wenn ≤ 50 % der Fragen fehlten, wurden Werte für die fehlenden Items anhand des Durchschnitts der nicht fehlenden Items imputiert und auf die nächste ganze Zahl gerundet.

Wie präspezifiziert, wurde die Veränderung des PROMIS Short Form – Pain Interference 8a nach 6 bzw. 12 Monaten im Vergleich zu Baseline untersucht. Baseline wurde als die letzte nicht fehlende Bewertung vor der ersten Dosis des Studienmedikaments definiert. Für die Auswertung zum jeweiligen Analysezeitpunkt wird ein Mittelwert über zwei bzw. drei Erhebungszeitpunkte gebildet (bspw. Woche 51 und 53 für Monat 12).

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Zu berücksichtigen sind die Unklarheiten zur Reihenfolge der PRO-Instrumente und möglicher Auswirkungen auf die Erhebung (siehe Abschnitt „PGIS“ in Kapitel 2.3.2).

Für die Nutzenbewertung werden die kontinuierlichen Auswertungen zu Monat 12 ergänzend dargestellt. Responderanalysen lagen nicht vor.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen.

In der Studie Balance wird die schmerzbedingte Beeinträchtigung bereits über den PROMIS-29+2 erhoben, wobei die Hälfte der Items (4 von 8) des PROMIS Short Form – Pain Interference 8a

identisch mit der Domäne „Schmerzbeeinträchtigung“ des PROMIS-29+2 erfasst werden und somit eine eindeutige Doppelerfassung darstellen. Für die vorliegende Nutzenbewertung wird angenommen, dass der PROMIS Short Form – Pain Interference 8a im Vergleich zum PROMIS-29+2 nur eingeschränkt zusätzliche Informationen erhebt. Der Endpunkt wird daher nur ergänzend in der Nutzenbewertung berücksichtigt.

Validität

Bei dem PROMIS Short Form – Pain Interference 8a handelt es sich um eine Kurzform des PROMIS – Pain Interference. Zur Entwicklung der Itembank des PROMIS – Pain Interference wurden zunächst 644 Items anhand einer Literaturrecherche und Rückmeldungen von Personen mit Schmerzen innerhalb von Fokusgruppen-Interviews (N = 24) und kognitiven Interviews (N = 44) identifiziert. Diese bildeten den Ausgangspunkt für die weitere Entwicklung der Itembank [3]. Aus diesem Pool wurden nach Befragung von Fachpersonen und Durchführung von Patienteninterviews 56 Items ausgewählt. Anhand der auch in der Veröffentlichung von Cella et al. (2010) verwendeten US-amerikanischen Population (Allgemeinbevölkerung; n = 21.133) wurden weitere Untersuchungen dieser Items durchgeführt, wobei 14.848 Personen verwertbare Antworten auf die 56 Items lieferten [9]. Zudem lieferten ein Convenience Sample aus 532 an Krebs erkrankten Personen sowie 523 Befragten mit chronischen Schmerzen Ergebnisse, sodass ein finaler Itempool mithilfe eines IRT-Modells („Item Response Theory“) kalibriert wurde. Eine endgültige Itembank mit 41 Items (v1.0) wurde in Bezug auf IRT-Annahmen, Modellanpassung, Reliabilität und Konstruktvalidität analysiert [3]. Die Eindimensionalität des PROMIS – Pain Interference wurde u. a. sowohl durch eine konfirmatorische als auch eine explorative Faktorenanalyse bestätigt. Wie gut die tatsächlich erhobenen Daten den Vorhersagen des Modells entsprechen, wurde durch eine „Goodness of Fit“-Statistik quantifiziert. Kein Item zeigte sich dabei auffällig [3].

Die interne Konsistenz der gesamten Itembank des PROMIS – Pain Interference mit 41 Items zeigt sich mit Cronbach's $\alpha = 0,99$ als sehr hoch, was auf eine hohe Reliabilität, aber gleichzeitig auch auf Redundanz einzelner Items hindeutet [3]. Die Untersuchung zur konvergenten Validität lieferte Ergebnisse in den durch die Autorengruppe erwarteten Richtungen: die stärksten Korrelationen zeigten Messungen schmerzbezogener Domänen der Brief Pain Inventory Interference Subscale (Spearman's $\rho = 0,90$), der SF-36 Bodily Pain Subscale (Spearman's $\rho = -0,84$) und der numerischen Bewertung der Schmerzintensität von 0 bis 10 (Spearman's $\rho = 0,48$) [3]. Weitere Untersuchungen der konvergenten Validität mehrerer Kurzformen des PROMIS – Pain Interference (4, 6 und 8 Items) anhand bereits existierender Datensätze aus der ersten Entwicklungsphase (Cella et al. (2010)) zeigen eine gute Korrelation dieser Kurzformen mit der Itembank ($> 0,9$). Anhand der graphischen Darstellung scheinen die untersuchten Kurzformen in der Lage zu sein, zwischen Menschen zu unterscheiden, die ihre Gesundheit und ihre Lebensqualität als „schlecht“, „mittelmäßig“, „gut“, „sehr gut“ und „ausgezeichnet“ bewerteten. Hinsichtlich der verwendeten Version des PROMIS Short Form – Pain Interference 8a fällt jedoch auf, dass die Fragen „How much did pain interfere with work around the home?“ und „How much did pain interfere with your household chores?“ in Teilen redundant erscheinen. Zudem muss berücksichtigt werden, dass der PROMIS Short Form – Pain Interference 8a sich speziell auf schmerzbedingte Beeinträchtigungen im Alltag fokussiert und somit primär die physischen Aspekte der Lebensqualität adressiert. Psychische und soziale Domänen werden nicht umfassend abgedeckt, sodass unklar ist, ob durch den PROMIS Short Form – Pain Interference 8a das multidimensionale Konstrukt der Lebensqualität ausreichend erfasst wird. Infolgedessen wäre auch eine Zuordnung zur Endpunktkategorie „Morbidity“ denkbar. Trotz dieser Unsicherheiten wird der der PROMIS Short Form – Pain Interference 8a ergänzend zum PROMIS-29+2 in der Domäne „Schmerzbeeinträchtigung“ unter der Kategorie gesundheitsbezogene Lebensqualität dargestellt. Die genannten Einschränkungen sind bei der Interpretation der Ergebnisse jedoch zu beachten.

Zusammenfassend scheint der PROMIS Short Form – Pain Interference 8a ein validiertes und reliables Instrument zur Erfassung der selbstberichteten Beeinträchtigung durch Schmerzen zu sein. Es ist jedoch zu beachten, dass durch die Auswahl der Items aus der Itembank für die Kurzform der Einfluss von Schmerz auf das multidimensionale Konstrukt der Lebensqualität möglicherweise nicht vollumfänglich abgedeckt ist und nicht-schmerzbedingte Auswirkungen auf die gesundheitsbezogene Lebensqualität unzureichend berücksichtigt sind.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt „Unerwünschte Ereignisse“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Ein unerwünschtes Ereignis (UE) war in der Studie Balance definiert als jedes unerwünschte und unbeabsichtigte Anzeichen (z. B. auch ein abnormaler Laborbefund), Symptom oder jede Krankheit, die vorübergehend mit der Anwendung des Studienmedikament zusammenhängt, unabhängig davon, ob ein kausaler Zusammenhang mit dem Studienmedikament bestand oder nicht.

Die standardisierte Kodierung erfolgte mittels „Medical Dictionary for Regulatory Activities“ (MedDRA; Version 26.0) nach Systemorganklasse (SOC) und Preferred Term (PT). Die Klassifikation der UE nach Schweregrad (mild, moderat bzw. schwer) erfolgte durch die Beurteilung des Prüfpersonal unter Berücksichtigung der „Common Terminology Criteria for Adverse Events“ (CTCAE; Version 5.0).

Ein Ereignis wurde als schwerwiegendes unerwünschtes Ereignis (SUE) definiert, wenn dieses

- zum Tode führte;
- unmittelbar lebensbedrohend war;
- eine Hospitalisierung oder deren Verlängerung erforderlich machte;
- zu einer anhaltenden oder schwerwiegenden Behinderung/Invalidität führte;
- eine kongenitale Anomalie oder einen Geburtsfehler nach sich zog;
- ein klinisch relevantes Ereignis darstellte (z. B. medizinische Intervention, um eines der zuvor genannten Ereignisse zu verhindern).

Alle während der Studie auftretenden UE und SUE, die mit einem schwerwiegenden unerwünschten kardiovaskulären Ereignis („Major Adverse Cardiovascular Event“ (MACE)) übereinstimmen, wurden zudem von einem verblindeten, unabhängigen Ausschuss gemäß der MACE Adjudication Charter beurteilt. Dazu gehörten folgende Ereignisse, welche nach Randomisierung auftraten:

- Tödliche Ereignisse
- Myokardinfarkt
- Krankenhausaufenthalt wegen instabiler Angina pectoris
- Herzinsuffizienz
- Geplante und elektive koronare Revaskularisierungsverfahren
- Nicht-koronare Revaskularisierungsverfahren
- Schlaganfall

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz und Validität

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen. Die Patientenrelevanz von Laborparametern ist unklar. Der pU legte keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Zwar wird die Erhebung der UE als valide eingeschätzt, jedoch ist nicht auszuschließen, dass auch Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen.

Dies betrifft ebenfalls den Endpunkt „MACE“. Dieser wurde in der Studie Balance im Rahmen der Sicherheit erfasst. Eine Rationale für eine Zuordnung zur Endpunktkategorie „Sicherheit“ liegt jedoch nicht vor. Eine Zuordnung zur Endpunktkategorie „Morbidity“ ist denkbar, kann jedoch nicht abschließend bewertet werden. Aufgrund der Erhebung des Endpunkts „MACE“ im Rahmen der Sicherheitsbewertung, werden die Ergebnisse für diesen Endpunkt in der vorliegenden Nutzenbewertung in der Kategorie „Sicherheit“ berichtet. Dabei werden jedoch keinerlei Inzidenzgrenzen für UE angesetzt, sondern alle Ereignisse aufgeführt.

2.3.5 Erhebungszeitpunkte

Tabelle 6: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie Balance

Endpunkte	Studienphase Woche	Screening	Randomisierte Behandlungsphase																EOT ¹)	Nachbeobachtung			
		-4 bis -1	1 ²⁾	5	9	13	17	21	23	25	27	29	33	37	41	45	49	51	53		57	61	66
Todesfälle ³⁾		kontinuierlich																					
Nüchtern-Triglyzeridwert ⁴⁾		x	x ⁵⁾	x	x	x	x	x	x	x	x	x		x		x		x	x	x	x	x	x
Bestätigte akute Pankreatitis		kontinuierlich																					
FCS-SIS (Symptom)		täglich																					
FCS-SIS (Impact)		x	x ⁶⁾	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
PROMIS-29+2			x			x			x	x	x							x	x	x			
PROMIS Short Form – Pain Interference 8a ⁷⁾		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Unerwünschte Ereignisse		kontinuierlich																					

¹⁾ Bei vorzeitiger Beendigung der Studie, wurden die Patientinnen und Patienten nachdrücklich angehalten, an den „Landmark“-Visiten zu Woche 23, 25, 27 sowie Woche 51 und 53 teilzunehmen.

²⁾ Die Baseline-Erhebung konnte bis zu 3 Tage (-3 Tage) vor dem eigentlichen Tag 1 (erste Dosis der Studienmedikation) durchgeführt werden.

³⁾ Todesfälle wurden im Rahmen der Sicherheit erhoben.

⁴⁾ Primärer Endpunkt wird ergänzend dargestellt.

⁵⁾ Baseline definiert als Mittelwert aus der letzten nicht fehlenden Messung während der Screening-Periode und der Messung vor Verabreichung der ersten Studienmedikation an Tag 1.

⁶⁾ Baseline definiert als der Durchschnitt während der Screening-Periode bis einschließlich Tag 1 der Studienmedikation.

⁷⁾ Endpunkt wird ergänzend dargestellt.

Abkürzungen: EOT: End of Treatment; FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; PROMIS: Patient-Reported Outcomes Measurement Information System.

2.4 Statistische Methoden

Die im Folgenden aufgeführten statistischen Auswertungen basieren auf dem SAP in der Version 4 vom 15.08.2023 und post hoc mit dem Dossier eingereichten Auswertungen.

Die Änderung des SAP auf Version 4 erfolgte nach Ende der Studienbehandlung am 14.07.2023, jedoch zeitlich vor Schließung der Datenbank am 30.08.2023 bzw. 30.11.2023.

Analysepopulationen

Full Analysis Set (FAS) / Sicherheitspopulation: Alle Patientinnen und Patienten, die randomisiert wurden und jegliche Dosis der Studienmedikation (Olezarsen oder Placebo) erhielten. Die Wirksamkeitsanalysen basieren auf dem FAS. Das FAS bzw. die Sicherheitspopulation entspricht der Intention-to-Treat (ITT)-Population.

Datenschnitte

Für die vorliegende Nutzenbewertung wurden 2 Datenschnitte vorgelegt:

- Datenschnitt zum Ende der Behandlungsphase vom 14.07.2023
- Korrigierter Datenschnitt vom 30.11.2023 für die AP-Ereignisse

Interimsdatenschnitte sind gemäß Studienprotokoll weder geplant gewesen noch wurden diese berichtet.

Gemäß Angaben des pU wurde der Datenschnitt vom 30.11.2023 aufgrund eines Fehlers bei der Anzahl der gemeldeten bestätigten AP-Ereignisse im Olezarsen-80mg-Arm durchgeführt. Aufgrund einer fehlerhaften Datumsangabe wurde ein Ereignis der Screening- und nicht der Behandlungsphase zugerechnet; dieser Fehler wurde mit dem Datenschnitt vom 30.11.2023 korrigiert. Nur die Daten für die bestätigten Pankreatitis-Episoden stammen vom Datenschnitt am 30.11.2023. Anderweitige Änderungen soll es gemäß pU zu diesem Datenschnitt nicht gegeben haben. Die Sicherheitsdaten seien zudem korrekt und vollständig gewesen sowie die Studie planmäßig beendet worden. Gleichzeitig beschreibt der pU jedoch, dass eine Aktualisierung des Studienberichts geplant sei, um die nach dem Ende der Behandlungsphase gesammelten Daten der Nachbeobachtungsphase zu berücksichtigen. Die Erhebung der Sicherheit sollte für Personen, die nicht in die offene, einarmige Extensionsstudie (ISIS-678354-CS13) übergangen, 13 Wochen über das Behandlungsende hinaus erfolgen. Ein weiterer Datenschnitt oder Studienbericht wurde jedoch nicht vorgelegt. Demnach ist unklar, ob die vorgelegten Sicherheitsdaten zum Datenschnitt vom 14.07.2023 vollständig sind. Auf Basis der Angaben zu Beendigung bzw. Abbruch der Nachbeobachtung (siehe Tabelle 9) wird vorliegend jedoch davon ausgegangen, dass sich zum Datenschnitt vom 14.07.2023 noch maximal 1 Person pro Behandlungsarm in der Nachbeobachtung befand. Darauf basierend ergeben sich keine Auswirkungen für die Höhe des Verzerrungspotentials der Endpunktkategorien „Mortalität“ und „Sicherheit“.

Zur Nutzenbewertung werden sowohl der Datenschnitt zum Ende der Behandlungsphase vom 14.07.2023 als auch der korrigierte Datenschnitt für den Endpunkt „Bestätigte akute Pankreatitis“ vom 30.11.2023 herangezogen.

Präspezifizierte und durchgeführte Subgruppenanalysen

Es waren ausschließlich für den primären Endpunkt Subgruppenanalysen (nach z. B. Geschlecht, Alter, Region, Ethnie etc.) präspezifiziert. Post hoc wurden Subgruppenanalysen für alle in Modul 4 dargestellten Endpunkte durchgeführt. Aufgrund des geringen Stichprobenumfangs ($N < 10$ in den jeweiligen Subgruppen der anderen Subgruppenanalysen) sind nur die Subgruppenanalysen nach Geschlecht bewertungsrelevant.

Es zeigten sich keine statistisch signifikanten Interaktionen zwischen einem Endpunkt und dem Merkmal „Geschlecht“.

Präspezifizierung geplanter und durchgeführter Analysen

Die präspezifizierten Analysen, welche in der vorliegenden Nutzenbewertung dargestellt werden, wurden weitestgehend wie im SAP geplant durchgeführt. Es wurden aufgrund von Konvergenzproblemen bei den Auswertungen des Endpunkts „Nüchtern-Triglyzeridwert“ geringfügige Modellanpassungen vorgenommen, die vorliegend als grundsätzlich nachvollziehbar bewertet werden.

Weder für die Wirksamkeits- noch für die Sicherheitsendpunkte wurden Auswertungen anhand des Relativen Risikos (RR) mit zugehörigem 95%-Konfidenzintervall (KI) und p-Wert präspezifiziert. Post hoc wurden für die Wirksamkeitsendpunkte RR und p-Werte unter Berücksichtigung der Strata mit der Mantel-Haenszel-Methode berechnet; für Sicherheitsendpunkte wurde der p-Wert hingegen mit dem exakten Test nach Fisher durchgeführt. Insbesondere bei geringen Populationsgrößen bzw. Ereignisraten werden exakte Tests im Allgemeinen bevorzugt, da sie robust sind, jedoch i. d. R. den Nachteil einer geringeren Power haben. Sind auf Basis des 95%-KI und des p-Werts unterschiedliche Schlussfolgerungen bezüglich der Signifikanz möglich, so sollte die Signifikanz mithilfe des exakten Tests (CSZ-Methode [4]) für die Nutzenbewertung bestimmt werden. Für den Endpunkt „MACE“ liegen keine statistischen Auswertungen und Effektschätzer vor. Die Darstellung erfolgt deskriptiv.

Für Modul 4 berichtet der pU für das PRO-Instrument „FCS-SIS“ Post-hoc-Responderanalysen zur Verbesserung und Verschlechterung um 15 % der Skalenspannweite. Für den Endpunkt „Bestätigte akute Pankreatitis“ werden Post-hoc-Auswertungen hinsichtlich der Personen mit mindestens einem AP-Ereignis sowie der Zeit bis zum erstem AP-Ereignis vorgelegt (siehe Kapitel 2.3).

Die Auswertungen erfolgten jeweils anhand des FAS zu Monat 6 bzw. 12.

Fehlende Werte und Daten-Imputation

In der Auswertung der prozentualen Veränderung des Nüchtern-TG-Werts wurden fehlende Werte unter Verwendung des „Jump to Reference“ (J2R)-Ansatzes bzw. für die Placebo-Gruppe unter der Annahme „Missing at Random“ (MAR) imputiert. Eine Imputation fehlender bzw. unvollständiger Werte in den anderen Endpunkten war nicht vorgesehen.

Sensitivitätsanalysen für fehlende Werte

Laut SAP war eine Tipping-Point-Analyse anhand des FAS vorgesehen, um die Auswirkungen fehlender Daten auf das Ergebnis der Analyse zu Monat 12 zu bewerten.

Einschätzung der statistischen Auswertungen

Für die Nutzenbewertung erfolgt die Darstellung der Ergebnisse zu Monat 12. Eine Ausnahme stellt der im Anhang dargestellte Endpunkt „PGIC“ dar (siehe Anhang für die Rationale). Die Auswertung anhand des FAS, welches der ITT-Population entspricht, wird als sachgerecht bewertet.

Bei FCS handelt es sich um ein heterogenes Erkrankungsbild, welches sich auch in der Studienpopulation widerspiegelt. Die eingeschlossenen Patientinnen und Patienten weisen mit Baseline-Werten des Nüchtern-TG-Werts von über 2.000 mg/dl gemäß EMA ein hohes Risiko für AP auf – unabhängig von der Vorgeschichte zu AP [13]. Etwa 31 % der Studienpopulation weisen zu Baseline bereits eine chronifizierte Pankreatitis auf. Gleichzeitig berichten jedoch etwa 29 % der Studienpopulation keinerlei AP in den vergangenen 10 Jahren und im Median wird in den vergangenen 5 Jahren eine AP berichtet. Insgesamt wird von den Patientinnen und Patienten zu Baseline eine geringe Symptomlast berichtet. Einige Personen berichten zu Baseline keinerlei Symptome im jeweiligen Endpunkt, sodass nicht für alle Patientinnen und Patienten eine Verbesserung ein mögliches Therapieziel darstellt. Daher werden vorliegende Responderanalysen zur Verschlechterung bevorzugt dargestellt. Aufgrund der Heterogenität der Erkrankung und der Studien-

population, kann für einen nicht unwesentlichen Anteil der Patientinnen und Patienten jedoch auch eine Verbesserung ein relevantes Therapieziel darstellen. Die Responderanalysen zur Verbesserung werden demnach ergänzend dargestellt. Diese sind gegenüber den Responderanalysen zur Verschlechterung jedoch als nachrangig zu betrachten, da die dabei berücksichtigte Studienpopulation auf Personen eingegrenzt ist, die zu Baseline Symptome bzw. Einschränkungen aufweisen.

Hinsichtlich des Endpunkts „Bestätigte akute Pankreatitis“ liegen sowohl post hoc berichtete Ereigniszeitanalysen mittels Cox-Proportional-Hazards-Modell mit Hazard Ratio als Effektschätzer als auch Auswertungen hinsichtlich der Anzahl der Personen mit ≥ 1 AP-Ereignis mit RR als Effektschätzer vor. Da vergleichbare Nachbeobachtungszeiten vorliegen, werden zur vorliegenden Nutzenbewertung die post hoc berichteten Analysen hinsichtlich der Anzahl der Personen mit mindestens einem AP-Ereignis und die präspezifizierten Analysen hinsichtlich der mittleren Ereignisrate herangezogen.

Für den ergänzend dargestellten Endpunkt „Nüchtern-Triglyzeridwert“ wird eine J2R-Imputation angewendet. Die J2R-Imputation stellt vorliegend einen eher konservativeren Ansatz im Umgang mit fehlenden Werten dar, da die Imputation mit Daten eines durchschnittlichen Profils von Teilnehmenden der Kontrollgruppe, unter MAR-Annahme, den Behandlungseffekt eher nicht überschätzt. Die Daten-Imputation für den primären Endpunkt wird entsprechend als sachgerecht bewertet.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 7: Verzerrungspotential der Studie Balance

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
Balance	Unklar ¹⁾²⁾	Ja	Ja	Ja	Nein	Unklar ³⁾	Unklar

¹⁾ Die Randomisierung erfolgte über ein Interactive-Response-Technology-System unter Berücksichtigung der definierten Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“.

²⁾ Vorgesehen war eine Randomisierung auf 2 Kohorten A und B (1:1) mit jeweils anschließender anteiliger Randomisierung (2:1) auf eine Olezarsen-Dosierung und eine Placebo-Dosierung. Daraus sollten 4 Behandlungsgruppen mit Olezarsen 50 mg (n = 20) bzw. Placebo 0,5 ml (n = 10) und Olezarsen 80 mg (n = 20) und Placebo 0,8 ml (n = 10) resultieren. Die Auswertung war jedoch Kohorten-übergreifend gegenüber dem gepoolten Placebo-Arm (n = 20) präspezifiziert. Da die Randomisierung innerhalb der jeweiligen Kohorte im Verhältnis 2:1 erfolgte und die Dosierung im Olezarsen-Arm mit 50 mg nicht der Fachinformation entspricht, wurde der dazugehörige Placebo-Arm 0,5 ml zusammen mit einer nicht für die Nutzenbewertung relevanten Studienpopulation randomisiert. Somit wurden vermutlich etwa 50 % des gepoolten Placebo-Arms nicht direkt mit dem für die Nutzenbewertung relevanten Behandlungssarm mit Olezarsen 80 mg stratifiziert randomisiert. Dieses gestaffelte Vorgehen bei der Randomisierung kann in Verbindung mit der geringen Studienpopulation zu beobachtbaren und nicht-beobachtbaren Imbalancen führen. Eine systematische einseitige Imbalance relevanter Prognosefaktoren zugunsten eines Behandlungsarms ist jedoch in den beobachteten Baseline-Charakteristika nicht erkennbar. Die Auswirkungen des gestaffelten Vorgehens bei der Randomisierung sind in der Gesamtschau unklar.

³⁾ Es ist unklar inwiefern die Diätvorgaben von ≤ 20 mg Fett pro Tag während des Studienverlaufs eingehalten wurden.

Das Verzerrungspotential auf Studienebene wird als unklar eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 8 dargestellt.

Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie Balance

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität					
Todesfälle ¹⁾	Ja	Ja ²⁾	Nein	Ja ³⁾	Unklar
Morbidität					
Nüchtern-Triglyzeridwert ⁴⁾	Ja	Nein ⁵⁾	Nein	Nein	Hoch
Bestätigte akute Pankreatitis	Ja	Nein ⁶⁾	Nein	Nein	Hoch
PGIS	Ja	Nein ⁷⁾	Nein	Ja ⁸⁾	Hoch
FCS-SIS (Symptom)	Ja	Nein ⁹⁾	Nein	Nein	Hoch
Lebensqualität					
FCS-SIS (Impact)	Ja	Nein ¹⁰⁾	Nein	Nein	Hoch
PROMIS-29+2	Ja	Nein ¹¹⁾	Nein	Nein	Hoch
PROMIS Short Form – Pain Interference 8a ⁴⁾	Ja	Nein ¹²⁾	Nein	Nein	Hoch
Sicherheit					
Schwere UE (Gesamtrate, SOC/PT)	Ja	Ja ²⁾	Nein	Ja ³⁾¹³⁾	Unklar
SUE (Gesamtrate, SOC/PT)	Ja	Ja ²⁾	Nein	Ja ³⁾¹³⁾	Unklar
Abbruch wegen UE (Gesamtrate)	Ja	Ja ²⁾	Nein	Ja ¹³⁾	Unklar
UE, jeglichen Schweregrads (SOC/PT)	Ja	Ja ²⁾	Nein	Ja ³⁾¹³⁾	Unklar

¹⁾ Todesfälle wurden im Rahmen der Sicherheit erhoben.

²⁾ Die Sicherheitspopulation entspricht der ITT-Population.

³⁾ Im Olezarsen-Arm haben 18 Personen (82 %) die 13-wöchige Nachbeobachtung abgebrochen, im Placebo-Arm waren es 20 Personen (87 %). Es wird davon ausgegangen, dass zum vorliegenden Datenschnitt die Nachbeobachtung von maximal 1 Person pro Behandlungsarm noch nicht abgeschlossen ist. Angaben zur Beobachtungsdauer der beiden Behandlungsarme liegen jedoch nicht vor. Ob und inwiefern systematische Unterschiede zwischen den Nachbeobachtungsdauern der Behandlungsarme bestehen, kann daher nicht abschließend bewertet werden.

⁴⁾ Der Endpunkt wird ergänzend dargestellt.

⁵⁾ Fehlende Werte von > 20 % im Olezarsen-Arm bzw. > 10 % im Placebo-Arm zu Monat 12 (Woche 51 bzw. 52) sowie ein Unterschied von 10 Prozentpunkten zwischen den beiden Studienarmen.

⁶⁾ 3 Personen im Olezarsen-Arm (13,6 %) und 1 Person im Placebo-Arm (4,3 %) haben die Behandlung bis Monat 12 abgebrochen, wobei die Person im Placebo-Arm bereits vor Abbruch der Studienmedikation ein AP-Ereignis aufwies. Wie viele von diesen Personen die Studie bzw. Nachbeobachtung abbrachen, ist nicht bekannt. Es kann sich bei etwaigen Studienabbrüchen um eine informative Zensierung zugunsten von Olezarsen handeln.

- ⁷⁾ Die Rücklaufquoten betragen 72,7 % im Olezarsen-Arm und 82,6 % im Placebo-Arm. Der Unterschied der Non-Responder zwischen den Behandlungsarmen beträgt somit 10 Prozentpunkte.
- ⁸⁾ 5 Personen (26,3 %) im Olezarsen-Arm bzw. 7 Personen (33,3 %) im Placebo-Arm wiesen zu Baseline im PGIS keine Symptomatik (Wert von 0) auf und konnten sich demnach im Verlauf der Studie nicht verbessern.
- ⁹⁾ Die Rücklaufquoten betragen 77,3 % im Olezarsen-Arm und 87,0 % im Placebo-Arm. Der Unterschied der Non-Responder zwischen den Behandlungsarmen beträgt somit annähernd 10 Prozentpunkte.
- ¹⁰⁾ Die Rücklaufquoten betragen 77,3 % im Olezarsen-Arm und 91,3 % im Placebo-Arm. Der Unterschied der Non-Responder zwischen den Behandlungsarmen beträgt somit > 10 Prozentpunkte.
- ¹¹⁾ Die Rücklaufquoten zur Veränderung von Monat 12 gegenüber Baseline liegen bei allen Domänen außer „Schmerzbeeinträchtigung“ bei ca. 70 % im Olezarsen-Arm und bei ca. 80 % im Placebo-Arm. In der Domäne „Schmerzbeeinträchtigung“ liegen die Rücklaufquoten im Olezarsen-Arm bei ca. 82 % und im Placebo-Arm bei ca. 92 %. Damit liegt zudem ein Unterschied von knapp bzw. mehr als 10 Prozentpunkten zwischen den beiden Studienarmen vor.
- ¹²⁾ Die Rücklaufquoten zur Veränderung von Monat 12 gegenüber Baseline liegen bei ca. 82 % im Olezarsen-Arm gegenüber ca. 96 % im Placebo-Arm. Damit liegt ein Unterschied von > 10 Prozentpunkten zwischen den beiden Studienarmen vor.
- ¹³⁾ Die Analysen (Effektschätzer und p-Werte) für die UE erfolgten post hoc und unstratifiziert.

Abkürzungen: AP: Akute Pankreatitis; FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; ITT: Intention to Treat; PGIS: Patient Global Impression of Severity; PROMIS: Patient-Reported Outcomes Measurement Information System; PT: Preferred Term; SOC: Systemorganklasse; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

Tabelle 9: Allgemeine Angaben; Studie Balance

Studie Balance Allgemeine Angaben	Olezarsen N = 22	Placebo N = 23
Randomisiert / FAS ¹⁾ / Sicherheitspopulation, n (%)	22 (100)	23 (100)
Abbruch Einnahme der Studienmedikation, n (%)	3 (13,6)	1 (4,30)
Aufgrund von:		
UE	2 (9,10)	0 (0)
Entscheidung Prüfpersonal	0 (0)	1 (4,30)
Freiwilliger Austritt	1 (4,50)	0 (0)
Einnahme Studienmedikation beendet, n (%)	19 (86,4)	22 (95,7)
Nachbeobachtung beendet, n (%)	2 (9,1)	1 (4,3)
Abbruch Nachbeobachtung, n (%)	18 (81,8)	20 (87,0)
Aufgrund von:		
Übergang in die offene Extensionsstudie	16 (72,7)	19 (82,6)
Freiwilliger Studienaustritt	2 (9,1)	0 (0)
Andere Gründe ²⁾	0 (0)	1 (4,3)
Mediane Behandlungsdauer in Wochen (min; max)	52 (4,0; 57,1)	52 (28,4; 52,6)
Mediane Anzahl der verabreichten Dosen (min; max)	13 (1; 13)	13 (6; 13)
Mediane Beobachtungsdauer in Wochen (min; max)	k. A.	k. A.

¹⁾ Entspricht der ITT-Population.

²⁾ Vom Prüfpersonal als „Proband nicht für die Nachbeobachtung verfügbar“ definiert.

Abkürzungen: FAS: Full Analysis Set; ITT: Intention-to-Treat; k. A.: keine Angabe; UE: Unerwünschtes Ereignis.

Tabelle 10: Charakterisierung der Studienpopulation; Studie Balance

Studie Balance Charakterisierung der Studienpopulation	Olezarsen N = 22	Placebo N = 23
Alter (Jahre)		
MW (SD)	47,7 (13,3)	44 (14,7)
Median (min; max)	47 (25; 78)	46 (21; 69)
Altersgruppe (Jahre), n (%)		
< 65	20 (90,9)	20 (87,0)
≥ 65	2 (9,1)	3 (13,0)
Geschlecht, n (%)		
männlich	11 (50,0)	11 (47,8)
weiblich	11 (50,0)	12 (52,2)
Gewicht (kg)		
MW (SD)	68,5 (16,7)	67,8 (16,1)
Median (min; max)	68,2 (35,8; 112,7)	64,7 (44,0; 101,9)
BMI (kg/m ²)		
MW (SD)	25,1 (6,0)	24,2 (4,12)
Median (min; max)	23,5 (14,9; 46,5)	24,4 (16,9; 31,5)

Studie Balance Charakterisierung der Studienpopulation	Olezarsen N = 22	Placebo N = 23
<i>Abstammung (genetisch), n (%)</i>		
kaukasisch	17 (77,3)	22 (95,7)
asiatisch	3 (13,6)	0 (0)
Hawaiianer oder andere Bewohner der Pazifikinseln	0 (0)	0 (0)
andere	2 (9,10)	1 (4,30)
<i>Region, n (%)</i>		
Europa (inklusive UK)	14 (63,6)	11 (47,8)
Nordamerika	8 (36,4)	12 (52,2)
<i>Land, n (%)</i>		
Kanada	1 (4,5)	5 (21,7)
Frankreich	2 (9,1)	1 (4,3)
Italien	3 (13,6)	3 (13,0)
Niederlande	2 (9,1)	1 (4,3)
Norwegen	1 (4,5)	0 (0)
Portugal	1 (4,5)	1 (4,3)
Slowakei	2 (9,1)	1 (4,3)
Spanien	2 (9,1)	3 (13)
Schweden	0 (0)	1 (4,3)
UK und Nordirland	1 (4,5)	0 (0)
USA	7 (31,8)	7 (30,4)
<i>Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening¹⁾, n (%)</i>		
ja	17 (77,3)	15 (65,2)
nein	5 (22,7)	8 (34,8)
<i>Anzahl der Pankreatitis-Episoden²⁾ in den letzten 10 Jahren, n (%)</i>		
MW (SD)	3,4 (6,01)	6,3 (16,3)
Median (min; max)	1 (0; 25)	2 (0; 78)
<i>Anzahl der Pankreatitis-Episoden²⁾ in den letzten 5 Jahren, n (%)</i>		
MW (SD)	1,5 (3,16)	3,6 (8,24)
Median (min; max)	1 (0; 15)	1 (0; 38)
Personen mit bekannter chronischer Pankreatitis, n (%)	8 (36,4)	6 (26,1)
<i>Vorbehandlung mit Volanesorsen¹⁾, n (%)</i>		
ja	8 (36,4)	10 (43,5)
nein	14 (63,6)	13 (56,5)
Krankheitsspezifische Charakteristika zu Baseline³⁾		
<i>Nüchtern-TG-Wert (mg/dl)</i>		
MW (SD)	2.613 (1.499)	2.596 (1.256)
Median (min; max)	2.086 (683; 6.898)	2.493 (334; 5.436)
<i>Nüchtern-TG-Wert (mg/dl), n (%)</i>		
≤ 880 ⁴⁾	1 (4,50)	2 (8,70)
≥ 880 bis < 1.000	0 (0)	0 (0)
≥ 1.000 bis < 2.000	9 (40,9)	7 (30,4)
≥ 2.000	12 (54,5)	14 (60,9)
<i>Nüchtern-ApoC3 (mg/dl)</i>		
MW (SD)	27,5 (11,6)	27,7 (11,7)
Median (min; max)	28,4 (9,94; 49,0)	25,6 (9,02; 55,3)

Studie Balance Charakterisierung der Studienpopulation	Olezarsen N = 22	Placebo N = 23
<i>Nüchtern-Chylomikron-TG-Wert (mg/dl)</i> Median (min; max)	1.873 (375; 9.867)	1.958 (227; 4.852)
<i>Diabetes-Status, n (%)</i> Typ 1 oder Typ 2 Nein	7 (31,8) 15 (68,2)	6 (26,1) 17 (73,9)
Funktionsverlust-Mutationen⁵⁾		
<i>LPL</i> homozygot hemizygot compound-heterozygot doppelt-heterozygot	10 (45,5) 1 (4,50) 5 (22,7) 1 (4,50)	16 (69,6) 0 (0) 5 (21,7) 0 (0)
<i>ApoA5</i> homozygot doppelt-heterozygot	1 (4,50) 1 (4,50)	1 (4,30) 0 (0)
<i>GPIHBP1</i> homozygot	1 (4,50)	1 (4,30)
<i>LMF1</i> compound-heterozygot	1 (4,50)	0 (0)
<i>ApoC2</i> homozygot	2 (9,10)	0 (0)

¹⁾ Stratifikationsfaktor in der Studie Balance.

²⁾ Akute Pankreatitis oder Episode mit starken Bauchschmerzen oder anderen Symptomen, die auf eine Pankreatitis hindeuten und zu einem Krankenhausaufenthalt geführt haben.

³⁾ Baseline war definiert als Mittelwert aus der letzten nicht fehlenden Messung während der Screening-Periode und aus der Messung vor Verabreichung der Studienmedikation an Tag 1.

⁴⁾ Das Einschlusskriterium „Nüchtern-TG-Wert $\geq$ 880 mg/dl“ wurde zum Datum des Screenings angewendet.

⁵⁾ Die genetische Bestätigung erfolgte entweder durch ein positives Testergebnis des zentralen genetischen Labors oder, wenn das Testergebnis des zentralen genetischen Labors unbestimmt war, durch die Beurteilung eines Fachexperten.

Abkürzungen: Apo: Apolipoprotein; GPIHBP1: Glycosylphosphatidylinositol-anchored High-Density Lipoprotein-Binding Protein 1; LMF1: Lipase-Maturation-Faktor 1; LPL: Lipoproteinlipase; MW: Mittelwert; SD: Standardabweichung; TG: Triglyzerid.

Protokollverletzungen

Insgesamt traten bei 22 von 45 Personen (48,9 %) schwerwiegende Protokollverletzungen auf, bei 12 Personen (54,5 %) im Olezarsen-Arm und bei 10 Personen (43,5 %) im Placebo-Arm. Die häufigsten schwerwiegenden Protokollverletzungen betrafen fehlerhafte bzw. unvollständige Studieneinwilligungen (31,8 % Olezarsen-Arm vs. 17,4 % Placebo-Arm) sowie verpasste Studienvisiten (18,2 % vs. 17,4 %). Darüber hinaus wurden bei 3 Personen (13,6 %) im Olezarsen-Arm und 4 Personen (17,4 %) im Placebo-Arm Protokollverletzungen in den Studienprozeduren berichtet, wobei eine genaue Definition in den Studienunterlagen nicht ersichtlich ist.

Auf Basis der berichteten Protokollverletzungen wird nicht von wesentlichen verzerrenden Aspekten für die Bewertung des Zusatznutzens ausgegangen.

Begleitmedikation

Anhand der in der Studie Balance dokumentierten Begleitmedikation ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens. Die lipidsenkenden Wirkstoffe, die vor der ersten Dosis der Studienbehandlung eingenommen und während des Studienverlaufs beibehalten wurden, werden nachfolgend in Tabelle 11 aufgeführt.

Tabelle 11: Begleitmedikation; Studie Balance

Studie Balance Begleitmedikation	Olezarsen N = 22 n (%)	Placebo N = 23 n (%)
Statine	5 (22,7)	7 (30,4)
Omega-3-Fettsäuren	12 (54,5)	7 (30,4)
Fibrate	11 (50,0)	11 (47,8)
Andere lipidsenkende Wirkstoffe ¹⁾	3 (13,6)	3 (13,0)

¹⁾ Ezetimib bzw. Evolocumab.

3.2 Mortalität

Während der Studie Balance aufgetretene Todesfälle wurden im Rahmen der Sicherheit erfasst. Während der Studie Balance sind keine Todesfälle in den beiden Behandlungsarmen aufgetreten.

3.3 Morbidität

Nüchtern-Triglyzeridwert (ergänzend)

Zu Baseline lagen die TG-Werte im Olezarsen-Arm im Mittel bei 2.613 mg/dl bzw. im Median bei 2.086 mg/dl (min; max: 683; 6.898). Im Placebo-Arm betrug der Baseline-TG-Wert im Mittel 2.596 mg/dl bzw. im Median 2.493 mg/dl (min; max: 334; 5.436). Während die Mittelwerte der jeweiligen Baseline-TG-Werte zwischen den beiden Behandlungsarmen nur wenig voneinander abwichen, ist hinsichtlich des Medians ein größerer Unterschied zwischen den beiden Behandlungsarmen ersichtlich. Die Veränderung der Nüchtern-TG-Werte von Baseline zu Monat 12 ist der nachfolgenden Tabelle zu entnehmen.

Tabelle 12: Nüchtern-Triglyzeridwert, Prozentuale Veränderung zu Monat 12 (ANCOVA); Studie Balance, FAS (DS: 14.07.2023)

Studie Balance Nüchtern-Triglyzeridwert (mg/dl)	Olezarsen N = 22	Placebo N = 23
Baseline ¹⁾ , n (%) MW (SD)	22 (100) 2.613 (1.499)	23 (100) 2.596 (1.256)
Woche 51, n (%) MW (SD)	15 (68,2) 1.498 (1.094)	19 (82,6) 2.766 (1.614)
Woche 53, n (%) MW (SD)	17 (77,3) 1.437 (867)	20 (87,0) 2.926 (1.431)
Monat 12 ²⁾ , n (%) MW (SD)	22 (100) ³⁾ 1.590 (914)	23 (100) ³⁾ 2.959 (1.430)
Prozentuale Veränderung zu Monat 12 ⁴⁾⁵⁾ , n (%) LS Mean [95%-KI]	22 (100) -38,5 [-58,2; -18,8]	23 (100) 20,89 [1,02; 40,8]
LS-Mean-Differenz [95%-KI]; p-Wert ⁶⁾	-59,4 [-90,7; -28,1]; 0,0002	

¹⁾ Baseline definiert als Mittelwert aus der letzten nicht fehlenden Messung während der Screening-Periode und aus der Messung vor Verabreichung der ersten Studienmedikation an Tag 1.

²⁾ Monat 12 definiert als Mittelwert der Wochen 51 und 53.

³⁾ Fehlende Werte wurden imputiert; es ist unklar wie viele Patientinnen und Patienten einen Wert (n) zu Monat 12 (Mittelwert der Wochen 51 und 53) vorwiesen.

⁴⁾ Fehlende Daten nach dem Tag der letzten Dosis für Patientinnen und Patienten im Olezarsen-Arm wurden unter Verwendung des J2R-Ansatzes imputiert. Fehlende Daten vor der letzten Dosis oder für Patientinnen und Patienten in der Placebo-Gruppe wurden unter MAR-Annahme imputiert.

⁵⁾ Prozentuale Veränderung zu Monat 12 mittels ANCOVA mit Behandlungsgruppe und den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“, als feste Effekte sowie dem logarithmierten TG-Wert zu Baseline als Kovariate.

⁶⁾ p-Werte und 95%-KI unter Verwendung des robusten Varianzschätzers nach Bell und McCaffrey.

Abkürzungen: ANCOVA: Analysis of Covariance; DS: Datenschnitt; FAS: Full Analysis Set; J2R: Jump to Reference; KI: Konfidenzintervall; LS: Least Squares; MAR: Missing at Random; MW: Mittelwert; SD: Standardabweichung; TG: Triglyzerid.

Bestätigte akute Pankreatitis

Bei der Interpretation ist zu berücksichtigen, dass 3 Personen im Olezarsen-Arm (13,6 %) und 1 Person im Placebo-Arm (4,3 %) die Behandlung bis Monat 12 abbrachen und entsprechend nicht vollständig in die Auswertung eingingen. Die Person im Placebo-Arm wies bereits vor dem Abbruch der Studienmedikation ein AP-Ereignis auf.

Tabelle 13: Bestätigte akute Pankreatitis bis Monat 12; Studie Balance; FAS (DS: 30.11.2023¹⁾)

Studie Balance Bestätigte akute Pankreatitis ²	Olezarsen N = 22	Placebo N = 23	Effektschätzer [95%-KI]; p-Wert
Personen mit mind. einem Ereignis (post hoc), n (%)	1 (4,50)	7 (30,4)	RR ³⁾ : 0,68 [0,49; 0,95]; 0,009
Gesamtzahl an Ereignissen	1	11	
Gesamtzahl der Patientenjahre	20,3	22,9	
Mittlere Ereignisrate/100 Patientenjahre [95%-KI]	4,93 [0,61; 40,2]	36,1 [14,5; 89,9]	MRR ⁴⁾ : 0,14 [0,01; 1,32]; 0,09

¹⁾ Aufgrund eines Fehlers bei der Anzahl der gemeldeten bestätigten AP-Ereignisse im Olezarsen-80mg-Arm, wurde zum Zeitpunkt des endgültigen Datenschnitts am 30.11.2023 eine Korrektur für dieses eine Ereignis vorgenommen. Daher stammen die Daten für die bestätigten Pankreatitis-Episoden vom Datenschnitt am 30.11.2023.

²⁾ Als „Bestätigte akute Pankreatitis“ für die Auswertung galten dokumentierte, wahrscheinliche und mögliche AP gemäß der Einschätzung des „Acute Pancreatitis Adjudication Committee“ (siehe Kapitel 2.3.2).

³⁾ RR und p-Wert stratifiziert nach beiden Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

⁴⁾ MRR, 95%-KI sowie p-Wert (Wald-Chi-Quadrat-Test) anhand Negativ-Binomial-Regressionsmodell mit Behandlungsgruppe und Vorbehandlung mit Volanesorsen (ja vs. nein) als Faktoren und der Anzahl bestätigter AP-Ereignisse innerhalb 5 Jahre vor Studienaufnahme als Kovariate.

Abkürzungen: AP: Akute Pankreatitis; DS: Datenschnitt; FAS: Full Analysis Set; KI: Konfidenzintervall; MRR: Mean Rate Ratio; RR: Relatives Risiko.

PGIS

Für den PGIS lag zu Baseline im Olezarsen-Arm ein Median von 1 (min; max: 0; 4) bzw. im Placebo-Arm von 1 (min; max: 0; 3) vor, wobei die Skalenspannweite von 0 bis 4 reicht und höhere Werte auf eine schwerere FCS-Symptomatik hindeuten. 5 Personen (26,3 %) im Olezarsen- bzw. 7 Personen (33,3 %) im Placebo-Arm wiesen zu Baseline im PGIS keine Symptomatik (Wert von 0) auf und konnten sich demnach im Verlauf der Studie nicht verbessern. Jeweils 14 Personen beider Arme wiesen mindestens milde Symptome (Wert von 1 bis 4) auf und konnten sich verbessern. Für 3 Personen (13,6 %) im Olezarsen-Arm und 2 Personen (8,7 %) im Placebo-Arm lag kein Baseline-Wert für den PGIS vor. In der Responderanalyse zur Verbesserung nach 12 Monaten (bzw. zu Woche 53) zeigt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.

Tabelle 14: Anteil der Personen mit einer Veränderung im PGIS zu Monat 12; Studie Balance, FAS (DS: 14.07.2023)

Studie Balance PGIS ¹⁾ Responderanalyse zu Monat 12 ²⁾	Olezarsen N = 22	Placebo N = 23
Verschlechterung ³⁾		
Verfügbare Werte, n (%)	k. A.	k. A.
Responder, n (%)	k. A.	k. A.
RR [95%-KI]; p-Wert	k. A.	
Verbesserung ⁴⁾ (ergänzend)		
Verfügbare Werte, n (%)	16 (72,7)	19 (82,6)
Responder, n (%)	3 (18,8)	5 (26,3)
RR [95%-KI]; p-Wert ⁵⁾	0,90 [0,64; 1,28]; 0,59	

¹⁾ Die Skalenspannweite beträgt 0 bis 4; höhere Werte zeigen schwerere FCS-Symptome an.

²⁾ Monat 12 ist präspezifiziert als Mittelwert der Wochen 51 und 53. Aufgrund der vorliegenden Daten wird davon ausgegangen, dass diese Auswertung einer Responderanalyse zu Woche 53 gegenüber Baseline entspricht (s. Kapitel 2.3.2).

³⁾ Da etwa ein Viertel der Personen im Olezarsen-Arm und ein Drittel der Personen im Placebo-Arm zu Baseline keine Symptome aufwiesen und sich demnach nicht verbessern konnten, werden die Responderanalysen zur Verschlechterung in dieser Studienpopulation gegenüber Auswertungen zur Verbesserung als vorrangig relevant eingeschätzt.

⁴⁾ Verbesserung definiert als Veränderung im Vergleich zu Baseline < 0.

⁵⁾ RR und p-Wert (Cochran-Mantel-Haenszel) unter Berücksichtigung der Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

Abkürzungen: DS: Datenschnitt; FAS: Full Analysis Set; k. A.: keine Angabe; KI: Konfidenzintervall; PGIS: Patient Global Impression of Severity; RR: Relatives Risiko.

FCS-SIS (Symptom)

In Tabelle 15 werden die Responderanalysen in der FCS-SIS (Symptom) zu Monat 12 dargestellt. Wie in Kapitel 2.4 beschrieben, erfolgt die Darstellung der Responderanalyse zur Verbesserung ergänzend. Für die FCS-SIS (Symptom) lag zu Baseline im Olezarsen-Arm ein Median von 0,9 (min; max: 0; 6) vor und im Placebo-Arm ein Median von 0,5 (min; max: 0; 4), wobei die Skalenspannweite von 0 bis 10 reicht und höhere Werte auf eine schwerere FCS-Symptomatik hindeuten.

Tabelle 15: Anteil der Personen mit einer Veränderung in der FCS-SIS (Symptom) zu Monat 12; Studie Balance, FAS (DS: 14.07.2023)

Studie Balance FCS-SIS (Symptom) ¹⁾ Responderanalyse zu Monat 12 ²⁾	Olezarsen N = 22	Placebo N = 23
Verschlechterung um mind. 15 % der Skalenspannweite ³⁾		
Verfügbare Werte, n (%)	17 (77,3)	20 (87,0)
Responder, n (%)	1 (5,90)	1 (5,0)
RR [95%-KI]; p-Wert ⁴⁾	1,05 [0,91; 1,20]; 0,60	
Verbesserung um mind. 15 % der Skalenspannweite ⁵⁾ (ergänzend)		
Verfügbare Werte, n (%)	17 (77,3)	20 (87,0)
Responder, n (%)	0 (0)	0 (0)
RR [95%-KI]; p-Wert ⁴⁾		

¹⁾ Die Skalenspannweite beträgt 0 bis 10; höhere Werte zeigen schwerere FCS-Symptome an.

²⁾ Monat 12 definiert als Mittelwert der nicht fehlenden Werte der Wochen 49 bis 53.

³⁾ Verschlechterung definiert als Erhöhung des Symptom-Gesamtwerts um ≥ 15 % der Skalenspannweite zu Monat 12. Ein entsprechender Punktwert wurde durch den pU nicht genannt.

⁴⁾ RR und p-Wert stratifiziert nach den beiden Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Fehlende Daten wurden nicht imputiert. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

⁵⁾ Verbesserung definiert als Verringerung des Symptom-Gesamtwerts um ≥ 15 % der Skalenspannweite zu Monat 12. Ein entsprechender Punktwert wurde durch den pU nicht genannt.

Abkürzungen: - : nicht berechenbar; DS: Datenschnitt; FAS: Full Analysis Set; FCS: Familiäres Chylomikronämie-Syndrom; FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

3.4 Lebensqualität

FCS-SIS (Impact)

In Tabelle 16 werden die Responderanalysen in der FCS-SIS (Impact) zu Monat 12 dargestellt. Wie in Kapitel 2.4 beschrieben, erfolgt die Darstellung der Responderanalyse zur Verbesserung ergänzend. Für die FCS-SIS (Impact) lag zu Baseline im Olezarsen-Arm ein Median von 2,2 (min; max: 0; 4) bzw. im Placebo-Arm von 2,0 (min; max: 0; 4) vor, wobei die Skalenspannweite von 0 bis 4 reicht und höhere Werte auf eine stärkere Belastung durch FCS hindeuten.

Tabelle 16: Anteil der Personen mit einer Veränderung in der FCS-SIS (Impact) zu Monat 12; Studie Balance, FAS (DS: 14.07.2023)

Studie Balance FCS-SIS (Impact) ¹⁾ Responderanalyse zu Monat 12 ²⁾	Olezarsen N = 22	Placebo N = 23
Verschlechterung um mind. 15 % der Skalenspannweite ³⁾		
Verfügbare Werte, n (%)	17 (77,3)	21 (91,3)
Responder, n (%)	2 (11,8)	4 (19,0)
RR [95%-KI]; p-Wert ⁴⁾	0,89 [0,66; 1,19]; 0,41	
Verbesserung um mind. 15 % der Skalenspannweite ⁵⁾ (ergänzend)		
Verfügbare Werte, n (%)	17 (77,3)	21 (91,3)
Responder, n (%)	2 (11,8)	4 (19,0)
RR [95%-KI]; p-Wert ⁴⁾	1,0 [0,73; 1,32]; 0,89	

¹⁾ Die Skalenspannweite beträgt 0 bis 4; höhere Werte zeigen stärkere Belastung durch FCS an.

²⁾ Monat 12 definiert als Mittelwert der nicht fehlenden Werte der Wochen 49, 51 und 53.

³⁾ Verschlechterung definiert als Erhöhung des Impact-Gesamtwerts um ≥ 15 % der Skalenspannweite zu Monat 12. Ein entsprechender Punktwert wurde durch den pU nicht genannt.

⁴⁾ RR und p-Wert stratifiziert nach beiden Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Fehlende Daten wurden nicht imputiert. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

⁵⁾ Verbesserung definiert als Verringerung des Impact-Gesamtwerts um ≥ 15 % der Skalenspannweite zu Monat 12. Ein entsprechender Punktwert wurde durch den pU nicht genannt.

Abkürzungen: DS: Datenschnitt; FAS: Full Analysis Set; FCS: Familiäres Chylomikronämie-Syndrom; FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

PROMIS-29+2

*Tabelle 17: Veränderung des PROMIS-29+2 zu Monat 12 (ANCOVA); Studie Balance, FAS
(DS: 14.07.2023)*

Studie Balance PROMIS-29+2	Olezarsen N = 22	Placebo N = 23
Körperlicher Summenscore	k. A.	k. A.
Psychischer Summenscore	k. A.	k. A.
Körperliche Funktionen ¹⁾		
Baseline ²⁾ , n (%) MW (SD)	19 (86,4) 47,8 (9,36)	20 (87,0) 48,5 (10,4)
Monat 12 ³⁾ , n (%) MW (SD)	17 (77,3) 49,2 (10,8)	21 (91,3) 50,6 (9,2)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD) LS-Mean-Differenz [95%-KI]; p-Wert	15 (68,2) 1,67 (9,30) 0,25 [-4,07; 4,56]; 0,91	18 (78,3) 0,66 (4,08)
Ängste / Sorgen ⁵⁾		
Baseline ²⁾ , n (%) MW (SD)	19 (86,4) 58,3 (11,9)	20 (87,0) 55,8 (11,6)
Monat 12 ³⁾ , n (%) MW (SD)	17 (77,3) 58,6 (13,5)	21 (91,3) 53,4 (11,1)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD) LS-Mean-Differenz [95%-KI]; p-Wert	15 (68,2) -1,80 (9,83) -0,12 [-6,01; 5,78]; 0,97	18 (78,3) -0,33 (5,55)
Depression ⁵⁾		
Baseline ²⁾ , n (%) MW (SD)	19 (86,4) 51,9 (12,1)	20 (87,0) 51,8 (10,3)
Monat 12 ³⁾ , n (%) MW (SD)	17 (77,3) 54,3 (13,4)	21 (91,3) 48,1 (10,2)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD) LS-Mean-Differenz [95%-KI]; p-Wert	15 (68,2) 0,30 (9,84) 3,32 [-2,56; 9,19]; 0,26	18 (78,3) -2,20 (7,67)
Müdigkeit ⁵⁾		
Baseline ²⁾ , n (%) MW (SD)	19 (86,4) 51,2 (11,9)	20 (87,0) 51,7 (12,6)
Monat 12 ³⁾ , n (%) MW (SD)	17 (77,3) 52,9 (13,5)	21 (91,3) 50,4 (12,9)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD) LS-Mean-Differenz [95%-KI]; p-Wert	15 (68,2) 0,17 (7,86) 1,14 [-3,23; 5,51]; 0,60	18 (78,3) -0,37 (3,76)

Studie Balance PROMIS-29+2	Olezarsen N = 22	Placebo N = 23
Schmerzbeeinträchtigung ⁵⁾		
Baseline ²⁾ , n (%) MW (SD)	22 (100) 53,0 (12,6)	23 (100) 52,5 (11,8)
Monat 12 ³⁾ , n (%) MW (SD)	18 (81,8) 52,2 (11,7)	21 (91,3) 49,3 (10,4)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD)	18 (81,8) -2,81 (10,5)	21 (91,3) -2,86 (7,94)
LS-Mean-Differenz [95%-KI]; p-Wert	0,83 [-4,28; 5,94]; 0,75	
Schlafstörungen ⁵⁾		
Baseline ²⁾ , n (%) MW (SD)	19 (86,4) 50,2 (8,25)	20 (87,0) 49,0 (8,21)
Monat 12 ³⁾ , n (%) MW (SD)	17 (77,3) 52,0 (10,5)	21 (91,3) 50,1 (10,0)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD)	15 (68,2) -2,86 (8,17)	18 (78,3) 1,84 (7,13)
LS-Mean-Differenz [95%-KI]; p-Wert	-2,79 [-8,77; 3,19]; 0,35	
Fähigkeit zur Teilnahme an sozialen Aktivitäten ¹⁾		
Baseline ²⁾ , n (%) MW (SD)	19 (86,4) 50,7 (10,6)	20 (87,0) 50,8 (12,4)
Monat 12 ³⁾ , n (%) MW (SD)	17 (77,3) 49,5 (10,7)	21 (91,3) 54,5 (13,2)
Veränderung zu Monat 12 ⁴⁾ , n (%) LS Mean (SD)	15 (68,2) 0,71 (7,48)	18 (78,3) 1,72 (6,20)
LS-Mean-Differenz [95%-KI]; p-Wert	-1,88 [-6,66; 2,91]; 0,43	
Schmerzintensität	k. A.	k. A.

¹⁾ Anhand t-transformierter Werte mit einem Mittelwert von 50 und einer Standardabweichung von 10; niedrigere Werte als 50 zeigen eine höhere Einschränkung an.

²⁾ Baseline definiert als die letzte nicht fehlende Messung vor Verabreichung der ersten Dosis des Studienmedikaments.

³⁾ Monat 12 definiert als Mittelwert der Wochen 51 und 53.

⁴⁾ Veränderung zu Monat 12 mittels ANCOVA mit Behandlungsgruppe und den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“ als feste Effekte sowie dem jeweiligen Domänenscore zu Baseline als Kovariate.

⁵⁾ Anhand t-transformierter Werte mit einem Mittelwert von 50 und einer Standardabweichung von 10; höhere Werte als 50 zeigen eine höhere Einschränkung an.

Abkürzungen: ANCOVA: Analysis of Covariance; DS: Datenschnitt; FAS: Full Analysis Set; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MW: Mittelwert; PROMIS: Patient-Reported Outcomes Measurement Information System; SD: Standardabweichung.

PROMIS Short Form – Pain Interference 8a (ergänzend)

Tabelle 18: Veränderung des PROMIS Short Form – Pain Interference 8a zu Monat 12 (ANCOVA); Studie Balance, FAS (DS: 14.07.2023)

Studie Balance PROMIS Short Form – Pain Interference 8a¹⁾	Olezarsen N = 22	Placebo N = 23
<i>Baseline²⁾, n (%)</i>	22 (100)	23 (100)
<i>MW (SD)</i>	53,1 (13,4)	52,7 (12,5)
<i>Monat 12³⁾, n (%)</i>	18 (81,8)	22 (95,7)
<i>MW (SD)</i>	53,4 (11,2)	48,6 (10,8)
<i>Veränderung zu Monat 12⁴⁾, n (%)</i>	18 (81,8)	22 (95,7)
<i>LS Mean (SD)</i>	-2,04 (11,5)	-3,16 (7,9)
LS-Mean-Differenz [95%-KI]; p-Wert	2,46 [-2,62; 7,53]; 0,34	

¹⁾ Anhand t-transformierter Werte mit einem Mittelwert von 50 und einer Standardabweichung von 10; höhere Werte als 50 eine zeigen höhere Einschränkung an.

²⁾ Baseline definiert als die letzte nicht fehlende Messung vor Verabreichung der ersten Dosis des Studienmedikaments.

³⁾ Monat 12 definiert als Mittelwert der Wochen 51 und 53.

⁴⁾ Veränderung zu Monat 12 mittels ANCOVA mit Behandlungsgruppe und den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“ als feste Effekte sowie dem PROMIS – Pain Interference zu Baseline als Kovariate.

Abkürzungen: ANCOVA: Analysis of Covariance; DS: Datenschnitt; FAS: Full Analysis Set; KI: Konfidenzintervall; LS: Least Squares; MW: Mittelwert; PROMIS: Patient-Reported Outcomes Measurement Information System; SD: Standardabweichung.

3.5 Sicherheit

Im Folgenden werden die Ergebnisse der Endpunktkategorie „Sicherheit“ der Studie Balance dargestellt. Die mediane Beobachtungszeit in den Behandlungsarmen betrug jeweils 52 Wochen (min; max: Olezarsen: 4,0; 57,1 vs. Placebo: 28,4; 52,6). Es ist anzumerken, dass keine Auswertungen unter Nichtberücksichtigung von allen erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vorlagen (siehe auch Kapitel 2.3.4). Daher ist nicht auszuschließen, dass weitere Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen. Die in den nachfolgenden Tabellen dargestellten Effektschätzer (RR inkl. 95%-KI) und p-Werte für die UE erfolgten post hoc und waren nicht stratifiziert.

Tabelle 19: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der verblindeten Behandlungsphase; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)

Balance Zusammenfassung der UE	Olezarsen N = 22 n (%)	Placebo N = 23 n (%)	RR [95%-KI]; p-Wert¹⁾
<i>Personen mit mindestens einem ...</i>			
UE (ergänzend dargestellt)	19 (86,4)	22 (95,7)	-
schwere UE (CTCAE-Grad ≥ 3)	1 (4,50)	9 (39,1)	0,12 [0,02; 0,84]; 0,01
SUE	3 (13,6)	9 (39,1)	0,35 [0,11; 1,12]; 0,09
UE, das zum Abbruch der Studienmedikation führte	2 (9,10)	0 (0)	5,22 [0,26; 102,9]; 0,23

¹⁾ Berechnung mittels exaktem Test nach Fisher. Die Stratifikationsfaktoren wurden nicht berücksichtigt.

Abkürzungen: - : nicht berechenbar; CTCAE: Common Terminology Criteria for Adverse Events; DS: Datenschnitt; KI: Konfidenzintervall; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Unerwünschte Ereignisse

Tabelle 20: UE mit Inzidenz $\geq 10\%$; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)

Studie Balance UE MedDRA-Systemorganklasse Preferred Term	Olezarsen N = 22 n (%)	Placebo N = 23 n (%)	RR [95%-KI]; p-Wert ¹⁾
Infektionen und parasitäre Erkrankungen	8 (36,4)	13 (56,5)	0,64 [0,33; 1,24]; 0,24
COVID-19	3 (13,6)	8 (34,8)	0,39 [0,12; 1,29]; 0,17
Cellulitis	0 (0)	3 (13,0)	0,15 [0,01; 2,73]; 0,23
Erkrankungen des Gastrointestinaltrakts	10 (45,5)	16 (69,6)	0,65 [0,38; 1,11]; 0,14
Abdominalschmerzen	4 (18,2)	8 (34,8)	0,52 [0,18; 1,49]; 0,31
Diarrhö	2 (9,10)	6 (26,1)	0,35 [0,08; 1,55]; 0,24
Pankreatitis	1 (4,50)	4 (17,4)	0,26 [0,03; 2,16]; 0,35
Akute Pankreatitis	0 (0)	3 (13,0)	0,15 [0,01; 2,73]; 0,23
Allgemeine Erkrankungen und Beschwerden an der Injektionsstelle	8 (36,4)	8 (34,8)	1,05 [0,48; 2,30]; 1,00
Fatigue	1 (4,50)	4 (17,4)	0,26 [0,03; 2,16]; 0,35
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	7 (31,8)	4 (17,4)	1,83 [0,62; 5,39]; 0,31
Schmerzen in den Extremitäten	3 (13,6)	1 (4,30)	3,14 [0,35; 27,9]; 0,35
Erkrankungen des Nervensystems	3 (13,6)	6 (26,1)	0,52 [0,15; 1,84]; 0,46
Kopfschmerzen	1 (4,50)	3 (13,0)	0,35 [0,04; 3,10]; 0,61
Erkrankungen der Atemwege, des Brustraums und Mediastinums	5 (22,7)	5 (21,7)	1,05 [0,35; 3,12]; 1,00
Verletzungen, Vergiftungen und Komplikationen bei Eingriffen	2 (9,10)	3 (13,0)	0,70 [0,13; 3,78]; 1,00
Untersuchungen	5 (22,7)	4 (17,4)	1,31 [0,40; 4,24]; 0,72
Reduktion der Thrombozytenzahl	3 (13,6)	1 (4,30)	3,14 [0,35; 27,9]; 0,35
Erkrankungen des Blutes und des lymphatischen Systems	3 (13,6)	2 (8,70)	1,57 [0,29; 8,51]; 0,67
Erkrankungen der Haut und des Unterhautgewebes	1 (4,50)	5 (21,7)	0,21 [0,03; 1,65]; 0,19
Stoffwechsel- und Ernährungsstörungen	2 (9,10)	5 (21,7)	0,42 [0,09; 1,94]; 0,41

¹⁾ Berechnung mittels exaktem Test nach Fisher. Die Stratifikationsfaktoren wurden nicht berücksichtigt.

Abkürzungen: DS: Datenschnitt; MedDRA: Medical Dictionary for Regulatory Activities; KI: Konfidenzintervall; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

Schwere unerwünschte Ereignisse (CTCAE-Grad ≥ 3)

Tabelle 21: Schwere UE (CTCAE-Grad ≥ 3), die bei ≥ 5 % der Personen in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)

Studie Balance Schwere UE MedDRA-Systemorganklasse Preferred Term	Olezarsen N = 22 n (%)	Placebo N = 23 n (%)	RR [95%-KI]; p-Wert ¹⁾
Erkrankungen des Gastrointestinaltrakts	1 (4,50)	7 (30,4)	0,15 [0,02; 1,12]; 0,047
Akute Pankreatitis	0 (0)	3 (13,0)	0,15 [0,01; 2,73]; 0,23
Nekrotisierende Pankreatitis	0 (0)	2 (8,70)	0,21 [0,01; 4,12]; 0,49
Nieren- und Harnwegserkrankungen	0 (0)	2 (8,70)	0,21 [0,01; 4,12]; 0,49
Proteinurie	0 (0)	2 (8,70)	0,21 [0,01; 4,12]; 0,49

¹⁾ Berechnung mittels exaktem Test nach Fisher. Die Stratifikationsfaktoren wurden nicht berücksichtigt.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; DS: Datenschnitt; MedDRA: Medical Dictionary for Regulatory Activities; KI: Konfidenzintervall; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

Schwerwiegende unerwünschte Ereignisse

Tabelle 22: SUE, die bei ≥ 5 % der Personen in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie Balance, Sicherheitspopulation (DS: 14.07.2023)

Studie Balance SUE MedDRA-Systemorganklasse Preferred Term	Olezarsen N = 22 n (%)	Placebo N = 23 n (%)	RR [95%-KI]; p-Wert ¹⁾
Erkrankungen des Gastrointestinaltrakts	2 (9,1)	9 (39,1)	0,23 [0,06; 0,96]; 0,04
Pankreatitis	1 (4,50)	4 (17,4)	0,26 [0,03; 2,16]; 0,35
Akute Pankreatitis	0 (0)	3 (13,0)	0,15 [0,01; 2,73]; 0,23
Nekrotisierende Pankreatitis	0 (0)	2 (8,70)	0,21 [0,01; 4,12]; 0,49
Stoffwechsel- und Ernährungsstörungen	0 (0)	2 (8,70)	0,21 [0,01; 4,12]; 0,49

¹⁾ Berechnung mittels exaktem Test nach Fisher. Die Stratifikationsfaktoren wurden nicht berücksichtigt.

Abkürzungen: DS: Datenschnitt; MedDRA: Medical Dictionary for Regulatory Activities; KI: Konfidenzintervall; RR: Relatives Risiko; SUE: Schwerwiegendes unerwünschtes Ereignis.

Major Adverse Cardiovascular Event

Die Zuordnung des Endpunkts „MACE“ zur Endpunktkategorie „Sicherheit“ ist unklar. Eine Erfassung von Ereignissen, die mit der Grunderkrankung einhergehen, kann vorliegend nicht ausgeschlossen werden (siehe auch Kapitel 2.3.4).

In der Studie Balance wurde jeweils ein MACE je Behandlungsarm berichtet. Welches Ereignis jeweils vorlag, wurde nicht berichtet.

4 Diskussion der Methodik und Ergebnisse

4.1 Design und Methodik der Studie

Studiendesign

Bei der pivotalen Studie Balance (ISIS-678354-CS3) handelt es sich um eine randomisierte, doppelblinde, Placebo-kontrollierte, multizentrische Phase-III-Studie zur Untersuchung der Wirksamkeit und Sicherheit von Olezarsen bei erwachsenen Patientinnen und Patienten mit genetisch bestätigtem familiärem Chylomikronämie-Syndrom (FCS).

Das Studiendesign umfasste zunächst eine Screening-Phase (4–8 Wochen), einschließlich einer mindestens 2-wöchigen Phase zur Diätstabilisierung für Patientinnen und Patienten, die zuvor keine stabile Diät einhielten (≤ 20 g Fett pro Tag). Anschließend folgte eine stratifizierte Randomisierung anhand von 2 Faktoren („Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“). Die Randomisierung erfolgte gestaffelt: Vorgesehen war zuerst eine Randomisierung auf 2 Kohorten A und B (1 : 1) mit jeweils anschließender anteiliger Randomisierung (2 : 1) auf eine Olezarsen-Dosierung und eine Placebo-Dosierung. Daraus resultierten 4 Behandlungsgruppen: Olezarsen 50 mg und Placebo 0,5 ml sowie Olezarsen 80 mg und Placebo 0,8 ml. Für die Nutzenbewertung ist die Fachinformation-konforme Olezarsen-Dosierung mit 80 mg relevant. Die Auswertung erfolgte präspezifiziert gegenüber den gepoolten Placebo-Armen (0,5 ml + 0,8 ml; siehe Kapitel 4.3). Nach Randomisierung traten die Patientinnen und Patienten in eine 53-wöchige doppelblinde Behandlungsphase ein, gefolgt von einer 13-wöchigen Nachbeobachtung oder alternativ dem Übergang in die offene, einarmige Extensionsstudie (ISIS-678354-CS13). Gemäß Angaben des pU wurde die Studie Balance planmäßig beendet.

Für die Nutzenbewertung werden die vergleichenden Ergebnisse der randomisierten Studie Balance von Olezarsen (80 mg) gegenüber Placebo (gepoolt) mit einer Behandlungsdauer von 53 Wochen herangezogen. Für die Nutzenbewertung wurden 2 Datenschnitte vorgelegt: ein Datenschnitt vom 14.07.2023 zum Ende der Behandlungsphase für alle Wirksamkeitsendpunkte (außer „Bestätigte akute Pankreatitis“) und Sicherheitsendpunkte sowie ein korrigierter Datenschnitt vom 30.11.2023 für den Endpunkt „Bestätigte akute Pankreatitis“ (siehe Kapitel 2.4). Beide Datenschnitte werden zur Nutzenbewertung herangezogen. Dabei liegen jedoch nicht nachvollziehbare Angaben des pU zur Vollständigkeit der Sicherheitsdaten vor. Einerseits wird vom pU beschrieben, dass die Sicherheitsdaten zum Datenschnitt vom 14.07.2023 korrekt und vollständig seien, andererseits sei eine Aktualisierung des Studienberichts geplant, der die nach Ende der Behandlungsphase gesammelten Daten der Nachbeobachtung berücksichtigt. Die Erhebung der Sicherheit sollte für Personen, die nicht in die offene Extensionsstudie übergangen, 13 Wochen über das Behandlungsende hinaus erfolgen. Ein weiterer Datenschnitt oder Studienbericht wurde jedoch nicht vorgelegt. Die letzte Behandlung in der Studie erfolgte am 20.06.2023, etwa 4 Wochen vor dem Datenschnitt. Demnach ist unklar, ob die vorgelegten Sicherheitsdaten zum Datenschnitt vom 14.07.2023 vollständig sind. Auf Basis der Angaben zu Beendigung bzw. Abbruch der Nachbeobachtung (siehe Tabelle 9) wird vorliegend jedoch davon ausgegangen, dass sich zum Datenschnitt vom 14.07.2023 noch maximal 1 Person pro Behandlungsarm in der Nachbeobachtung befand. Darauf basierend ergeben sich keine Auswirkungen für die Höhe des Verzerrungspotentials der Endpunktkategorien „Mortalität“ und „Sicherheit“.

Studienpopulation

Die Studie Balance umfasst Patientinnen und Patienten im Alter von mindestens 18 Jahren mit genetischem FCS und einer homozygoten, compound-heterozygoten oder doppelt-heterozygoten Mutation sowie einem Funktionsverlust im Typ-1-verursachenden Gen.

Von den insgesamt 144 gescreenten Patientinnen und Patienten aus 29 Studienzentren wurden 66 Personen in 4 Behandlungsarme eingeschlossen: 2 Olezarsen-Dosierungen und 2 dazugehörige Placebo-Dosierungen. Diese gingen aus einer gestaffelten Randomisierung hervor (siehe Abschnitt „Studiendesign“), wobei für die Placebo-Arme nur gepoolte Angaben vorliegen. Die für die Nutzenbewertung relevanten Behandlungsarme sind Olezarsen 80 mg (n = 22) und die beiden gepoolten Placebo-Arme (n = 23). Da die Randomisierung zuerst in 2 Kohorten und anschließend in 4 Behandlungsarme (jeweils im Verhältnis 2 : 1 (Olezarsen : Placebo)) erfolgte und für die finale Auswertung 3 Behandlungsarme berücksichtigt wurden (Olezarsen 80 mg und die beiden gepoolten Placebo-Arme), wurden vermutlich etwa 50 % des gepoolten Placebo-Arms nicht direkt mit dem für die Nutzenbewertung relevanten Behandlungsarm mit Olezarsen 80 mg stratifiziert randomisiert. Die Auswirkungen des gestaffelten Vorgehens bei der Randomisierung sind unklar. In Verbindung mit der geringen Stichprobengröße kann dies zu beobachtbaren und nicht beobachtbaren (da nicht erhoben) Imbalancen in den Baseline-Charakteristika geführt haben. Zwar liegen Imbalancen in manchen Baseline-Charakteristika vor, wie z. B. bei dem Stratifikationsfaktor „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening“ (77,3 % im Olezarsen-Arm vs. 65,2 % im Placebo-Arm), bei der Anzahl der Pankreatitis-Episoden in den letzten 10 Jahren (im Median 1 vs. 2), bei der Chronifizierung der Pankreatitis (36,4 % vs. 26,1 %), bei den Nüchtern-TG-Werten zu Baseline (im Median 2.086 mg/dl vs. 2.493 mg/dl) oder bei der Abstammung (z. B. kaukasisch: 77,3 % vs. 95,7 %). Da jedoch keine systematische einseitige Ungleichverteilung relevanter Prognosefaktoren für einen Behandlungsarm erkennbar ist und die beobachteten Unterschiede in Anbetracht der Stichprobengröße nicht als groß eingeschätzt werden, folgt daraus ein unklares Verzerrungspotential auf Studienebene.

Gemäß Einschlusskriterien sollte der überwiegende Anteil an Patientinnen und Patienten eine Pankreatitis-Vorgeschichte innerhalb von 10 Jahren vor dem Screening aufweisen. Gezählt wurden dabei dokumentierte Diagnosen akuter Pankreatitiden oder Hospitalisierungen wegen starker Bauchschmerzen, die mit einer AP übereinstimmten und für die keine alternative Diagnose gestellt wurde. Die Rekrutierung von Patientinnen und Patienten ohne vorherige Pankreatitis-Anamnese war zulässig, aber auf maximal 35 % der geplanten Population begrenzt. Diese Begrenzung lag bei ursprünglich 20 %, wurde jedoch mit Amendment 6 des Studienprotokolls vom 25.10.2021, rund 10 Monate nachdem die erste Person die erste Dosis der Studienmedikation erhielt, auf 35 % erhöht. Insgesamt hatten zu Baseline 17 Personen (77,3 %) im Olezarsen-Arm und 15 Personen (65,2 %) im Placebo-Arm eine Vorgeschichte bzgl. Pankreatitis-Episoden, die eine Hospitalisierung innerhalb von 10 Jahren vor dem Screening erforderlich machten: im Median wurde 1 Pankreatitis-Episode (min; max: 0; 25) im Olezarsen-Arm berichtet und 2 Pankreatitis-Episoden (min; max: 0; 78) im Placebo-Arm. Zudem weisen 8 Personen (36,4 %) im Olezarsen-Arm und 6 Personen (26,1 %) im Placebo-Arm eine bekannte chronische Pankreatitis auf. Patientinnen und Patienten mit einer aktiven Pankreatitis innerhalb von 4 Wochen vor dem Screening waren von der Studienteilnahme ausgeschlossen. „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen“ (jeweils ja vs. nein) sind die Stratifikationsfaktoren der Studie Balance. Ein weiteres Einschlusskriterium war ein Nüchtern-TG-Wert ≥ 880 mg/dl (10 mmol/l) zum Screening. Die Patientinnen und Patienten beider Studienarme wiesen zu Baseline stark erhöhte TG-Werte von im Median 2.086 mg/dl im Olezarsen-Arm bzw. 2.493 mg/dl im Placebo-Arm auf. Da bei TG-Werten > 1.000 mg/dl ein erhöhtes Risiko für AP besteht [20], das bei Werten > 2.000 mg/dl weiter zunimmt [6], ist die betrachtete Population auch unabhängig von der Pankreatitis-Vorgeschichte als Hochrisiko-

population für eine AP zu betrachten. Aufgrund der Heterogenität der Erkrankung und der vorliegenden Studienpopulation ist dies jedoch nicht gleichbedeutend mit einer hohen Symptomlast. Insgesamt wird von den Patientinnen und Patienten zu Baseline eine geringe Symptomlast berichtet, wobei einige Personen zu Baseline keinerlei Symptome im jeweiligen Endpunkt berichten.

Personen mit klinisch bedeutsamen Auffälligkeiten in der Anamnese oder der körperlichen Untersuchung sollten zudem ausgeschlossen werden. Als Beispiele wurden ein früheres akutes Koronarsyndrom innerhalb von 6 Monaten vor dem Screening oder größere chirurgische Eingriffe innerhalb von 3 Monaten vor dem Screening genannt. Die Beispiele können zwar nachvollzogen werden, jedoch kann aufgrund einer fehlenden Definition von „Klinische bedeutsame Auffälligkeiten“ nicht abschließend eingeschätzt werden, inwiefern die Studienpopulation insgesamt eingeschränkt wurde.

Studienmedikation

Olezarsen (80 mg bzw. 50 mg) wurde subkutan einmal monatlich von Woche 1 bis Woche 49 verabreicht. Die Placebo-Verabreichung war vergleichbar, nur ohne den aktiven Wirkstoff. Die Patientinnen und Patienten mussten eine Diät mit ≤ 20 g Fett pro Tag einhalten und durften die zuvor erhaltene lipidsenkende Therapie wie Statine, Omega-3-Fettsäuren (verschreibungspflichtig und nicht-verschreibungspflichtig), Fibrate oder andere lipidsenkende Medikamente in gleicher Dossierung fortführen (siehe Tabelle 11). Dabei ist auffällig, dass der Anteil der Patientinnen und Patienten, die Omega-3-Fettsäuren begleitend erhielten, im Olezarsen-Arm um etwa 15 % höher war als im Placebo-Arm. Laut EMA ist es jedoch unwahrscheinlich, dass die Gabe lipidsenkender Medikamente einen Einfluss auf die therapeutische Wirkung von Olezarsen bei Patientinnen und Patienten mit FCS hat [13]. Entsprechend ergeben sich anhand der dokumentierten Begleitmedikation keine verzerrenden Aspekte für die Bewertung des Zusatznutzens. Mit 3 Personen (13,6 %) aus dem Olezarsen-Arm und 1 Person (4,3 %) aus dem Placebo-Arm war der Anteil der Personen mit einem Abbruch der Studienmedikation zwar relativ gering, jedoch beträgt der Unterschied zwischen den beiden Behandlungsarmen annähernd 10 %. Gründe für den Abbruch der Studienmedikation im Olezarsen-Arm waren bei 2 Personen (9,1 %) ein UE sowie ein freiwilliger Studienaustritt. Der Abbruch der Studienmedikation im Placebo-Arm erfolgte durch Entscheidung des Prüfpersonals, da die Person nicht verfügbar für die Nachbeobachtung war. Die mediane Behandlungsdauer mit je 52 Wochen und die Anzahl verabreichter Injektionen war zwischen den betrachteten Behandlungsarmen identisch. Die Nachbeobachtungsdauer wurde jedoch nicht vorgelegt.

4.2 Zulassungsstatus und Zulassungspopulation

Zulassungsstatus und Anwendungsgebiet

Olezarsen ist zugelassen zur Behandlung von erwachsenen Patientinnen und Patienten mit genetisch bestätigtem FCS ergänzend zu einer Diät. Gemäß Fachinformation beträgt die zugelassene Dosierung 80 mg, verabreicht als subkutane Injektion einmal monatlich [23]. Die Dosierung von Olezarsen in Kohorte A (Olezarsen 50 mg) entspricht laut Fachinformation nicht der zugelassenen Dosierung und wurde daher nicht für die vorliegende Nutzenbewertung berücksichtigt.

Das Anwendungsgebiet scheint weitestgehend durch die in die Studie Balance eingeschlossenen Patientinnen und Patienten abgedeckt. Jedoch ist zu berücksichtigen, dass aufgrund der Ausschlusskriterien die eingeschlossenen Personen zu Baseline vermutlich keine aktive Pankreatitis aufwiesen. Da eine Definition von „Klinisch bedeutsame Auffälligkeiten“ (siehe Abschnitt „Studienpopulation“ in Kapitel 4.1) fehlt, kann zudem nicht abschließend eingeschätzt werden, inwiefern die Studienpopulation insgesamt eingeschränkt wurde.

Übertragbarkeit auf deutschen Versorgungskontext

Bei der Studie Balance handelt es sich um eine multizentrische Studie, bei der 63,6 % der Personen im Olezarsen-Arm und 47,8 % der Personen im Placebo-Arm an Studienzentren in Europa behandelt wurden, jedoch an keinem Studienzentrum in Deutschland. Weitere eingeschlossene Patientinnen und Patienten wurden in Nordamerika (Kanada und USA) behandelt, wobei in den USA gleichverteilt zwischen den Studienarmen ein knappes Drittel der Studienpopulation eingeschlossen wurde. Regional sind je nach Versorgungskontext Unterschiede zur Notwendigkeit und Möglichkeit eines Besuchs in der Notaufnahme bzw. einer Hospitalisierung gegenüber dem deutschen Versorgungskontext möglich. Diese werden vorliegend jedoch als nicht schwerwiegend eingeschätzt.

In der Gesamtschau wird für den Großteil der Studienpopulation überwiegend ein vergleichbarer Versorgungskontext angenommen. Unklar ist jedoch, ob die Gabe von Placebo im Kontrollarm übertragbar auf den deutschen Versorgungskontext ist, da im Anwendungsgebiet das Arzneimittel Volanesorsen zugelassen ist.

4.3 Zusammenfassende Beurteilung zum Verzerrungspotential und zur Ergebnissicherheit

Das Verzerrungspotential auf Studienebene wird als unklar eingeschätzt. Hauptsächlich basiert das unklare Verzerrungspotential auf den Unsicherheiten hinsichtlich des gestaffelten Vorgehens bei der Randomisierung in Verbindung mit einer geringen Stichprobengröße und einer Poolung der Placebo-Arme. Zudem sind fehlende Angaben, inwiefern die Diätvorgaben während des Studienverlaufs in den jeweiligen Behandlungsarmen eingehalten wurden, zu berücksichtigen.

In Bezug auf die Ergebnissicherheit ist einschränkend die geringe Größe der Studienpopulation (N = 22 vs. 23) zu berücksichtigen. Zudem ist die externe Validität der Ergebnisse für die Gruppe der über 65-Jährigen eingeschränkt, da nur eine geringe Anzahl an Personen dieser Alterskohorte eingeschlossen wurde (n = 5). Eine ähnliche Limitation ergibt sich durch den Ausschluss von Patientinnen und Patienten mit aktiver Pankreatitis innerhalb von 4 Wochen vor dem Screening. Zuletzt ist die unklare Einschränkung der Studienpopulation durch das nicht eindeutig definierte Ausschlusskriterium „Klinisch bedeutsame Auffälligkeiten“ zu berücksichtigen. Trotz der benannten Limitationen wird vorliegend jedoch nicht von erheblichen Einschränkungen der externen Validität der Ergebnisse durch die genannten Populationseingrenzungen ausgegangen.

4.4 Mortalität

Todesfälle wurden im Rahmen der Sicherheit erfasst. Im Verlauf der 53-wöchigen randomisierten Behandlungsphase der Studie Balance ist keine Person verstorben. Das Verzerrungspotential wird aufgrund von fehlenden Informationen zur Nachbeobachtungsdauer in beiden Studienarmen als unklar eingeschätzt (siehe Tabelle 8).

4.5 Morbidität

Für die Endpunktkategorie „Morbidität“ legte der pU Daten zu den Endpunkten „Nüchtern-Triglyzeridwert“ (primärer Endpunkt), „Bestätigte akute Pankreatitis“ und „Krankenhausaufenthalte“ sowie zu den patientenberichteten Endpunkten „PGIS“, „PGIC“ und „FCS-SIS (Symptom)“ vor. Die detaillierte Bewertung der Endpunkte ist Kapitel 2.3 zu entnehmen. Die Ergebnisse der Endpunkte „Bestätigte akute Pankreatitis“ und „FCS-SIS (Symptom)“ wurden für die Nutzenbewertung herangezogen. Für den Endpunkt „PGIS“ wurde eine für die Nutzenbewertung als relevanter erachtete Responderanalyse zur Verschlechterung nicht vorgelegt; der Endpunkt und die vorliegende Responderanalyse hinsichtlich einer Verbesserung werden daher ergänzend darge-

stellt. Ergebnisse zum PGIC sind dem Anhang zu entnehmen. Der primäre Endpunkt „Nüchtern-Triglyzeridwert“ wird als nicht unmittelbar patientenrelevant eingeschätzt. Da es sich jedoch um einen klinisch relevanten Parameter für FCS handelt, welcher der Therapiesteuerung dienen kann, wird der Endpunkt somit ergänzend dargestellt.

Bei hohem Verzerrungspotential liegt für den ergänzend dargestellten Endpunkt „Nüchtern-Triglyzeridwert“ ein statistisch signifikanter Unterschied für die prozentuale Veränderung zu Monat 12 zugunsten von Olezarsen gegenüber Placebo vor (LS-Mean-Differenz: -59,3 (95%-KI: [-90,7; -28,1])). Dabei weisen die Patientinnen und Patienten in beiden Behandlungsarmen nach 12-monatiger Behandlung weiterhin stark erhöhte mittlere TG-Werte auf (Olezarsen: 1.590 mg/dl; Placebo: 2.959 mg/dl), die mit einem erhöhten Risiko für AP assoziiert sind. Es ist jedoch zu berücksichtigen, dass die TG-Werte im Olezarsen-Arm zu Monat 12 unter 2.000 mg/dl liegen, während sie im Placebo-Arm annähernd doppelt so hoch sind wie im Olezarsen-Arm. Gemäß Berglund et al. steigt das Risiko einer Pankreatitis deutlich an, wenn die TG-Werte über 2.000 mg/dl liegen [6]. Bei der Analyse des primären Endpunkts „Nüchtern-Triglyzeridwert“ wurden die fehlenden Werte zu Monat 12 unter Verwendung eines J2R-Ansatzes bzw. unter MAR-Annahme imputiert. Da mittels der Sensitivitätsanalyse (Tipping-Point-Analyse) die Robustheit des berichteten Effekts nachgewiesen werden konnte, ist die Ergebnissicherheit trotz des erhöhten Verzerrungspotentials nicht eingeschränkt.

Im Endpunkt „Bestätigte akute Pankreatitis“ zeigt sich ein statistisch signifikanter Unterschied zugunsten von Olezarsen in den post hoc durchgeführten Responderanalysen der AP-Inzidenz (RR: 0,68 (95%-KI: [0,49; 0,95])). In der präspezifizierten Analyse der mittleren Ereignisrate konnte kein signifikanter Unterschied zwischen den Behandlungsarmen festgestellt werden. Das Verzerrungspotential des Endpunkts wird aufgrund möglicher informativer Zensurierung als hoch bewertet, da insgesamt 3 Personen (13,6 %) im Olezarsen-Arm gegenüber 1 Person (4,30 %) im Placebo-Arm die Behandlung abbrachen und somit möglicherweise nicht bzw. nicht vollständig in die Auswertungen eingingen. Angaben zu Studienabbrüchen fehlen.

In der post hoc durchgeführten Responderanalyse des Endpunkts „FCS-SIS (Symptom)“ zeigte sich hinsichtlich der Verschlechterung um mindestens 15 % der Skalenspannweite bis Monat 12 kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen. Das Verzerrungspotential des Endpunkts wird aufgrund differenzieller Rücklaufquoten von annähernd 10 Prozentpunkten als hoch eingestuft.

4.6 Lebensqualität

Für die Endpunktkategorie (gesundheitsbezogene) „Lebensqualität“ wurden Daten für die patientenberichteten Instrumente „FCS-SIS (Impact)“, „PROMIS-29+2“ und „PROMIS Short Form – Pain Interference 8a“ vorgelegt. Der „FCS-SIS (Impact)“ und „PROMIS-29+2“ wurden für die Nutzenbewertung herangezogen. Der Endpunkt „PROMIS Short Form – Pain Interference 8a“ wurde aufgrund einer Mehrfacherfassung (Hälfte der Fragebogen-Items) mit dem PROMIS-29+2 ergänzend dargestellt. Die detaillierte Bewertung der Endpunkte ist in Kapitel 2.3.3 abgebildet.

In der post hoc durchgeführten Responderanalyse des Endpunkts „FCS-SIS (Impact)“ zeigte sich hinsichtlich der Verschlechterung um mindestens 15 % der Skalenspannweite bis Monat 12 kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen. Das Verzerrungspotential des Endpunkts wird aufgrund differenzieller Rücklaufquoten von > 10 Prozentpunkten als hoch eingestuft.

Ebenso konnten bezüglich der Veränderung in den Domänen des PROMIS-29+2 zu Monat 12 (ANCOVA) keine signifikanten Unterschiede zwischen den Studienarmen aufgezeigt werden. Auswertungen bzgl. der Veränderung in der Domäne „Schmerzintensität“ sowie des körperlichen

bzw. psychischen Summenscores (kontinuierliche Auswertung bzw. Responderanalysen) lagen mit den eingereichten Studienunterlagen nicht vor. Das Verzerrungspotential wird als hoch bewertet, da die Rücklaufquoten zur Veränderung von Monat 12 gegenüber Baseline bei allen Domänen – außer der Domäne „Schmerzbeeinträchtigung“ – bei 68,2 % im Olezarsen-Arm und ca. 80 % im Placebo-Arm liegen. In der Domäne „Schmerzbeeinträchtigung“ liegen die Rücklaufquoten mit ca. 82 % im Olezarsen-Arm und ca. 92 % im Placebo-Arm zwar höher, jedoch besteht auch hier ein Unterschied von knapp bzw. mehr als 10 Prozentpunkten zwischen den beiden Studienarmen.

4.7 Sicherheit

Die mediane Behandlungsdauer betrug in beiden Behandlungsarmen 52 Wochen (min; max: Olezarsen: 4,0; 57,1 vs. Placebo: 28,4; 52,6). Angaben zu den Beobachtungsdauern liegen jedoch nicht vor. Das Verzerrungspotential wird, aufgrund von fehlenden Informationen zur Nachbeobachtungsdauer in den beiden Studienarmen, als unklar eingeschätzt (siehe Tabelle 8). Der post hoc berichtete Effektschätzer (RR) und p-Wert erfolgte zudem unstratifiziert.

In der Studie Balance trat bei 19 Personen (86,4 %) im Olezarsen-Arm und bei 22 (95,7 %) im Placebo-Arm mindestens ein UE auf. Schwere UE wurden bei 1 Person (4,50 %) im Olezarsen-Arm und bei 9 Personen (39,1 %) im Placebo-Arm erfasst, wobei es sich um einen statistisch signifikanten Unterschied zugunsten von Olezarsen handelt (RR: 0,12 (95%-KI: [0,02; 0,84])). SUE traten bei 3 Personen (13,6 %) im Olezarsen-Arm und bei 9 (39,1 %) im Placebo-Arm auf. Während im Olezarsen-Arm 2 Personen (9,10 %) die Studienmedikation aufgrund von UE abbrachen, betraf dies im Placebo-Arm keine Person.

Für jegliche UE zeigten sich auf Ebene von SOC und PT keine signifikanten Unterschiede zwischen den Behandlungsarmen. Die häufigsten UE fanden sich in den SOC „Erkrankungen des Gastrointestinaltrakts“ (Olezarsen: 45,5 % vs. Placebo: 69,6 %), „Infektionen und parasitäre Erkrankungen“ (36,4 % vs. 56,5 %) und „Allgemeine Erkrankungen und Beschwerden an der Injektionsstelle“ (36,4 % vs. 34,8 %). Das häufigste UE auf PT-Ebene war in beiden Behandlungsarmen „Abdominalschmerzen“ bei 4 Personen (18,2 %) im Olezarsen-Arm und 8 Personen (34,8 %) im Placebo-Arm.

Bei den schweren UE zeigte sich auf Ebene von SOC und PT ein statistisch signifikanter Unterschied zugunsten von Olezarsen in der SOC „Erkrankungen des Gastrointestinaltrakts“ (RR: 0,15 (95%-KI: [0,02; 1,12])), wobei jeweils im Placebo-Arm 3 Personen (13 %) eine „Akute Pankreatitis“ und 2 Personen (8,70 %) eine „Nekrotisierende Pankreatitis“ aufwiesen. Ähnliches zeigt sich für die SUE, hier zeigte sich ebenfalls ein statistisch signifikanter Unterschied zugunsten von Olezarsen in der SOC „Erkrankungen des Gastrointestinaltrakts“ (RR: 0,23 (95%-KI: [0,06; 0,96])), wobei jeweils im Placebo-Arm 4 Personen (17,4 %) eine „Pankreatitis“, 3 Personen (13,0 %) eine „Akute Pankreatitis“ und 2 Personen (8,70 %) eine „Nekrotisierende Pankreatitis“ aufwiesen. Bei den berichteten PT der schweren UE und SUE ist zu berücksichtigen, dass diese Ereignisse auch Teil der Grunderkrankung sind.

Der Endpunkt „MACE“ wurde im Rahmen der Studie Balance in der Endpunktkategorie „Sicherheit“ erfasst. Eine Rationale für die Zuordnung liegt jedoch nicht vor. Eine Zuordnung zur Endpunktkategorie „Morbidity“ ist denkbar, kann jedoch nicht abschließend bewertet werden. Je Behandlungsarm wurde 1 Ereignis berichtet. Es kann nicht ausgeschlossen werden, dass es sich dabei um Ereignisse der Grunderkrankung handelt.

Die EMA merkt im „European Public Assessment Report“ (EPAR) studienübergreifend an, dass nur sehr wenige Patientinnen und Patienten über 65 Jahre an den Studien teilnahmen und dass die Studiendauer etwas weniger als 1 Jahr betrug. Daher ist die Langzeitsicherheit nicht ausreichend bekannt. Die wichtigsten Sicherheitsbedenken betrafen laut EPAR die Überempfindlichkeitsreaktionen. Insgesamt war die Häufigkeit von Nebenwirkungen bei Olezarsen und Placebo ähnlich,

es traten jedoch unter Olezarsen häufiger verringerte Thrombozytenwerte, Schmerzen in den Gliedmaßen, abdominale Beschwerden und Reaktionen an der Injektionsstelle auf. Hepatotoxizität, Nierentoxizität und Thrombozytopenie gelten gemäß EMA als potentiell wichtige Risiken. [13]

Zusammenfassend zeigten sich bei unklarem Verzerrungspotential in der Studie Balance statistisch signifikante Unterschiede zugunsten von Olezarsen gegenüber Placebo sowohl für die Gesamtraten der schweren UE als auch für die SOC „Erkrankungen des Gastrointestinaltrakts“ bezüglich schwerer UE und SUE. Es lagen jedoch keine Auswertungen vor, bei denen mögliche erkrankungsbezogene Ereignisse bzw. mögliche Ereignisse der Grunderkrankung rausgerechnet bzw. nicht berücksichtigt wurden. Bei einigen der berichteten SOC bzw. PT handelt es sich um Ereignisse, welche der Grunderkrankung zugeschrieben werden können. Die berichteten Effekte in der Sicherheit zugunsten von Olezarsen scheinen durch Aspekte der Morbidität getrieben. Die Effekte werden daher als nicht bewertbar eingestuft.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Olezarsen ist zugelassen für erwachsenen Patientinnen und Patienten ergänzend zu einer Diät zur Behandlung des genetisch bestätigten familiären Chylomikronämie-Syndroms (FCS). Die Nutzenbewertung von Olezarsen basiert auf der zulassungsbegründenden Studie Balance (ISIS-678354-CS3), eine multizentrische, doppelblinde, randomisiert kontrollierte Phase-III-Studie zur Untersuchung der Wirksamkeit und Sicherheit von Olezarsen gegenüber Placebo bei Patientinnen und Patienten mit genetisch bestätigtem familiärem Chylomikronämie-Syndrom (FCS) ergänzend zu einer Diät (≤ 20 g Fett pro Tag).

Die Ergebnisse der in der Nutzenbewertung berücksichtigten Endpunkte der Studie werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Olezarsen ein statistisch signifikant positiver Effekt bei hohem oder unklarem ($\uparrow$) oder niedrigem ($\uparrow\uparrow$) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem ($\downarrow$) oder niedrigem ($\downarrow\downarrow$) Verzerrungspotential bzw. kein Unterschied ($\leftrightarrow$) gezeigt werden konnte.

Weitergehende, über die Angaben in der nachfolgenden zusammenfassenden Tabelle hinausgehende Erläuterungen zur Methodik und den Ergebnissen der Endpunkte, finden sich in den jeweiligen Kapiteln 2.3 bis 2.5 und 3.2 bis 3.5 der Nutzenbewertung.

Tabelle 23: Zusammenfassende Darstellung der Ergebnisse der Studie Balance

Studie Balance Darstellung der Ergebnisse	Olezarsen N = 22			Placebo N = 23			Olezarsen vs. Placebo	Effekt
Mortalität								
Todesfälle	N ¹⁾	Personen mit Ereignis, n (%)		N ¹⁾	Personen mit Ereignis, n (%)		Effektschätzer [95%-KI]; p-Wert	
Gesamtüberleben	22	0 (0)		23	0 (0)		-	↔
Morbidität								
Bestätigte akute Pankreatitis ²⁾								
Bestätigte akute Pankreatitis bis Monat 12	N	Personen mit mind. einem Ereignis, n (%)	Gesamt- zahl an Ereignissen	N	Personen mit mind. einem Ereignis, n (%)	Gesamt- zahl an Ereignissen	RR ³⁾ [95%-KI]; p-Wert	
	22 ⁴⁾	1 (4,50)	1	23 ⁴⁾	7 (30,4)	11	0,68 [0,49; 0,95]; 0,009	↑
	N	Gesamtzahl der Patienten- jahre	Mittlere Ereignis- rate / 100 Patienten- jahre [95%-KI]	N	Gesamtzahl der Patienten- jahre	Mittlere Ereignis- rate / 100 Patienten- jahre [95%-KI]	MRR ⁵⁾ [95%-KI]; p-Wert	
	22	20,3	4,93 [0,61; 40,2]	23	22,9	36,1 [14,5; 89,9]	0,14 [0,01; 1,32]; 0,09	↔
FCS-SIS (Symptom) ⁶⁾	N ⁷⁾	Personen mit Ereignis, n (%)		N ⁷⁾	Personen mit Ereignis, n (%)		RR [95%-KI]; p-Wert ⁸⁾	
Verschlechterung um mind. 15 % der Skalenspannweite ⁹⁾	17	1 (5,90)		20	1 (5,0)		1,05 [0,91; 1,20]; 0,60	↔

Studie Balance Darstellung der Ergebnisse	Olezarsen N = 22			Placebo N = 23			Olezarsen vs. Placebo	Effekt
Lebensqualität								
FCS-SIS (Impact) ¹⁰⁾	N ⁷⁾	Personen mit Ereignis, n (%)		N ⁷⁾	Personen mit Ereignis, n (%)		RR [95%-KI]; p-Wert ⁸⁾	
Verschlechterung um mind. 15 % der Skalenspannweite ¹¹⁾	17	2 (11,8)		21	4 (19,0)		0,89 [0,66; 1,19]; 0,41	↔
PROMIS-29+2 ¹²⁾	N ¹³⁾	Baseline ¹⁴⁾ MW (SD)	Änderung Monat 12 ¹⁵⁾ zu Baseline LS Mean (SD)	N ¹³⁾	Baseline ¹⁴⁾ MW (SD)	Änderung Monat 12 ¹⁵⁾ zu Baseline LS Mean (SD)	LS-Mean- Differenz [95%-KI]; p-Wert	
Körperliche Funktionen ¹⁶⁾	15	47,8 (9,36)	1,67 (9,30)	18	48,5 (10,4)	0,66 (4,08)	0,25 [-4,07; 4,56]; 0,91	↔
Ängste / Sorgen ¹⁷⁾	15	58,3 (11,9)	-1,80 (9,83)	18	55,8 (11,6)	-0,33 (5,55)	-0,12 [-6,01; 5,78]; 0,97	↔
Depression ¹⁷⁾	15	51,9 (12,1)	0,30 (9,84)	18	51,8 (10,3)	-2,20 (7,67)	3,32 [-2,56; 9,19]; 0,26	↔
Müdigkeit ¹⁷⁾	15	51,2 (11,9)	0,17 (7,86)	18	51,7 (12,6)	-0,37 (3,76)	1,14 [-3,23; 5,51]; 0,60	↔
Schmerz- beeinträchtigung ¹⁷⁾	18	53,0 (12,6)	-2,81 (10,5)	21	52,5 (11,8)	-2,86 (7,94)	0,83 [-4,28; 5,94]; 0,75	↔
Schlafstörungen ¹⁷⁾	15	50,2 (8,25)	-2,86 (8,17)	18	49,0 (8,21)	1,84 (7,13)	-2,79 [-8,77; 3,19]; 0,35	↔
Fähigkeit zur Teil- nahme an sozialen Aktivitäten ¹⁶⁾	15	50,7 (10,6)	0,71 (7,48)	18	50,8 (12,4)	1,72 (6,20)	-1,88 [-6,66; 2,91]; 0,43	↔
Sicherheit ¹⁸⁾								
Unerwünschte Ereignisse	N ¹⁾	Personen mit mind. einem Ereignis, n (%)		N ¹⁾	Personen mit mind. einem Ereignis, n (%)		RR [95%-KI]; p-Wert ¹⁹⁾	
Schwere UE (CTCAE-Grad ≥ 3)	22	1 (4,50)		23	9 (39,1)		0,12 [0,02; 0,84]; 0,01	n.b. ²⁰⁾
SUE	22	3 (13,6)		23	9 (39,1)		0,35 [0,11; 1,12]; 0,09	↔
UE, das zum Abbruch der Studien- medikation führte	22	2 (9,10)		23	0 (0)		5,22 [0,26; 102,9]; 0,23	↔

Studie Balance Darstellung der Ergebnisse	Olezarsen N = 22		Placebo N = 23		Olezarsen vs. Placebo	Effekt
MACE ²¹⁾	22	1 (4,50)	23	1 (4,35)	k. A.	n. b. ²²⁾

¹⁾ Die Anzahl entspricht der Sicherheitspopulation.

²⁾ Als „Bestätigte akute Pankreatitis“ für die Auswertung galten dokumentierte, wahrscheinliche und mögliche AP gemäß der Einschätzung des „Acute Pancreatitis Adjudication Committee“ (siehe Kapitel 2.3.2).

³⁾ RR und p-Wert stratifiziert nach den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

⁴⁾ 3 Personen im Olezarsen-Arm (13,6 %) und 1 Person im Placebo-Arm (4,3 %) brachen die Behandlung bis Monat 12 ab und gingen nicht vollständig in die Auswertung ein.

⁵⁾ MRR, 95%-KI sowie p-Wert (Wald-Chi-Quadrat-Test) anhand Negativ-Binomial-Regressionsmodell mit Behandlungsgruppe und Vorbehandlung mit Volanesorsen (ja vs. nein) als Faktoren und der Anzahl bestätigter AP-Ereignisse innerhalb 5 Jahre vor Studienaufnahme als Kovariate.

⁶⁾ Die Skalenspannweite beträgt 0 bis 10; höhere Werte zeigen schwerere FCS-Symptome an.

⁷⁾ Rücklaufe zu Monat 12.

⁸⁾ RR und p-Wert stratifiziert nach den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Fehlende Daten wurden nicht imputiert. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

⁹⁾ Verschlechterung definiert als Erhöhung des Symptom-Gesamtwerts um ≥ 15 % der Skalenspannweite zu Monat 12.

¹⁰⁾ Die Skalenspannweite beträgt 0 bis 4; höhere Werte zeigen stärkere Belastung durch FCS an.

¹¹⁾ Verschlechterung definiert als Erhöhung des Impact-Gesamtwerts um ≥ 15 % der Skalenspannweite zu Monat 12.

¹²⁾ Für den körperlichen und psychischen Summenscore liegen keine Ergebnisse vor.

¹³⁾ Die Anzahl entspricht denjenigen Personen, welche bis Monat 12 in die Analyse der Veränderung eingehen.

¹⁴⁾ Baseline definiert als letzte nicht fehlende Messung vor Verabreichung der ersten Dosis des Studienmedikaments.

¹⁵⁾ Veränderung zu Monat 12 mittels ANCOVA mit Behandlungsgruppe und den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“ als feste Effekte sowie dem jeweiligen Domänenscore zu Baseline als Kovariate.

¹⁶⁾ Anhand t-transformierter Werte mit einem Mittelwert von 50 und einer Standardabweichung von 10; niedrigere Werte als 50 zeigen eine höhere Einschränkung an.

¹⁷⁾ Anhand t-transformierter Werte mit einem Mittelwert von 50 und einer Standardabweichung von 10; höhere Werte als 50 zeigen eine höhere Einschränkung an.

¹⁸⁾ Weitere Ergebnisse zur Sicherheit (UE nach Organsystemen und Einzelereignisse) sind in Kapitel 3.5 dargestellt.

¹⁹⁾ Berechnung mittels exaktem Test nach Fisher. Die Stratifikationsfaktoren wurden nicht berücksichtigt.

²⁰⁾ Bei einigen der berichteten SOC bzw. PT handelt es sich um Ereignisse, welche der Grunderkrankung zugeschrieben werden können. Die berichteten Effekte in der Sicherheit zugunsten von Olezarsen scheinen durch Aspekte der Morbidität getrieben. Der Vorteil wird daher als nicht bewertbar eingestuft.

²¹⁾ Für den Endpunkt „MACE“ ist die Zuordnung zur Endpunktkategorie „Sicherheit“ unklar. Aufgrund der Erhebung in der Studie Balance im Rahmen der Sicherheit wird diese für die Nutzenbewertung beibehalten (siehe Kapitel 2.3.4).

²²⁾ Es wurden keine Effektschätzer (mit dazugehörigem p-Wert) im Dossier vorgelegt. Aufgrund der Anzahl an Ereignissen wird von keinem statistisch signifikanten Unterschied zwischen den Behandlungsarmen ausgegangen.

Abkürzungen: ANCOVA: Analysis of Covariance; AP: Akute Pankreatitis; CTCAE: Common Terminology Criteria for Adverse Events; FCS: Familiäres Chylomikronämie-Syndrom; FCS-SIS: Familial Chylomicronemia Syndrome – Symptom and Impact Scale; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MACE: Major Adverse Cardiovascular Event; MRR: Mean Rate Ratio; MW: Mittelwert; n. b.: nicht bewertbar; PROMIS: Patient-Reported Outcomes Measurement Information System; PT: Preferred Term; RR: Relatives Risiko; SD: Standardabweichung; SOC: Systemorganklasse; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. PROMIS adult profile instruments; scoring manual [online]. Evanston (USA): Northwestern University Health Measures; 2025. [Zugriff: 16.02.2026].
URL: https://www.healthmeasures.net/images/PROMIS/manuals/Scoring_Manual_Only/PROMIS_Adult_Profile_Scoring_Manual_15July2025.pdf.
2. **Akcea Therapeutics Ireland**. Waylivra 285 mg Injektionslösung in einer Fertigspritze [online]. 11.2022. Frankfurt/Main. [Zugriff: 16.02.2026]. URL: <https://www.fachinfo.de/>.
3. **Amtmann D, Cook KF, Jensen MP, Chen WH, Choi S, Revicki D, et al.** Development of a PROMIS item bank to measure pain interference. Pain 2010;150(1):173-182. <https://dx.doi.org/10.1016/j.pain.2010.04.025>.
4. **Andrés AM, Mato AS.** Choosing the optimal unconditioned test for comparing two independent proportions. Computational Statistics & Data Analysis 1994;17(5):555-574. [https://dx.doi.org/10.1016/0167-9473\(94\)90148-1](https://dx.doi.org/10.1016/0167-9473(94)90148-1).
5. **Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, et al.** Classification of acute pancreatitis - 2012: revision of the Atlanta classification and definitions by international consensus. Gut 2013;62(1):102-111. <https://dx.doi.org/10.1136/gutjnl-2012-302779>.
6. **Berglund L, Brunzell JD, Goldberg AC, Goldberg IJ, Sacks F, Murad MH, et al.** Evaluation and treatment of hypertriglyceridemia: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2012;97(9):2969-2989. <https://dx.doi.org/10.1210/jc.2011-3213>.
7. **Brown TM, Vera-Llonch M, Kanu C, Sikora Kessler A, Yarlus A, Fehnel SE.** Cross-sectional quantitative evaluation of a novel patient-reported outcome measure in familial chylomicronemia syndrome. Patient Relat Outcome Meas 2024;15:45-59. <https://dx.doi.org/10.2147/PROM.S441583>.
8. **Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al.** PROMIS adult health profiles: efficient short-form measures of seven health domains. Value Health 2019;22(5):537-544. <https://dx.doi.org/10.1016/j.ival.2019.02.004>.
9. **Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al.** The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010;63(11):1179-1194. <https://dx.doi.org/10.1016/j.jclinepi.2010.04.011>.
10. **Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al.** The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH roadmap cooperative group during its first two years. Med Care 2007;45(5 Suppl 1):S3-S11. <https://dx.doi.org/10.1097/01.mlr.0000258615.42478.55>.
11. **Davidson D, Slota C, Vera-Llonch M, Brown TM, Hsieh A, Fehnel S.** Development of a novel PRO instrument for use in familial chylomicronemia syndrome. Journal of Patient-Reported Outcomes 2021;5(1):72. <https://dx.doi.org/10.1186/s41687-021-00347-5>.
12. **Davidson M, Stevenson M, Hsieh A, Ahmad Z, Roeters van Lennep J, Crowson C, et al.** The burden of familial chylomicronemia syndrome: results from the global IN-FOCUS study. J Clin Lipidol 2018;12(4):898-907 e892. <https://dx.doi.org/10.1016/j.jacl.2018.04.009>.

13. **European Medicines Agency (EMA).** Tryngolza (olezarsen): European public assessment report EMEA/H/C/006477/0000 [online]. 24.07.2025. Amsterdam (NED): EMA. [Zugriff: 01.12.2025]. URL: https://www.ema.europa.eu/en/documents/assessment-report/tryngolza-epar-public-assessment-report_en.pdf.
14. **Ionis Pharmaceuticals.** A randomized, double-blind, placebo-controlled, phase 3 study of AKCEA-APOCIII-L administered subcutaneously to patients with familial chylomicronemia syndrome (FCS), ISIS-678354-CS3; clinical study report [unveröffentlicht]. 15.03.2024.
15. **Ionis Pharmaceuticals.** A randomized, double-blind, placebo-controlled, phase 3 study of AKCEA-APOCIII-L administered subcutaneously to patients with familial chylomicronemia syndrome (FCS), ISIS-678354-CS3; clinical study report, additional analyses 2025 [unveröffentlicht]. 14.11.2025.
16. **Ionis Pharmaceuticals.** A randomized, double-blind, placebo-controlled, phase 3 study of AKCEA-APOCIII-L administered subcutaneously to patients with familial chylomicronemia syndrome (FCS), ISIS-678354-CS3; clinical study report, appendix 1.4: list and description of investigators and other important participants in the study [unveröffentlicht]. 15.03.2024.
17. **Ionis Pharmaceuticals.** A randomized, double-blind, placebo-controlled, phase 3 study of AKCEA-APOCIII-L administered subcutaneously to patients with familial chylomicronemia syndrome (FCS), ISIS-678354-CS3; statistical analysis plan, version 4.0 [unveröffentlicht]. 21.05.2020.
18. **Ionis Pharmaceuticals.** A randomized, double-blind, placebo-controlled, phase 3 study of AKCEA-APOCIII-L administered subcutaneously to patients with familial chylomicronemia syndrome (FCS); ISIS-678354-CS3; original protocol [unveröffentlicht]. 21.05.2020.
19. **Ionis Pharmaceuticals.** A randomized, double-blind, placebo-controlled, phase 3 study of AKCEA-APOCIII-L administered subcutaneously to patients with familial chylomicronemia syndrome (FCS); ISIS-678354-CS3; study protocol, amendment 9 [unveröffentlicht]. 02.08.2023.
20. **Kassner U, Dippel M, Steinhagen-Thiessen E.** Schwere Hypertriglyzeridämie: Diagnostik und neue Therapieprinzipien. Internist (Berl) 2017;58(8):866-876. <https://dx.doi.org/10.1007/s00108-017-0234-z>.
21. **Moulin P, Dufour R, Aversa M, Arca M, Cefalu AB, Noto D, et al.** Identification and diagnosis of patients with familial chylomicronaemia syndrome (FCS): expert panel recommendations and proposal of an "FCS score". Atherosclerosis 2018;275:265-272. <https://dx.doi.org/10.1016/j.atherosclerosis.2018.06.814>.
22. **Swedish Orphan Biovitrum.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 A: Olezarsen (Tryngolza), Anwendung bei erwachsenen Patienten ergänzend zu einer Diät zur Behandlung des genetisch bestätigten familiären Chylomikronämie-Syndroms (FCS); Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 28.11.2025.
23. **Swedish Orphan Biovitrum.** Tryngolza 80 mg Injektionslösung im Fertigpen [online]. 11.2025. Frankfurt/Main. [Zugriff: 01.12.2025]. URL: <https://www.fachinfo.de/>.
24. **Williams K, Tickler G, Valdivielso P, Alonso J, Vera-Llonch M, Cubells L, et al.** Symptoms and impacts of familial chylomicronemia syndrome: a qualitative study of the patient experience. Orphanet J Rare Dis 2023;18(1):316. <https://dx.doi.org/10.1186/s13023-023-02927-8>.

Anhang

Krankenhausaufenthalte (Hospitalisierungen aus jeglichen Gründen)

Bei geringen Populationsgrößen bzw. Ereignisraten sind exakte Tests bevorzugt darzustellen (siehe Kapitel 2.4). Das nachfolgend in Tabelle 24 dargestellte RR und der p-Wert wurden lediglich unter Berücksichtigung der Strata mit der Mantel-Haenszel-Methode berechnet; ein exakter Test zur Berechnung des p-Werts wurde nicht angewendet. Demnach sollte vorliegend das 95%-KI berücksichtigt werden, welches nicht auf ein signifikantes Ergebnis hindeutet.

*Tabelle 24: Hospitalisierungen aus jeglichen Gründen bis Monat 12; Studie Balance, FAS
(DS: 14.07.2023)*

Studie Balance Hospitalisierungen aus jeglichen Gründen	Olezarsen N = 22	Placebo N = 23	Effektschätzer [95%-KI]; p-Wert
Personen mit mind. einem Ereignis (post hoc), n (%)	3 (13,6)	9 (39,1)	RR ¹⁾ : 0,66 [0,42; 1,06]; 0,04
Gesamtzahl an Ereignissen	3	17	
Gesamtzahl der Patientenjahre	20,3	22,9	
Mittlere Ereignisrate / Patientenjahr [95%-KI]	0,10 [0,03; 0,39]	0,56 [0,26; 1,20]	MRR ²⁾ : 0,182 [0,05; 0,72]; 0,02

¹⁾ RR und p-Wert (Cochran-Mantel-Haenszel) unter Berücksichtigung der Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

²⁾ MRR, 95%-KI sowie p-Wert (Wald-Chi-Quadrat-Test) mittels Negativ-Binomial-Regressionsmodell mit Behandlungsgruppe und den Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“ als Faktoren. Logarithmus der Zeit in Jahren, die jede Person beobachtet wurde als Offset-Variable.

Abkürzungen: DS: Datenschnitt; FAS: Full Analysis Set; KI: Konfidenzintervall; MRR: Mean Rate Ratio; RR: Relatives Risiko.

PGIC

Für den PGIS und den PGIC wurden keine Responderanalysen zur Verschlechterung eingereicht. Da etwa ein Viertel der Personen im Olezarsen-Arm und ein Drittel der Personen im Placebo-Arm zu Baseline keine Symptome aufwiesen und sich demnach nicht verbessern konnten, werden die Responderanalysen zur Verschlechterung in dieser Studienpopulation gegenüber Auswertungen zur Verbesserung als vorrangig relevant eingeschätzt (siehe auch Kapitel 2.4).

Die Ergebnisse für die Responderanalyse zur Verbesserung sind für den PGIS zu Monat 12 nicht signifikant (siehe Kapitel 3.3), während sich im PGIC zu Monat 6 ein statistisch signifikanter Vorteil für Olezarsen zeigt, der bei stark variierenden Rücklaufquoten zu Monat 12 von über 15 % bestehen bleibt. Bei der Interpretation der Ergebnisse ist jedoch zu berücksichtigen, dass sich gemäß PGIS etwas weniger Personen im Placebo-Arm verbessern konnten als im Olezarsen-Arm und die Effektrichtung im PGIS zu Monat 12 dennoch leicht gegenläufig zu den Ergebnissen des PGIC ist. Zudem ist die Mehrfacherfassung mit der FCS-SIS (Symptom) zu berücksichtigen (siehe Kapitel 3.3).

Im PGIC liegt jeweils für die Responderanalyse zu Monat 6 und zu Monat 12 ein statistisch signifikanter Vorteil für Olezarsen gegenüber Placebo vor. Zu Monat 12 liegen im Olezarsen-Arm jedoch nur noch 16 (72 %) verfügbare Werte vor, während im Placebo-Arm 21 (91 %) verfügbare Werte vorliegen. Damit liegen stark differenzielle Rücklaufquoten zu Monat 12 vor und die Auswertungen des PGIC zu Monat 12 werden nicht dargestellt.

*Tabelle 25: Anteil der Personen mit einer Verbesserung im PGIC zu Monat 6; Studie Balance, FAS
(DS: 14.07.2023)*

Studie Balance PGIC ¹⁾ Responderanalyse zu Monat 6 ²⁾	Olezarsen N = 22	Placebo N = 23
Verbesserung³⁾		
Verfügbare Werte, n (%)	19 (86,4)	22 (95,7)
Responder, n (%)	15 (78,9)	9 (40,9)
RR [95%-KI]; p-Wert ⁴⁾	2,89 [1,015; 8,21]; 0,018	

¹⁾ Die Skalenspannweite beträgt 1–5; höhere Werte zeigen einen schlechteren FCS-Gesundheitszustand seit Behandlungsbeginn an.

²⁾ Monat 12 definiert als Mittelwert der Wochen 51 und 53.

³⁾ Verbesserung definiert als: viel besser (PGIC-Score < 1,5), etwas besser (PGIC-Score 1,5 bis < 2,5).

⁴⁾ RR und p-Wert (Cochran-Mantel-Haenszel) unter Berücksichtigung der Stratifikationsfaktoren „Pankreatitis in der Vorgeschichte innerhalb von 10 Jahren vor dem Screening (ja vs. nein)“ und „Vorbehandlung mit Volanesorsen (ja vs. nein)“. Eine exakte Berechnung des p-Werts ist nicht beschrieben.

Abkürzungen: DS: Datenschnitt; FAS: Full Analysis Set; FCS: Familiäres Chylomikronämie-Syndrom; KI: Konfidenzintervall; PGIC: Patient Global Impression of Change; RR: Relatives Risiko.