

Belumosudil (chronische Graft-versus-Host-Krankheit)

1. Addendum zum Projekt A25-34 (AbD-Konzept)

A decorative horizontal bar spanning the width of the page, composed of 18 squares of varying shades of blue and grey. The text 'ADDENDUM (ABD-KONZEPT)' is centered in white on a dark blue background.

ADDENDUM (ABD-KONZEPT)

Projekt: A25-110

Version: 1.0

Stand: 26.09.2025

IQWiG-Berichte – Nr. 2098

DOI: 10.60584/A25-110

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Belumosudil – 1. Addendum zum Projekt A25-34

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

26.08.2025

Interne Projektnummer

A25-110

DOI-URL

<https://doi.org/10.60584/A25-110>

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Siegburger Str. 237
50679 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Belumosudil (chronische Graft-versus-Host-Krankheit); 1. Addendum zum Projekt A25-34 (AbD-Konzept) [online]. 2025 [Zugriff: TT.MM.JJJJ]. URL: <https://doi.org/10.60584/A25-110>.

Schlagwörter

Belumosudil, Graft-versus-host-Krankheit, Adoleszent, Erwachsene, Register, Nutzenbewertung, Konzept

Keywords

Belumosudil, Graft vs Host Disease, Adolescent, Adult, Registries, Benefit Assessment, Concept

An dem Addendum beteiligte Mitarbeiterinnen und Mitarbeiter des IQWiG

- Jana Göbel
- Anna-Katharina Barnert
- Lars Beckmann
- Felix Schwarz
- Volker Vervölgyi

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abbildungsverzeichnis	vi
Abkürzungsverzeichnis.....	vii
1 Hintergrund.....	1
2 Bewertung.....	2
2.1 Beschreibung der Studie ROCKreal.....	2
2.2 Bewertung der Studie ROCKreal.....	4
2.3 Einordnung der Auswirkungen der ROCKreal-Studiendaten auf die Durchführbarkeit einer orientierenden Fallzahlschätzung und auf die orientierende Betrachtung von Fallzahlszenarien	6
2.4 Einordnung der im Beteiligungsverfahren vorgebrachten Ausführungen zur Patientenzahl im Anwendungsgebiet auf die orientierende Betrachtung von Fallzahlszenarien für die AbD.....	8
2.4.1 Beschreibung des Vorgehens zur Ermittlung der Patientenzahlen.....	8
2.4.2 Bewertung	10
2.4.3 Auswirkungen auf die orientierende Betrachtung von Fallzahlszenarien	11
2.5 Zusammenfassung und Fazit	14
3 Literatur	15
Anhang A Aufdeckbarer Effekt für verschiedene Stichprobengrößen und verschiedene Ereignisanteile unter Intervention und Vergleichstherapie bei unterschiedlichen Verteilungsverhältnissen	17

Tabellenverzeichnis

	Seite
Tabelle 1: Orientierende Schätzung der Patientenzahlen auf Grundlage der einbezogenen epidemiologischen Kriterien	11
Tabelle 2: Aufdeckbarer Effekt (Effektmaß: RR) bei verschiedenen Stichprobengrößen und Ereignisanteilen in beiden Behandlungsgruppen.....	19

Abbildungsverzeichnis

	Seite
Abbildung 1: Aufdeckbarer Effekt (Effektmaß: relatives Risiko) in Abhängigkeit von der Stichprobengröße und den Ereignisanteilen in beiden Behandlungsgruppen (Verhältnis Intervention zu Vergleich 5:1).....	13
Abbildung 2: Aufdeckbarer Effekt (Effektmaß: relatives Risiko) in Abhängigkeit von der Stichprobengröße und den Ereignisanteilen in beiden Behandlungsgruppen (Verhältnis Intervention zu Vergleich 1:1).....	17
Abbildung 3: Aufdeckbarer Effekt (Effektmaß: relatives Risiko) in Abhängigkeit von der Stichprobengröße und den Ereignisanteilen in beiden Behandlungsgruppen (Verhältnis Intervention zu Vergleich 3:1).....	18

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AbD	Anwendungsbegleitende Datenerhebung
AESI	unerwünschte Ereignisse von besonderem Interesse
BAT	Best available Therapy (bestverfügbare Therapie)
cGvHD	chronische Graft-versus-Host-Krankheit
DAG-HSZT	Deutsche Arbeitsgemeinschaft für Hämatopoetische Stammzelltransplantation und Zelluläre Therapie
DGHO	Deutsche Gesellschaft für Hämatologie und medizinische Onkologie
DRST	Deutsches Register für hämatopoetische Stammzelltransplantation und Zelltherapie
EU	Europäische Union
FDA	Food and Drug Administration
FFS	Failure-free-survival
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
NIH	National Institutes of Health
pU	pharmazeutischer Unternehmer
RR	relatives Risiko
SAP	statistischer Analyseplan
SGB	Sozialgesetzbuch

1 Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) am 26.08.2025 mit ergänzenden Bewertungen zum Konzept für eine anwendungsbegleitende Datenerhebung (AbD) gemäß § 35a Absatz 3b SGB V für den Wirkstoff Belumosudil (AbD-Konzept A25-34 vom 18.06.2025 [1]) beauftragt.

Der Auftrag umfasst die folgende Bewertung und Einbindung der im Beteiligungsverfahren vorgebrachten Aspekte:

- Prüfung der nachgereichten Unterlagen zur Studie ROCKreal in Bezug darauf, ob durch die Studiendaten die Evidenzlücke, auf die sich die Fragestellung der AbD bezieht, geschlossen werden kann.
- Bei weiterhin bestehender Evidenzlücke in Bezug auf die Fragestellung der AbD:
 - Prüfung der Durchführbarkeit einer orientierenden Fallzahlschätzung unter Einbeziehung der Studiendaten zur Studie ROCKreal
 - Bei Nichtdurchführbarkeit einer orientierenden Fallzahlschätzung:
 - Einordnung der Auswirkungen der Studiendaten auf die orientierende Betrachtung von Fallzahlszenarien für die AbD
 - Einordnung der im Beteiligungsverfahren vorgebrachten Ausführungen zur Patientenzahl im Anwendungsgebiet auf die orientierende Betrachtung von Fallzahlszenarien für die AbD

2 Bewertung

Gemäß Auftrag des G-BA sollen die vom pU nachgereichten Unterlagen zur Studie ROCKreal in Bezug darauf geprüft werden, ob durch die Studiendaten die vorliegende Evidenzlücke geschlossen werden kann. Im Nachgang zum Beteiligungsverfahren legt der pU neben einer weiteren schriftlichen Stellungnahme [2] die Publikation Hall 2025 inklusive Supplement [3] sowie den statistischen Analyseplan (SAP) [4] zur Studie ROCKreal vor. Zum Zeitpunkt der Konzepterstellung wurde eine Informationsbeschaffung nach laufenden, geplanten als auch abgeschlossenen Datenerhebungen zu Belumosudil durchgeführt. Da zu diesem Zeitpunkt keine veröffentlichten Daten zur Studie ROCKreal vorlagen, wurde die Studie im Rahmen der Informationsbeschaffung nicht identifiziert.

Im Folgenden erfolgt auftragsgemäß die Bewertung der vorgelegten Evidenz zur Studie ROCKreal sowie die Prüfung der Durchführbarkeit einer orientierenden Fallzahlschätzung. Abschließend werden die im Beteiligungsverfahren vorgebrachten Ausführungen zur Patientenzahl eingeordnet und die Auswirkung auf die orientierende Betrachtung von Fallzahlszenarien beschrieben.

2.1 Beschreibung der Studie ROCKreal

Bei der Studie ROCKreal handelt es sich um eine retrospektive Studie anhand von versorgungsnahen Daten aus Patientenakten zum Vergleich von Belumosudil mit einer bestverfügbaren Therapie (BAT). Der pU beschreibt, die Studie gemäß einer Target Trial Emulation zur Nachbildung einer randomisiert kontrollierten Studie geplant zu haben. Eingeschlossen wurden Patientinnen und Patienten im Alter von ≥ 12 Jahren mit chronischer Graft-versus-Host-Krankheit (cGvHD) nach Erhalt von 2 bis 5 vorherigen Therapielinien. Patientinnen und Patienten durften zuvor keine Therapie mit Belumosudil erhalten haben. Ausgeschlossen wurden Patientinnen und Patienten mit einem dokumentierten Wiederauftreten der malignen Grunderkrankung. Gemäß Publikation wurden auch Patientinnen und Patienten mit einer 1-jährigen Historie einer cGvHD-Behandlung im Rahmen einer klinischen Studie ausgeschlossen. Diese Angabe findet sich nicht im SAP.

Der pU wählt, abweichend von der Emulation einer randomisierten kontrollierten Studie, als Beobachtungseinheit einzelne Therapielinien statt Patientinnen und Patienten. Dabei wurde der Beginn einer neuen Therapielinie definiert als (1) Behandlung mit einem neuen systemischen Wirkstoff; (2) Wiederaufnahme eines zuvor verwendeten Wirkstoffs, nachdem in der Zwischenzeit andere Wirkstoffe eingesetzt wurden; (3) Beginn einer Kortikosteroid-Therapie (wenn es sich um die erstmalige Anwendung bei cGvHD handelte) oder wenn die Kortikosteroid-Dosis auf ≥ 1 mg/kg/Tag erhöht wurde. Therapielinien mit einer Dauer von < 2 Wochen wurden aus der Analyse ausgeschlossen. Therapielinien nach einer erfolgten Behandlung mit Belumosudil wurden ebenfalls ausgeschlossen, da das Einschlusskriterium

„keine vorherige Exposition mit Belumosudil“ nicht erfüllt wurde. Eine Patientin bzw. ein Patient konnte zu mehreren Zeitpunkten die Voraussetzungen für einen Studieneinschluss erfüllen (z. B. Beginn der 3., 4., 5. oder 6. Therapielinie) und mehrfach (pro Therapielinie) in die Auswertung eingehen. Es war somit möglich, dass eine Patientin bzw. ein Patient sowohl im Interventionsarm als auch (ggf. mehrfach) im Vergleichsarm ausgewertet wurde. Durch die Wahl der Therapielinien als Beobachtungseinheit liegen für den Großteil der Daten keine Angaben zu den jeweiligen Anteilen der Patientinnen und Patienten mit bestimmten Merkmalen oder Ereignissen vor. Während für die Interventionsseite ein Rückschluss von der Therapielinie auf den Patienten möglich ist, da Patientinnen und Patienten nur bei der 1. Exposition mit Belumosudil die Einschlusskriterien erfüllen, ist dies für die Vergleichsseite nicht möglich. Zudem kann eine Patientin bzw. ein Patient mehrfach in die Auswertung eingehen, wodurch diese eingehenden Beobachtungen nicht mehr unabhängig voneinander sind.

Im Interventionsarm der Studie ROCKreal wurde Belumosudil in 9,7 % der Therapielinien als Monotherapie und in 21,2 % in Kombination mit Kortikosteroiden verabreicht. Insgesamt wurden 23 weitere Wirkstoffkombinationen mit Belumosudil dokumentiert.

Die bestverfügbare Therapie (BAT) umfasste folgende Wirkstoffe bzw. nicht medikamentöse Verfahren: Abatacept, Alemtuzumab, Ciclosporin, extrakorporale Photopherese, Etanercept, Hydroxychloroquin, Ibrutinib, Imatinib, Interleukin-2, Methotrexat, Mycophenolatmofetil, Pentostatin, Rituximab, Ruxolitinib, Sirolimus, Tacrolimus und nodale Bestrahlung. In den Therapielinien des Vergleichsarms wurden insgesamt 89 verschiedene Wirkstoffkombinationen dokumentiert. Ruxolitinib wurde in 49 % der BAT Therapielinien eingesetzt. Die Vergleichstherapie in der Studie ROCKreal umfasst 5 Wirkstoffe (Abatacept, Alemtuzumab, Etanercept, Interleukin 2, nodale Bestrahlung), welche nicht in der patientenindividuellen Therapie der zweckmäßigen Vergleichstherapie des G-BA gelistet sind. Da abgesehen von Abatacept, welcher in 2 Therapielinien verabreicht wurde, keiner der genannten Wirkstoffe verwendet wurde, hat dies jedoch keine Konsequenz.

Gemäß SAP war kein fester Beobachtungszeitraum präspezifiziert. Entsprechend der Publikation sollten Patientinnen und Patienten nach Beginn einer Therapielinie insgesamt für 1 Jahr beobachtet werden. Im Median lag die Dauer einer Therapielinie bei 11,3 Monaten für Belumosudil und bei 8,5 Monaten für BAT. Da der pU eine Therapielinie als Beobachtungseinheit definiert, hing die tatsächliche Beobachtungsdauer von der Dauer der jeweiligen Therapielinie ab.

Entsprechend der Angaben aus der Publikation von Hall 2025 wurden insgesamt 196 Patientinnen und Patienten aus 8 Transplantationszentren in den USA eingeschlossen und gingen mit 358 Therapielinien (113 Belumosudil- und 245 BAT-Therapielinien) in die Auswertungen ein. Von den 196 Patientinnen und Patienten erhielten 83 in keiner

Therapielinie Belumosudil. Von den 113 Patientinnen und Patienten, die als Belumosudil-Therapielinie analysiert wurden, gingen 60 Patientinnen und Patienten ebenfalls mit ≥ 1 Therapielinie auf der Vergleichsseite in die Auswertung ein.

Als Endpunkte wurden das Gesamtansprechen (ORR) nach 6 Monaten, Failure Free Survival (FFS) nach 1 Jahr und einige unerwünschte Ereignisse von besonderem Interesse (AESI) erhoben.

Als Datenquelle dienten Versorgungsdaten aus Transplantationszentren in dem Zeitraum März 2015 bis März 2024. Zur retrospektiven Erhebung sowie zur Sicherstellung der Richtigkeit und Vollständigkeit der Daten wurde ein elektronisches Datenerfassungssystem, welches mithilfe eines Auftragsforschungsinstitut für diese Studie implementiert wurde, verwendet.

Da es sich bei der Studie ROCKreal um einen nicht randomisierten Vergleich handelt, ist die systematische Identifikation und Adjustierung von Confoundern essenziell. Der pU beschreibt potenzielle Confounder mithilfe von klinischen Expertinnen und Experten identifiziert und in einer Liste zusammengefasst zu haben. Neben Schlüsselfaktoren, welche die Behandlungszuweisung beeinflussen, wurden Informationen über Kovariaten gesammelt, die das Outcome vorhersagen. Nachfolgend wurde eine statistikbasierte Bewertung vorgenommen, in der nicht informative Kovariaten (z. B. binäre Variablen mit geringer Variation) ausgeschlossen oder hochprädiktive Kovariaten für die Behandlungszuweisung über eine Lasso-Regression identifiziert wurden. Die Auswertung der erhobenen Daten erfolgte mittels Targeted Maximum Likelihood Estimation (TMLE).

2.2 Bewertung der Studie ROCKreal

Die vom pU vorgelegte Studie ROCKreal ist nicht geeignet, die bestehende Evidenzlücke der vorliegenden Fragestellung zu schließen. Dies wird im Folgenden begründet.

Beobachtungseinheit Therapielinien

Wie in Abschnitt 2.1 beschrieben, wählt der pU als Beobachtungseinheit die einzelnen Therapielinien. Da die Therapielinien in den vorliegenden Auswertungen isoliert und ohne Bezug zu den Patientinnen und Patienten betrachtet werden, ist die Beantwortung der Fragestellung der AbD nicht möglich. Um den patientenrelevanten Nutzen bzw. Schaden einer Intervention gegenüber der zweckmäßigen Vergleichstherapie bestimmen zu können, braucht es Angaben dazu, bei wie vielen Patientinnen und Patienten relevante Ereignisse auftreten.

Präspezifizierung der Analysen unklar

Wie oben beschrieben wurden in die Studienpopulation der Studie ROCKreal Patientinnen und Patienten eingeschlossen, welche zwischen März 2015 und März 2024 in 8 Transplantationszentren in den USA mit Belumosudil oder der bestverfügbaren Therapie behandelt wurden. Laut dem SAP des pU wurden die meisten Daten im 1. Quartal 2024 in das

elektronische Datenerfassungssystem eingegeben und Ende März 2024 zur Analyse extrahiert. Der durch den pU zur Verfügung gestellte SAP wurde im April 2024 erstellt. Im SAP ist beschrieben, dass interessierte Transplantationszentren und diejenigen, die die Teilnahmebedingungen erfüllten, nach Unterzeichnung einer Vertraulichkeitsvereinbarung eine Kopie des Protokolls erhielten. Ein Studienprotokoll zur Studie ROCKreal wurde nicht vorgelegt.

Es ist somit unklar, ob die durchgeführten Analysen präspezifiziert waren und anhand welcher Studienunterlagen die Datenextraktion erfolgte. Im vorliegenden finalen SAP liegt keine Änderungsübersicht zu ggf. vorangegangenen SAP-Versionen vor. Zur Operationalisierung des primären Endpunkts zum Gesamtansprechen wird Folgendes ausgeführt: „The final algorithm reflects an attempt at operationalizing an objective and clinically relevant primary endpoint from the data.“. Die Operationalisierung von Endpunkten anhand bereits extrahierter Daten ist nicht sachgerecht. Es ist zu beachten, dass auch bei retrospektiven Datenerhebungen das Studienprotokoll bzw. der SAP vor Zugriff auf die Daten finalisiert werden.

Zum Zeitpunkt der Konzepterstellung konnte die Studie ROCKreal nicht identifiziert werden (siehe Konzept A25-34). Die Studie wurde somit nicht vorab registriert.

Identifikation der Confounder potenziell unvollständig und keine transparente Dokumentation

In nicht randomisierten Studien ist die notwendige Strukturgleichheit zwischen den Behandlungsgruppen nicht gewährleistet, daher müssen Gruppenunterschiede bei möglichen Confoundern (Störgrößen), d. h. Faktoren, die sowohl mit der Behandlung als auch mit Endpunkten in Beziehung stehen und somit die Schätzung des Behandlungseffekts verfälschen können, in der Schätzung berücksichtigt werden. Voraussetzung dafür ist zunächst, dass relevante Confounder systematisch identifiziert werden [5,6], wie beispielsweise in Pufulete 2022 [6] und im IQWiG-Arbeitspapier GA23-02 [7] beschrieben. Anschließend muss sichergestellt werden, dass der verwendete Datensatz die notwendigen Angaben zu den identifizierten Confoundern enthält. Darauf basierend ist dann mittels geeigneter Adjustierungsmethoden (z. B. Propensity Score-Gewichtung) ein möglicher verzerrender Effekt von Confoundern adäquat zu berücksichtigen [5].

Das Vorgehen des pU zur Identifikation von potenziellen Confoundern ist nicht sachgerecht. Die Identifikation einer zunächst erstellten Liste potenziell relevanter Confounder basiert allein auf der Einbindung klinischer Expertinnen und Experten. Eine systematische Literaturrecherche zur Confounderidentifikation wurde nicht durchgeführt, sodass durch das Vorgehen des pU eine vollständige Identifikation aller potenziell relevanten Confounder nicht sichergestellt ist. Zudem legt der pU auch keine Angabe dazu vor, wie viele Klinikerinnen und Kliniker eingebunden wurden. In seiner Stellungnahme stützt der pU sein Vorgehen zur

anschließenden Confounderselektion auf die Auswertungsmethode der Targeted Maximum Likelihood Estimation (TMLE). Diese ist charakterisiert durch die Kombination eines Propensity-Score-Modell für die Modellierung der Wahrscheinlichkeit der Behandlungszuweisung und einem Outcome-Modell für die Modellierung des Zusammenhangs zwischen Outcome und Kovariablen, welches auch die Effektschätzung beinhaltet. Die Anwendung des TMLE führt zu unverzerrten Effektschätzungen, wenn eines der beiden Modelle korrekt spezifiziert ist. Die systematische Confounderidentifikation ist somit auch für die vorliegenden Adjustierungsmethoden eine notwendige Voraussetzung. Da eine adäquate Confounderidentifikation für mindestens eines der beiden Modelle die Basis für die Effektschätzung bildet und die Confounderidentifikation nicht systematisch erfolgte, wird im Folgenden nicht näher auf die vom pU verwendete Methodik zur Adjustierung der Confounder eingegangen.

Darüber hinaus sind die durchgeführten Arbeitsschritte zur Confounderidentifikation und die nachfolgende Confounderselektion im TMLE-Framework unzureichend dokumentiert. Es ist beispielsweise nicht nachvollziehbar, welche Confounder ins Outcome- und welche ins Propensity-Score Modell eingehen. Zusätzlich fehlen Angaben dazu, welche Confounder zunächst mittels klinischer Expertinnen und Experten identifiziert wurden und welche Confounder nachfolgend durch die statistikbasierte Bewertung als nicht informative Variablen ausgeschlossen wurden.

Fazit

Auf Basis der nachgereichten Unterlagen des pU zur Studie ROCKreal lässt sich die Evidenzlücke zu Belumosudil nicht schließen. Dies begründet sich insbesondere in der Unsicherheit bezüglich der Präspezifizierung der Analysen sowie der nicht systematisch durchgeführten Confounderidentifikation. Darüber hinaus liegen in den Analysen nur Angaben bezogen auf die jeweilige Therapielinie, nicht jedoch zum Anteil von Patientinnen und Patienten mit bestimmten Merkmalen, bzw. Ereignissen vor.

2.3 Einordnung der Auswirkungen der ROCKreal-Studiendaten auf die Durchführbarkeit einer orientierenden Fallzahlschätzung und auf die orientierende Betrachtung von Fallzahlsszenarien

Operationalisierung des primären Endpunkts nicht sachgerecht

Im AbD-Konzept zu Belumosudil [1] wird die Operationalisierung des Gesamtansprechens gemäß den National-Institutes-of-Health(NIH)-Konsensus-Kriterien [8] empfohlen. Diese Kriterien werden zur Beurteilung der Beeinträchtigung der einzelnen Organe und zugleich auch zur Beurteilung des Gesamtansprechens verwendet und basieren auf patientenrelevanten Symptomen. Lediglich die Beurteilung der Leber erfolgt auf Basis von Laborwerten. Wie zuvor im AbD-Konzept beschrieben, sollte ein weiteres Therapieziel neben dem Gesamtansprechen sein, die Dosis der Kortikosteroide zu reduzieren, im Idealfall auf eine Dosis unterhalb der Cushing-Schwelle. Hierbei ist es sinnvoll, eine relevante Verbesserung der

Symptomatik mit dem gleichzeitigen Erreichen einer Kortikosteroid-Dosis unterhalb einer relevanten Schwelle (z. B. der Cushing-Schwelle) zu koppeln.

In der Studie ROCKreal erhebt der pU als primären Endpunkt das Gesamtansprechen nach 6 Monaten und operationalisiert diesen mithilfe von 4 Elementen, die in folgender Hierarchie (je nach Verfügbarkeit der Daten) zwischen Therapieansprechen und -versagen differenzieren und berücksichtigt wurden:

- Tod oder Wiederauftreten der Grunderkrankung (Therapieversagen)
- Unterbrechung der Therapielinie vor Woche 26 (Therapieversagen), durch
 - Beginn einer neuen systematischen Therapie,
 - Absetzen systemischer Wirkstoffe aufgrund mangelnder Wirksamkeit oder unerwünschter Ereignisse,
 - Erhöhung der Kortikosteroid-Dosis auf ≥ 1 mg/kg/Tag für ≥ 2 Wochen
- Ansprechen pro Therapielinie basierend auf der Veränderung im NIH-Organscore und bei Bedarf ergänzt durch NIH-Response-Assessment und / oder Arztbewertung (Therapieansprechen oder Therapieversagen)
- Reduktion der Kortikosteroid-Dosis (keine der zuvor genannten Bedingungen ist erfüllt und es liegen Informationen zur verabreichten Steroid-Dosis vor):
 - Kortikosteroid-Reduktion um ≥ 50 % (Therapieansprechen),
 - keine Dosisreduktion der Kortikosteroide (d. h. Erhöhung der Dosis, keine Veränderung und Reduktion um < 50 %) (Therapieversagen).

Falls mehrere Einträge in den Versorgungsdaten vorhanden waren, wurde der NIH-Organscore dem NIH-Response-Assessment und beides der Arztbewertung vorgezogen. Gemäß der Publikation Hall 2025 [3] basieren 49 % des Therapieansprechens in den Therapielinien auf einer Reduktion im NIH-Organscore, 22 % auf einem NIH-Response-Assessment, 20 % auf einer Arztbewertung und 8 % auf einer Reduktion der Kortikosteroid-Dosis. Der pU legt keine differenzierten Angaben dazu vor, aus welchen einzelnen Komponenten sich das Therapieansprechen im Interventions- bzw. im Vergleichsarm zusammensetzt, sondern legt lediglich eine zusammenfassende Darstellung über beide Arme vor. Es ist unklar, ob zu dem NIH-Organ-Score eine vollständige Dokumentation zu allen Organen vorlag oder ob von einem nicht vorhandenen Eintrag für einzelne Organe auf das Nichtvorhandensein von Symptomen bzw. einer stabilen Erkrankung geschlossen wurde.

Ein weiterer Kritikpunkt an der Operationalisierung des Endpunkts Gesamtansprechen ist, dass die Kortikosteroid-Dosis als eigenständige Komponente ohne Kopplung an die Symptomatik in den Auswertungsalgorithmus einbezogen wird. Eine alleinige Betrachtung der

Kortikosteroid-Dosis kann, wie im AbD-Konzept zu Belumosudil [1] beschrieben, ein Therapieziel darstellen, allerdings sollte hierfür eine relevante Schwelle, wie z. B. die Cushing-Schwelle, definiert werden, die anhaltend unterschritten wird. Die Berücksichtigung einer rein prozentualen Reduktion der Kortikosteroid-Dosis ist nicht adäquat. Für die AbD ist es sinnvoll, einen Endpunkt zu operationalisieren, bei dem eine relevante Verbesserung der Symptomatik mit dem gleichzeitigen Erreichen einer Kortikosteroid-Dosis unterhalb einer relevanten Schwelle (z. B. der Cushing-Schwelle) gekoppelt wird.

Eine weitere Unsicherheit entsteht durch die retrospektive Datenerhebung, welche nicht zu einem festen Zeitpunkt erfolgte. Die Erhebung konnte zwischen Woche 2 und 26 nach Beginn einer Therapielinie erfolgen, wobei die zuletzt dokumentierte Angabe vor Woche 26 verwendet wurde. Als Baselinewert konnten sowohl Werte vor Initiierung einer neuen Therapielinie, als auch danach herangezogen werden (Woche –13 bis 2). Angaben zur tatsächlichen Beobachtungsdauer des Endpunkts legt der pU nicht vor.

Insgesamt bestehen relevante Unsicherheiten in der Erhebung und Auswertung des primären Endpunkts Gesamtansprechen, die mithilfe der vorliegenden Dokumente nicht abschließend geklärt werden können.

Aufgrund der beschriebenen Gründe zur Operationalisierung des ORR liefern die vom pU vorgelegten Daten zur Studie ROCKreal keine geeigneten Informationen, die eine orientierende Fallzahlschätzung ermöglichen. Dieses Ergebnis resultiert vor allem aus den identifizierten Unsicherheiten für den primären Endpunkt Gesamtansprechen. Daher ergeben sich aus der Studie ROCKreal keine neuen Daten, anhand derer die zuvor im AbD-Konzept beschriebenen Annahmen für die orientierende Betrachtung von Fallzahlszenarien aktualisiert werden könnten. Aus diesem Grund erfolgt die orientierende Betrachtung von Fallzahlszenarien analog zum AbD-Konzept zu Belumosudil [1] anhand des Endpunkts Lee Symptom Scale Score. Insgesamt kann basierend auf den Daten der Studie ROCKreal somit weder eine orientierende Fallzahlschätzung durchgeführt werden, noch die im AbD-Konzept für die Intervention und Vergleichstherapie beschriebenen Annahmen für eine orientierende Betrachtung von Fallzahlszenarien aktualisiert werden.

2.4 Einordnung der im Beteiligungsverfahren vorgebrachten Ausführungen zur Patientenzahl im Anwendungsgebiet auf die orientierende Betrachtung von Fallzahlszenarien für die AbD

2.4.1 Beschreibung des Vorgehens zur Ermittlung der Patientenzahlen

Im Rahmen des Beteiligungsverfahrens wurden 2 Stellungnahmen mit alternativen Herleitungen für eine orientierende Schätzung der Patientenzahlen übermittelt [9,10]. Das zugrunde liegende Vorgehen in den Stellungnahmen des pU sowie der Deutschen Gesellschaft für Hämatologie und medizinische Onkologie (DGHO) gemeinsam mit der Deutschen

Arbeitsgemeinschaft für Hämatopoetische Stammzelltransplantation und Zelluläre Therapie (DAG-HSZT) wird an dieser Stelle zusammenfassend dargestellt:

Stellungnahme des pU

Der pU bezeichnet die orientierende Schätzung der Patientenzahlen als insgesamt deutlich zu hoch. Demnach sei die Herleitung in Teilen plausibel, einige der Annahmen entsprächen jedoch nicht dem aktuellen deutschen Versorgungskontext. Für Kriterium 1 (Patientinnen und Patienten mit allogener Transplantation) berücksichtigt der pU die gemeldeten Fallzahlen aus dem Jahresbericht des Deutschen Registers für hämatopoetische Stammzelltransplantation und Zelltherapie (DRST) für das Berichtsjahr 2023 [11]. Für Kriterium 2 (Patientinnen und Patienten mit cGvHD) zieht der pU die retrospektive Analyse von Langer 2024 [12] heran, welche auf Daten aus europäischen Zentren (u. a. aus Deutschland) zu Patientinnen und Patienten mit allogener Ersttransplantation im Jahr 2017 basieren: 43,0 % entwickeln eine cGvHD. Dieses Vorgehen deckt sich mit der orientierenden Schätzung der Patientenzahlen im Rahmen des AbD-Konzepts. Für Kriterium 3 weicht der pU vom Vorgehen aus dem AbD-Konzept ab und berücksichtigt erneut die Publikation von Langer 2024 [12], der sich ein Anteilswert von 28,6 % für Patientinnen und Patienten mit cGvHD, die eine Zweitlinienbehandlung benötigen, entnehmen lässt. Wie der pU korrekt anmerkt, wurde dieser Anteilswert auch im Rahmen der Anmerkungen und Limitationen des AbD-Konzepts bereits adressiert. Für die Bestimmung der Patientinnen und Patienten, die eine Drittlinientherapie benötigen zieht der pU die abschließende 3-Jahres-Analyse der Phase-III-Studie REACH3 von Zeiser 2025 [13] mit 329 Patientinnen und Patienten heran, die entweder Ruxolitinib oder alternativ die beste verfügbare Therapie (BAT) erhalten haben. Dieser Publikation entnimmt der pU Raten für das FFS (Beobachtungszeitraum 36 Monate) bei einer Zweitlinienbehandlung mit Ruxolitinib (56,5 %) bzw. mit BAT (18,2 %). Im Umkehrschluss geht er davon aus, dass 43,5 % derjenigen, die mit Ruxolitinib bzw. 81,8 % derjenigen, die mit BAT behandelt wurden, eine Drittlinientherapie benötigen. Für die Versorgungsanteile in der Zweitlinientherapie geht der pU auf Basis einer nicht weiter überprüfbaren Expertenschätzung davon aus, dass 80 % aller Patientinnen und Patienten eine Therapie mit Ruxolitinib erhalten und die verbleibenden 20 % mit BAT behandelt werden. Neben der Expertenschätzung verweist der pU auf eine Umfragestudie zu Veränderungen in der Behandlung der cGvHD unter allogenen transplantierenden Zentren in Deutschland, Österreich und der Schweiz [14], der sich zwar grundsätzlich entnehmen lässt, dass Ruxolitinib eine wichtige Rolle in der Zweitlinientherapie zukommt, der sich hierzu jedoch keine spezifischen Anteilswerte entnehmen lassen. Für den Anteil derjenigen, die eine Drittlinientherapie benötigen berechnet der pU einen Anteilswert von $51,16\% (= 80\% \times 43,5\% + 20\% \times 81,8\%)$. Unter Berücksichtigung eines GKV-Anteils von 88,3 % nimmt der pU somit eine Patientenpopulation 185 bis 198 Patientinnen und Patienten pro Jahr an [9].

Stellungnahme der DGHO und DAG-HSZT

Laut der Stellungnahme der Fachgesellschaften ist die orientierende Schätzung der Patientenzahlen deutlich überschätzt. Hierfür wird auf eine Vielzahl an Unsicherheitsfaktoren ohne genaue Quantifizierung der Auswirkung auf die Fallzahl verwiesen, demnach sei u. a. aufgrund der aktuellen Standard-GvHD-Prophylaxe von einer niedrigeren Inzidenz der cGvHD auszugehen als im AbD-Konzept angenommen. Zudem sei davon auszugehen, dass in den gemeldeten Fallzahlen des DRST (Kriterium 1) auch pädiatrischen Transplantationen von Patientinnen und Patienten im Alter < 12 Jahre enthalten sind. Zur Bestimmung der Patientenzahlen gehen DGHO und DAG-HSZT letztlich von ca. 3500 Transplantationen aus und verweisen für den Drittlinietherapiebedarf zur Behandlung der cGvHD im Jahr 2024 auf Erfahrungen des UKHD, wonach 10 der 122 Patientinnen und Patienten (ca. 8 %; unveröffentlichte Daten) ein Drittlinietherapie erhalten. In der Gesamtschau des deutschen Versorgungskontext wird daraufhin mit einem Anteil von 10 % für den Drittlinietherapiebedarf gerechnet. Darauf basierend wird eine Spanne von 300 bis 400 Patientinnen und Patienten im Anwendungsgebiet geschätzt. Zur Fallzahlplanung empfehlen DGHO und DAG-HSZT aufgrund der vorgebrachten Unsicherheitsfaktoren eine konservative Schätzung und nehmen eine Anzahl von maximal 300 Patientinnen und Patienten an [10].

2.4.2 Bewertung

In der Gesamtschau der verfügbaren Literatur sowie den im Beteiligungsverfahren eingereichten Stellungnahmen ist für den deutschen Versorgungskontext vermutlich von einer Überschätzung der Anzahl der Patientinnen und Patienten im Alter von ≥ 12 Jahren mit cGvHD, die mindestens 2 Vortherapien erhalten haben, auszugehen. Dies zeigt sich, wie von den Stellungnehmenden thematisiert, u. a. an der Publikation von Langer 2024 [12], die bereits für die Zweitlinientherapie einen niedrigeren Anteilswert nahelegt als die bisher herangezogene US-amerikanische Publikation von Bachier 2021 [15] für die Drittlinietherapie. Jedoch enthalten die im Beteiligungsverfahren vorgebrachten Angaben und Publikationen weiterhin keine Anteilswerte, die für eine Berechnung der Patientenzahlen herangezogen werden können, für Patientinnen und Patienten, die eine Drittlinietherapie benötigen. Wesentlicher Grund hierfür ist, dass nicht alle Annahmen bzw. Angaben zur Bestimmung der Anteile in der Drittlinietherapie (vollständig) durch Quellenbelege überprüfbar sind. Für die Stellungnahme des pU betrifft dies den Anteil der Patientinnen und Patienten in der Zweitlinientherapie, die Ruxolitinib bzw. BAT erhalten haben, der auf einer nicht überprüfbaren Expertenschätzung basiert. Die Stellungnahme der DGHO und DAG-HSZT weist einen Anteilswert von etwa 10 % für Patientinnen und Patienten mit allogener Transplantation aus, die eine Drittlinietherapie benötigen. Dieser Wert basiert primär auf unveröffentlichten Ergebnissen des UKHD. Eine Bewertung des hervorgebrachten Anteils ist daher zum aktuellen Stand nicht möglich.

Folglich fehlt weiterhin eine Angabe für den Anteil der Drittlinientherapie im deutschen Versorgungskontext. Der Anteilswert (28,6 %) für die Zweitlinientherapie von Langer 2024 [12], der bereits im AbD-Konzept dahin gehend genannt wurde, dass die US-amerikanische Publikation auch eine Überschätzung darstellen könnte, kann alternativ zu Herleitung der Patientenzahl herangezogen werden, unter der Voraussetzung, dass der deutsche Versorgungstext im Rahmen der orientierenden Schätzung der Patientenzahlen stärker gewürdigt wird. Es bleibt jedoch festzuhalten, dass der Anteilswert aus Langer 2024 nicht die gesuchte Population (Patientinnen und Patienten mit cGvHD, die eine Drittlinientherapie benötigen) darstellt.

Mittels des Anteilswerts aus der Publikation von Langer 2024 für die Zweitlinientherapie (28,6 % [12]) lässt sich folgende Anzahl der Patientinnen und Patienten – wie in Tabelle 1 dargestellt – im Rahmen der orientierenden Schätzung der Patientenzahl herleiten. Dabei bleiben die bereits beschriebenen Unsicherheitsaspekte der Kriterien 1 und 2 bestehen. Für Kriterium 3 ist insbesondere unklar, wie viele der Patientinnen und Patienten mit einer Zweitlinientherapie in der Folge tatsächlich für eine Drittlinientherapie infrage kommen. Dies führt zu einer Überschätzung, da insbesondere Patientinnen und Patienten, die auf die Zweitlinientherapie ansprechen oder, die im Rahmen der Zweitlinie verstorben sind, nicht dafür infrage kommen.

Tabelle 1: Orientierende Schätzung der Patientenzahlen auf Grundlage der einbezogenen epidemiologischen Kriterien

Kriterium	Vorgehen zur orientierenden Einschätzung der Patientenzahlen	Quelle	Anteil [%]	Ergebnis (Patientenzahl) ^a
1	Patientinnen und Patienten im Alter von ≥ 12 Jahren mit allogener Transplantation	DRST [11]	-	3327–3566
2	Patientinnen und Patienten mit cGvHD	Langer 2024 [12]	43,0	1431–1533
3	Patientinnen und Patienten mit mindestens 2 Vortherapien	Langer 2024 [12]	28,6	409–438
a. eigene Berechnung cGvHD: chronische Graft-versus-Host-Disease; DRST: Deutsches Register für hämatopoetische Stammzelltransplantation und Zelltherapie				

Insgesamt kann dem Sachverhalt im Kontext der Fallzahlplanung in höherem Maße Rechnung getragen werden, indem als orientierende Schätzung eine niedrigere Gesamtspanne von 409 bis 438 Patientinnen und Patienten (siehe Tabelle 1) veranschlagt wird.

2.4.3 Auswirkungen auf die orientierende Betrachtung von Fallzahlsszenarien

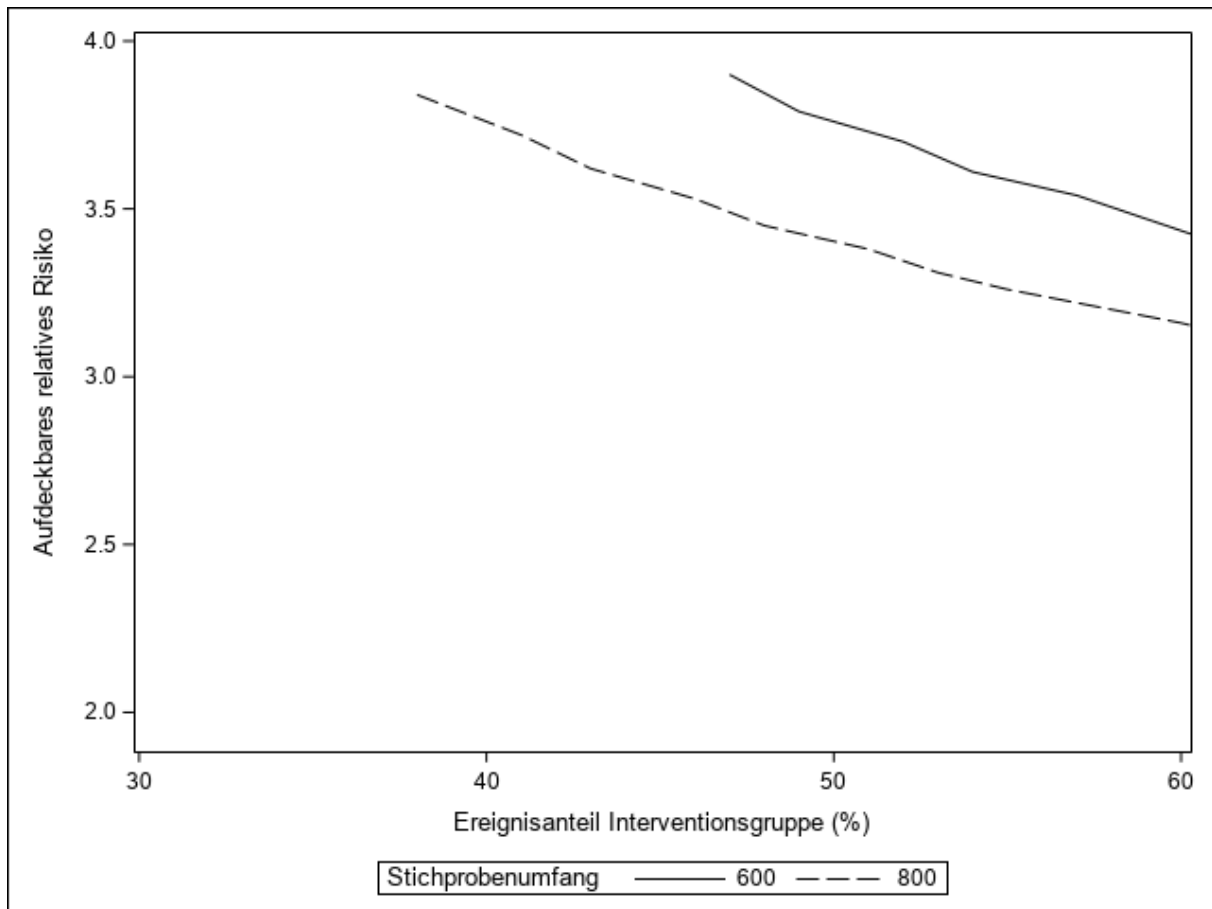
Ausgehend von der angepassten Gesamtspanne an Patientinnen und Patienten (409 bis 438 in Deutschland pro Jahr, siehe Abschnitt 2.4.2) wurde die orientierende Betrachtung von Fallzahlsszenarien aktualisiert. Basierend auf einer Rekrutierungszeit von 2 Jahren wird eine

Stichprobengröße von $N = 800$ betrachtet. Da die geschätzten Patientenzahlen von 409 bis 438 im vorliegenden Anwendungsgebiet ggf. überschätzt sind, wird zusätzlich ein Szenario mit einer Stichprobengröße von $N = 600$ dargestellt. Auf die erneute Darstellung der Stichprobengröße von $N = 1000$, $N = 1250$ und $N = 1500$ aus dem Konzept wird im vorliegenden Addendum verzichtet.

Um eine Einschätzung über die Durchführbarkeit der AbD treffen zu können, werden im Folgenden die Effekte eines mittels Responderanalyse ausgewerteten Endpunkts (Effektmaß: relatives Risiko [RR]) – im vorliegenden Fall die Symptomatik, erhoben mit der Lee Symptom Scale – dargestellt, die bei den im vorliegenden Anwendungsgebiet ermittelten Patientenzahlen mit einer Power von 80 % aufdeckbar sind. Analog zum AbD-Konzept A25-34 [1] werden für die Interventionsgruppe basierend auf den gesichteten Daten Anteile an Patientinnen und Patienten mit einer Verbesserung im Lee Symptom Scale Score (im Folgenden als „Ereignisanteil“ bezeichnet) von 30 % bis 60 % zu Monat 18 angenommen. Für die Kontrollgruppe werden die sich ergebenden Ereignisanteile von 10 % bis 23 % dargestellt (für detaillierte Ausführungen siehe AbD-Konzept A25-34 [1]).

Weiterhin angenommen werden übliche Werte für das Signifikanzniveau ($\alpha = 2,5 \%$, 1-seitiger Test) sowie eine verschobene Nullhypothese (H_0 : relatives Risiko $\leq 2,0$). Die Analysen (zu den orientierenden Betrachtungen von Fallzahlsszenarien) wurden mittels SAS (9.4, Prozedur POWER mit Statement TWOSAMPLEFREQ) für den Score-Test nach Farrington und Manning [16] durchgeführt.

Es werden Rekrutierungsverhältnisse von 5:1, 3:1 und 1:1 zwischen Intervention und Vergleichstherapie betrachtet (für 3:1 und 1:1 siehe Anhang A). Die folgende Abbildung 1 zeigt den aufdeckbaren Effekt (RR: 3,16 bis 3,84) mit einer Power von 80 % zum Vorteil von Belumosudil gegenüber der Vergleichstherapie unter Verwendung festgelegter Stichprobengrößen und Ereignisanteile in beiden Behandlungsgruppen (siehe auch Anhang A).



1-seitiger Test mit verschobener Nullhypothese (H_0 : relatives Risiko $\leq 2,0$), Power 80% und Signifikanzniveau $\alpha = 2,5\%$

Abbildung 1: Aufdeckbarer Effekt (Effektmaß: relatives Risiko) in Abhängigkeit von der Stichprobengröße und den Ereignisanteilen in beiden Behandlungsgruppen (Verhältnis Intervention zu Vergleich 5:1)

Aus Abbildung 1 sowie Tabelle 2 geht z. B. hervor, dass bei einer Stichprobengröße von 800 Patientinnen und Patienten bei einem Verteilungsverhältnis von 5:1 zwischen Interventions- und Vergleichstherapie, einem Ereignisanteil von 55 % in der Interventionsgruppe und einem Ereignisanteil von 17 % in der Kontrollgruppe ein relatives Risiko von 3,26 mit einer Power von 80 % aufgedeckt werden kann.

2.5 Zusammenfassung und Fazit

Basierend auf den vom pU vorgelegten Daten zur Studie ROCKreal kann die bestehende Evidenzlücke, auf die sich die Fragestellung der AbD bezieht, nicht geschlossen werden. Darüber hinaus können aufgrund der beschriebenen Unsicherheiten die vorgelegten Daten weder für eine orientierende Fallzahlschätzung, noch für eine Aktualisierung der im AbD-Konzept beschriebenen Annahmen, auf welcher die orientierende Betrachtung von Fallzahlszenarien beruht, genutzt werden. Unter Berücksichtigung der im Beteiligungsverfahren vorgebrachten Ausführungen wurden die weiteren Stichprobengrößen (N = 800 und N = 600) zur orientierenden Betrachtung von Fallzahlszenarien einbezogen und der aufdeckbare Effekt für die angepasste Stichprobengröße und die zuvor im Konzept getroffenen Annahmen dargestellt.

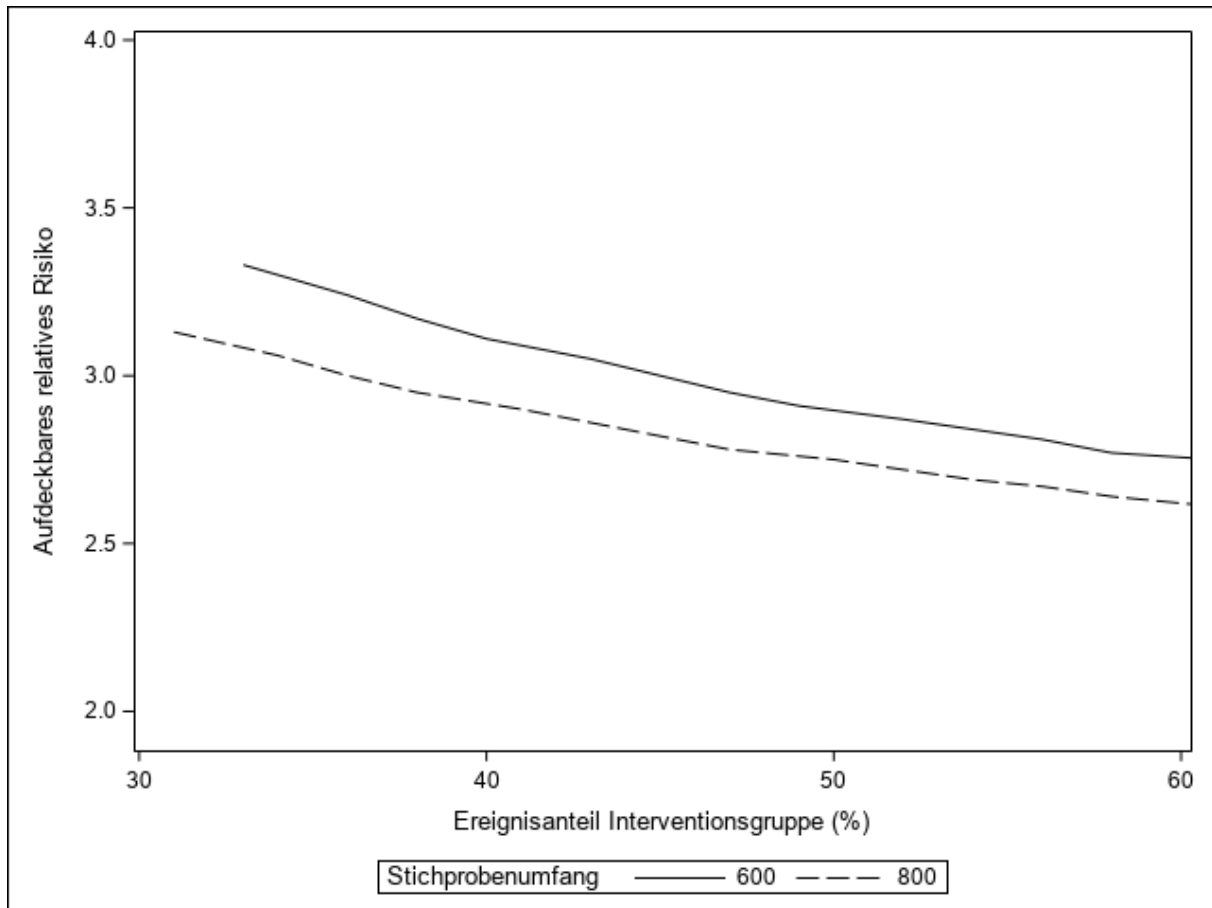
Vor diesem Hintergrund kommt der Durchführung einer AbD zur Generierung vergleichender Daten zu patientenrelevanten Endpunkten ein besonderer Stellenwert zu, um die bestehende Evidenzlücke zu füllen. Mithilfe einer AbD könnten prospektive Daten zu wesentlichen Endpunkten erhoben werden. Die Erhebung würde in der AbD durch ein präspezifiziertes Protokoll einheitlich erfolgen und sicherstellen, dass die Daten in den für die AbD relevanten Operationalisierungen erhoben werden.

3 Literatur

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Belumosudil (chronische Graft-versus-Host-Krankheit) – Bewertung gemäß § 35a SGB V; AbD-Konzept [online]. 2025. URL: <https://doi.org/10.60584/A25-34>.
2. Sanofi-Aventis Deutschland. Stellungnahme zu Einleitung eines Verfahrens zur Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen nach § 35a Absatz 3b SGB V: Belumosudil (chronische Graft-versus-Host-Erkrankung) [unveröffentlicht]. 2025.
3. Hall K, Lazaryan A, Der Laan MV et al. Efficacy and Safety of Belumosudil as Compared with Best Available Therapy for the Treatment of cGVHD in the US. Blood Adv 2025. <https://doi.org/10.1182/bloodadvances.2025015832>.
4. Sanofi-Aventis Deutschland. Efficacy And Safety of Belumosudil as Compared with Physician’s Choice of Best Available Therapy for The Treatment of Chronic Graft Versus Host Disease; The ROCKreal Study; Statistical Analysis Plan; Final Version [unveröffentlicht]. 2024.
5. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; Rapid Report [online]. 2020 [Zugriff: 02.09.2024]. URL: https://www.iqwig.de/download/a19-43_versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_rapid-report_v1-1.pdf.
6. Pufulete M, Mahadevan K, Johnson TW et al. Confounders and co-interventions identified in non-randomized studies of interventions. J Clin Epidemiol 2022; 148: 115-123. <https://doi.org/10.1016/j.jclinepi.2022.03.018>.
7. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS); Arbeitspapier [online]. 2025 [Zugriff: 26.06.2025]. URL: <https://doi.org/10.60584/GA23-02>.
8. Lee SJ, Wolff D, Kitko C et al. Measuring therapeutic response in chronic graft-versus-host disease. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: IV. The 2014 Response Criteria Working Group report. Biol Blood Marrow Transplant 2015; 21(6): 984-999. <https://doi.org/10.1016/j.bbmt.2015.02.025>.

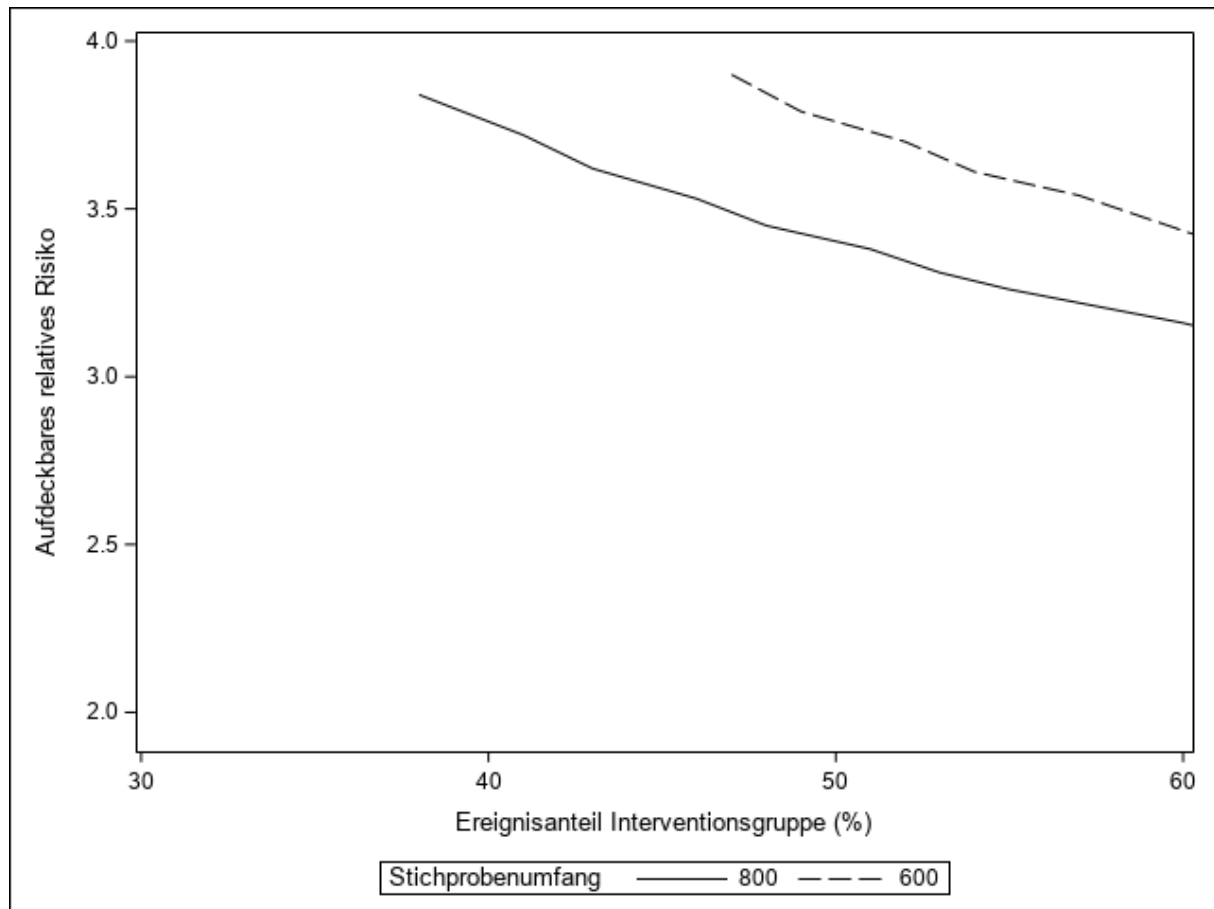
9. Sanofi-Aventis Deutschland. Stellungnahme im Beteiligungsverfahren im Rahmen des Verfahrens zur Forderung einer anwendungsbegleitenden Datenerhebung nach § 35a Absatz 3b SGB V Belumosudil zur Behandlung von Patientinnen und Patienten ab 12 Jahren mit chronischer Graft-versus-Host-Erkrankung nach Versagen von mindestens zwei vorherigen systemischen Therapien (2024-AbD-008). [Demnächst verfügbar unter: <https://www.g-ba.de/anwendungsbegleitende-datenerhebung-verfahren/21/> im Dokument "Zusammenfassende Dokumentation"].
10. Deutsche Gesellschaft für Hämatologie und medizinische Onkologie, Deutsche Arbeitsgemeinschaft für Hämatopoetische Stammzelltransplantation und Zelluläre Therapie. Stellungnahme Beteiligungsverfahren an dem Verfahren zur Forderung einer anwendungsbegleitenden Datenerhebung nach § 35a Absatz 3b SGB V zu einem Konzept des IQWiG zu einer anwendungsbegleitenden Datenerhebung zu Belumosudil zur Behandlung der chronischen Graft-versus-Host-Krankheit. [Demnächst verfügbar unter: <https://www.g-ba.de/anwendungsbegleitende-datenerhebung-verfahren/21/> im Dokument "Zusammenfassende Dokumentation"].
11. Deutsches Register für hämatopoetische Stammzelltransplantation und Zelltherapie. Jahresbericht 2023 [online]. 2024 [Zugriff: 16.05.2025]. URL: <https://drst.de/wp-content/uploads/2024/08/jb2023.pdf>.
12. Langer R, Lelas A, Rittenschober M et al. Retrospective analysis of the incidence and outcome of late acute and chronic graft-versus-host disease-an analysis from transplant centers across Europe. Front Transplant 2024; 3: 1332181. <https://doi.org/10.3389/frtra.2024.1332181>.
13. Zeiser R, Russo D, Ram R et al. Ruxolitinib in Patients With Corticosteroid-Refractory or Corticosteroid-Dependent Chronic Graft-Versus-Host Disease: 3-Year Final Analysis of the Phase III REACH3 Study. J Clin Oncol 2025; 43(23): 2566-2571. <https://doi.org/10.1200/JCO-24-02477>.
14. Denk A, Fante MA, Heidenreich S et al. Changes in Chronic Graft-versus-Host Disease Treatment Over Time: A 15-Years Survey Within Allogeneic Hematopoietic Stem Cell Transplant Centers in Germany, Austria, and Switzerland. Transplantation and Cellular Therapy 2025; 31(7): 450.e1-450.e10. <https://doi.org/10.1016/j.itct.2025.04.017>.
15. Bachier CR, Aggarwal SK, Hennegan K et al. Epidemiology and Treatment of Chronic Graft-versus-Host Disease Post-Allogeneic Hematopoietic Cell Transplantation: A US Claims Analysis. Transplant Cell Ther 2021; 27(6): 504.e1-504.e6. <https://doi.org/10.1016/j.itct.2020.12.027>.
16. Farrington CP, Manning G. Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. Stat Med 1990; 9(12): 1447-1454. <https://doi.org/10.1002/sim.4780091208>.

Anhang A Aufdeckbarer Effekt für verschiedene Stichprobengrößen und verschiedene Ereignisanteile unter Intervention und Vergleichstherapie bei unterschiedlichen Verteilungsverhältnissen



1-seitiger Test mit verschobener Nullhypothese (H_0 : relatives Risiko $\leq 2,0$), Power 80% und Signifikanzniveau $\alpha = 2,5\%$

Abbildung 2: Aufdeckbarer Effekt (Effektmaß: relatives Risiko) in Abhängigkeit von der Stichprobengröße und den Ereignisanteilen in beiden Behandlungsgruppen (Verhältnis Intervention zu Vergleich 1:1)



1-seitiger Test mit verschobener Nullhypothese (H_0 : relatives Risiko $\leq 2,0$), Power 80% und Signifikanzniveau $\alpha = 2,5\%$

Abbildung 3: Aufdeckbarer Effekt (Effektmaß: relatives Risiko) in Abhängigkeit von der Stichprobengröße und den Ereignisanteilen in beiden Behandlungsgruppen (Verhältnis Intervention zu Vergleich 3:1)

Tabelle 2: Aufdeckbarer Effekt (Effektmaß: RR) bei verschiedenen Stichprobengrößen und Ereignisanteilen in beiden Behandlungsgruppen; Verhältnis der Behandlungsgruppen von 1:1, 3:1 und 5:1 (Intervention:Kontrolle); Signifikanzniveau $\alpha = 2,5\%$ (1-seitiger Test); Power 80 %; verschobene Nullhypothese ($H_0: RR \leq 2$)

Ereignis- anteil (%) in der Kontroll- gruppe	Verteilungsverhältnis 1:1				Verteilungsverhältnis 3:1				Verteilungsverhältnis 5:1			
	Stichprobenumfang N = 600		Stichprobenumfang N = 800		Stichprobenumfang N = 600		Stichprobenumfang N = 800		Stichprobenumfang N = 600		Stichprobenumfang N = 800	
	Ereignisanteil (p _i in %) in Interventionsgruppe und aufdeckbares RR											
	p _i	RR	p _i	RR	p _i	RR	p _i	RR	p _i	RR	p _i	RR
10	33	3,33	31	3,13	–	–	38	3,84	–	–	38	3,84
11	36	3,24	34	3,06	–	–	41	3,72	–	–	41	3,72
12	38	3,17	36	3	47	3,9	43	3,62	47	3,9	43	3,62
13	40	3,11	38	2,95	49	3,79	46	3,53	49	3,79	46	3,53
14	43	3,95	41	2,9	52	3,7	48	3,45	52	3,7	48	3,45
15	45	3	43	2,86	54	3,61	51	3,38	54	3,61	51	3,38
16	47	2,95	45	2,82	57	3,54	53	3,31	57	3,54	53	3,31
17	49	2,91	47	2,78	59	3,47	55	3,26	59	3,47	55	3,26
18	52	2,87	50	2,75	–	–	58	3,2	–	–	58	3,2
19	54	2,84	52	2,72	–	–	60	3,16	–	–	60	3,16
20	56	2,81	54	2,69	–	–	–	–	–	–	–	–
21	58	2,77	56	2,67	–	–	–	–	–	–	–	–
22	–	–	58	2,64	–	–	–	–	–	–	–	–
23	–	–	60	2,62	–	–	–	–	–	–	–	–
N: Stichprobenumfang; p _i : Ereignisanteil in der Interventionsgruppe; RR: relatives Risiko												