

Nutzenbewertung



Gemeinsamer
Bundesausschuss

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: Iptacopan

**Anwendungsgebiet: Paroxysmale nächtliche
Hämoglobinurie (PNH), nicht vorbehandelte Patienten**

Datum der Veröffentlichung: 1. Oktober 2024

Version 1.0

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abkürzungsverzeichnis	5
Hintergrund.....	6
1 Fragestellung.....	7
2 Auswahl und Methodik der relevanten Studien	8
2.1 Studienbasis für die Nutzenbewertung	8
2.2 Beschreibung der eingeschlossenen Studie	8
2.3 Endpunkte	12
2.3.1 Mortalität.....	13
2.3.2 Morbidität.....	13
2.3.3 Lebensqualität	20
2.3.4 Sicherheit	21
2.3.5 Erhebungszeitpunkte	22
2.4 Statistische Methoden.....	24
2.5 Verzerrungspotential auf Studien- und Endpunktebene	26
2.6 Indirekter Vergleich	26
2.6.1 Zusammenfassende Bewertung der Methodik des indirekten Vergleichs	32
3 Ergebnisse der eingeschlossenen Studien	34
3.1 Studiencharakteristika und Studienmedikation	34
3.2 Mortalität	36
3.3 Morbidität	36
3.3.1 Transfusionsvermeidung (ergänzend dargestellt)	36
3.3.2 Durchbruchhämolyse (ergänzend dargestellt).....	36
3.3.3 Fatigue	37
3.3.4 EORTC QLQ-C30 – Symptomskalen und Symptome	38
3.3.5 EQ-5D-VAS	39
3.3.6 MAVe	39
3.4 Lebensqualität.....	39
3.5 Sicherheit	40
4 Diskussion der Methodik und Ergebnisse.....	43
4.1 Zulassungsstatus und Zulassungspopulation von Iptacopan	43
4.2 Design und Methodik der Studie.....	44
4.3 Mortalität	44
4.4 Morbidität	44
4.5 Lebensqualität.....	45
4.6 Sicherheit	45
5 Anforderung an eine qualitätsgesicherte Anwendung	47
6 Zusammenfassung der Nutzenbewertung	48
Referenzen	50
Anhang	52

Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis.....	8
Tabelle 2:	Charakterisierung der Studie APPOINT.....	8
Tabelle 3:	Charakterisierung der Intervention in der Studie APPOINT	11
Tabelle 4:	Zusammenfassung der Endpunktbewertung der Studie APPOINT.....	12
Tabelle 5:	Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte in der Studie APPOINT	23
Tabelle 6:	Übersicht über die durch systematische Literaturrecherche und Experteninterviews ermittelten Confounder	30
Tabelle 7:	Übersicht über die präspezifizierte Confounder-Wahl	30
Tabelle 8:	Übersicht über verwendete Confounder im Propensity-Score-Modell.....	30
Tabelle 9:	Allgemeine Angaben; Studie APPOINT	34
Tabelle 10:	Charakterisierung der Studienpopulation; Studie APPOINT	34
Tabelle 11:	Transfusionsvermeidung während der gesamten Studiendauer (Woche 48); Studie APPOINT; FAS (Datenschnitt: 18.04.2023).....	36
Tabelle 12:	Auftreten einer Durchbruchhämolyse während der gesamten Studiendauer (Woche 48) (ergänzend dargestellt); Studie APPOINT, FAS (Datenschnitt: 18.04.2023).....	36
Tabelle 13:	Verbesserung des FACIT-Fatigue zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022).....	37
Tabelle 14:	Verbesserung des PGIS zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022).....	37
Tabelle 15:	Verbesserung der Symptomatik mittels EORTC QLQ-C30 zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022).....	38
Tabelle 16:	Verbesserung der EQ-5D-VAS um ≥ 15 Punkte zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022)	39
Tabelle 17:	Verbesserung der Lebensqualität mittels EORTC QLQ-C30 um ≥ 10 Punkte zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS; Datenschnitt: 02.11.2022	39
Tabelle 18:	Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der gesamten Studie (Woche 48); Studie APPOINT, SAS (Datenschnitt: 18.04.2023) ..	40
Tabelle 19:	UE mit Inzidenz ≥ 10 % während der gesamten Studiendauer (Woche 48); Studie APPOINT, SAS (Datenschnitt: 18.04.2023).....	41
Tabelle 20:	Schwere UE (Grad 3 oder höher), die bei ≥ 5 % der Personen, nach Systemorgan-klasse und Preferred Term, während der gesamten Studiendauer (Woche 48) aufgetreten sind; Studie APPOINT, SAS (Datenschnitt: 18.04.2023).....	41
Tabelle 21:	SUE, die bei ≥ 5 % der Personen, nach Systemorganklasse und Preferred Term, während der gesamten Studiendauer (Woche 48) aufgetreten sind; Studie APPOINT; SAS (Datenschnitt: 18.04.2023)	42

Tabelle 22:	UE von besonderem Interesse während der gesamten Studiendauer (Woche 48); Studie APPOINT, SAS (Datenschnitt: 18.04.2023)	42
Tabelle 23:	Zusammenfassende Darstellung der Ergebnisse der Studie APPOINT	48
Tabelle 24:	Hämoglobinwert-assoziierte Endpunkte; Studie APPOINT, FAS (Datenschnitte: 02.11.2022 (Woche 24) und 18.04.2023 (Woche 48))	52
Tabelle 25:	Veränderung des FACIT-Fatigue von Baseline bis Ende Extensionsphase (Woche 48); Studie APPOINT, FAS (Datenschnitt: 18.04.2023)	52
Tabelle 26:	Deskriptive Auswertungen der Symptomatik mittels EORTC QLQ-C30 von Baseline bis Ende Extensionsphase (Woche 48); Studie APPOINT, FAS (Datenschnitt: 18.04.2023)	53
Tabelle 27:	Deskriptive Auswertungen der Lebensqualität mittels EORTC QLQ-C30 von Baseline bis Ende Extensionsphase (Woche 48); Studie APPOINT, FAS (Datenschnitt: 18.04.2023)	54

Abkürzungsverzeichnis

AA	Aplastische Anämie
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
AIPW	Augmented Inverse Probability Weighting
ATT	Average Treatment Effect on the Treated
EK	Erythrozytenkonzentrat
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30
EPAR	European Public Assessment Report
EQ-5D-VAS	Visuelle Analogskala des European Quality of Life 5-Dimension
FACIT-Fatigue	Functional Assessment of Chronic Illness Therapy – Fatigue Scale
FAS	Full Analysis Set
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
GLMM	Generalisiertes lineares gemischtes Modell (Generalized Linear Mixed Model)
Hb	Hämoglobin
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
LDH	Laktatdehydrogenase
MAVE	Schweres unerwünschtes vaskuläres Ereignis (Major Adverse Vascular Event)
MedDRA	Medical Dictionary for Regulatory Activities
N	Anzahl
PGIS	Patient Global Impression of Severity
PNH	Paroxysmale nächtliche Hämoglobinurie
PRO	Patient-Reported Outcome
pU	pharmazeutischer Unternehmer
RIME	Französische Register für Knochenmarkversagen
SAS	Safety Analysis Set
SAP	Statistischer Analyseplan
SGB	Sozialgesetzbuch
SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
ULN	Upper Limit of Normal
VerfO	Verfahrensordnung des G-BA

Hintergrund

Iptacopan 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 Iptacopan 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 Iptacopan in seiner Sitzung am 24. September 2024 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 27. Juni 2024 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 1. Oktober 2024 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

Iptacopan (Fabhalta®) 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 [15]:

Behandlung erwachsener Patienten mit paroxysmaler nächtlicher Hämoglobinurie (PNH), die eine hämolytische Anämie aufweisen, nicht vorbehandelte Patienten.

Gemäß Fachinformation beträgt die zugelassene Dosierung 200 mg, die zweimal täglich oral eingenommen wird. Um das Risiko einer Hämolyse zu minimieren, sollten Patientinnen und Patienten auf die Einhaltung des Dossierungsschema hingewiesen werden.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studiename	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Studien zum Wirkstoff				
APPOINT ¹⁾	Ja	Ja	Ja	-
Studien zu externen Kontrollen				
APPEX	Ja	Ja	Nein	Siehe Kapitel 2.6.

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

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

Zur Nutzenbewertung für Iptacopan herangezogene Unterlagen

- Dossier und Zusatzanalysen des pU zu Iptacopan [13,14]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [3]
- Studienberichte inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie APPOINT [10,11,12]
- Studienbericht inkl. Studienprotokoll und SAP der Studie APPEX [8,9,19]
- Confounder-Recherche des pU zum Anwendungsgebiet [7]
- Fachinformation zu Iptacopan [15]

2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Iptacopan im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie APPOINT. Die Studie und die Intervention werden in Tabelle 2 und Tabelle 4 charakterisiert.

Tabelle 2: Charakterisierung der Studie APPOINT

Charakteristikum	Beschreibung
Design und Studienablauf	<p>Design Multizentrische, einarmige, offene klinische Phase-III-Studie in Erwachsenen mit PNH, welche behandlungsnaiv gegenüber der Behandlung mit Komplementinhibitoren sind.</p> <p>Studienablauf</p> <ul style="list-style-type: none"> • Screening-Periode (8 Wochen) • Behandlungsphase (24 Wochen; bis Woche 24) Erhebung des primären Endpunkts • Extensionsphase (24 Wochen; Woche 24 bis Woche 48) Nach Abschluss der Behandlungsphase zu Woche 24 konnten die teilnehmenden Personen Iptacopan im Rahmen der Extensionsphase weiter erhalten. <p>Die Studie ist zum Zeitpunkt der Nutzenbewertung abgeschlossen. Nach Abschluss der Studie konnten die teilnehmenden Personen an einer Langzeitstudie teilnehmen.</p>

Charakteristikum	Beschreibung
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Erwachsene ≥ 18 Jahre mit PNH-Diagnose, bestätigt durch den Nachweis einer PNH-Klonggröße von $\geq 10\%$ bezogen auf Erythrozyten und Granulozyten/Monozyten mittels hochsensibler Durchflusszytometrie. • Mittlerer Hb-Wert < 10 g/dl zur Screening-Visite und vor Beginn der Studienbehandlung bestätigt durch: <ul style="list-style-type: none"> ○ mind. 2 Messungen (mittlerer Hb-Wert < 10 g/dl im Abstand von 2 bis 8 Wochen bei Erwachsenen ohne EK-Transfusion von während des Screenings); ○ 1 Messung (Hb-Wert < 10 g/dl) bei einer Screening-Visite für Erwachsene mit EK-Transfusion während des Screenings. • LDH $> 1,5 \times$ ULN bei mind. 2 Messung des Zentrallabors im Abstand von 2 bis 8 Wochen während des Screenings. • Schutzimpfungen gegen <i>Neisseria meningitidis</i>, <i>Streptococcus pneumoniae</i> und <i>Haemophilus influenzae</i>. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Vorherige Behandlung mit einem Komplementinhibitor. • Bekannter oder vermuteter hereditärer Komplementdefekt. • Hämatopoetische Stammzelltransplantation in der Anamnese. • Nachgewiesene Knochenmarkinsuffizienz (Retikulozyten $< 100 \times 10^9/l$; Thrombozyten $< 30 \times 10^9/l$; Neutrophile Granulozyten $< 0,5 \times 10^9/l$). • Vorliegen einer Infektion¹⁾. • Vorliegen von Fieber innerhalb von 7 Tagen vor Verabreichen der Studienmedikation. • Schwere Komorbidität, einschl. aber nicht beschränkt auf: <ul style="list-style-type: none"> ○ Schwere Nierenerkrankung (z. B. dialysepflichtig); ○ fortgeschrittene Herzerkrankung; ○ schwere Lungenerkrankung; ○ Lebererkrankung, die nach Ermessen des Prüfpersonals die Teilnahme an der Studie ausschließt. • Instabile Erkrankung zum Zeitpunkt des Screenings, einschl. aber nicht beschränkt auf myokardiale Ischämie, aktive gastrointestinale Blutung, koexistente chronische Anämie ohne Verbindung zur PNH oder Thrombose.
Intervention und Zahl der Patientinnen und Patienten	<p>Gescreent gesamt: N = 52 Randomisiert/Eingeschlossen: N = 40</p>
Ort und Zeitraum der Durchführung; Datenschnitte	<p>Studienzentren 16 Studienzentren in 10 Ländern: Deutschland, Frankreich, Großbritannien, Italien, Tschechien, China, Japan, Malaysia, Singapur, Südkorea.</p> <p>Studienzeitraum</p> <ul style="list-style-type: none"> • Erste Person erste Visite: 19.07.2021 <p>Datenschnitte</p> <ul style="list-style-type: none"> • Datenschnitt primäre Analyse: 02.11.2022 (präspezifiziert) • Datenschnitt finale Analyse: 18.04.2023 (präspezifiziert)

Charakteristikum	Beschreibung
Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt Erhöhung des Baseline-Hb-Level um ≥ 2 g/dl (bewertet zwischen Woche 18 und Woche 24) bei gleichzeitiger Transfusionsvermeidung²⁾ zwischen Tag 14 und Woche 24.</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Ansprechen definiert als Hb-Level ≥ 12 g/dl zwischen Woche 18 und Woche 24 ohne Transfusion²⁾ zwischen Tag 14 und Woche 24. • Anteil der Personen mit Transfusionsvermeidung zwischen Tag 14 und Woche 24, die keine Transfusion erhielten oder keines der Transfusionskriterien²⁾ erfüllten. • Veränderung des Hb-Werts zwischen Baseline und Ende der Behandlungsphase (Mittelwert der Visiten zwischen Woche 18 und Woche 24). • Veränderung des LDH zwischen Baseline und Ende der Behandlungsphase (Mittelwert der Visiten zwischen Woche 18 und Woche 24). • Durchbruchhämolyse-Rate³⁾ im Zeitraum von Tag 1 bis Woche 24. • Veränderung der Retikulozytenzahl zwischen Baseline und Ende der Behandlungsphase (Mittelwert der Visiten zwischen Woche 18 und Woche 24). • Veränderung des FACIT-Fatigue-Scores zwischen Baseline und Ende der Behandlungsphase (Mittelwert der Visiten zwischen Woche 18 und Woche 24). • Rate der MAVE⁴⁾. • Sicherheit und Verträglichkeit.

¹⁾ Aktive systemische bakterielle Infektion oder Pilzinfektion, bekannte HIV-Infektion oder positiver HIV-Test; wiederkehrende invasive Infektion mit bekapselten Erregern.

²⁾ Eine EK-Transfusion wurde verabreicht, wenn eines der folgenden Kriterien erfüllt war: Hb-Wert > 7 und ≤ 9 g/dl (> 6 und ≤ 8 g/dl für die chinesische Population) sowie Vorliegen klinischer Anzeichen oder Symptome eines Schweregrads, der eine Transfusion rechtfertigt (z. B. schwere oder sich verschlechternde Fatigue, schwere oder sich verschlechternde Dyspnoe, Palpitationen oder Angina oder eine Veränderung des mentalen Zustands (z. B. Ohnmacht, Benommenheit, Verwirrtheit, Schlaganfall, transitorische ischämische Attacke)) oder Hb-Wert ≤ 7 g/dl (≤ 6 g/dl für die chinesische Population), unabhängig vom Vorliegen klinischer Anzeichen oder Symptome.

³⁾ Das Vorliegen einer klinisch relevanten Durchbruchhämolyse war definiert als LDH-Wert $> 1,5 \times$ ULN sowie Erhöhung des LDH-Werts im Vergleich zu den letzten beiden Messungen UND Vorliegen eines klinischen Kriteriums, d. h. massive Hämoglobinurie, Schmerzkrise, Dysphagie oder andere klinische Anzeichen oder Symptome in Verbindung mit der PNH ODER Absinken des Hb-Werts um mind. 2 g/dl im Vergleich zur vorherigen Messung oder einer Messung innerhalb der letzten 15 Tage.

⁴⁾ Siehe Operationalisierung in Kapitel 2.3.

Abkürzungen: EK: Erythrozytenkonzentrat; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; Hb: Hämoglobin; LDH: Laktatdehydrogenase; MAVE: Schweres unerwünschtes vaskuläres Ereignis; PNH: Paroxysmale nächtliche Hämoglobinurie; ULN: Upper Limit of Normal.

Protokolländerungen

Es wurden 4 Änderungen des Originalprotokolls vom 16.02.2021 vorgenommen. Es konnten keine für die Nutzenbewertung relevante Protokolländerung identifiziert werden.

Charakterisierung der Intervention

Tabelle 3: Charakterisierung der Intervention in der Studie APPOINT

Intervention
<p>Iptacopan 200 mg, 2 x täglich oral. Dosisanpassungen bzw. Behandlungsunterbrechungen waren gemäß den Vorgaben des Studienprotokolls nicht vorgesehen.</p>
<p>Nicht erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • Komplementinhibitoren • Lebendimpfstoffe <p>Erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • ESA und HIF-Prolylhydroxylase-Inhibitoren nur bei stabiler Dosierung über mind. 8 Wochen vor dem Screening (während der Behandlung sollte die Dosis von ESA und/oder HIF-Prolylhydroxylase-Inhibitoren in Abhängigkeit des Hb-Werts angepasst bzw. die Behandlung ggf. beendet werden). • Systemische Kortikosteroide zur Behandlung hämatologischer Erkrankungen nur bei stabiler Dosierung über mind. 4 Wochen vor dem Screening (weniger als 0,5 mg/kg). Während der Behandlungsphase sollte die Dosis nicht geändert werden. • Sensitive OATP1B1-Substrate sowie OATP1B1-Substrate mit geringer therapeutischer Breite. • Direkte orale Antikoagulanzen, die gleichzeitig P-gp-Substrate sind (z. B. Apixaban, Rivaroxaban und Edoxaban). • Immunsuppressiva, die Substrate von P-Glykoprotein sind (z. B. Cyclosporin, Sirolimus o. Tacrolimus), sollten, wenn keine Behandlungsalternative bestand, zeitlich versetzt verabreicht werden.

Abkürzungen: ESA: Erythropoese-stimulierende Wirkstoffe; Hb: Hämoglobin; HIF: Hypoxie-induzierter Faktor; OATP1B1: Leberspezifischer Organic-Anion-Transporter 1B1.

2.3 Endpunkte

In diesem Kapitel wird 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 4 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 4: Zusammenfassung der Endpunktbewertung der Studie APPOINT

Studie APPOINT Endpunktbewertung	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Gesamtmortalität ¹⁾	Mortalität	Ja	Ja
Transfusionsvermeidung	Morbidität	Ja	Ergänzend
Hämoglobinwert-assoziierte Endpunkte ²⁾		Ja	Nein ³⁾
Fatigue • FACIT-Fatigue • PGIS		Ja Ergänzend	Ja
EORTC QLQ-C30 – Symptomskalen und Symptome			
EQ-5D-VAS		Ja	Ja
Durchbruchhämolyse		Ja	Ergänzend
Retikulozytenzahl		Ergänzend	Nein
MAVE		Ja	Ja
EORTC QLQ-C30 • Funktionsskalen • Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“	Lebensqualität	Ja	Ja
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

¹⁾ Todesfälle wurden im Rahmen der Sicherheitsanalysen erhoben.

²⁾ Primärer Endpunkt.

³⁾ Der Endpunkt wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt. Der primäre Endpunkt wird aus Transparenzgründen im Anhang dargestellt.

⁴⁾ Die Symptomskala „Übelkeit und Erbrechen“ sowie die Symptome „Appetitverlust“, „Obstipation“ und „Diarrhö“ werden aufgrund unklarer Patientenrelevanz lediglich ergänzend dargestellt (siehe Kapitel 2.3.2).

Abkürzungen: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; EQ-5D-VAS: Visuelle Analogskala des European Quality of Life 5-Dimension; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; MAVE: Schweres unerwünschtes vaskuläres Ereignis; PGIS: Patient Global Impression of Severity; pU: pharmazeutischer Unternehmer.

2.3.1 Mortalität

Gesamtmortalität

Der Endpunkt „Gesamtmortalität“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Todesfälle jeglicher Ursache wurden im Rahmen der Sicherheitserfassung bis zum Ende der Studie (Woche 48) erfasst.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz und Validität

Mortalität stellt einen patientenrelevanten Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) dar. Die Erhebung des Endpunkts wird als valide erachtet.

2.3.2 Morbidität

Transfusionsvermeidung

Der Endpunkt „Transfusionsvermeidung“ wird in der Nutzenbewertung aufgrund von Einschränkungen in der Operationalisierung und daraus resultierender unklarer Patientenrelevanz ergänzend dargestellt.

Operationalisierung

Beschreibung

In der Studie APPOINT war der Endpunkt „Transfusionsvermeidung“ definiert als der Anteil der Personen, die zwischen Tag 14 (Woche 2) und Woche 24 bzw. Woche 48 keine Transfusion von Erythrozytenkonzentrat (EK) erhielten sowie keines der folgenden Kriterien für die Notwendigkeit einer Transfusion erfüllten:

- Hämoglobin (Hb)-Wert > 7 und ≤ 9 g/dl (> 6 und ≤ 8 g/dl für die chinesische Population) sowie Vorliegen von symptomatischen Transfusionskriterien mit ausreichendem Schweregrad, die eine Transfusion rechtfertigen.
- Hb-Wert < 7 g/dl (≤ 6 g/dl für die chinesische Population) unabhängig vom Auftreten der symptomatischen Transfusionskriterien.

Als symptomatische Transfusionskriterien wurde das Auftreten folgender Symptome spezifiziert:

- Schwere oder sich verschlechternde Fatigue.
- Schwere oder sich verschlechternde Dyspnoe/Kurzatmigkeit.
- Palpitationen/Angina (oder sich verschlechternde Symptome).
- Veränderung des mentalen Zustands (Ohnmacht, Benommenheit, Verwirrtheit, Schlaganfall, transitorische ischämische Attacke).

Die Messung des Hb-Werts, welcher zur Bestimmung der Transfusionskriterien herangezogen wurde, kann aus dem lokalen Labor erfolgen. Das Prüfpersonal musste eine weitere Probe dem Zentrallabor zur Verfügung stellen. Sollte eine Person die Transfusionskriterien erfüllen, sich das Prüfpersonal aber gegen eine Transfusion entscheiden, sollten die Gründe festgehalten werden.

Es liegen Auswertungen von Tag 1 bis zum Ende der Studie (Woche 48) vor. Es liegen Auswertungen dazu vor, wie viele Personen die Transfusionskriterien erfüllt haben, sowie zum Anteil der Personen, die tatsächlich eine Transfusion erhalten haben. Die Symptome, die zur Transfusion geführt haben, wurden ebenfalls berichtet.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Die Gründe für die unterschiedlichen Transfusionskriterien für die chinesische Population sowie die Auswertungsstrategie mit Start der Analyse zu Tag 14 konnten in den Unterlagen nicht identifiziert werden. Insbesondere der Beginn des Analysezeitraums zu Tag 14 wird kritisch gesehen. Analysen über die gesamte Behandlungsperiode beginnend mit Tag 1 werden für die Nutzenbewertung bevorzugt.

Patientenrelevanz und Validität

Leitlinien empfehlen die Substitution von EK als unterstützende Therapie bei der hämolytischen PNH [2]. Transfusionen können jedoch mit schwerwiegenden, unerwünschten Transfusionsreaktionen (z. B. akute allergische/anaphylaktische Transfusionsreaktionen, transfusionsassoziierte Volumenüberladung, hämolytische Transfusionsreaktionen und Fehltransfusionen) einhergehen [1]. Zudem besteht das Risiko einer Eisenüberladung. Für Patientinnen und Patienten bedeutet daher eine längerfristige Transfusionsfreiheit, dass sie keine Symptome aufweisen, für deren Behandlung eine Transfusion notwendig wird und dass Transfusionskomplikationen vermieden werden.

Jedoch wurden bei der vorgelegten Operationalisierung der Transfusionsvermeidung Personen als Non-Responder („transfundiert“) gewertet, wenn sie die Kriterien zur Transfusion erfüllten, aber tatsächlich keine Transfusion erhielten. Für diese Personen ergibt sich jedoch kein erhöhtes Risiko für Transfusionskomplikationen, weshalb die Patientenrelevanz für dieses Kriterium als unklar bewertet wird.

Die Verabreichung von Transfusionen erfolgte in der Studie APPOINT nach präspezifizierten Kriterien. Die Kriterien entsprechen weitestgehend den Empfehlungen der Querschnittsleitlinie der Bundesärztekammer zur Therapie mit Blutkomponenten und Plasmaderivaten (2020) [1]. Es kann nicht abschließend beurteilt werden, inwieweit im Versorgungsalltag Transfusionen bei Hb-Werten < 7 g/dl regelhaft ohne Vorhandensein von Symptomen verabreicht werden.

Hämoglobinwert-assoziierte Endpunkte

Der Endpunkt „Hämoglobinwert-assoziierte Endpunkte“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt und wird in der Nutzenbewertung aus Transparenzgründen im Anhang dargestellt.

Operationalisierung

Beschreibung

Der Endpunkt bezüglich des Hb-Werts wurde für die Auswertung des primären Endpunkts mit der Transfusionsvermeidung kombiniert:

- Die Erhöhung des Hb-Werts um ≥ 2 g/dl im Vergleich zu Baseline in 3 von 4 Erhebungen in den Visiten zum Ende der Hauptbehandlungszeit (Woche 18 bis Woche 24).
- Keine EK-Transfusion erhalten oder die Kriterien zur Verabreichung zwischen Tag 14 und Woche 24 erfüllt.

Der Hb-Wert zu Baseline ist definiert als der Mittelwert von 2 Erhebungen, die das Einschlusskriterium von Personen bestätigten, die keine Transfusionen zwischen den beiden Erhebungen benötigten. Bei Personen, die eine Transfusion nach Erhebung des ersten Werts erhalten haben, wurde der ersterhobene Wert herangezogen.

Es liegen ebenfalls Auswertungen zum Studienende (Woche 48) vor.

Bewertung

Die Operationalisierung ist nachvollziehbar beschrieben. Länderspezifische Normwerte und Unterschiede in den Hb-Werten zwischen asiatischen [6,17] und europäischen Personen wurden nicht berücksichtigt.

Patientenrelevanz

Der Endpunkt stellt einen Laborparameter ohne direkten Symptombefund dar und ist nicht per se patientenrelevant. Die Ursachen für eine Veränderung im Laborparameter können vielfältig sein und sind nicht ausschließlich begrenzt auf bzw. spezifisch für die PNH. Hämoglobin ist neben weiteren Faktoren ein Parameter, der in der klinischen Praxis zur Beurteilung der Krankheitsaktivität und zur Therapiesteuerung herangezogen wird. Untersuchungen zur Surrogatvalidierung werden durch den pU nicht vorgelegt.

Validität

Der Endpunkt wurde valide erhoben.

Fatigue

FACIT-Fatigue

Der Endpunkt „FACIT-Fatigue“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die „Fatigue Scale“ des „Functional Assessment of Chronic Illness Therapy“ (FACIT-Fatigue) umfasst 13 Items, mit denen das Ausmaß der Fatigue und die Beeinträchtigung alltäglicher Aktivitäten und Funktionen in der vergangenen Woche selbstberichtet erhoben werden.

Die Patientinnen und Patienten beantworten die Items auf einer 5-Punkte-Likert-Skala (0 = „Überhaupt keine Beeinträchtigung“ bis 4 = „Sehr schwere Beeinträchtigung“). Der FACIT-Fatigue-Score berechnet sich aus dem Mittelwert der vorhandenen Antworten multipliziert mit 13. Für den FACIT-Fatigue-Score ergibt sich ein Bereich zwischen 0 und 52, wobei höhere Werte geringere Beeinträchtigungen anzeigen.

Im Rahmen der Studie APPOINT wurde ein Ansprechen als eine Verbesserung um ≥ 5 Punkte geplant. Es liegen Post-hoc-Auswertungen mit der relevanten Responderschwelle von ≥ 8 Punkten (15 % der Skalenspannweite) zum Ende der Behandlungsphase (Woche 24) vor. Zum Ende der Extensionsphase (Woche 48) liegen Auswertungen zu den Änderungen des FACIT-Fatigue-Scores gegenüber Baseline vor. Als Baseline-Wert wurde der Mittelwert zwischen der Erhebung zum Screening der Studie sowie die Erhebung zu Tag 1 herangezogen.

Bewertung

Die Operationalisierung ist nachvollziehbar. Es fehlen Angaben, in welcher Reihenfolge die „Patient-Reported Outcomes“ (PRO) administriert wurden. Es ist daher unklar, welchen Effekt das möglicherweise zuvor erfolgte Ausfüllen des PGIS auf das Ausfüllen des FACIT-Fatigue hatte.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen. Die patientenberichtete Erfassung der Fatigue einschließlich Veränderungen wird als patientenrelevant erachtet.

Validität

Die Validität des Endpunkts „FACIT-Fatigue“ wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt. In der Studie APPOINT wird die Fatigue über mehrere Instrumente (FACIT-Fatigue, PGIS, Symptomskala „Fatigue“ des EORTC QLQ-C30) erhoben und stellt somit eine inhaltliche Überschneidung (Doppelerfassung) dar.

PGIS

Der Endpunkt „PGIS“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Der „Patient Global Impression of Severity“ (PGIS) wurde in der Studie APPOINT zusammen mit dem FACIT-Fatigue erhoben. Diese selbstberichtete Skala besteht aus einem Item: „Wie würden Sie ihre Gesamtsymptome bezüglich der Fatigue innerhalb der letzten 7 Tage beschreiben?“ Die Einschätzung erfolgt auf einer 5-Punkte-Likert-Skala von 0 („keine Symptome“) bis 4 („sehr schwere Symptome“).

Als Responder wurden Personen gewertet, die zum jeweiligen Zeitpunkt eine Verbesserung des PGIS um mindestens eine Kategorie erreicht hatten.

Als Baseline-Wert wird Tag 1 oder die letztmalige Erhebung vor Behandlungsstart definiert. Es liegen Auswertungen zum Ende der Behandlungsphase (Woche 24) sowie der Extensionsphase (Woche 48) vor.

Bewertung

Die Operationalisierung ist nachvollziehbar. Es fehlen Angaben, in welcher Reihenfolge die PRO administriert wurden. Es ist daher unklar, welchen Effekt das möglicherweise zuvor erfolgte Ausfüllen des FACIT-Fatigue auf die Bewertung des PGIS hat.

Patientenrelevanz und Validität

Die Einschätzung der Patientinnen und Patienten zum Schweregrad der jeweils eigenen erkrankungsspezifischen Symptomatik ist als patientenrelevant zu bewerten, wobei eine Erhebung per PGIS grundsätzlich als geeignet eingeschätzt wird. Mit dem Instrument wird die Symptomschwere zum Zeitpunkt der Visite bewertet. Es ist einschränkend zu berücksichtigen, dass keine Überprüfung der Gütekriterien des Items und der Skalenausprägungen (0–4) bezüglich der PNH vorliegt. In der Studie APPOINT wird die Fatigue über mehrere Instrumente (FACIT-Fatigue, PGIS, Symptomskala „Fatigue“ des EORTC QLQ-C30) erhoben und stellt somit eine inhaltliche Überschneidung (Doppelerfassung) dar.

EORTC QLQ-C30 – Symptomskalen und Symptome

Die Symptome „Dyspnoe“ und „Schlaflosigkeit“ sowie die Symptomskalen „Schmerz“ und „Fatigue“ des „Quality of Life Questionnaire Core 30“ der „European Organisation for Research and Treatment of Cancer“ (EORTC QLQ-C30) werden in der Nutzenbewertung berücksichtigt. Die Symptomskala „Übelkeit und Erbrechen“ und die Symptome „Appetitverlust“, „Obstipation“ und „Diarrhö“ werden für die Kategorie der Morbidität in der Nutzenbewertung aufgrund unklarer Patientenrelevanz und Validität im vorliegenden Anwendungsgebiet ergänzend dargestellt.

Operationalisierung

Beschreibung

Der EORTC QLQ-C30 ist ein generisches Instrument zur Erfassung der gesundheitsbezogenen Lebensqualität und der Symptome bei Personen mit onkologischen Erkrankungen. Der Fragebogen gliedert sich in

- 5 Funktionsskalen
 - Physische Funktion (5 Items)
 - Rollenfunktion (2 Items)
 - Emotionale Funktion (4 Items)
 - Kognitive Funktion (2 Items)
 - Soziale Funktion (2 Items)

- 3 Symptomskalen
 - Fatigue (3 Items)
 - Schmerz (2 Items)
 - Übelkeit und Erbrechen (2 Items)
- 1 globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“

Zusätzlich umfasst der Fragebogen einzelne Items zu Symptomen („Dyspnoe“, „Appetitverlust“, „Schlaflosigkeit“, „Obstipation“, „Diarrhö“) und zu finanziellen Beeinträchtigungen.

Die Fragen werden auf 4- bzw. 7-Punkte-Skalen beantwortet. Der Bezugszeitraum ist die vergangene Woche. Gemäß Manual werden der Gesamtscore und die Scores für Funktions-, Symptom- und den anderen Einzelskalen auf eine Skala von 0 bis 100 Punkten transformiert.

Die Berechnung der Multi-Item-Skalenwerte erfolgte, wenn mindestens 50 % der Items beantwortet wurden.

Höhere Werte auf den Funktionsskalen und der globalen Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ bedeuten eine bessere Funktion bzw. Gesundheit/Lebensqualität. Höhere Werte auf den Symptomskalen bzw. den einzelnen Symptom-Items bedeuten eine schwerere Symptomatik.

Als Baseline-Wert wurde der Mittelwert zwischen der Erhebung zum Screening der Studie sowie die Erhebung zu Tag 1 herangezogen. Es liegen Responderanalysen mit der relevanten Verbesserung um ≥ 10 Punkte (Responseschwelle von 15 %) vor.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz und Validität

Der EORTC QLQ-C30 ist ein validiertes generisches Messinstrument zur Untersuchung der gesundheitsbezogenen Lebensqualität und der Symptomatik bei an Krebs erkrankten Personen [4]. Es liegt eine Querschnittsstudie mit 29 Patientinnen und Patienten mit PNH aus Großbritannien, USA, Frankreich und Spanien zum EORTC QLQ-C30 vor, in der die Relevanz der Items des EORTC QLQ-C30 bei PNH untersucht wurde [18]. Basierend auf diesen Daten wurde geschlussfolgert, dass das Instrument grundsätzlich geeignet ist, um Daten zu Morbidität und Lebensqualität bei an PNH erkrankten Personen zu erheben. Die Relevanz der Items des EORTC QLQ-C30 wurde jedoch etwas geringer eingeschätzt als die ebenfalls erhobene Relevanz der FACIT-Fatigue-Items. Die Befragten schlugen zusätzliche Fragen vor, die die Relevanz des EORTC QLQ-C30 bei PNH erhöhen könnten.

Weitere Informationen zur Relevanz bzw. Wichtigkeit der Items des QLQ-C30 bei PNH liefert die Untersuchung von Groth et al. (2017) [5]. In dieser Studie wurden für die Entwicklung eines neuen Fragebogens Personen mit PNH und/oder aplastischer Anämie (AA) zu Aspekten der Erkrankung befragt (N = 19, davon 53 % mit PNH oder PNH/AA) und gebeten, deren Wichtigkeit einzuschätzen (N = 30, davon 66 % mit PNH oder PNH/AA). Es wurde überprüft, ob die von den Befragten genannten Aspekte der Erkrankung im EORTC QLQ-C30 abgebildet sind. Im Ergebnis wurden 16 Items des EORTC QLQ-C30 als Aspekt der Erkrankung genannt und als wichtig von den Befragten erachtet (u. a. Dyspnoe, Schlaflosigkeit und Schmerz, Konzentrationsprobleme und Müdigkeit), während 12 Items des EORTC QLQ-C30 entweder als nicht wichtig bewertet wurden und von den Befragten nicht als Aspekt der Erkrankung genannt wurden. Zu diesen 12 Items gehörten Items zu Symptomen (Diarrhö, Nausea, Appetitverlust, Erbrechen und Verstopfung) und einzelne Items der Funktionsskalen zur körperlichen, emotionalen, kognitiven und sozialen Funktion. Die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ des EORTC QLQ-C30 wurde nicht bewertet. In der Studie wurden zudem weitere bei PNH/AA wichtige Aspekte zur Lebensqualität identifiziert, die nicht über den EORTC QLQ-C30 erfasst werden (z. B. ständige Angst vor Infektionen, Abhängigkeit von zeitaufwändigen Therapien).

Eine Leitlinie der Deutschen Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO) benennt als vordergründige, klassische Charakteristika einer chronischen hämolytischen Anämie die Symptome „Schwäche“, „Fatigue“ und „Belastungsdyspnoe“ [2]. Darüber hinaus werden abdominale und thorakale Schmerzen sowie Dyspnoe als Frühindikatoren für thromboembolische Ereignisse benannt.

Basierend auf den Ergebnissen dieser Untersuchung werden die Symptomskala „Übelkeit und Erbrechen“ sowie die Symptome „Appetitverlust“, „Obstipation“ und „Diarrhö“ des EORTC QLQ-C30 aufgrund unklarer Patientenrelevanz und Validität im vorliegenden Anwendungsgebiet ergänzend dargestellt. Die Symptome „Dyspnoe“ und „Schlaflosigkeit“ sowie die Symptomskalen „Schmerz“ und „Fatigue“ werden berücksichtigt. Bei der Fatigue handelt es sich um ein Hauptsymptom der PNH. In der Studie APPOINT wird Fatigue über mehrere Instrumente (FACIT-Fatigue, PGIS, Symptomskala „Fatigue“ des EORTC QLQ-C30) erhoben und stellt somit eine inhaltliche Überschneidung (Doppelerfassung) dar.

EQ-5D-VAS

Der Endpunkt „EQ-5D-VAS“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Der selbsteingeschätzte allgemeine Gesundheitszustand wurde in der Studie APPOINT anhand der visuellen Analogskala des European Quality of Life 5-Dimension (EQ-5D-VAS) erhoben. Die EQ-5D-VAS erfasst den aktuellen durch die Patientin / den Patienten bewerteten Gesundheitszustand auf einer Skala von 0 bis 100, wobei „0“ für den schlechtesten vorstellbaren Gesundheitszustand und „100“ für den besten vorstellbaren Gesundheitszustand steht. Die Erhebung erfolgte mittels ePRO.

Zur Auswertung wurde eine Responderanalyse herangezogen und ein Schwellenwert von 15 Punkten (entspricht 15 % der Skalenspannweite) als Responsekriterium präspezifiziert.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Validität

Die Validität der EQ-5D-VAS wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt.

Durchbruchhämolyse

Der Endpunkt „Durchbruchhämolyse“ wird in der Nutzenbewertung aufgrund unklarer Patientenrelevanz ergänzend dargestellt.

Operationalisierung

Beschreibung

In der Studie APPOINT war das Vorliegen einer Durchbruchhämolyse definiert als

- Vorliegen eines laboratorischen Kriteriums, d. h. Laktatdehydrogenase (LDH)-Wert $> 1,5 \times \text{ULN}$ sowie Erhöhung des LDH-Werts im Vergleich zu den letzten beiden Messungen UND
- Vorliegen eines klinischen Kriteriums, d. h.

- massive Hämoglobinurie, Schmerzkrise, Dysphagie oder andere klinische Anzeichen oder Symptome in Verbindung mit der PNH ODER
- Absinken des Hb-Werts um ≥ 2 g/dl im Vergleich zur vorherigen Messung oder einer Messung innerhalb der letzten 15 Tage.

Es liegen Auswertungen zum Ende der Behandlungsphase sowie zum Ende der Extensionsphase vor. Die Auswertungen beinhalten Informationen über die Personen, die sowohl die Kriterien des Laborparameters einer Durchbruchhämolyse erfüllen sowie die aufgetretenen klinischen Symptome berichten. Weiterhin wurden die verabreichten Transfusionen im Kontext einer Durchbruchhämolyse berichtet.

Die Bewertung konnte von lokalen Laboren vorgenommen werden. Eine weitere Probe zur Bestimmung des LDH-Werts und des Hb-Werts sollte dem Zentrallabor zur Verfügung gestellt werden.

Bewertung

Die Operationalisierung ist nachvollziehbar. Es ist unklar, ob alle patientenrelevanten Symptome der Durchbruchhämolyse erfasst wurden sind.

Patientenrelevanz und Validität

Eine Durchbruchhämolyse reflektiert den zeitweiligen Verlust der Krankheitskontrolle. Es handelt sich bei den Auswertungen des pU um die sogenannte „klinische Durchbruchhämolyse“. Durchbruchhämolysen können mit schwerwiegenden Komplikationen wie Thrombosen oder Nierenversagen verbunden sein. Das Auftreten von Symptomen, die mit der PNH in Verbindung stehen, ist als patientenrelevant zu bewerten. Es ist unklar, ob durch die Operationalisierung alle Symptome, die im Kontext einer Durchbruchhämolyse auftreten können, vollumfänglich erfasst worden sind. Die Gabe von Transfusionen in Folge von mit PNH assoziierten Symptomen wird als patientenrelevant angesehen. Die Operationalisierung des Endpunkts verknüpft die Erhebung der Symptome an eine gleichzeitige Erhöhung des LDH-Werts um $> 1,5 \times \text{ULN}$. Zudem müssen bei einem erhöhten LDH-Wert und dem gleichzeitigen Absinken des Hb-Werts um ≥ 2 g/dl, keine der genannten klinischen Anzeichen bzw. Symptome vorliegen. Eine Durchbruchhämolyse kann somit entkoppelt von klinisch relevanten Symptomen, lediglich basierend auf den Laborwerten (Laktatdehydrogenase und Hämoglobin) auftreten. Es ist unklar, ob in diesem Fall noch von einer Patientenrelevanz ausgegangen werden kann. Die Ursachen für einen Anstieg von Laborwerten können jedoch vielfältig sein und sind nicht ausschließlich begrenzt auf bzw. spezifisch für eine PNH. Die Patientenrelevanz im vorliegenden Anwendungsgebiet wird daher als unklar bewertet und der Endpunkt ergänzend dargestellt.

Retikulozytenzahl

Der Endpunkt „Retikulozytenzahl“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt.

Operationalisierung

Beschreibung

Der pU legt ergänzend Auswertungen zu der Veränderung der Retikulozytenzahl im Studienverlauf vor. Es liegen Ergebnisse zum Ende der Behandlungsphase sowie der Extensionsphase vor. Die Auswertungen zum Ende der Behandlungsphase sind als Mittelwerte der Visiten zwischen Woche 18 bis Woche 24 dargestellt.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Der Endpunkt stellt einen Laborparameter ohne direkten Symptombefund dar und ist nicht per se patientenrelevant. Die Retikulozytenzahl wird in der klinischen Praxis zur Einschätzung der Knochenmarkaktivität zur Regeneration der zerstörten Blutzellen verwendet, wobei ein Anstieg der Retikulozytenzahl nach intravasaler und extravasaler Hämolyse zu beobachten ist.

Validität

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt

MAVE

Der Endpunkt „MAVE“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Ein schweres unerwünschtes vaskuläres Ereignis (MAVE) war in der Studie APPOINT definiert als eines der folgenden Ereignisse:

- Akuter peripherer Gefäßverschluss
- Amputation (nicht traumatisch, nicht diabetisch)
- Zerebraler Arterienverschluss / Schlaganfall
- Zerebraler Venenverschluss
- Dermale Thrombose
- Gangrän (nicht traumatisch, nicht diabetisch)
- Lebervenen-/Pfortaderthrombose (Budd-Chiari-Syndrom)
- Mesenteriale/Viszerale arterielle Thrombose oder Infarkt
- Mesenteriale/Viszerale Venenthrombose oder -infarkt
- Myokardinfarkt
- Lungenembolie
- Nierenarterienthrombose
- Nierenvenenthrombose
- Thrombophlebitis / Tiefe Venenthrombose
- Transitorische ischämische Attacke
- Instabile Angina pectoris
- Andere

Es liegen Auswertungen zum Ende der Behandlungsphase sowie der Extensionsphase vor.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz und Validität

Das Auftreten von schweren unerwünschten vaskulären Ereignissen ist patientenrelevant. Die Erhebung wird als valide erachtet. Es wird angemerkt, dass entsprechende Ereignisse sowohl über den Endpunkt „MAVE“ als auch über den Endpunkt „Unerwünschte Ereignisse“ abgebildet werden; eine Doppelerfassung kann nicht ausgeschlossen werden.

2.3.3 Lebensqualität

EORTC QLQ-C30 – Funktionsskalen und Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“

Die Funktionsskalen und die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ wurden in vorherigen Verfahren bereits als patientenrelevant eingeschätzt und werden für die Kategorie der Lebensqualität in der Nutzenbewertung berücksichtigt.

Beschreibung der Operationalisierung und Bewertung der Validität siehe Kapitel 2.3.2.

2.3.4 Sicherheit

Unerwünschte Ereignisse

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

Operationalisierung

Beschreibung

Ein unerwünschtes Ereignis (UE) war definiert als jedes ungünstige und unbeabsichtigte Ereignis (einschließlich anormaler Laborbefunde, Symptome oder Krankheit), unabhängig davon, ob ein kausaler Zusammenhang mit dem Studienmedikament besteht oder nicht. Erhoben wurden die UE nach Einwilligung zur Studienteilnahme. UE wurden bei jeder Visite erhoben ohne, dass das Prüfpersonal explizit nach diesen fragt.

Abnormale Laborwerte oder Testergebnisse stellen nur dann ein UE dar, wenn sie mindestens zu klinischen Anzeichen oder Symptomen führten, als klinisch bedeutsam angesehen wurden oder eine Therapie erforderlich machten. Es wurden Häufigkeit und Schweregrad von UE erfasst, die während der Behandlung mit der Studienmedikation auftraten oder sich in Bezug auf den Schweregrad verschlechterten. In die Auswertung gingen alle UE ein, die in der Zeit von der ersten Behandlung mit der Studienmedikation bis zu 7 Tage nach der letzten Dosis oder dem Zeitpunkt der Visite zu Woche 48 (je nachdem, was früher eintrat) auftraten. Die standardisierte Kodierung erfolgte mittels Medical Dictionary for Regulatory Activities (MedDRA®, Version 26.0) nach Systemorganklasse und Preferred Term.

Die Klassifikation der UE nach Schweregrad erfolgte anhand folgender Kriterien:

- Leicht: Vorübergehende Beschwerden, ohne die Beeinträchtigung von Alltagsaktivitäten.
- Moderat: Beschwerden in einem Ausmaß, das die Alltagsaktivitäten beeinträchtigt.
- Schwer: Verhinderung von Alltagsaktivitäten.

Schwerwiegende unerwünschte Ereignisse

Als schwerwiegendes unerwünschtes Ereignis (SUE) wurden Ereignisse definiert, die:

- zum Tode führen;
- unmittelbar lebensbedrohend waren;
- zu einer anhaltenden oder schwerwiegenden Behinderung/Invalidität führen;
- eine kongenitale Anomalie oder einen Geburtsfehler nach sich ziehen;
- eine Hospitalisierung oder deren Verlängerung erforderlich machen, ausgenommen:
 - Routinebehandlung oder Überwachung der Studienindikation, die nicht im Zusammenhang mit einer Verschlechterung der Erkrankung stand;
 - elektive oder vorausgeplante Behandlung für eine bereits bestehende Erkrankung, welche nicht im Zusammenhang mit der Studienindikation stand und sich nach Unterzeichnung der Einwilligungserklärung nicht verschlechterte;
 - soziale Hintergründe oder Kurzzeitpflege bei gleichzeitiger Abwesenheit einer Verschlechterung des allgemeinen Gesundheitszustands;
 - ambulante Notfallbehandlung für ein Ereignis, das nicht die Kriterien für ein SUE erfüllte und nicht in einer Hospitalisierung resultierte;
- anderweitig als medizinisch bedeutsam betrachtet wurde, d. h. die Patientin / den Patienten gefährdete oder medizinische oder chirurgische Maßnahmen zur Vermeidung eines der oben genannten Vorfälle erforderlich machte.

Die Erhebung der SUE endete 30 Tage nach der letzten Studienvisite und SUE mussten innerhalb von 24 Stunden berichtet werden.

Unerwünschte Ereignisse von besonderem Interesse

Laut Studienbericht zum Ende der Extensionsphase basieren die ausgewählten UE von besonderem Interesse auf dem Wirkmechanismus sowie präklinischen und klinischen Studien. Zu diesen gehören:

- Infektionen mit bekapselten Bakterien
- Infektionen
- Hämolyse und thrombotische Ereignisse
- Testikuläre Effekte
- Veränderungen im Zusammenhang mit der Schilddrüse

Weiterhin zählte eine verminderte Zahl an Thrombozyten zu den UE von besonderem Interesse. Nach Woche 24 wurde „Hypersensitivität“ als UE von besonderem Interesse hinzugefügt.

Bewertung

Die Operationalisierung ist nachvollziehbar. Der pU legte keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen.

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. Die Erhebung der UE wird als valide bewertet.

2.3.5 Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 5.

Tabelle 5: Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte in der Studie APPOINT

Studienphase	Screening	Behandlungsphase												Extensionsphase					
	Tag	1	7	14	28	42	56	84	112	126	140	154	168	196	224	252	280	308	336 / EOS
Woche	-8 bis -1	1	1	2	4	6	8	12	16	18	20	22	24	28	32	36	40	44	48
Endpunkt																			
Gesamtmortalität	kontinuierlich																		
Hämoglobin ¹⁾	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Transfusionsvermeidung ²⁾	kontinuierlich												kontinuierlich						
Durchbruchhämolyse ²⁾	kontinuierlich												kontinuierlich						
FACIT-Fatigue / PGIS	x	x	x	x		x		x		x	x	x	x	x	x	x	x	x	x
EORTC QLQ-C30	x	x		x		x		x		x	x	x	x	x	x	x	x	x	x
EQ-5D-VAS	x	x		x		x		x		x	x	x	x	x	x	x	x	x	x
MAVE	kontinuierlich												kontinuierlich						
Unerwünschte Ereignisse	kontinuierlich												kontinuierlich						

¹⁾ Der primäre Endpunkt wird aus Transparenzgründen im Anhang dargestellt.

²⁾ Endpunkt wird ergänzend dargestellt.

Abkürzungen: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; EOS: End of Study; EQ-5D-VAS: Visuelle Analogskala des European Quality of Life 5-Dimension; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; MAVe: Schweres unerwünschtes vaskuläres Ereignis; PGIS: Patient Global Impression of Severity.

2.4 Statistische Methoden

Analysepopulationen

- Full Analysis Set (FAS): Alle Personen mit bestätigter Eignung zur Teilnahme an der Studie, die die Studienmedikation zugewiesen bekommen haben. Das FAS wird für die Analyse von allen Wirksamkeitsendpunkten herangezogen.
- Safety Analysis Set (SAS): Alle Personen, die mindestens eine Dosis der Studienmedikation erhalten haben.

Datenschnitte

Die primäre Analyse erfolgte präspezifiziert zu Woche 24 (Tag 168) und erfolgte am 02.11.2022 zum Ende der Behandlungsphase. Es liegen ebenfalls Auswertungen zu Woche 48 (Tag 336) mit Datenschnitt vom 18.04.2023 zum Studienende (Extensionsphase) vor.

Präspezifizierte Subgruppenanalysen

Für den primären Endpunkt wurden a priori geplante Subgruppenanalysen zu den folgenden Subgruppenmerkmalen dargestellt:

- Zeit seit der Diagnose (< 3 Jahre; ≥ 3 Jahre)
- Alter (< 45 Jahre; ≥ 45 Jahre)
- Geschlecht (männlich; weiblich)
- Hb-Wert zu Baseline (< 8 g/dl; ≥ 8 g/dl)
- Vorgeschichte von MAVE vor dem Screening (ja; nein)
- Transfusionshistorie in den letzten 6 Monaten vor Beginn mit der Studienmedikation (Transfusion erhalten; keine Transfusion erhalten)
- Anzahl Transfusionen in den letzten 6 Monaten vor Beginn mit der Studienmedikation (< 2; ≥ 2)
- China; andere Länder als China

Präspezifizierung geplanter und durchgeführter Analysen

Im Rahmen der Studie APPOINT sollten deskriptive Statistiken erstellt werden, welche auf den beobachteten und nicht-imputierten Daten basieren.

Die primäre Analyse für den primären Endpunkt zu Woche 24 wurde als logistische Regression geplant, um die durchschnittliche Wahrscheinlichkeit des Erreichens des Endpunkts zu schätzen. Die ins Modell eingeschlossenen Kovariaten waren „Geschlecht“, „Alter“, „Baseline-Hb-Wert“ und „Transfusionsabhängigkeit“. Das Verhältnis der Responder wurde durch marginale Wahrscheinlichkeit geschätzt. Da das präspezifizierte Modell nicht konvergierte, wurden marginale Ansprechraten mittels „einfacher Proportion“ geschätzt. Die Auswertungen beinhalten also nicht die tatsächlichen Responder, sondern die durchschnittliche Wahrscheinlichkeit auf Basis der herangezogenen Kovariaten ein Responder zu sein. Die Berechnung und Darstellung der marginalen Responsewahrscheinlichkeit war auch für die weiteren Endpunkte vorgesehen.

Für PRO-Instrumente liegen Responderanalysen zu Woche 24 mit der relevanten Schwelle von 15 % vor. Die marginalen Responderwahrscheinlichkeiten wurden mit einem präspezifizierten „Generalisierten linearen gemischten Modell“ (GLMM) berechnet. Zu Woche 48 liegen deskriptive Auswertungen vor.

Fehlende Werte und Daten-Imputation

Es liegen durchgängig hohe Rücklaufquoten und Daten vor, sodass auf eine Beschreibung des Umgangs mit fehlenden Daten verzichtet wird.

Einschätzung der statistischen Auswertungen

Die Endpunkterhebung und Auswertung erfolgte für die Endpunkte „Transfusionsvermeidung“ und „Hämoglobin“ ab Woche 2 bis Woche 24 bzw. Woche 48. Eine Rationale für die fehlende Berücksichtigung der ersten beiden Wochen der Behandlungsphase konnte in den Studienunterlagen nicht identifiziert werden. Es werden Auswertungen ab Beginn der Behandlungsphase (Tag 1) präferiert.

Zur Auswertung des primären Endpunkts und weiterer sekundärer Endpunkte reicht der pU ein präspezifiziertes Modell zur logistischen Regression ein, welches „marginale Responderanteile“ unter Einbezug von Einflussfaktoren (Kovariaten) berechnet. Dieses Modell sollte die Wahrscheinlichkeit einer Person berechnen auf Basis der gegebenen Kovariaten ein Responder zu sein. Aus der personenspezifischen Responsewahrscheinlichkeit wird der marginale Responderanteil mittels logistischer Regression berechnet, der der durchschnittlichen Wahrscheinlichkeit entspricht, ein Responder zu sein. Das Modell unter Einbezug der Kovariaten konvergierte nicht, sodass eine einfache logistische Regression durchgeführt worden ist. Weiterhin musste die Erhöhung des Hb-Werts bei 3 von 4 Visiten gegenüber dem Baseline-Wert vorliegen. Es ist unklar, ob die logistische Regression in dem konkreten Anwendungsfall geeignet ist die marginalen Responderanteile zu schätzen, da die erhobenen Beobachtungen der jeweiligen Person im Zeitintervall zwischen Woche 18 und Woche 24 nicht mehr unabhängig voneinander sind. Wiederholte Messungen an den gleichen Personen sind nicht unabhängig, sodass Modelle benötigt werden, die diese Unterschiede und spezifischen Abhängigkeitsstrukturen berücksichtigen. In der Gesamtschau werden die „marginalen Responderanteile“ für die primären und sekundären Endpunkte nicht herangezogen und deskriptive Auswertungen (beobachtete Responder) der Erhebungen im Zeitintervall zwischen Woche 18 und Woche 24 ergänzend oder im Anhang dargestellt.

Die eingereichten Auswertungen zu den PRO-Instrumenten unter Einbeziehung des Mittelwerts der Erhebungen im Zeitintervall zwischen Woche 18 und Woche 24 berücksichtigen die wiederholte Messung derselben Person mittels GLMM, sodass die Abhängigkeitsstrukturen der Erhebungen an einer Person zwischen Woche 18 und Woche 24 und die beiden Baseline-Erhebungen berücksichtigt werden. In den PNH-assoziierten Symptomen zu Baseline wiesen 12 von 40 Personen (30 %) „keine“ und weitere 15 (37,5 %) „milde“ Fatigue-Symptome auf. Der Baseline-Wert und der Zeitpunkt der Visite gingen als linearer bzw. quadratischer Term ins Modell ein und nicht als Kovariate.

Ein GLMM schätzt in einer ersten Stufe zuerst die personenspezifische Response, um in einem zweiten Schritt Aussagen über die populationsspezifische Response treffen zu können. Durch die Schätzung der personenspezifischen Response bspw. einer nicht von Fatigue betroffenen Person wird also möglicherweise der populationsspezifische Behandlungseffekt einer Intervention auf eine hypothetische Fragestellung bei der Berechnung der personenspezifischen Response der Person ohne Fatigue extrapoliert. Bei der Berechnung von beobachteten Respondern gehen Personen ohne Fatigue zu Baseline und ohne Verbesserung zu Woche 24 als Non-Responder ein.

Für die patientenberichteten Responderanalysen zu Woche 24 legt der pU präspezifizierte Auswertungen vor, bei denen die Erhebungen der letzten 6 Wochen (Woche 18 bis Woche 24) zu Mittelwerten als Zeitintervall zusammengefasst werden. Das Heranziehen von Mittelwerten über 6 Wochen zur Responderanalyse kann anfällig für Ausreißerbeobachtungen sein, falls ein fluktuierender oder anfallsartiger Krankheitsverlauf vorliegt. Als weitere Limitation wird gesehen, dass nur eine Erhebung von 4 möglichen Erhebungen zwischen Woche 18 und Woche 24 vorliegen musste, um in die Auswertung der PRO-Instrumente einzugehen. Auf Basis der vorliegenden Daten ist unklar, bei wie vielen Personen alle 4 Beobachtungen in diesem Zeitraum vorlagen. Die Aussagekraft der Mittelwerte ist deshalb möglicherweise eingeschränkt. Ein möglicher Grund für die Verwendung des Zeitintervalls zur Responderanalyse ist eine mögliche Adjustierung im GLMM für unbalancierte personenspezifische Daten. Beispielsweise falls Unterschiede in der personen-

spezifischen Anzahl der Erhebungen vorliegen, kann es sinnvoll sein nicht zu einem spezifischen Zeitpunkt, sondern in einem spezifizierten Zeitintervall die Analyse durchzuführen. Eine Rationale für dieses Vorgehen konnte in den Studienunterlagen jedoch nicht identifiziert werden. Das Vorgehen wird im konkreten Fall aufgrund der hohen Rücklaufquoten und der Präspezifizierung als hinreichend valide und nicht ergebnisgesteuert angesehen.

Somit beinhalten die dargestellten Auswertungen des GLMM komplexe Annahmen an die vorliegende Datensituation. Die Herleitung des Schätzers und des Konfidenzintervalls erfolgte mittels Bootstrap. Der Schätzer ist durch die genannten Limitationen möglicherweise verzerrt durch eine unterschätzte Variabilität der personenspezifischen und populationspezifischen Response.

Die Auswertungsmethodik mittels GLMM und Bootstrapping zum Schätzen der marginalen Responderanteile der PRO-Instrumente kann zwar zu einem präzisen Schätzer, aber auch zu einer eingeschränkten Verallgemeinerbarkeit und Interpretierbarkeit des gesehenen Behandlungseffekts führen. Deshalb werden diese ergänzend dargestellt. Für die patientenberichteten Endpunkte werden die deskriptiven Auswertungen (beobachtete Anzahl an Respondern) zu Woche 24 zur Nutzenbewertung herangezogen und deskriptive, kontinuierliche Auswertungen zu Woche 48 im Anhang dargestellt.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Da es sich bei der Studie APPOINT um eine unkontrollierte Studie handelt, wird von einem hohen Verzerrungspotential ausgegangen.

2.6 Indirekter Vergleich

Für einen indirekten Vergleich zur Wirksamkeit von Iptacopan gegenüber „Standard of Care“ bei therapie-naiven Personen legt der pU individuelle Patientendaten aus 2 verschiedenen Registern als externe Kontrollkohorte für die Endpunkte „Transfusionsvermeidung“, „Hb-Wert-Erhöhung bei gleichzeitiger Transfusionsvermeidung“ und „Durchbruchhämolyse“ vor. Der pU reicht einen SAP, ein Studienprotokoll sowie einen Studienbericht zum indirekten Vergleich ein. Genauere Beschreibungen der Register konnten nicht identifiziert werden. Die Auswertungen zum indirekten Vergleich beziehen sich auf die 24-Wochen-Auswertungen zur primären Analyse.

Es liegt eine Auswertung mittels Propensity-Score-Weighting durch „Augmented Inverse Probability Weighting“ (AIPW) vor, um den „Average Treatment Effect on the Treated“ (ATT) zu schätzen.

Datenquellen

Im Rahmen einer Literaturrecherche und eines Advisory Boards wurden mehrere potentielle Register (insgesamt 15) vom pU identifiziert. Um als Datenquelle in Frage zu kommen mussten folgende Kriterien erfüllt werden:

- Nationales oder regionales PNH-Register
- Kontinuierlicher Einschluss und Patienten-Follow-up
- Krankheitsbasiertes Register
- Verfügbarkeit von Daten bzgl. der Behandlung mit Anti-C5-Antikörpern von Personen mit PNH
- Verfügbarkeit von hämatologischen/biochemischen Erhebungen

Für den indirekten Vergleich wurden das Französische Register für Knochenmarkversagen (RIME) sowie Daten vom St. James University Hospital in Leeds herangezogen. Für das RIME-Register liegen Daten ab 2017 vor. Labordaten wurden im Rahmen des Versorgungskontexts erhoben und entsprechen laut Angaben des pU der klinischen Praxis. Im St. James University Hospital begann

die Datenerhebung ab 2007. Für die Analyse herangezogen wurden Daten von 85 Personen, wobei 47 aus Frankreich und 38 aus dem Vereinigten Königreich eingeschlossen worden sind.

Aufgrund einer fehlenden Freigabe einer Ethikkommission konnte ein Register des Universitätsklinikums Essen nicht in der Auswertung berücksichtigt werden. Ausschlussgründe für die weiteren Register, die entweder durch die Literaturrecherche oder ein Advisory Board identifiziert worden sind, konnten den Unterlagen nicht entnommen werden.

Auswahl der Vergleichspopulation

Für die Daten aus dem Versorgungsalltag wurden für die retrospektive Kohorte APPEX Ein- und Ausschlusskriterien verwendet, die an die Kriterien der klinischen Studie APPOINT angelehnt sind. Es musste eine bestätigte PNH-Diagnose bei Erwachsenen vorliegen, bestätigt durch den Nachweis einer PNH-Klonggröße von $\geq 10\%$, sowie ein Hb-Wert von $< 10\text{ g/dl}$. Um dem Anwendungsgebiet zu entsprechen, mussten Personen in der Vergleichskohorte naiv gegenüber Komplementinhibitoren sein. Zu den Ausschlusskriterien zählten eine Knochenmarktransplantation innerhalb eines Jahres vor Beginn der Therapie mit Anti-C5-Antikörpern, hereditärer Komplementmangel sowie eine Retikulozytenzahl von $< 60 \times 10^9/\text{l}$, um ein Knochenmarkversagen auszuschließen, welches in der klinischen Studie über Retikulozyten $< 100 \times 10^9/\text{l}$, Thrombozyten $< 30 \times 10^9/\text{l}$ und Neutrophile Granulozyten $< 0,5 \times 10^9/\text{l}$ operationalisiert wurde. Da die Hämolyse Voraussetzung zur Verschreibung eines Anti-C5-Antikörpers sei, wurde auf den LDH-Wert ($> 1,5 \times \text{ULN}$) als Einschlusskriterium verzichtet.

Einschätzung zu den Registern als Kontrollkohorte (Registerdaten)

Auf Basis der vorliegenden Unterlagen liegen begrenzt Informationen zur Eignung der Register als Kontrollkohorte vor. Die beiden Register unterscheiden sich von der Verfügbarkeit der Daten voneinander (z. B. „Vorliegen einer Infektion“ fehlend in UK und „Zeichen und Symptome“ sowie „relevante medizinische Historie und weitere Erkrankungen“ fehlend in Frankreich). Wie viele Fallberichte in den Datenbanken der Register die Ein- und Ausschlusskriterien für die Grundpopulation erfüllten ist unklar. Es konnte keine Grundpopulation, also ein Patientenpool aus den herangezogenen Registern, für den indirekten Vergleich identifiziert werden. Insgesamt wurden 91 individuelle Patientendaten für den indirekten Vergleich herangezogen, wovon 85 für die Analyse berücksichtigt worden sind. Personen in Frankreich wurden durch Screening identifiziert, während die Stichprobenauswahl für die Vergleichspopulation in UK auf Basis der Screeningkriterien sowie einer Zufallsauswahl basiert. Eine Rationale oder nähere Erläuterung zu diesem Vorgehen konnte nicht identifiziert werden. Die Fallzahl von 85 basiert auf einer Fallzahlschätzung des pU, wobei die Größe der möglichen Gesamtpopulation bzw. Grundpopulation auf Basis der vorliegenden Unterlagen unklar ist.

Auf Basis der vorliegenden Unterlagen ist unklar, ob die herangezogenen Daten von angemessener Qualität für die Nutzenbewertung sind. Laut Angaben des pU seien Follow-up-Werte unregelmäßig erhoben worden. Die Zeitpunkte der Laborerhebungen seien so heterogen, dass die gesammelten Patientendaten nicht in definierte Zeitabstände unterteilt werden konnten. In der französischen Datenquelle wurden Laborwerte (u. a. Hämoglobin) häufiger erhoben als in der Datenquelle aus UK. In beiden Datenquellen seien LDH- und Retikulozyten-Werte nur unzureichend vorhanden gewesen. Die Erhebung von Baseline-Werten in den retrospektiven Erhebungen konnte nicht systematisch erfolgen und Zeitpunkte, Verhältnisse oder lokale Referenzintervalle konnten nicht adäquat erhoben werden. Als Einschlussdatum wurde der Tag der Initiierung der Anti-C5-Antikörper-Therapie herangezogen (Index-Datum). Als Baseline-Werte wurden Werte herangezogen, die innerhalb von 6 Monaten vor dem Index-Datum erhoben worden sind. In der klinischen Studie APPOINT dagegen wurden die Baseline-Daten innerhalb von 8 Wochen vor Baseline erhoben.

Das unterschiedliche regionale Setting mit einem großen Anteil einer asiatischen Population in der klinischen Studie APPOINT wird nicht berücksichtigt. Für die chinesische Population der Studie APPOINT liegen unterschiedliche Kriterien für eine Transfusion im Vergleich zur europäischen Population vor. Weiterhin kann sich der Zeitpunkt einer Therapieinitiierung zwischen europäischen und asiatischen Patientinnen und Patienten aufgrund unterschiedlicher regionaler Verfügbarkeit von Anti-C5-Antikörpern und weiteren Faktoren unterscheiden. Der Zeitraum für die Datenerhebung unterscheidet sich in beiden Registern. Die Datenerhebung in der UK-Datenbank begann ab 2007, in Frankreich in 2017. Die fehlende zeitliche Parallelität der klinischen Studien (ab 2021) und des Registers kann Auswirkungen auf die Positivität haben.

Die fehlende separate Darstellung von Baseline-Charakteristika der Register, fehlende Sensitivitätsanalysen bezüglich der einzelnen Register, eingeschränkt vergleichbare Datenverfügbarkeit und Erhebungen zwischen den Registern sowie unklare Fragestellungen und Ziele der Register führen dazu, dass ein Poolen der individuellen Patientendaten ohne weitere Berücksichtigung der Datenquelle zu einer Verzerrung führen kann.

Einschätzung der hergeleiteten Vergleichspopulation APPEX für den indirekten Vergleich

Die Ein- und Ausschlusskriterien des indirekten Vergleichs weichen von denen der klinischen Studie ab. Der häufigste Ausschlussgrund der Studie APPOINT im Screening war die labordiagnostische Evidenz von Knochenmarkversagen, welches in der Studie APPEX mit einem der drei in der Studie APPOINT verwendeten Laborwerten operationalisiert worden ist. Die Eignungsschwelle der Studie APPEX weicht von der der Studie APPOINT ab. Wenn der Baseline-Wert von $< 100 \times 10^9/l$ auf Basis des Befunds des Zentrallabors unterschritten worden ist, konnte in der Studie APPOINT der Befund des lokalen Labors herangezogen werden. Auf Basis der vorliegenden Informationen kann nicht ausgeschlossen werden, dass Personen in die Studie APPEX eingeschlossen worden sind, bei denen ein Knochenmarkversagen vorliegt. Dies ist jedoch ein eindeutiger Ausschlussgrund der Studie APPOINT. Ein Hinweis auf unterschiedliche Ein- und Ausschlusskriterien in dieser Hinsicht ist das Vorhandensein eines „Myelodysplastischen Syndroms“, welches bei 2 Personen in der Vergleichskohorte vorliegt.

In der Vergleichskohorte liegt bei 4 Personen zum Einschlusszeitpunkt eine bestehende schwere aplastische Anämie vor, während Angaben über den Schweregrad in der APPOINT-Kohorte nicht identifiziert werden konnten. Es ist auf Basis der herangezogenen Einschlusskriterien der Studie APPOINT unklar, ob Personen mit schwerer aplastischer Anämie in die Studie eingeschlossen werden konnten. Weiterhin ist unklar, ob die dargestellte relevante medizinische Historie sowie die relevanten Begleiterkrankungen unterschätzt werden, da sich in den Unterlagen nur Daten für die UK-Kohorte identifizieren lassen.

Die für den indirekten Vergleich herangezogenen Ein- und Ausschlusskriterien decken weiterhin nicht die relevanten (schweren) Komorbiditäten ab, die für die klinische Studie APPOINT als Ausschlusskriterium spezifiziert waren. In die Studie APPEX ist bspw. eine Person mit bestehendem „Budd-Chiari-Syndrom“ eingeschlossen worden, welches eine schwere Lebererkrankung mit erhöhtem Thromboserisiko darstellt und ein Ausschlussgrund für die Studie APPOINT war. Laut Angaben des pU seien in den Datenquellen auch keine Informationen über Nierenerkrankungen bzw. glomeruläre Filtrationsrate oder Kreatinin vorhanden gewesen.

Die Unterschiede in den Ein- und Ausschlusskriterien gehen mit einem hohen Risiko für einen Selektionsbias einher und resultieren potentiell in systematisch unterschiedlichen Populationen für den indirekten Vergleich. Weiterhin bestehen systematische Unterschiede in den Erhebungszeitpunkten zwischen den zu vergleichenden Populationen/Studienarmen. Dies erschwert die Vergleichbarkeit zusätzlich, da unklar ist, ob der Behandlungsbeginn in einem vergleichbaren

Krankheitsstadium erfolgte (unterschiedlicher „Time Zero“). Die fehlende zeitliche Parallelität führt zu weiteren Unsicherheiten.

Die Studie APPEX berücksichtigt retrospektive Fallberichte von Personen aus Europa. Der Großteil der Teilnehmenden der Studie APPOINT kam aus dem asiatischen Raum und gehörte zu 50 % einer chinesischen Population an. Auf Basis der vorliegenden Unterlagen kann nicht ausgeschlossen werden, dass strukturelle Unterschiede in den Baseline-Charakteristika und in der Initiierung einer Behandlung im Versorgungsalltag zwischen der asiatischen und der europäischen Population vorliegen könnten. Dies kann Auswirkungen auf die Positivität haben.

Die statistische Positivität der historischen Kontrollpopulation und der klinischen Studie (d. h. die Möglichkeit, dass beide Patientenpopulationen (theoretisch) gleichermaßen für die Therapien bzw. beide Studienpopulationen in Frage kommen und keine Gründe wie bspw. Kontraindikationen gegen eine entsprechende Zuteilung sprechen) ist insgesamt auf Basis der eingereichten Unterlagen nicht dargelegt („Strong Ignorable Treatment Assignment“-Bedingung).

Erhebung von Endpunkten in der Studie APPEX

Es liegt eine unklare Patientenrelevanz bei den für den indirekten Vergleich vorgelegten Endpunkten vor. Die Endpunkte „Hb-Wert-Erhöhung und gleichzeitige Transfusionsvermeidung“, „Transfusionsvermeidung“ sowie „Durchbruchhämolyse“ werden im indirekten Vergleich nicht im Kontext zu PNH-assoziierten Symptomen dargestellt, sodass die Patientenrelevanz weiterhin als unklar bzw. nicht dargelegt bewertet wird. Für PRO oder UE liegen keine vergleichenden Analysen vor. Weiterhin unterscheidet sich bspw. der Umgang mit fehlenden Daten und Imputation zwischen APPOINT und APPEX, sodass dies zu struktureller Verzerrung in den Endpunkten führen kann. Es wurden unterschiedliche Erhebungszeitfenster für die beiden Kohorten spezifiziert, um für die unterschiedlichen Häufigkeiten der Erhebungen zu adjustieren (siehe Einschätzung der Register).

Die strukturellen Unterschiede in der Erhebung von Labordaten in der Versorgungspraxis im Vergleich zu klinischen Studien können zu systematischen Unterschieden in der Endpunkterhebung führen („Ascertainment-Bias“). Aufgrund der eingeschränkten Datenverfügbarkeit kann nicht davon ausgegangen werden, dass die herangezogenen Mechanismen zur Eingrenzung des Ascertainment-Bias und der vorliegenden strukturellen Unsicherheiten ausreichen.

Identifikation, Auswahl und Erhebung von Confoundern

Der pU legt eine systematische Literaturrecherche und eine Einordnung durch Experten, welche zugleich die Register betreuen, sowie des Principal Investigators vor.

Literaturrecherche und Auswahl

Der pU führte eine systematische Literaturrecherche nach Leitlinien, systematischen Reviews und Beobachtungsstudien im Anwendungsgebiet durch. Es konnten 37 relevante Quellen identifiziert werden. Es liegen keine näheren Informationen zum durchgeführten Expertengespräch vor. Die nachfolgenden Tabellen 6 bis 8 geben eine Übersicht der durch die systematischen Literaturrecherche und Experteninterviews ermittelten, der präspezifizierten sowie der im Propensity-Score-Modell verwendeten Confounder.

Tabelle 6: Übersicht über die durch systematische Literaturrecherche und Experteninterviews ermittelten Confounder

Confounder-Relevanz ermittelt durch Experteninterviews			
Nicht wichtig	Weniger wichtig	Wichtig	Sehr wichtig
Schwangerschaft	Alter	Zytopenie	Performance Status
Myelodysplastisches Syndrom	Geschlecht	Knochenmark-erkrankung	Thromboembolische Ereignisse
Akute myeloische Leukämie	Ethnie	Retikulozytenzahl	Nierenerkrankung
Bedarf einer SZT	Jahr der Diagnose	Laktatdehydrogenase	Infektionen
PNH-Klonggröße		Hämoglobin	Bestehende Aplastische Anämie
Anämie			Transfusionsbedarf
Leukopenie			Knochenmarkversagen
Neutropenie			
Hämolyse			
Genetische Faktoren			
Subkategorie der PNH			
Anti-Komplement-Therapie			

Abkürzungen: PNH: Paroxysmale nächtliche Hämoglobinurie; SZT: Stammzelltransplantation.

Tabelle 7: Übersicht über die präspezifizierte Confounder-Wahl

Präspezifizierte Confounder-Wahl
<ul style="list-style-type: none"> • Transfusionsbedarf: Anzahl der in den letzten 6 Monaten transfundierten Einheiten. • Hb-Wert (g/dl) zu Baseline: Durchschnitt der verfügbaren Messungen. • Retikulozytenzahl $\times 10^9$ zu Baseline: Durchschnitt der verfügbaren Messungen. • Medizinische Historie von Infektionen: Ja / Nein. • Zeichen und Symptome: Kodiert als 0, 1, 2, 3; Schmerzen und Müdigkeit • Aplastische Anämie, Neutropenie, Knochenmarkversagen: Kodiert als 0 = nein; 1 = medizinische Historie; 2 = bestehend.

Abkürzungen: Hb: Hämoglobin.

Tabelle 8: Übersicht über verwendete Confounder im Propensity-Score-Modell

Verwendete Confounder im Propensity-Score-Modell
<ul style="list-style-type: none"> • Alter. • Geschlecht. • Transfusionsbedarf: Anzahl der in den letzten 6 Monaten transfundierten Einheiten. • Hb-Wert (g/dl) zu Baseline: Durchschnitt der verfügbaren Messungen. • Retikulozytenzahl $\times 10^9/l$ zu Baseline: Durchschnitt der verfügbaren Messungen. • Historie von MAVE. • Aplastische Anämie, Neutropenie, Knochenmarkversagen: Kodiert als 0 = nein/behoben; 1 = bestehend.

Abkürzungen: Hb: Hämoglobin; MAVE: Schweres unerwünschtes vaskuläres Ereignis.

Bewertung der verwendeten Confounder

Die Durchführung einer systematischen Literaturrecherche und einer anschließenden Befragung zur Einordnung der Confounder nach Relevanz wird insgesamt als sachgerecht angesehen. Das Verfahren zur Einordnung und Konsentierung der Relevanz der Confounder ist auf Basis der eingereichten Unterlagen nicht nachvollziehbar und stellt aufgrund der Expertenauswahl (siehe oben) keine von der Datenverfügbarkeit unabhängige Bewertung der Relevanz sicher. Mehrere Confounder, die häufig in Leitlinien als relevant identifiziert worden sind, sind im Expertengespräch als „nicht wichtig“ klassifiziert worden. Ein Großteil der als „sehr wichtig“ oder „wichtig“ eingestuften Confounder ging nicht ins Propensity-Score-Modell ein.

Die Operationalisierung der Aplastischen Anämie / Neutropenie / Knochenmarkversagen aus drei unterschiedlichen Confoundern zu einem kombinierten Confounder ist auf Basis der im vorherigen Abschnitt genannten Limitationen als fraglich anzusehen. Der als wichtig identifizierte Confounder „Nierenerkrankung“ konnte nicht ins Modell aufgenommen werden aufgrund eingeschränkter Datenverfügbarkeit. Als Perturbationsvariable, um ungemessenem Confounding vorzubeugen, wurden „Alter“ und „Geschlecht“ ins Modell aufgenommen. Der als „sehr wichtig“ identifizierte Confounder „Performance Status“ wurde ebenfalls nicht ins Modell aufgenommen.

Auf Basis der eingereichten Studienunterlagen und Änderungen hinsichtlich der Relevanz der Confounder kann eine durch die Datenverfügbarkeit getriebene Auswahl der Confounder nicht ausgeschlossen werden. Die Auswahl der verwendeten Confounder für das Propensity-Score-Modell für die Hauptanalyse wird als unvollständig und daher nicht sachgerecht angesehen.

Statistische Methoden für den indirekten Vergleich

Der pU verwendet fortgeschrittene statistische Methoden zur Herleitung der Propensity Scores, die an Machine-Learning-Methoden angelehnt sind. Als Adjustierungsmethode wurde eine AIPW verwendet mit einem ATT-Schätzer als „Target of Inference“. IPTW-Analysen erzeugen eine Pseudopopulation, mit dem Ziel potentiell Confounding aufzuheben.

Für die Durchführung der Analyse und um interpretierbare Schätzer des Therapieeffekts zu erhalten, muss u. a. die Bedingung der Positivität gegeben sein und es muss eine ausreichende Überlappung sowie nach Anwendung des Propensity-Score-Verfahrens Balanciertheit vorliegen.

Die genannten Limitationen zur Erfüllung der Grundbedingungen (Vollständigkeit der Confounder, Positivität, „Time Zero“) führen dazu, dass der berechnete Effektschätzer auf Basis eines ATT-Schätzers für Aussagen zur Wirksamkeit nicht sinnvoll interpretierbar ist. Dieser bezieht sich nicht mehr auf die gesamte hergeleitete Vergleichspopulation, sondern auf eine konstruierte Population, die nicht klar beschreibbar ist. Es ist unklar, ob der ATT-Ansatz für den Rahmen einer Nutzenbewertung geeignet ist. Dieser gewichtet die herangezogene Vergleichspopulation der APPEX-Kohorte so, als ob diese für die Intervention geeignet gewesen wäre. Der geschätzte Effekt bezieht sich auf die Personen, die tatsächlich die Intervention erhalten haben und nicht auf alle Personen, die dafür geeignet gewesen wären (Zielpopulation).

Es konnte keine Verteilung der Propensity Scores in den Unterlagen identifiziert werden, sodass die Überlappung der Studienpopulationen unklar ist. Eine zum Großteil überlappende Verteilung der jeweiligen Propensity Scores der beiden zu vergleichenden Kohorten auf Basis der erhobenen Variablen, ist, neben der vollständigen Erhebung der Confounder, eine notwendig zu erfüllende Bedingung. Weiterhin wurden die Gewichte in der Vergleichskohorte nicht berichtet. Aufgrund des als unvollständig bewerteten Propensity-Score-Modells wird auf die Beschreibung der Balanciertheit verzichtet.

Die AIPW-Schätzung wurde als doppelt-robustes Modell mit einem Outcome-Modell spezifiziert. Bei einem doppelt-robusten Modell wird das Propensity-Score-Modell so erweitert, dass das

gesehene Outcome einer Person auf Basis des geschätzten Propensity Scores einer Person gewichtet wird. Das Outcome-Modell wurde analog zur Schätzung des primären Endpunkts der klinischen Studie APPOINT im SAP der Studie APPEX mit den Variablen „Alter“, „Geschlecht“, „Baseline-Hb-Wert“ und „Transfusionsabhängigkeit“ spezifiziert. Im Outcome-Modell des Studienberichts für die APPEX-Kohorte finden sich abweichend folgende Variablen im Modell: Alter, Geschlecht, Baseline-Hb-Wert, Retikulozytenzahl, Anzahl der verabreichten Transfusionseinheiten innerhalb von 6 Monaten vor Studieneinschluss, bestehende aplastische Anämie. Das Outcome-Modell für den primären Endpunkt zur Berechnung der marginalen Responderanteile konvergierte in der klinischen Studie APPOINT nicht, sodass dort eine einfache logistische Regression berechnet worden ist. Nähere Informationen zur Auswertung der Studie APPOINT finden sich in Kapitel 2.4. Sensitivitätsanalysen mit einem weiteren Adjustierungsansatz oder „Average Treatment Effect“-Schätzer konnte in den Studienunterlagen nicht identifiziert werden.

Anhand der vorgelegten Unterlagen kann nicht abschließend beurteilt werden, inwiefern sich die erzeugte Pseudopopulation von der Gesamtpopulation unterscheidet und ob sich die Ergebnisse für die Pseudopopulation auf die Gesamtpopulation übertragen lassen. Als weitere Limitation wird die geringe Stichprobengröße der klinischen Studie APPOINT gesehen, welche die Validität von Propensity-Score-Verfahren mittels Gewichtung (durch eine mögliche Unterschätzung der Standardfehler, Konfidenzintervalle und p-Werte) einschränkt [16].

Die herangezogenen statistischen Methoden zur Berechnung der Propensity Scores erfordern z. T. sehr starke Annahmen an die Datensituation. Die fehlende Berücksichtigung von relevanten Confoundern und deren eingeschränkte Vergleichbarkeit bezüglich „Time Zero“ sowie nicht dargelegte Positivität können nicht durch komplexe statistische Methoden mit unbegründeten Entscheidungsstrukturen (Effektive Fallzahl, Fallzahlkalkulation, Divergenz, aufgetretene Gewichte, SMD) zur Modellwahl aufgewogen werden. Auf Basis der vorliegenden Unterlagen führen die eingereichten Analysen zu dem Modell, welches am besten zu der vorliegenden Datensituation und der herangezogenen Entscheidungsstruktur passt. Eine datengetriebene Modellauswahl kann auf Basis der vorliegenden Unterlagen nicht ausgeschlossen werden. Im Kontext der Nutzenbewertung ist ein verallgemeinerbares Modell von Relevanz.

2.6.1 Zusammenfassende Bewertung der Methodik des indirekten Vergleichs

Der pU legt einen indirekten Vergleich von Iptacopan gegenüber „Standard of Care“ bei therapie-naiven Patientinnen und Patienten in Endpunkten mit unklarer Patientenrelevanz vor. Dazu wurde unter Einbezug von Registerdaten mittels retrospektiven Fallberichten aus einer nicht näher beschriebenen Grundpopulation ein Vergleichsarm konstruiert. Die herangezogenen Register weisen erhebliche Unterschiede in der Datenerhebung von Baseline-Charakteristika und Follow-up-Werten und in der Datenverfügbarkeit zwischen den Registern auf. Auf Basis der vorliegenden Unterlagen kann nicht davon ausgegangen werden, dass die herangezogenen Register als Datenquelle zur Konstruktion einer Vergleichspopulation im Rahmen der Nutzenbewertung geeignet sind.

Die Unterschiede in den Ein- und Ausschlusskriterien sowie Region der Register gehen mit einem hohen Risiko für einen Selektionsbias einher und resultieren potentiell in systematisch unterschiedlichen Populationen für den indirekten Vergleich. Weiterhin bestehen systematische Unterschiede in den Erhebungszeitpunkten zwischen den zu vergleichenden Populationen/Studienarmen. Dies erschwert die Vergleichbarkeit zusätzlich, da unklar ist, ob der Behandlungsbeginn in einem vergleichbaren Krankheitsstadium erfolgte (unterschiedlicher „Time Zero“). Die eingeschränkte zeitliche Parallelität führt zu weiteren Unsicherheiten.

Die Durchführung einer systematischen Literaturrecherche und einer anschließenden Befragung zur Einordnung der Confounder nach Relevanz wird prinzipiell als sachgerecht angesehen. Das Verfahren zur Einordnung und Konsentierung der Relevanz der Confounder ist auf Basis der eingereichten Unterlagen jedoch nicht nachvollziehbar und stellt aufgrund der Expertenauswahl (die zugleich Registerbetreiber sind) keine von der Datenverfügbarkeit unabhängige Bewertung der Relevanz sicher. Eine durch die Datenverfügbarkeit getriebene Auswahl der Confounder kann daher nicht ausgeschlossen werden. Darüber hinaus werden nicht alle als sehr wichtig benannten Confounder in der Analyse berücksichtigt. Zudem liegt eine hohe Anzahl von fehlenden Werten bei einigen wichtigen Baseline-Charakteristika und Confoundern vor. Eine ausreichende Strukturgleichheit der beiden Kohorten kann auf Basis der vorgelegten Daten und Analysen nicht erreicht werden.

Insgesamt führen die limitierte Datenverfügbarkeit, die unklare Strukturgleichheit und Vergleichbarkeit der Population der Studie APPOINT und der externen Vergleichspopulation sowie die unzureichende Confounder-Adjustierung mittels der verwendeten Propensity-Score-Methoden dazu, dass der indirekte Vergleich nicht zur Nutzenbewertung herangezogen wird.

3 Ergebnisse der eingeschlossenen Studien

3.1 Studiencharakteristika und Studienmedikation

Tabelle 9: Allgemeine Angaben; Studie APPOINT

Studie APPOINT Allgemeine Angaben	Iptacopan
FAS, n (%)	40 (100)
SAS, n (%)	40 (100)
Abbruch der Einnahme der Studienmedikation, n (%)	0
Behandlungsphase: Mediane Behandlungsdauer in Wochen (min; max)	24 (23; 25)
Extensionsphase: Mediane Behandlungsdauer in Wochen (min; max)	48 (17; 54)

¹⁾ Definition siehe Kapitel 2.4.

Abkürzungen: FAS: Full Analysis Set; SAS: Safety Analysis Set.

Tabelle 10: Charakterisierung der Studienpopulation; Studie APPOINT

Studie APPOINT Charakterisierung der Studienpopulation	FAS N = 40
Alter (Jahre) MW (SD) Median (min; max)	42,1 (15,9) 38,5 (18; 81)
Altersgruppe (Jahre), n (%) 18 bis < 65 65 bis < 75 > 75	37 (92,5) 2 (5,0) 1 (2,5)
Geschlecht, n (%) männlich weiblich	23 (57,5) 17 (42,5)
BMI (kg/m ²) MW (SD) Median (min, max)	24,6 (3,3) 24,4 (18,9; 35,7)
Abstammung (genetisch), n (%) kaukasisch/weiß afroamerikanisch asiatisch	12 (30,0) 1 (2,5) 27 (67,5)
Patientenklassifikation, n (%) chinesisch nicht-chinesisch	20 (50,0) 20 (50,0)
Krankheitsdauer (Jahre) MW (SD) Median (min; max)	4,7 (5,5) 3,63 (0,01; 23,2)
Baseline-Hämoglobin (g/dl) ¹⁾ MW (SD) Median (min, max)	8,2 (1,1) 8,05 (5,8; 10)
Baseline-Hämoglobin, n (%) < 8 g/dl ≥ 8 g/dl	15 (37,5) 25 (62,5)

Studie APPOINT Charakterisierung der Studienpopulation	FAS N = 40
<i>Baseline-LDH-Level (U/l), n (%)</i> MW (SD) Median (min, max)	1.698,8 (683,3) 1.581,5 (522; 3.244)
<i>Transfusion innerhalb von 12 Monaten vor Screening, n (%)</i> ja nein	27 (67,5) 13 (32,5)
<i>Transfusion innerhalb von 6 Monaten vor Behandlungsbeginn, n (%)</i> ja nein	28 (70,0) 12 (30,0)
<i>Anzahl Transfusionen innerhalb von 6 Monaten vor Behandlungsstart, n (%)</i> < 2 ≥ 2	19 (47,5) 21 (52,5)
<i>Anzahl Transfusionen innerhalb von 6 Monaten vor Behandlungsstart bei Personen mit Transfusion</i> MW (SD) Median (min, max)	N = 28 3,1 (2,1) 2,0 (1,8)
<i>Thrombozyten im Blut (10⁹/l)</i> MW (SD) Median (min, max)	159,4 (61,1) 150,5 (35; 316)
<i>Retikulozytenzahl im Blut (10⁹/l)</i> MW (SD) Median (min, max)	154,33 (63,7) 139,2 (59,4; 324,8)
<i>PNH-Klongröße in %</i> MW (SD) Median (min, max)	42,7 (21,2) 40,3 (9; 92,87)
<i>FACIT-Fatigue-Score zu Baseline²⁾</i> MW (SD) Median (min, max)	32,8 (10,2) 34,3 (13,0; 50,5)
<i>Vorgeschichte von mind. einem MAVE, n (%)</i> ja nein	5 (12,5) 35 (87,5)
<i>Fortbestehende aplastische Anämie, n (%)</i> ja nein	16 (40,0) 24 (60,0)
<i>Personen mit mind. einem PNH-assoziierten Symptom, n (%)</i>	39 (97,5)
<i>PNH-assoziierte Symptome zu Baseline, n (%)</i> Hämoglobinurie Empfundene Schwäche oder Müdigkeit Kurzatmigkeit / Dyspnoe Dysphagie / Schluckbeschwerden Brustschmerzen Abdomineller Schmerz Erektile Dysfunktion	29 (72,5) 28 (70,0) 12 (30,0) 10 (25,0) 7 (17,5) 11(27,5) 8 (20,0)

¹⁾ Bei den Hb-Werten zu Baseline handelt es sich um den Mittelwert aus 2 Erhebungen. Wenn eine Transfusion gegeben wurde nach erstmaliger Erhebung, geht der ersterhobene Wert in die Berechnung ein.

²⁾ Skala von 0 bis 52; höhere Werte zeigen geringere Beeinträchtigungen an.

Abkürzungen: FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; FAS: Full Analysis Set; Hb: Hämoglobin; LDH: Laktatdehydrogenase; MAVE: Schweres unerwünschtes vaskuläres Ereignis; MW: Mittelwert; PNH: Paroxysmale nächtliche Hämoglobinurie; ULN: Upper Limit of Normal; SD: Standardabweichung.

Protokollverletzungen

Anhand der in der Studie APPOINT dokumentierten Protokollverletzungen ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Begleitmedikation

Anhand der in der Studie APPOINT dokumentierten Begleitmedikation ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Folgetherapien

In der Studie APPOINT wurden keine Folgetherapien verabreicht.

3.2 Mortalität

Gesamtmortalität

In der Studie APPOINT wurden Todesfälle im Rahmen der Sicherheitserhebung erhoben. Während der gesamten Studiendauer sind keine Todesfälle aufgetreten.

3.3 Morbidität

3.3.1 Transfusionsvermeidung (ergänzend dargestellt)

Tabelle 11: Transfusionsvermeidung während der gesamten Studiendauer (Woche 48); Studie APPOINT; FAS (Datenschnitt: 18.04.2023)

Studie APPOINT Transfusionsvermeidung (ergänzend dargestellt)	Iptacopan N = 40 n (%)
Personen mit Transfusionsvermeidung während der gesamten Studiendauer (Woche 48)	33 (82,5)
Personen mit Transfusionsvermeidung zwischen Tag 14 und Woche 48	39 (97,5)
Personen, die im Studienverlauf mind. einmal die Transfusionskriterien erfüllt haben	7 (17,5)
Personen, die tatsächlich eine Transfusion erhalten haben	6 (15)

Abkürzungen: FAS: Full Analysis Set.

3.3.2 Durchbruchhämolyse (ergänzend dargestellt)

Tabelle 12: Auftreten einer Durchbruchhämolyse während der gesamten Studiendauer (Woche 48) (ergänzend dargestellt); Studie APPOINT, FAS (Datenschnitt: 18.04.2023)

Studie APPOINT Durchbruchhämolyse (ergänzend dargestellt)	Iptacopan N = 40 n (%)
Personen mit mind. einem Ereignis	2 (5,0)
Personen, die in Folge einer Durchbruchhämolyse die Transfusionskriterien erfüllt haben	2 (5,0)
Personen, die eine Transfusion in Folge einer Durchbruchhämolyse erhalten haben	1 (2,5)
Aufgetretene Symptome, die mit einer Durchbruchhämolyse assoziiert sind	2 (5,0)
Massive Hämoglobinurie	2 (5,0)
Schmerzkrisen	1 (2,5)
Fatigue	2 (5,0)

Abkürzungen: FAS: Full Analysis Set.

3.3.3 Fatigue

FACIT-Fatigue

Tabelle 13: Verbesserung des FACIT-Fatigue zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022)

Studie APPOINT FACIT-Fatigue ¹⁾²⁾	Iptacopan N = 40
Anteil der Personen mit einer Verbesserung, n (%)	22 ³⁾ (55,0)
Marginaler Responderanteil (%) [95%-KI] ⁴⁾ (ergänzend dargestellt)	61,5 [47,6; 76,0]

¹⁾ Eine Zunahme des Scores um ≥ 8 Punkte im Vergleich zu Baseline wird als Verbesserung angesehen, dies entspricht 15 % der Skalenspannweite. Die Skalenspannweite beträgt 0 bis 52; höhere Werte zeigen geringere Beeinträchtigungen an.

²⁾ Es musste zu mind. einer Visite zwischen Woche 18 und Woche 24 (Ende Behandlungsphase) eine Erhebung vorliegen.

³⁾ Es wurden die Mittelwerte der beobachteten Werte für die Visiten zwischen Woche 18 und Woche 24 berücksichtigt. Aufgrund der insgesamt hohen Rücklaufquoten für die in Berechnung der Mittelwerte einbezogenen Visiten wird das Vorgehen als adäquat befunden.

⁴⁾ Die Auswertungen zu den Responderanalysen basieren auf einem GLMM unter Verwendung einer Logit-Link-Funktion mit der Beobachtungszeit (Visite) in Wochen als kontinuierliche Kovariate. Das Modell enthält den Baseline-Wert und den Zeitpunkt der Visite als linearen und quadratischen Term. Interaktionsterme für den Baseline-Wert sowie die linearen und quadratischen Zeitpunkte der Visite sind Teil des Modells. Der marginale Responderanteil inkl. 95%-KI wurde durch Bootstrapping ermittelt. Für die Analyse wurde der Mittelwert der Schätzer für die Visiten zwischen Woche 18 und Woche 24 herangezogen.

Abkürzungen: FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; FAS: Full Analysis Set; GLMM: Generalisiertes lineares gemischtes Modell; KI: Konfidenzintervall.

PGIS

Tabelle 14: Verbesserung des PGIS zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022)

Studie APPOINT PGIS ¹⁾²⁾³⁾	Iptacopan N = 40
Anteil der Personen mit einer Verbesserung, n (%)	24 ⁴⁾ (60,0)
Marginaler Responderanteil, in % [95%-KI] ⁵⁾ (ergänzend dargestellt)	71,2 [58,1; 82,9]

¹⁾ Die Einschätzung erfolgt auf einer 5-Punkte-Likert-Skala von 0 („keine Symptome“) bis 4 („sehr schwere Symptome“).

²⁾ Als Verbesserung ist der Wechsel in ≥ 1 tiefere Kategorie definiert.

³⁾ Es musste zu mind. einer Visite zwischen Woche 18 und Woche 24 (Ende Behandlungsphase) eine Erhebung vorliegen.

⁴⁾ Es wurden die Mittelwerte der beobachteten Werte für die Visiten zwischen Woche 18 und Woche 24 berücksichtigt. Aufgrund der insgesamt hohen Rücklaufquoten für die in Berechnung der Mittelwerte einbezogenen Visiten wird das Vorgehen als adäquat befunden.

⁵⁾ Die Auswertungen zu den Responderanalysen basieren auf einem GLMM unter Verwendung einer Logit-Link-Funktion mit der Beobachtungszeit (Visite) in Wochen als kontinuierliche Kovariate. Das Modell enthält den Baseline-Wert und den Zeitpunkt der Visite als linearen und quadratischen Term. Interaktionsterme für den Baseline-Wert sowie die linearen und quadratischen Zeitpunkte der Visite sind Teil des Modells. Der marginale Responderanteil inkl. 95%-KI wurde durch Bootstrapping ermittelt. Für die Analyse wurde der Mittelwert der Schätzer für die Visiten zwischen Woche 18 und Woche 24 herangezogen.

Abkürzungen: FAS: Full Analysis Set; GLMM: Generalisiertes lineares gemischtes Modell; KI: Konfidenzintervall; PGIS: Patient Global Impression of Severity.

3.3.4 EORTC QLQ-C30 – Symptomskalen und Symptome

Die Symptomskala „Übelkeit und Erbrechen“ sowie die Symptome „Appetitverlust“, „Obstipation“ und „Diarrhö“ werden aufgrund unklarer Patientenrelevanz und Validität im vorliegenden Anwendungsgebiet ergänzend dargestellt. Die Symptomskalen „Schmerz“ und „Fatigue“ sowie die Symptome „Dyspnoe“ und „Schlaflosigkeit“ werden berücksichtigt.

Tabelle 15: Verbesserung der Symptomatik mittels EORTC QLQ-C30 zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022)

Studie APPOINT Verbesserung im EORTC QLQ-C30¹⁾	Iptacopan N = 40
Symptomskala „Fatigue“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	27 (67,5) 70,3 [57,3; 81,5]
Symptomskala „Schmerz“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	16 (40,0) 41,8 [28,3; 54,9]
Symptom „Dyspnoe“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	22 (55,0) 57,9 [43,7; 71,1]
Symptom „Schlaflosigkeit“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	13 (32,5) 37,4 [25,2; 50,2]
Symptomskala „Übelkeit und Erbrechen“ (ergänzend dargestellt)	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	10 (25,0) 28,6 [16,0; 42,8]
Symptom „Appetitverlust“ (ergänzend dargestellt)	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	22 (55,0) 58,3 [43,0; 71,6]
Symptom „Obstipation“ (ergänzend dargestellt)	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	4 (10,0) 19,7 [9,7; 31,6]
Symptom „Diarrhö“ (ergänzend dargestellt)	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	6 (15,0) Nicht konvergiert

¹⁾ Werte von 0 bis 100; höhere Werte entsprechen einer schwereren Krankheitssymptomatik.

²⁾ Eine Abnahme des Scores um ≥ 10 Punkte im Vergleich zu Baseline wurde als Verbesserung angesehen.

³⁾ Es wurden die Mittelwerte der beobachteten Werte für die Visiten zwischen Woche 18 und Woche 24 berücksichtigt. Aufgrund der insgesamt hohen Rücklaufquoten für die in Berechnung der Mittelwerte einbezogenen Visiten wird das Vorgehen als adäquat befunden.

⁴⁾ Die Auswertungen zu den Responderanalysen basieren auf einem GLMM unter Verwendung einer Logit-Link-Funktion mit der Beobachtungszeit (Visite) in Wochen als kontinuierliche Kovariate. Das Modell enthält den Baseline-Wert und den Zeitpunkt der Visite als linearen und quadratischen Term. Interaktionsterme für den Baseline-Wert sowie die linearen und quadratischen Zeitpunkte der Visite sind Teil des Modells. Der marginale Responderanteil inkl. 95%-KI wurde durch Bootstrapping ermittelt. Für die Analyse wurde der Mittelwert der Schätzer für die Visiten zwischen Woche 18 und Woche 24 herangezogen.

Abkürzungen: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; FAS: Full Analysis Set; GLMM: Generalisiertes lineares gemischtes Modell; KI: Konfidenzintervall.

3.3.5 EQ-5D-VAS

Tabelle 16: Verbesserung der EQ-5D-VAS um ≥ 15 Punkte zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS (Datenschnitt: 02.11.2022)

Studie APPOINT EQ-5D-VAS ¹⁾	Iptacopan N = 40
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%)	20 (50)
Marginaler Responderanteil (%) [95%-KI] ⁴⁾ (ergänzend dargestellt)	55,2 [39,2; 68,8]

¹⁾ Werte von 0 bis 100; höhere Werte entsprechen einem besseren Gesundheitszustand.

²⁾ Eine Abnahme des Scores um ≥ 10 Punkte im Vergleich zu Baseline wurde als Verbesserung angesehen.

³⁾ Es wurden die Mittelwerte der beobachteten Werte für die Visiten zwischen Woche 18 und Woche 24 berücksichtigt. Aufgrund der insgesamt hohen Rücklaufquoten für die in Berechnung der Mittelwerte einbezogenen Visiten wird das Vorgehen als adäquat befunden.

⁴⁾ Die Auswertungen zu den Responderanalysen basieren auf einem GLMM unter Verwendung einer Logit-Link-Funktion mit der Beobachtungszeit (Visite) in Wochen als kontinuierliche Kovariate. Das Modell enthält den Baseline-Wert und den Zeitpunkt der Visite als linearen und quadratischen Term. Interaktionsterme für den Baseline-Wert sowie die linearen und quadratischen Zeitpunkte der Visite sind Teil des Modells. Der marginale Responderanteil inkl. 95%-KI wurde durch Bootstrapping ermittelt. Für die Analyse wurde der Mittelwert der Schätzer für die Visiten zwischen Woche 18 und Woche 24 herangezogen.

Abkürzungen: EQ-5D-VAS: Visuelle Analogskala des European Quality of Life 5-Dimension; FAS: Full Analysis Set; GLMM: Generalisiertes lineares gemischtes Modell; KI: Konfidenzintervall.

3.3.6 MAVE

In der Studie APPOINT trat im gesamten Studienzeitraum (Tag 1 bis Woche 48) kein MAVE auf.

3.4 Lebensqualität

EORTC QLQ-C30 – Funktionskalen und Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“

Tabelle 17: Verbesserung der Lebensqualität mittels EORTC QLQ-C30 um ≥ 10 Punkte zum Ende der Behandlungsphase (Woche 24); Studie APPOINT, FAS; Datenschnitt: 02.11.2022

Studie APPOINT Verbesserung im EORTC QLQ-C30 ¹⁾	Iptacopan N = 40
Skala „Körperliche Funktion“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%)	24 (60,0)
Marginaler Responderanteil (%) [95%-KI] ⁴⁾	64,1 [50,4; 76,6]
Skala „Rollenfunktion“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%)	27 (67,5)
Marginaler Responderanteil (%) [95%-KI] ⁴⁾	60,6 [46,4; 73,8]
Skala „Emotionale Funktion“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%)	22 (55,0)
Marginaler Responderanteil (%) [95%-KI] ⁴⁾	50,7 [37,0; 64,6]
Skala „Kognitive Funktion“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%)	13 (32,5)
Marginaler Responderanteil (%) [95%-KI] ⁴⁾	28,3 [16,7; 41,5]

Studie APPOINT Verbesserung im EORTC QLQ-C30 ¹⁾	Iptacopan N = 40
Skala „Soziale Funktion“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	20 (50,0) 51,5 [37,8; 64,4]
Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“	
Anteil der Personen mit einer Verbesserung ²⁾³⁾ , n (%) Marginaler Responderanteil (%) [95%-KI] ⁴⁾	30 (75,0) 76,3 [63,5; 87,1]

¹⁾ Werte von 0 bis 100; höhere Werte entsprechen einer besseren Funktion bzw. Gesundheit/Lebensqualität.

²⁾ Eine Zunahme des Scores um ≥ 10 Punkte im Vergleich zu Baseline wurde als Verbesserung angesehen.

³⁾ Es wurden die Mittelwerte der beobachteten Werte für die Visiten zwischen Woche 18 und Woche 24 berücksichtigt. Aufgrund der insgesamt hohen Rücklaufquoten für die in Berechnung der Mittelwerte einbezogenen Visiten wird das Vorgehen als adäquat befunden. Es musste zu mind. einer Visite zwischen Woche 18 und Woche 24 (Ende der Behandlungsphase) eine Erhebung vorliegen.

⁴⁾ Die Auswertungen zu den Responderanalysen basieren auf einem GLMM unter Verwendung einer Logit-Link-Funktion mit der Beobachtungszeit (Visite) in Wochen als kontinuierliche Kovariate. Das Modell enthält den Baseline-Wert und den Zeitpunkt der Visite als linearen und quadratischen Term. Interaktionsterme für den Baseline-Wert sowie die linearen und quadratischen Zeitpunkte der Visite sind Teil des Modells. Der marginale Responderanteil inkl. 95%-KI wurde durch Bootstrapping ermittelt. Für die Analyse wurde der Mittelwert der Schätzer für die Visiten zwischen Woche 18 und Woche 24 herangezogen.

Abkürzungen: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; FAS: Full Analysis Set; GLMM: Generalisiertes lineares gemischtes Modell; KI: Konfidenzintervall.

3.5 Sicherheit

Tabelle 18: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der gesamten Studie (Woche 48); Studie APPOINT, SAS (Datenschnitt: 18.04.2023)

Studie APPOINT Zusammenfassung der UE Personen mit mindestens einem ...	Iptacopan N = 40 n (%)
UE (ergänzend dargestellt)	37 (92,5)
UE CTCAE-Grad ≥ 3	4 (10,0)
SUE	8 (20,0)
UE, das zum Abbruch der Studienmedikation führte	0

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; SAS: Safety Analysis Set; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Der pU legte keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen.

Unerwünschte Ereignisse

*Tabelle 19: UE mit Inzidenz ≥ 10 % während der gesamten Studiendauer (Woche 48);
Studie APPOINT, SAS (Datenschnitt: 18.04.2023)*

Studie APPOINT UE <i>MedDRA-Systemorganklasse</i> <i>Preferred Term</i>	Iptacopan N = 40 n (%)
Augenerkrankungen	4 (10,0)
Erkrankungen des Gastrointestinaltrakts	13 (32,5)
Diarrhö	5 (12,5)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	9 (22,5)
Infektionen¹⁾	14 (35,0)
COVID-19	8 (20,0)
Infektionen der oberen Atemwege	6 (15,0)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	6 (15,0)
Untersuchungen	12 (30,0)
Stoffwechsel- und Ernährungsstörungen	9 (22,5)
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen	5 (12,5)
Erkrankungen des Nervensystems	12 (30,0)
Kopfschmerzen	11 (27,5)
Erkrankungen der Nieren und Harnwege	5 (12,5)
Erkrankungen der Atemwege, des Brustraums und des Mediastinums	6 (15,0)
Erkrankungen der Haut und des Unterhautzellgewebes	5 (12,5)

¹⁾ UE von besonderem Interesse.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; SAS: Safety Analysis Set; UE: Unerwünschtes Ereignis.

*Tabelle 20: Schwere UE (Grad 3 oder höher), die bei ≥ 5 % der Personen, nach Systemorganklasse und Preferred Term, während der gesamten Studiendauer (Woche 48) aufgetreten sind;
Studie APPOINT, SAS (Datenschnitt: 18.04.2023)*

Studie APPOINT Schwere UE <i>MedDRA-Systemorganklasse</i> <i>Preferred Term</i>	Iptacopan N = 40 n (%)
Infektionen¹⁾	5 (12,5)
COVID-19	2 (5,0)

¹⁾ UE von besonderem Interesse.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; SAS: Safety Analysis Set; UE: Unerwünschtes Ereignis.

Tabelle 21: SUE, die bei $\geq 5\%$ der Personen, nach Systemorganklasse und Preferred Term, während der gesamten Studiendauer (Woche 48) aufgetreten sind; Studie APPOINT; SAS (Datenschnitt: 18.04.2023)

Studie APPOINT SUE MedDRA-Systemorganklasse Preferred Term	Iptacopan N = 40 n (%)
Infektionen¹⁾	5 (12,5)
COVID-19	2 (5,0)

¹⁾ UE von besonderem Interesse

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; SAS: Safety Analysis Set; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Tabelle 22: UE von besonderem Interesse während der gesamten Studiendauer (Woche 48); Studie APPOINT, SAS (Datenschnitt: 18.04.2023)

Studie APPOINT UE von besonderem Interesse	Iptacopan N = 40 n (%)
PNH-Hämolyse und Thrombose	3 (7,5)
Hypersensitivität	4 (10,0)
Schwere oder schwerwiegende Infektionen	5 (12,5)
Infektionen mit bekapselten Bakterien	2 (5,0)
Testikuläre Effekte	1 (2,5)
Veränderung im Zusammenhang mit der Schilddrüse	1 (2,5)

Abkürzungen: PNH: Paroxysmale nächtliche Hämoglobinurie; SAS: Safety Analysis Set; UE: Unerwünschtes Ereignis.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Iptacopan

Gemäß Fachinformation wird Iptacopan (Fabhalta®) angewendet als Monotherapie zur Behandlung erwachsener Personen mit PNH, die eine hämolytische Anämie aufweisen. Für die vorliegende Nutzenbewertung sind die Personen mit PNH relevant, die eine hämolytische Anämie aufweisen sowie therapienaiv gegenüber Komplementinhibitoren sind. Die empfohlene Dosierung beträgt 200 mg Iptacopan, die zweimal täglich oral eingenommen wird. Es handelt sich bei der Gabe von Iptacopan um eine Langzeittherapie.

Für die Nutzenbewertung relevant ist die klinische Studie APPOINT, eine einarmige Studie mit 40 erwachsenen Personen mit PNH (RBC-Klonggröße $\geq 10\%$ zum Studieneinschluss), die einen mittleren Hb-Spiegel von < 10 g/dl und LDH-Spiegel von $> 1,5 \times$ ULN während der Screening-Periode aufwiesen und nicht mit einem Komplementinhibitor vorbehandelt waren. Die in der Studie eingesetzte Dosierung stimmt mit der empfohlenen Iptacopan-Anwendung gemäß Fachinformation überein.

Gemäß Einschlusskriterien handelt es sich bei der Population der Studie APPOINT um erwachsene Personen. Der Großteil der eingeschlossenen Personen war bei Studieneinschluss unter 65 Jahre alt, somit liegen limitierte Daten für ältere Alterskohorten vor.

Bei fast allen eingeschlossenen Personen (97,5 %) lag mindestens ein PNH-assoziiertes Symptom zu Baseline vor. Die Schweregrade der Symptome wurden mit der UE-Klassifikation ermittelt. Die am häufigsten aufgetretenen Symptome waren „Verfärbter (Morgen)-Urin / Hämoglobinurie“ bei 72,5 % und „Gefühl der Schwäche / Müdigkeit“ bei 70 % der eingeschlossenen Personen.

Von der Studie ausgeschlossen wurden u. a. Personen mit PNH und gleichzeitiger labordiagnostischer Evidenz von Knochenmarkversagen. Weiterhin wurden Personen mit schwerer Nierenfunktionsstörung von der Studie ausgeschlossen. Die Übertragbarkeit der gesehenen Ergebnisse auf diese Populationen kann eingeschränkt sein. Der Einfluss auf die Verallgemeinerbarkeit der gesehenen Ergebnisse ist unklar.

Ein Großteil der Studienzentren lag in asiatischen Ländern, sodass 67,5 % der eingeschlossenen Personen asiatischer Abstammung waren. Die Hälfte der Studienpopulation wurde dabei als „chinesisch“ klassifiziert. Es liegen Studien vor, die Unterschiede in den Charakteristika von asiatischen und nicht-asiatischen Patientinnen und Patienten mit PNH zeigen sowie Hinweise zu einer möglicherweise unterschiedlichen Pathogenese bezüglich Hämolyse und deren Begleitkomplikationen [17,20]. Der hohe Anteil einer asiatischen Population kann die Übertragbarkeit auf den deutschen Versorgungskontext einschränken.

Die European Medicines Agency (EMA) stufte den hohen Anteil der asiatischen Population als ein „Major Objective“ ein. Im Rahmen des Zulassungsprozesses wurde dargelegt, dass die asiatische Population ebenfalls an einer klassischen PNH erkrankt sei, unter der Einschränkung, dass der Anteil an Personen mit aplastischer Anämie in der asiatischen Population höher sei. Die Extrapolation von einer nicht europäischen auf eine europäische Population wurde daher als akzeptabel eingestuft.

4.2 Design und Methodik der Studie

Grundlage für die vorliegende Nutzenbewertung ist die pivotale, offene, unkontrollierte und multizentrische Phase-III-Studie APPOINT. Die Studie wurde zum Großteil in Asien sowie in Europa durchgeführt. Als primärer Endpunkt war „Hb-Wert-Erhöhung um ≥ 2 g/dl bei gleichzeitiger Transfusionsvermeidung“ definiert. Es wurden 52 Personen gescreent, insgesamt wurden 40 Personen in die Studie eingeschlossen. Der häufigste Ausschlussgrund war das Vorliegen einer labordiagnostischen Evidenz für Knochenmarkversagen. Die Studienpopulation entspricht der Zulassungspopulation und somit liegen für alle in die Studie eingeschlossenen Personen Auswertungen im FAS für die Wirksamkeitsendpunkte und im SAS für die Sicherheitsendpunkte vor.

Die Studie war in eine Screening-Periode (8 Wochen), eine Behandlungsphase (24 Wochen) und eine Extensionsphase (24 Wochen) unterteilt. Nach erfolgreicher Aufnahme in die Studie nach der Screening-Periode begann die Behandlungsphase. Die Behandlungsphase diente der Erhebung des primären Endpunkts zu Woche 24, weiteren sekundären (patientenberichteten) Endpunkten, der Erhebung von weiteren PRO-Instrumenten und der Erhebung der UE. In der Behandlungsphase erfolgten die Studienvisiten alle 2 Wochen. Danach gingen alle 40 Personen in die Extensionsphase über, in der die genannten Endpunkte weiterhin erhoben wurden. Hier erfolgten die Studienvisiten im Abstand von 4 Wochen. Es liegen somit Auswertungen zum Studienende (Woche 48) vor. Alle in die Studie eingeschlossenen Personen schlossen die Studie ab.

Die vom pU auf Basis der marginalen Responderanteile vorgelegten Schätzer werden aufgrund einer möglicherweise eingeschränkten Verallgemeinerbarkeit und Interpretierbarkeit (siehe Kapitel 2.4) nicht zur Nutzenbewertung herangezogen und stattdessen die einfachen relativen Häufigkeiten berichtet.

Der vorgelegte indirekte Vergleich wird nicht für die Nutzenbewertung herangezogen; Limitationen siehe Kapitel 2.6.

Das Verzerrungspotential der Studie APPOINT wird aufgrund des unkontrollierten Studiendesigns als hoch eingeschätzt.

4.3 Mortalität

Die Gesamtmortalität wurde im Rahmen der Sicherheitserhebung erhoben. Bis zum Ende der Extensionsphase sind keine Todesfälle aufgetreten.

Eine Interpretation und Bewertung der Gesamtmortalität ist aufgrund der fehlenden Kontrollgruppe nicht möglich. Der Effekt von Iptacopan auf die Mortalität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.4 Morbidität

Die Endpunkte „FACIT-Fatigue“, „PGIS“, „EQ-5D-VAS“, „MAVE“ sowie die Symptomskalen und Symptome des EORTC QLQ-C30“ wurden als patientenrelevant eingestuft und in der Nutzenbewertung dargestellt. Die Endpunkte „Transfusionsvermeidung“, „Hämoglobinwert-assoziierten Endpunkte“ (primärer Endpunkt) und „Durchbruchhämolyse“ werden ergänzend dargestellt. Nähere Ausführungen zur Bewertung finden sich in Kapitel 2.3. Es liegen im gesamten Studienverlauf hohe Rücklaufquoten der PRO-Instrumente vor.

Ein Hauptaugenmerk der Studie lag auf der Erhebung der Fatigue. Zum Ende der Behandlungsphase nach 24 Wochen lag bei 22 von 40 Personen (55,0 %) eine Verbesserung um ≥ 8 Punkte für den FACIT-Fatigue vor. Ähnliche Responderanteile zeigen sich nach dieser Zeit für den PGIS mit 24 von 40 Personen (60 %) sowie für die Symptomskala „Fatigue“ des EORTC QLQ-C30 mit 27 von

40 Personen (67,5 %) mit einer Verbesserung um ≥ 10 Punkte. Responderanalysen zu Woche 48 liegen nicht vor. Im EORTC QLQ-C30 wurden ebenfalls Verbesserungen um ≥ 10 Punkte in den Symptomen „Dyspnoe“ und „Appetitverlust“ bei jeweils 55 % der Studienpopulation beobachtet. In der Symptomskala „Schmerz“ lag bei 40 % der Studienteilnehmenden eine Verbesserung vor. Im Endpunkt „EQ-5D-VAS“ erreichten 50 % der Personen eine relevante Verbesserung von 15 %. In der Studie APPOINT trat im gesamten Studienverlauf kein MAVE auf.

Bei der Interpretation der Ergebnisse zur Morbidität muss beachtet werden, dass es sich nicht um die Erhebung der Verbesserung zu Woche 24 handelt, sondern um den Durchschnittswert der Erhebungen im Zeitintervall zwischen Woche 18 und Woche 24. Eine Rationale für dieses Vorgehen liegt nicht vor.

Das Modell zur Erhebung des primären Endpunkts konvergierte nicht. Die deskriptiven Auswertungen mit Hämoglobinwert-assoziierten Endpunkten sind im Anhang dargestellt.

Da es sich um eine offene, einarmige Studie handelt, sind die Ergebnisse, insbesondere der subjektiven, patientenberichteten Endpunkte, als potentiell hochverzerrt anzusehen, da auch „Placeboeffekte“ nicht ausgeschlossen werden können. Der Effekt von Iptacopan auf die Morbidität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.5 Lebensqualität

In der Studie APPOINT wurden die Funktionsskalen sowie die Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ des EORTC QLQ-C30 in der Kategorie „Lebensqualität“ erhoben. Es liegen im gesamten Studienverlauf hohe Rücklaufquoten vor. Bei der Interpretation der Ergebnisse zur Lebensqualität muss beachtet werden, dass es sich nicht um die Verbesserung zur Erhebung zu Woche 24 handelt, sondern um den Durchschnittswert der Erhebungen zwischen Woche 18 und Woche 24.

In der Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ zeigte sich bei 75 % der Personen eine Verbesserung um ≥ 10 Punkte zum Ende der Behandlungsphase zu Woche 24. Zudem zeigten sich auch in den Funktionsskalen des EORTC QLQ-C30 Verbesserungen.

Da es sich um eine offene, einarmige Studie handelt, sind die Ergebnisse, insbesondere der subjektiven, patientenberichteten Endpunkte, als potentiell hochverzerrt anzusehen, da auch „Placeboeffekte“ nicht ausgeschlossen werden können. Der Effekt von Iptacopan auf die Lebensqualität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.6 Sicherheit

In der Studie APPOINT wurden UE über den gesamten Studienzeitraum bis Woche 48 erfasst. Für die Nutzenbewertung werden die Ergebnisse zum Studienende (Woche 48) dargestellt. Die Erhebung begann mit der erstmaligen Gabe der Studienmedikation und endete 7 Tage nach der letzten Dosierung oder dem Zeitpunkt der letzten Visite, je nachdem was früher eintrat. In der Behandlungsphase fanden Visiten alle 2 Wochen statt und in der Extensionsphase alle 4 Wochen. Keine Person brach die Studie während der Einnahme der Studienmedikation ab, die mediane Behandlungsdauer beträgt 337 Tage (48 Wochen).

Bei fast allen teilnehmenden Personen (92,5 %) trat ein UE auf. Bei 4 Personen (10 %) wurde ein UE CTCAE-Grad ≥ 3 oder höher erfasst, bei 8 Personen (20 %) traten SUE auf. Es ist zu beachten, dass Ereignisse der Grunderkrankung in den Ergebnissen abgebildet sind. Auf Ebene der Systemorganklassen wurden „Infektionen“ bei 14 Personen (35 %) sowie „Erkrankungen des Gastrointestinaltrakts“ bei 13 Personen (32,5 %) am häufigsten berichtet. Der am häufigsten aufgetretene Preferred Term war „Kopfschmerzen“ bei 11 Personen (27,5 %).

„Infektionen“ waren die am häufigsten aufgetretenen schweren UE (Grad 3 oder höher) und SUE, die bei jeweils 5 Personen (12,5 %) auftraten. Bei den UE von besonderem Interesse traten am häufigsten „Schwere oder schwerwiegende Infektion“ bei 5 Personen (12,5 %) auf. Weitere beobachtete UE von besonderem Interesse betrafen „Hypersensitivität“ bei 4 Personen (10 %) sowie „PNH-Hämolyse und Thrombose“ bei 3 Personen (7,5 %). Die Sicherheit von Iptacopan kann aufgrund der fehlenden Kontrollgruppe nicht abschließend beurteilt werden.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Iptacopan ist zugelassen als Monotherapie zur Behandlung erwachsener Personen mit PNH, die eine hämolytische Anämie aufweisen. Die Nutzenbewertung von Iptacopan basiert auf der zulassungsbegründenden Studie APPOINT, einer einarmigen, offenen und multizentrischen Phase-III-Studie.

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 APPOINT

Studie APPOINT Darstellung der Ergebnisse	Iptacopan N = 40 ¹⁾
Mortalität	
Todesfälle	Kein Ereignis aufgetreten.
Morbidität	
FACIT-Fatigue (Verbesserung um ≥ 8 Punkte)²⁾	
Personen mit Ereignis, n (%)	22 (55,0)
Fatigue mittels PGIS (Verbesserung um ≥ 1 Kategorie)³⁾	
Personen mit Ereignis, n (%)	24 (60,0)
Symptomatik mittels EORTC QLQ-C30 (Verbesserung um ≥ 10 Punkte)⁴⁾	
<i>Symptomskala „Fatigue“</i> Personen mit Ereignis, n (%)	27 (67,5)
<i>Symptomskala „Schmerz“</i> Personen mit Ereignis, n (%)	16 (40,0)
<i>Symptom „Dyspnoe“</i> Personen mit Ereignis, n (%)	22 (55,0)
<i>Symptom „Schlaflosigkeit“</i> Personen mit Ereignis, n (%)	13 (32,5)
EQ-5D-VAS (Verbesserung um ≥ 15 Punkte)⁵⁾	
Personen mit Ereignis, n (%)	20 (50,0)
MAVE	
Personen mit Ereignis, n (%)	Kein Ereignis aufgetreten.
Lebensqualität	
Lebensqualität mittels EORTC QLQ-C30 (Verbesserung ≥ 10 Punkte)⁶⁾	
<i>Skala „Körperliche Funktion“</i> Personen mit Ereignis, n (%)	24 (60,0)
<i>Skala „Rollenfunktion“</i> Personen mit Ereignis, n (%)	27 (67,5)
<i>Skala „Emotionale Funktion“</i> Personen mit Ereignis, n (%)	22 (55,0)
<i>Skala „Kognitive Funktion“</i> Personen mit Ereignis, n (%)	13 (32,5)
<i>Skala „Soziale Funktion“</i> Personen mit Ereignis, n (%)	20 (50,0)

Studie APPOINT Darstellung der Ergebnisse	Iptacopan N = 40¹⁾
<i>Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“</i> Personen mit Ereignis, n (%)	30 (75,0)
Sicherheit⁷⁾⁸⁾	
Unerwünschte Ereignisse	
<i>UE CTCAE-Grad ≥ 3</i> Personen mit Ereignis, n (%)	4 (10,0)
<i>SUE</i> Personen mit Ereignis, n (%)	8 (20,0)
<i>UE, das zum Abbruch der Studienmedikation führte</i> Personen mit Ereignis, n (%)	0

¹⁾ Bezogen auf das FAS.

²⁾ Skala von 0 bis 52; höhere Werte zeigen geringere Beeinträchtigungen an.

³⁾ Als Verbesserung ist der Wechsel in ≥ 1 tiefere Kategorie definiert.

⁴⁾ Werte von 0 bis 100; höhere Werte entsprechen einer schwereren Krankheitssymptomatik.

⁵⁾ Werte von 0 bis 100; höhere Werte entsprechen einem besseren Gesundheitszustand.

⁶⁾ Werte von 0 bis 100; höhere Werte entsprechen einer besseren Funktion bzw. Gesundheit/Lebensqualität.

⁷⁾ Weitere Ergebnisse zur Sicherheit (UE nach Organsystemen und Einzelereignisse) sind in Kapitel 3.5 dargestellt.

⁸⁾ Entspricht dem SAS.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; EQ-5D-VAS: Visuelle Analogskala des European Quality of Life 5-Dimension; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; FAS: Full Analysis Set; MAVE: Schweres unerwünschtes vaskuläres Ereignis; PGIS: Patient Global Impression of Severity; SAS: Safety Analysis Set; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. **Bundesärztekammer (BÄK)**. Querschnitts-Leitlinien zur Therapie mit Blutkomponenten und Plasmaderivaten; Gesamtnovelle 2020 [online]. Berlin (GER): BÄK; 2020. [Zugriff: 17.09.2024]. URL: https://www.bundesaerztekammer.de/fileadmin/user_upload/old-files/downloads/pdf-Ordner/MuE/Querschnitts-Leitlinien_BAEK_zur_Therapie_mit_Blutkomponenten_und_Plasmaderivaten-Gesamtnovelle_2020.pdf
2. **Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO), Österreichische Gesellschaft für Hämatologie und Medizinische Onkologie (OEGHO), Schweizerische Gesellschaft für Medizinische Onkologie (SGMO), Schweizerische Gesellschaft für Hämatologie (SGH+SSH)**. Paroxysmale nächtliche Hämoglobinurie (PNH) [online]. Last update: 06.2023. Berlin (GER): DGHO; 2023. [Zugriff: 03.09.2024]. URL: <https://www.onkopedia.com/de/onkopedia/guidelines/paroxysmale-naechtlische-haemoglobinurie-pnh/@guideline/html/index.html>.
3. **European Medicines Agency (EMA)**. Fabhalta: European public assessment report EMEA/H/C/005764/0000 [online]. Last update: 21.03.2024. Amsterdam (NED): EMA; 2024. [Zugriff: 02.07.2024]. URL: https://www.ema.europa.eu/en/documents/assessment-report/fabhalta-public-assessment-report_en.pdf
4. **European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Group**. EORTC quality of life questionnaire - core questionnaire (EORTC QLQ-C30), version 3 [online]. In: PROQOLID. Lyon (FRA): Mapi Research Trust; 2001. [Zugriff: 24.07.2024]. URL: <https://eprovide.mapi-trust.org/instruments/eortc-quality-of-life-questionnaire-core-questionnaire>.
5. **Groth M, Singer S, Niedeggen C, Petermann-Meyer A, Röth A, Schrezenmeier H, et al**. Development of a disease-specific quality of life questionnaire for patients with aplastic anemia and/or paroxysmal nocturnal hemoglobinuria (QLQ-AA/PNH) - report on phases I and II. *Annals of Hematology* 2017;96:171-181.
6. **Lim E, Miyamura J, Chen JJ**. Racial/ethnic-specific reference intervals for common laboratory tests: a comparison among Asians, Blacks, Hispanics and White. *Hawaii J Med Public Health* 2015;74(9):302-310.
7. **Novartis**. Documentation of literature research: potential confounder in paroxysmal nocturnal hemoglobinuria (PNH) [unveröffentlicht]. 2023.
8. **Novartis**. Hematological response in patients with paroxysmal nocturnal hemoglobinuria treated with anti-C5 antibody: an external control arm study for iptacopan use in anti-C5 naïve patients (APPEX/CLNP023N12002R); final report of NIS-SUD with no Novartis drug in focus, version 1.0 [unveröffentlicht]. 2023.
9. **Novartis**. Hematological response in patients with paroxysmal nocturnal hemoglobinuria treated with anti-C5 antibody: an external control arm study for iptacopan use in anti-C5 naïve patients (APPEX/CLNP023N12002R); statistical analysis plan for NIS SUD Research Collaboration, SAP amendment 5, final 1.0 [unveröffentlicht]. 2023.
10. **Novartis**. A multicenter, single-arm, open-label trial to evaluate efficacy and safety of oral, twice daily iptacopan in adult PNH patients who are naïve to complement inhibitor therapy (CLNP023C12301); amended clinical trial protocol, version 04 (clean) [unveröffentlicht]. 2022.

11. **Novartis.** A multicenter, single-arm, open-label trial to evaluate efficacy and safety of oral, twice daily iptacopan in adult PNH patients who are naive to complement inhibitor therapy (CLNP023C12301); clinical study report, final 48-week analysis, version 1.0 [unveröffentlicht]. 2023.
12. **Novartis.** A multicenter, single-arm, open-label trial to evaluate efficacy and safety of oral, twice daily iptacopan in adult PNH patients who are naive to complement inhibitor therapy (CLNP023C12301); clinical study report, primary endpoint analysis, version 1.0. [unveröffentlicht]. 2023.
13. **Novartis Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 B – Anhang 4-G: Iptacopan (Fabhalta), nicht vorbehandelte erwachsene Patienten mit PNH, die eine hämolytische Anämie aufweisen; ergänzende Analysen zu der Studie APPOINT-PNH (CLNP023C12301) [unveröffentlicht]. 2024.
14. **Novartis Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 B: Iptacopan (Fabhalta), Nicht vorbehandelte erwachsene Patienten mit paroxysmaler nächtlicher Hämoglobinurie (PNH), die eine hämolytische Anämie aufweisen; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 2024.
15. **Novartis Pharma.** Fabhalta 200 mg Hartkapseln [online]. 05.2024. [Zugriff: 17.09.2024]. URL: <https://www.fachinfo.de/>.
16. **Raad H, Cornelius V, Chan S, Williamson E, Cro S.** An evaluation of inverse probability weighting using the propensity score for baseline covariate adjustment in smaller population randomised controlled trials with a continuous outcome. BMC Med Res Methodol 2020;20(1):70.
17. **Sakurai M, Jang JH, Chou WC, Kim JS, Wilson A, Nishimura JI, et al.** Comparative study on baseline clinical characteristics of Asian versus non-Asian patients with paroxysmal nocturnal hemoglobinuria. Int J Hematol 2019;110(4):411-418.
18. **Weitz I, Meyers G, Lamy T, Cahn JY, Uranga M, García Vela J, et al.** Cross-sectional validation study of patient-reported outcomes in patients with paroxysmal nocturnal haemoglobinuria. Internal Medicine Journal 2012;43(3):298-307.
19. **Worldwide Real World Evidence, Novartis Oncology, Hospital Saint-Lois, St. James's University Hospital.** Hematological response in patients with paroxysmal nocturnal hemoglobinuria treated with anti-C5 antibody: an external control arm study for iptacopan use in anti-C5 naïve patients (APPEX/CLNP023N12002R); research collaboration protocol, version 2.0. [unveröffentlicht]. 2023.
20. **Yu F, Du Y, Han B.** A comparative analysis of clinical characteristics of patients with paroxysmal nocturnal hemoglobinuria between Asia and Europe/America. Int J Hematol 2016;103(6):649-654.

Anhang

Hämoglobinwert-assoziierte Endpunkte

Tabelle 24: Hämoglobinwert-assoziierte Endpunkte; Studie APPOINT, FAS
(Datenschnitte: 02.11.2022 (Woche 24) und 18.04.2023 (Woche 48))

Studie APPOINT Hämoglobinwert-assoziierte Endpunkte	Iptacopan N = 40 n (%)
Zum Ende der Behandlungsphase (Woche 24)	
Hb-Erhöhung um ≥ 2 g/dl bei gleichzeitiger Transfusionsvermeidung ¹⁾	31 (77,5)
Hb-Erhöhung auf ≥ 12 g/dl bei gleichzeitiger Transfusionsvermeidung ¹⁾	19 (47,5)
Zum Ende der Extensionsphase (Woche 48)	
Hb-Erhöhung um ≥ 2 g/dl	38 (95,0)
Hb-Erhöhung auf ≥ 12 g/dl	31 (77,5)

¹⁾ Primärer Endpunkt.

Abkürzungen: FAS: Full Analysis Set; Hb: Hämoglobin.

FACIT-Fatigue

Tabelle 25: Veränderung des FACIT-Fatigue von Baseline bis Ende Extensionsphase (Woche 48);
Studie APPOINT, FAS (Datenschnitt: 18.04.2023)

Studie APPOINT FACIT-Fatigue	Iptacopan N = 40	
	Deskriptive Statistiken	Veränderung ggü. Baseline
<i>Baseline</i> n (%) MW (SD) Median (min; max)	40 (100) 32,8 (10,2) 34,3 (13; 51)	
<i>Zum Ende der Behandlungsphase</i> n (%) MW (SD) Median (min; max)	37 43,9 (6,2) 46 (29; 52)	37 12,1 (9,9) 11,5 (-9; 31)
<i>Zum Ende der Extensionsphase</i> n (%) MW (SD) Median (min; max)	39 42,9 (7,2) 45 (21; 52)	39 10,4 (10,1) 12,5 (-12; 32)

Abkürzungen: FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; FAS: Full Analysis Set; MW: Mittelwert; SD: Standardabweichung.

EORTC QLQ-C30 – Symptomskalen und Symptome
Tabelle 26: Deskriptive Auswertungen der Symptomatik mittels EORTC QLQ-C30 von Baseline bis Ende Extensionsphase (Woche 48); Studie APPOINT, FAS (Datenschnitt: 18.04.2023)

Studie APPOINT Symptomatik mittels EORTC QLQ-C30 ¹⁾	Iptacopan N = 40		
	MW (SD)	Median (min; max)	MW-Veränderung ggü. Baseline (SD)
Symptomskala „Fatigue“			
Baseline	45,0 (24,4)	41,7 (0; 100)	-
Zum Ende der Behandlungsphase	21,9 (14,9)	22,2 (0; 56)	-25,1 (24,4)
Zum Ende der Extensionsphase	22,5 (22,2)	22,2 (0; 89)	-22,2 (25,0)
Symptomskala „Schmerz“			
Baseline	20,6 (18,5)	25,0 (0; 58)	-
Zum Ende der Behandlungsphase	10,8 (13,7)	0 (0; 50)	-10,4 (18,8)
Zum Ende der Extensionsphase	13,2 (18,8)	0 (0; 67)	-7,9 (20,5)
Symptom „Dyspnoe“			
Baseline	29,6 (24,0)	33,3 (0; 100)	-
Zum Ende der Behandlungsphase	11,7 (17,9)	0 (0; 67)	-18,9 (28,6)
Zum Ende der Extensionsphase	12,0 (17,9)	0 (0; 67)	-17,5 (28,6)
Symptom „Schlaflosigkeit“			
Baseline	24,2 (25,9)	16,7 (0; 100)	-
Zum Ende der Behandlungsphase	14,4 (20,0)	0 (0; 67)	-10,8 (24,6)
Zum Ende der Extensionsphase	20,5 (27,2)	0 (0; 100)	-4,3 (31,9)
Symptomskala „Übelkeit und Erbrechen“ (ergänzend dargestellt)			
Baseline	9,0 (12,4)	0 (0; 42)	-
Zum Ende der Behandlungsphase	1,8 (6,5)	0 (0; 33)	-7,7 (12,2)
Zum Ende der Extensionsphase	3,8 (8,9)	0 (0; 33)	-5,3 (12,8)
Symptom „Appetitverlust“ (ergänzend dargestellt)			
Baseline	22,1 (23,1)	16,7 (0; 100)	-
Zum Ende der Behandlungsphase	5,4 (12,5)	0 (0; 33)	-17,6 (26,0)
Zum Ende der Extensionsphase	10,3 (15,6)	0 (0; 33)	-12,4 (25,3)
Symptom „Obstipation“ (ergänzend dargestellt)			
Baseline	9,2 (13,6)	0 (0; 50)	-
Zum Ende der Behandlungsphase	12,6 (24,0)	0 (0; 100)	2,7 (20,2)
Zum Ende der Extensionsphase	11,1 (20,7)	0 (0; 67)	1,7 (13,7)
Symptom „Diarrhö“ (ergänzend dargestellt)			
Baseline	6,7 (15,9)	0 (0; 83)	-
Zum Ende der Behandlungsphase	6,3 (15,4)	0 (0; 67)	-0,9 (21,9)
Zum Ende der Extensionsphase	8,5 (19,8)	0 (0; 100)	1,7 (16,6)

¹⁾ Werte von 0 bis 100; höhere Werte entsprechen einer schwereren Krankheitssymptomatik.

Abkürzungen: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; FAS: Full Analysis Set; MW: Mittelwert; SD: Standardabweichung.

EORTC QLQ-C30 – Funktionsskalen und Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“
Tabelle 27: Deskriptive Auswertungen der Lebensqualität mittels EORTC QLQ-C30 von Baseline bis Ende Extensionsphase (Woche 48); Studie APPOINT, FAS (Datenschnitt: 18.04.2023)

Studie APPOINT Lebensqualität mittels EORTC QLQ-C30 ¹⁾	Iptacopan N = 40		
	MW (SD)	Median (min; max)	MW-Veränderung ggü. Baseline (SD)
Skala „Körperliche Funktion“			
Baseline	72,7 (16,7)	73,3 (43; 100)	-
Zum Ende der Behandlungsphase	88,3 (11,9)	93,3 (60; 100)	17,2 (16,5)
Zum Ende der Extensionsphase	88,4 (12,0)	93,3 (40; 100)	16,2 (17,5)
Skala „Rollenfunktion“			
Baseline	64,8 (24,4)	66,7 (8; 100)	-
Zum Ende der Behandlungsphase	86,5 (15,1)	83,3 (50; 100)	23,9 (23,3)
Zum Ende der Extensionsphase	81,2 (20,3)	83,3 (0; 100)	16,9 (22,9)
Skala „Emotionale Funktion“			
Baseline	69,4 (18,7)	66,7 (0; 100)	-
Zum Ende der Behandlungsphase	80,4 (22,3)	83,3 (0; 100)	13 (23,7)
Zum Ende der Extensionsphase	78,0 (23,8)	83,3 (0; 100)	8,3 (23,4)
Skala „Kognitive Funktion“			
Baseline	72,7 (20,4)	75,0 (33; 100)	-
Zum Ende der Behandlungsphase	78,8 (21,8)	83,3 (33; 100)	7 (18,2)
Zum Ende der Extensionsphase	74,4 (26,7)	83,3 (0; 100)	2,4 (22,8)
Skala „Soziale Funktion“			
Baseline	62,1 (24,0)	66,7 (0; 100)	-
Zum Ende der Behandlungsphase	78,4 (24,8)	83,3 (0;100)	18,7 (27,7)
Zum Ende der Extensionsphase	75,2 (26,5)	83,3 (0;100)	13,2 (28,1)
Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“			
Baseline	52,9 (18,7)	52,1 (21; 88)	-
Zum Ende der Behandlungsphase	76,1 (14,9)	83,3 (25; 100)	25,1 (18,7)
Zum Ende der Extensionsphase	74,6 (16,9)	83,3 (25; 100)	22 (21,3)

¹⁾ Werte von 0 bis 100; höhere Werte entsprechen einer besseren Funktion bzw. Gesundheit/Lebensqualität.

Abkürzungen: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30; FAS: Full Analysis Set; MW: Mittelwert; SD: Standardabweichung.