

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: Pegzilarginase

Datum der Veröffentlichung: 15. April 2024

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abkürzungsverzeichnis	4
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	9
2.3 Endpunkte	13
2.3.1 Mortalität	15
2.3.2 Morbidität	15
2.3.3 Lebensqualität	30
2.3.4 Sicherheit	32
2.3.5 Erhebungszeitpunkte	34
2.4 Statistische Methoden	34
2.5 Verzerrungspotential auf Studien- und Endpunktebene	36
3 Ergebnisse der eingeschlossenen Studie	38
3.1 Studiencharakteristika und Studienmedikation	38
3.2 Mortalität	43
3.3 Morbidität	43
3.4 Lebensqualität	45
3.5 Sicherheit	45
4 Diskussion der Methodik und Ergebnisse	49
4.1 Zulassungsstatus und Zulassungspopulation von Pegzilarginase	49
4.2 Design und Methodik der Studie	50
4.3 Mortalität	52
4.4 Morbidität	52
4.5 Lebensqualität	54
4.6 Sicherheit	54
5 Anforderung an eine qualitätsgesicherte Anwendung	56
6 Zusammenfassung der Nutzenbewertung	57
Referenzen	59
Anhang	63

Tabellenverzeichnis

Tabelle 1: Übersicht über die Studienbasis	8
Tabelle 2: Charakterisierung der Studie PEACE.....	9
Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie PEACE.....	12
Tabelle 4: Charakterisierung der Intervention Pegzilarginase	13
Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie PEACE.....	14
Tabelle 6: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie PEACE	34
Tabelle 7: Verzerrungspotential der RCT-Phase der Studie PEACE.....	36
Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie PEACE	37
Tabelle 9: Allgemeine Angaben; Studie PEACE	38
Tabelle 10: Charakterisierung der Studienpopulation der Studie PEACE	39
Tabelle 11: Angaben zur Begleitbehandlung mit Stickstoffängern, Aminosäurehaltige Ergänzungsmittel und Ernährungsanpassungen; Studie PEACE	42
Tabelle 12: Veränderung der Gehstrecke in Metern zu Studienwoche 24 (RCT) in der Studie PEACE, FAS	43
Tabelle 13: Veränderung im GMFM Dimension E zu Studienwoche 24 (RCT) in der Studie PEACE, FAS.....	44
Tabelle 14: Ergebnisse für den GFAQ zu Studienwoche 24 (RCT) in der Studie PEACE, FAS.....	44
Tabelle 15: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der RCT-Phase der Studie PEACE, FAS	45
Tabelle 16: UE mit Inzidenz $\geq 10\%$ nach Systemorganklasse und Preferred Term in der RCT-Phase der PEACE-Studie, FAS	46
Tabelle 17: Schwerwiegende UE, die bei $\geq 5\%$ der Patientinnen und Patienten in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term in der RCT-Phase der PEACE-Studie, FAS.....	47
Tabelle 18: UE von besonderem Interesse (AESI) in der RCT-Phase der PEACE-Studie, FAS	48
Tabelle 19: Zusammenfassende Darstellung der Ergebnisse der Studie PEACE	57
Tabelle 20: Veränderung der Argininkonzentration im Plasma von Baseline zu Studienwoche 24 (RCT) der log-transformierten Arginin-Daten in der Studie PEACE, FAS	63
Tabelle 21: Veränderung im GMFM Dimension D zu Studienwoche 24 (RCT-Phase) in der Studie PEACE, FAS.....	64
Tabelle 22: Sensitivitätsanalyse der Veränderung im GMFM Dimension D zu Studienwoche 24 (RCT-Phase) in der Studie PEACE, FAS	65
Tabelle 23: Ergebnisse für die FMS – der RCT-Phase der PEACE-Studie, FAS	65
Tabelle 24: Veränderung im CaGI-C zu Studienwoche 24, deskriptive Darstellung der Caregiver-Bewertung in der Studie PEACE, FAS.....	66
Tabelle 25: Ergebnisse für das Diättagebuch – der RCT-Phase der PEACE-Studie, FAS.....	68

Abkürzungsverzeichnis

9HPT	9-Hole Pegboard Test
2MWT	2 Minutes Walking Test
6MWT	6 Minutes Walking Test
ABCS	Adaptive Behaviour Composite Score
ADA	Anti-Drug Antibodies
AESI	Adverse events of special interest (Unerwünschte Ereignisse von besonderem Interesse)
ALT	Alanin-Aminotransferase
ANCOVA	Analysis of covariance (Kovarianzanalyse)
AST	Aspartat-Aminotransferase
ARG1-D	Arginase-1-Mangel
ATS	American Thoracic Society
BT	Botulinumtoxin
CaGI-C/CGI-C	Caregiver/Clinician Global Impressions of Change
CaGI-S/CGI-S	Caregiver/Clinician Global Impressions of Severity
COVID-19	Coronavirus SARS-CoV-2
DCS	Domain Composite Score
EAA	Essential amino acids (essentielle Aminosäuren)
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FAS	Full Analysis Set
FMS	Functional Mobility Scale
FSIQ	Full Scale Intelligence Quotient
GAA	Guanidinoacetic Acid (Guanidinoessigsäure)
G-BA	Gemeinsamer Bundesausschuss
GC	Guanidino Compund (Guanidinoverbindung)
GFAQ	Gillette Functional Assessment Questionnaire
GGT	Gamma-Glutamyltransferase
GKV	Gesetzliche Krankenversicherung
GMFCS	Gross Motor Function Classification System
GMFM	Gross Motor Function Measure-88
HLT	High Level Term nach MedDRA
ICC	Intraklassen-Korrelation
IDM	Individuelles Disease Management
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to Treat
i.v.	Intravenös
IXRS	Interactive Web/Voice Response System
KG	Körpergewicht
KI	Konfidenzintervall

LTE	Long Term Extension (Langzeit-Extensionsphase)
MAS	Modified Ashworth Scale
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Difference (Minimaler relevanter Unterschied)
µM	Mikromolar
MMRM	Mixed Models for Repeated Measure
MW	Mittelwert
N	Anzahl
NAA	N-Acetyl-L-Arginin
NIH	National Institutes of Health
PedsQL	Pediatric Quality of Life Inventory
PT/INR	Prothrombin Time/International Normalized Ratio
PT	Preferred Term/s
PK/PD	Pharmakokinetik/Pharmakodynamik
PP	Per Protocol
pU	Pharmazeutischer Unternehmer
RBC	Red Blood Cells
RCP	Randomisiert kontrollierte Periode
RCT	Randomized controlled trial (randomisierte kontrollierte Studie)
RR	Risk Ratio
SAP	Statistischer Analyseplan
SD	Standardabweichung
SF-36	36-Item Short Form Health Survey
SGB	Sozialgesetzbuch
SOC	Systemorganklasse
SUE	Schwerwiegendes unerwünschtes Ereignis
TEAE	Treatment-emergent adverse event (behandlungsbedingte unerwünschte Ereignisse)
UE	Unerwünschtes Ereignis
VABS-II	Vineland Adaptive Behavior Scale, Second Edition
VerfO	Verfahrensordnung des G-BA
WAIS-IV	Wechsler Adult Intelligence Scale, Fourth Edition
WIS	Wechsler Intelligence Scales
WISC-V	Wechsler Intelligence Scale for Children, Fifth Edition
WPPSI-IV	Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition
ZBI-12	Zarit Burden Interview (12-Item)

Hintergrund

Pegzilarginase 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 Pegzilarginase 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 Pegzilarginase in seiner Sitzung am 09. April 2024 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 15. Januar 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 15. April 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

Pegzilarginase (Loargys®) 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 [20]:

Zur Behandlung von Arginase-1-Mangel (ARG1-D), auch bekannt als Hyperargininämie, bei Erwachsenen, Jugendlichen und Kindern im Alter ab 2 Jahren.

Gemäß Fachinformation ist Pegzilarginase bestimmt für die chronische Behandlung von Patientinnen und Patienten mit ARG1-D in Verbindung mit einer individualisierten Behandlung der Erkrankung, wie z. B. der Proteinrestriktion in der Ernährung, Aminosäure-haltigen Ergänzungsmitteln und pharmakologischer Behandlung, einschließlich Stickstoffängern.

Die empfohlene Anfangsdosis beträgt gemäß Fachinformation 0,1 mg/kg pro Woche. Nach Beginn der Behandlung sollte die wöchentliche Dosis auf der Grundlage der Arginin-Konzentrationen vor der Gabe angepasst werden, um die Plasmakonzentration von Arginin innerhalb des Normbereichs zu halten. Die Dosis kann in Schritten von 0,05 mg/kg erhöht oder verringert werden, um die therapeutischen Ziele zu erreichen. Dosen über 0,2 mg/kg/Woche wurden in klinischen Studien zu ARG1-D nicht untersucht. Pegzilarginase sollte durch intravenöse Infusion oder subkutane Injektion gegeben werden, wobei die Dosis gleich ist.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Studien zum Wirkstoff				
Studie PEACE (CAEB1102-300A) ¹⁾	Ja	Ja	Ja	-
CAEB-1102-101A ²⁾	Ja	Nein	Nein	<ul style="list-style-type: none"> • einarmige Dosisfindungsstudie (PK/PD und Sicherheit; N = 16) • Dosierung erfolgt nicht gemäß der Vorgabe und Empfehlung der FI³⁾ • Kürzere Beobachtungsdauer als im bewertungsrelevanten RCT (PEACE)
CAEB1102-102A ²⁾	Ja	Nein	Nein	<ul style="list-style-type: none"> • Einarmige Langzeitnachbeobachtung (N = 14) zu 101A mit identifizierten Einschränkungen in der Dosierung nach FI (s. o.)

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

²⁾ Im Rahmen des Zulassungsverfahrens bei der EMA ergänzend zu der Zulassungsstudie für die Sicherheit berücksichtigt.

³⁾ Dosisgabefrequenz nicht über gesamte Studiendauer zulassungskonform und Dosierung entspricht nicht für alle Studienteilnehmenden den Empfehlungen der FI.

Abkürzungen: EMA: European Medicines Agency; EPAR: European Public Assessment Report; FI: Fachinformation; PK/PD: Pharmakokinetik/Pharmakodynamik; pU: pharmazeutischer Unternehmer; RCT: randomisierte kontrollierte Studie

Die für die Nutzenbewertung herangezogene Studie entspricht derjenigen Studie, auf der die Einschätzungen des pU zum Zusatznutzen basieren. Ergebnisse zur Langzeitextensionsphase von 150 Wochen der PEACE-Studie, welche laut Dossier [22] am 01. Februar 2023 abgeschlossen wurde, wurden beim pU angefragt, jedoch mit der Begründung eines fehlenden Studienberichtes nicht nachgereicht. Für die Nutzenbewertung wird daher lediglich die randomisiert kontrollierte Phase der Studie berichtet.

Zur Nutzenbewertung für Pegzilarginase herangezogene Unterlagen

- Dossier des pU zu Pegzilarginase [21,22]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [15,16]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie PEACE (CAEB1102-300A) [3,4,5]
- Dosisfindungsstudie CAEB-1102-101A mit Langzeitextensionsphase 102A [2,6]
- Fachinformation/Produktinformation zu Pegzilarginase [20]

2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Pegzilarginase im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie PEACE (CAEB1102-300A). Diese wird im Folgenden nur noch als PEACE-Studie bezeichnet. Die Studie und die Intervention werden in den Tabellen 2 und 4 charakterisiert.

Tabelle 2: Charakterisierung der Studie PEACE

Charakteristikum	Beschreibung
Design und Studienablauf	<p>Design Die Studie PEACE ist eine parallele, doppelblinde, multizentrische, Placebo-kontrollierte Phase-III-Studie zur Untersuchung der Wirksamkeit und Verträglichkeit von Pegzilarginase im Vergleich zu Placebo, jeweils in Kombination mit individuellem Krankheitsmanagement, bei Kindern (≥ 2 Jahre), Jugendlichen und Erwachsenen mit Arginase-1-Mangel (ARG1-D).</p> <p>Die Patientinnen und Patienten wurden im Verhältnis 2:1 randomisiert (Pegzilarginase, Placebo) und nach der Schwere der Hyperammonämie in der Vorgeschichte stratifiziert.</p> <p><u>Strata der Randomisierung¹⁾:</u></p> <ul style="list-style-type: none"> • ≥ 1 hyperammonämische Episode innerhalb von 90 Tagen vor der Einwilligung oder ≥ 2 hyperammonämische Episoden innerhalb von 365 Tagen vor der Einwilligung • Keine hyperammonämische Episode innerhalb von 365 Tagen vor der Einwilligung oder 1 hyperammonämische Episode, die > 90 Tage vor der Einwilligung auftrat <p>Studienablauf Die PEACE-Studie besteht aus 3 Phasen:</p> <ol style="list-style-type: none"> 1. Screening-Phase (3 bis 4 Wochen) 2. Doppelt verblindete RCT-Phase (24 Wochen) 3. Unverblindete, einarmige Langzeit-Extensions-Phase (LTE) (bis zu 150 Wochen).²⁾
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Vorliegen einer aktuellen ARG1-D-Diagnose, die in der Krankenakte dokumentiert ist und eine der folgenden Bedingungen erfüllen muss: <ul style="list-style-type: none"> ○ Erhöhte Arginin-Konzentration im Plasma, ○ Mutationsanalyse, mit Diagnose einer pathogenen Mutation oder ○ reduzierte Arginase-Aktivität in den Erythrozyten. <p>Des Weiteren muss das folgende Kriterium zutreffen: Der Durchschnitt aller gemessenen Plasma-Arginin-Werte während der Screening-Phase vor der Randomisierung (Visite 1, Studientag 1) muss $\geq 250 \mu\text{M}$ sein. Wenn ein Patient / eine Patientin erneut gescreent wird, werden nur die Werte aus der aktuellen Screening-Phase berücksichtigt.</p> <ul style="list-style-type: none"> • Alter ≥ 2 Jahre • Der Studienteilnehmende muss bewertbar sein. D. h. die Patientin / der Patient muss in der Lage sein, die Untersuchung durchzuführen, und muss zu Baseline ein Defizit in mindestens einer Komponente (2MWT, GMFM-D, GMFM-E) aufweisen (siehe Abbildung S. 12). • Bestätigung des ärztlichen Prüfpersonals und/oder des Ernährungsberaters / der Ernährungsberaterin, dass der Studienteilnehmende seine Ernährung gemäß den im Protokoll dargelegten Ernährungsvorgaben beibehalten kann, d. h. bzgl. aktuellem Niveau des Proteinkonsums (einschließlich natürlichem Eiweiß sowie Ergänzung essentieller Aminosäuren)

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Bei Therapie mit Stickstoffängern, Antiepileptika und/oder Medikamenten gegen Spastizität muss ≥ 4 Wochen vor Randomisierung eine stabile Dosis des Medikaments vorliegen und der Patient / die Patientin muss bereit sein, diese Dosis während der verblindeten Studiendauer beizubehalten. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Hyperammonämische Episode innerhalb der letzten 6 Wochen vor der ersten Studienmedikationsgabe. Eine hyperammonämische Episode ist definiert als ein Ereignis, bei dem eine Person einen Ammoniakspiegel $\geq 100 \mu\text{M}$ sowie ein oder mehrere Symptome, die im Zusammenhang mit einer Hyperammonämie stehen, aufweist und deshalb hospitalisiert oder in der Notaufnahme behandelt wird. • Eine aktive Infektion, die eine antiinfektiöse Therapie benötigt, innerhalb von 3 Wochen vor der ersten Studienmedikationsgabe. • Extreme Mobilitätsdefizite, definiert als jene Personen, welche ungeeignet für eine Beurteilung mittels des Gillette Functional Assessment Questionnaire (GFAQ) sind bzw. welche einen GFAQ-Gesamtwert von 1 aufweisen. • Ein Gesundheitszustand oder Komorbiditäten, welche nach Ansicht des ärztlichen Prüfpersonals die Einhaltung der Studie bzw. die Interpretation der Studiendaten beeinträchtigen (z. B. schwere geistige Beeinträchtigungen). • Teilnahme an einer früheren Pegzilarginase-Interventionsstudie. • Bekannte Überempfindlichkeit gegenüber Polyethylenglykol, welche nach Einschätzung des ärztlichen Prüfpersonals für die Studienteilnehmenden mit inakzeptablen Risiken für Nebenwirkungen verbunden ist. • Patientinnen/Patienten, welche mit Botulinumtoxin(BT)-haltigen Therapien behandelt werden oder eine entsprechende Behandlung während der doppelt verblindeten Phase der Studie sowie der verblindeten Nachbeobachtungsphase planen oder welche eine chirurgische bzw. BT-Behandlung aufgrund von spastikbedingten Komplikationen innerhalb von 16 Wochen³⁾ vor der ersten Dosis der Studienmedikation erhalten haben. • Frühere Leber- oder hämatopoetische Stammzelltransplantation
Intervention und Zahl der Patientinnen und Patienten	<p>Gescreent gesamt: N = 44⁵⁾ Randomisiert/Eingeschlossen: N = 32 Intervention: N = 21 Kontrolle: N = 11 Übergang in die LTE-Phase der Studie: N = 31</p>
Ort und Zeitraum der Durchführung; Datenschnitte	<p>Studienzentren 19 Studienzentren in 7 Ländern (USA 9, Vereinigtes Königreich 4, Österreich 1, Kanada 1, Frankreich 2, Deutschland 1, Italien 1)</p> <p>Studienzeitraum</p> <ul style="list-style-type: none"> • Erste/r Patient/in, erste Visite: 01. Mai 2019 • Letzte/r Patient/in, letzte Visite in der RCT-Periode: k. A. • Letzte/r Patient/in, letzte Visite der LTE: k. A. • Studienende: 01. Februar 2023 <p>Datenschnitte/Interimsanalysen Es wurden keine Interimsanalysen für den Zeitraum vor Abschluss der RCT-Phase präspezifiziert oder durchgeführt. Interimsanalysen zu Ende der RCT-Phase (präspezifiziert): 14. Oktober 2021 Studienende nach Woche 150 der LTE-Phase: 24. Januar 2023</p>

Charakteristikum	Beschreibung
	Die Interimsanalyse am Ende der 24-wöchigen RCT-Phase ist Gegenstand der vorliegenden Nutzenbewertung. Ergebnisse zum Datenschnitt vom 24. Januar 2023 wurden nicht vorgelegt (siehe Kapitel 2.1).
Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt Veränderung der Arginin-Konzentration im Plasma von Baseline zu Woche 24</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • 2-Minute Walk Test (2MWT) (Mittlere Veränderung von Baseline zu Woche 24) • Gross Motor Function Measure-88 (GMFM) Teil E und Teil D (Mittlere Veränderung von Baseline zu Woche 24) • Functional Mobility Scale (FMS)-5, FMS-50 und FMS-500 (Mittlere Veränderung von Baseline zu Woche 24) • Gilette Functional Assessment Questionnaire (GFAQ) (Mittlere Veränderung von Baseline zu Woche 24) • Vineland Adaptive Behavior Scales, Second Edition (VABS)-II (Mittlere Veränderung von Baseline zu Woche 24) • Anti-Arzneimittel-Antikörper (Anti-Drug Antibodies, ADA) • Pharmakokinetik und Pharmakodynamik • Unerwünschte Ereignisse (UE) • Ornithin-Werte sowie Guanidinverbindungen (Guanidino Compounds, GC) <p>Tertiäre Endpunkte</p> <ul style="list-style-type: none"> • Responderanalyse des kombinierten Mobilitätsendpunkts, bestehend aus: 2MWT, GMFM Teil E und GMFM Teil D zu Woche 24 • Responderanalyse zu Woche 24 des: <ul style="list-style-type: none"> ○ FMS-5, FMS-50 und FMS-500 ○ GFAQ ○ VABS-II • Modified Ashworth Scale (MAS)⁴⁾ <p>Deskriptive tertiäre Endpunkte</p> <ul style="list-style-type: none"> • Caregiver/Clinician Global Impressions of Change (CaGi-C/CGI-C) • Caregiver/Clinician Global Impressions of Severity (CaGi-S/CGI-S) • z-Scores des 9-Hole Pegboard Test (9HPT) • Pediatric Quality of Life Inventory (PedsQL) • 36-Item Short Form Health Survey (SF-36) • 12-Item Short Form Zarit Burden Interview (ZBI-12) • Neurokognitive Fähigkeiten mittels altersspezifischer Fragebögen: <ul style="list-style-type: none"> ○ Wechsler Adult Intelligence Scale, Fourth Edition (WAIS-IV) ○ Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V) ○ Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV) ○ Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) • 3-Tages-Ernährungsprotokolle • Wachstumsparameter: z-Scores von Größe, Gewicht und BMI • Ammoniak-Konzentration im Plasma

¹⁾ Ein Ausschlusskriterium lautet: Hyperammonämische Episode innerhalb der letzten 6 Wochen bevor die erste Dosis des Studienmedikaments verabreicht wird.

²⁾ Die ersten 8 Wochen der LTE-Phase waren weiterhin verblindet. Probandinnen und Probanden, die ursprünglich auf Pegzilarginase randomisiert wurden, erhielten weiterhin ihre jeweils optimierte Dosis. Probandinnen und Probanden aus dem Placeboarm begannen die 8-wöchige DB-Periode mit einer Dosis von 0,10 mg/kg, die während der LTE-Periode je nach Argininspiegel angepasst werden konnte. Nach den ersten 8 Wochen hatten die Probandinnen und Probanden die Möglichkeit, Pegzilarginase in Form von s.c.

Injektionen zu erhalten (4 Dosen am Prüfartzstandort). Die s.c. Dosen entsprechen dabei den ermittelten optimalen i.v. Dosen.

- 3) Mit Protokollversion 3 vom 25.07.2019 wurde der Zeitraum von 6 Monaten auf 16 Wochen verkürzt mit der Begründung, dass eine wiederholte Anwendung von BT zur Behandlung der Spastizität alle 12 Wochen erfolgt und durch eine Reduzierung der Zeit mehr Patientinnen und Patienten eingeschlossen werden können.
- 4) Wurde erst mit Protokollversion 5 vom 08.06.2020 eingeführt, wodurch nur 7 Personen im Pegzilarginasearm und 5 Personen im Placeboarm ausgewertet wurden.
- 5) Informationen zu den 12 Personen, die nicht für die Studie PEACE geeignet waren (Screen failures) liegen nicht vor.

Abkürzungen: 2MWT: 2-Minuten-Gehtest (2-Minute Walking Test); 9HPT: 9-Hole Pegboard Test; BSID-III: Bayley Scales of Infant and Toddler Development, Third Edition; CaGI-S: Caregiver Global Impressions of Severity; CGI-S: Clinician Global Impressions of Severity; DB: Doppelblind; FMS: Functional Mobility Scale; GFAQ: Gillette Functional Assessment Questionnaire; GMFM: Gross Motor Function Measure-88; PedsQL: Pediatric Quality of Life Inventory; LTE: Langzeit-Extension; i.v.: intravenös; s.c.: subkutan; SF-36: 36-Item Short Form Health Survey; VABS-II: Vineland Adaptive Behavior Scale, Second Edition; WAIS-IV: Wechsler Adult Intelligence Scale, Fourth Edition; WIS: Wechsler Intelligence Scales; WISC-V: Wechsler Intelligence Scale for Children, Fifth Edition; ZBI-12: Zarit Burden Interview (12-Item)

Zielkriterium	Definition der Baseline-Defizite		
2MWT	Die Definition eines Baseline-Defizites für den 2MWT variiert je nach Alter und Geschlecht		
	Alter [Jahre]	Weiblich	Männlich
	3–5	< 112,9	< 110,6
	6–8	< 155,8	< 154,9
	9–11	< 172,0	< 169,9
	12–15	< 168,7	< 172,1
	16–17	< 167,5	< 173,4
	≥ 18	< 142,4	< 148,8
GMFM-D	< 35		
GMFM-E	< 68		

Abbildung: Einschlusskriterium: Aufzuweisende Defizite

Protokolländerungen

Es wurden 6 Änderungen des Originalprotokolls vom 11.12.2018 vorgenommen. Der Einschluss des ersten Studienteilnehmenden in die Studie PEACE war der 01. Mai 2019.

Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie PEACE

Amendment	Wesentliche Änderungen
Version 3 vom 25.07.2019 (nur für USA und Österreich), (k. A. über die Anzahl der bis dahin eingeschlossenen Personen)	<ul style="list-style-type: none"> • Erfordernis einer stabilen, konsistenten Ernährung während der verblindeten Studienphasen (RCT und 8 Wochen der LTE) • Erhebung des GMFM-Fragebogens auf die Teile D und E beschränkt. • Überempfindlichkeitsreaktionen und hyperammonämische Episoden werden als AESI betrachtet.
Version 5 ¹⁾ vom 08.06.2020, (k. A. über die Anzahl der bis dahin eingeschlossenen Personen)	<p>Version 5 umfasst für alle Studienzentren außerhalb der USA und Österreich die Änderungen aus Version 3 (s. o.).</p> <ul style="list-style-type: none"> • Es wurden keine weiteren für die Nutzenbewertung relevanten Protokolländerungen identifiziert.
Version 6 vom 11.12.2020, (k. A. über die Anzahl der bis dahin eingeschlossenen Personen)	<ul style="list-style-type: none"> • Klarstellung, dass ein Krankenhausaufenthalt nur dann als SUE gewertet wird, wenn dieser mindestens 24 Stunden beträgt.

¹⁾ Version 4 vom 02.03.2020 wurde zurückgezogen, da Version 5 entwickelt wurde, bevor Version 4 eingereicht war.

Abkürzungen: AESI: Unerwünschte Ereignisse von besonderem Interesse; GMFM: Gross Motor Function Measure-88; k. A.: keine Angabe; LTE: Langzeit-Extension; SUE: Schwerwiegendes unerwünschtes Ereignis

Charakterisierung der Intervention

Tabelle 4: Charakterisierung der Intervention Pegzilarginase

Intervention	Kontrolle
Pegzilarginase + IDM <ul style="list-style-type: none"> • Dosierung: Einmal wöchentlich 0,05–0,20 mg/kg Körpergewicht intravenös • Startdosis: 0,10 mg/kg Körpergewicht (Dosierungsstufe 2)¹⁾ • Dosisanpassung ab Visite 5 • Dosiserhöhung um 0,1 mg/kg Körpergewicht, wenn die Plasma-Arginin-Konzentration 168 Stunden nach einer Dosis > 150 µM liegt, die Dosis während der letzten beiden Visiten nicht verändert wurde und die letzten beiden Visiten nicht verpasst wurden • Maximale Dosis: 0,2 mg/kg Körpergewicht (Dosierungsstufe 4)¹⁾ • Dosisreduktion um 0,05 mg/kg Körpergewicht, wenn Plasma-Arginin 168 Stunden nach einer Dosis bei 2 aufeinander folgenden Visiten < 50 µM ist • Minimale Dosis: 0,05 mg/kg Körpergewicht (Dosierungsstufe 1)¹⁾ • Prämedikation mit einem nicht-sedierenden Antihistaminikum ist 30 Minuten vor Start der Infusion empfohlen. 	Placebo + IDM Wie bei Intervention
Nicht erlaubte Begleitmedikation bzw. -therapien <ul style="list-style-type: none"> • Botulinumtoxin bis zum Ende der verblindeten Phase (LTE09) Erlaubte Begleitmedikation Erlaubte Begleittherapien wurden – mit Ausnahme der IMD (s. u.) – nicht näher spezifiziert.	
IMD in der PEACE-Studie konnte folgende Therapiemöglichkeiten beinhalten: <ul style="list-style-type: none"> • Diätetische Proteinrestriktion • Supplementierung essentieller Aminosäuren und Vitamine • Symptomatische Behandlung der zahlreichen Komplikationen des ARG1-D²⁾ 	

¹⁾ Dosierungsstufe 1: 0,05 mg/kg; 2: 0,1 mg/kg; 3: 0,15 mg/kg; 4: 0,2 mg/kg

²⁾ Schließt auch eine medikamentöse Therapie mit Stickstofffängern ein. Diese mussten mindestens 4 Wochen vor der Randomisierung stabil eingestellt sein und über die RCT-Phase und 8 Wochen der verblindeten LTE-Phase stabil gehalten werden.

Abkürzungen: ARG1-D: Arginase-1-Mangel; LTE: Long Term Extension; IDM: Individuelles Disease Management

2.3 Endpunkte

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

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie PEACE

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Todesfälle ¹⁾	Mortalität	Ja	Ja
Krankheitsspezifische Laborparameter ³⁾ <ul style="list-style-type: none"> • Arginin-Konzentration im Plasma • Guanidinverbindungen • Ornithin • Ammoniak im Plasma 	Morbidität	Ja	Nein ²⁾
2MWT		Ja	Ja
GMFM <ul style="list-style-type: none"> • Dimension D • Dimension E 		Ja Ja	Nein ⁴⁾ Ja
Kombinationsendpunkt aus 2MWT und GMFM		Ja	Nein ⁵⁾
FMS		Ja	Nein ⁶⁾
GFAQ		Ja	Ja
VABS-II		Ja	Nein
CaCI-S und CGI-S		Ja	Nein ⁷⁾
CaGI-C und CGI-C		Ja	Nein ⁶⁾
9HPT		Ja	Nein ⁸⁾
BSID-III		Nein	Nein ⁹⁾
WIS (FSIQ)		Ja	Nein ¹⁰⁾
PedsQL		Ja	Ja ¹¹⁾
SF-36		Nein	Nein ¹²⁾
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

¹⁾ Im Rahmen der Sicherheit erhoben.

²⁾ Die Arginin-Konzentration wird aufgrund fehlender Patientenrelevanz nicht in der Nutzenbewertung berücksichtigt. Der primäre Endpunkt wird aus Transparenzgründen im Anhang dargestellt.

³⁾ Aminosäuren wurden vom pU im Dossier nicht dargestellt und werden nicht in der Nutzenbewertung berücksichtigt.

⁴⁾ Der GMFM Dimension D wird aufgrund von Unklarheiten zu den Auswertungen im Anhang dargestellt.

⁵⁾ Kombination aus 2MWT und GMFM. Die verwendeten Schwellenwerte liegen für alle Teilendpunkte < 15 %. Darüber hinaus werden für die Analyse nicht alle Patientinnen und Patienten der FAS-Population berücksichtigt.

⁶⁾ Endpunkt wird im Anhang der Nutzenbewertung dargestellt (siehe Kapitel 2.3.2.)

⁷⁾ CaGI-S wird in der Nutzenbewertung nicht dargestellt, da wesentliche Unsicherheiten zur Validität bestehen. Zusätzlich liegen keine adäquaten Analysen mit Rücklaufquoten > 70 % vor.

⁸⁾ Der Endpunkt 9HPT wird in der Nutzenbewertung nicht berücksichtigt, da die Rücklaufquoten zwischen den Studienarmen sich um mehr als 15 % unterscheiden.

⁹⁾ Die BSID-III sollte in der PEACE-Studie lediglich für die Altersgruppe 2–3,5 Jahre eingesetzt werden und es wurde vom ärztlichen Prüfpersonal entschieden, ob diese oder die WIS (WPPSI-IV) angewendet werden. Nur bei 1 Person wurde die BSID-III eingesetzt. Aufgrund der geringen Rückläufe wird der Endpunkt nicht in der Nutzenbewertung berücksichtigt.

¹⁰⁾ Die Rücklaufquoten liegen bereits zu Baseline im Pegzilarginasearm bei 71,4 % und im Placeboarm bei 54,5 %. Diesbezügliche Gründe sind nicht berichtet. Aufgrund der geringen Rückläufe wird der Endpunkt nicht in der Nutzenbewertung berücksichtigt.

¹¹⁾ Für die Nutzenbewertung liegen zum PedsQL keine validen Auswertungen vor, weshalb die Ergebnisse nicht dargestellt werden.

¹²⁾ Der SF-36 wurde in der PEACE-Studie für die Erhebung der Lebensqualität für Patientinnen und Patienten > 19 Jahre eingesetzt. Es befanden sich insgesamt 2 Personen im Pegzilarginasearm und 1 Person im Placeboarm, die älter als 19 Jahre waren. Aufgrund der geringen Rückläufe wird der Endpunkt nicht in der Nutzenbewertung berücksichtigt.

Abkürzungen: 2MWT: 2-Minuten-Gehtest; 9HPT: 9-Hole Pegboard Test; BSID-III: Bayley Scales of Infant and Toddler Development, Third Edition; CaGI-C/CGI-C: Caregiver/Clinician Global Impressions of Change; CaGI-S/CGI-S: Caregiver/Clinician Global Impressions of Severity; FAS: Full Analysis Set; FMS: Functional Mobility Scale; FSIQ: Full Scale Intelligence Quotient; GFAQ: Gillette Functional Assessment Questionnaire; GMFM: Gross Motor Function Measure-88; PedsQL: Pediatric Quality of Life Inventory; pU: pharmazeutischer Unternehmer; SF-36: 36-Item Short Form Health Survey; VABS-II: Vineland Adaptive Behavior Scale, Second Edition; WIS: Wechsler Intelligence Scales; WPPSI-IV: Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition

2.3.1 Mortalität

Der Endpunkt Mortalität wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Alle Todesfälle wurden innerhalb der Sicherheitserfassung erhoben.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

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

Validität

Der Endpunkt wird als valide bewertet.

2.3.2 Morbidität

Krankheitsspezifische Laborwerte

Arginin

Bei der Arginin-Konzentration im Plasma handelt es sich um den primären Endpunkt. Dieser wird als nicht unmittelbar patientenrelevant eingeschätzt, dient jedoch der Therapiesteuerung. Der Endpunkt wird ergänzend im Anhang dargestellt.

Operationalisierung

Beschreibung

Die Arginin-Konzentration im Plasma wird jeweils 168 Stunden nach der letzten und unmittelbar vor der nächsten Gabe der Studienmedikation erhoben. Die Bestimmung des Plasma-Arginins erfolgte standardisiert in einem Zentrallabor. Für die optimale Dosisfindung werden ab Protokollamendment 3 vom 25.07.2019 die Arginin-Daten durch das integrierte Web-/Sprachdialogsystem (IXRS) ausgewertet und dem/der unverblindeten Pharmazeuten/Pharmazeutin und/oder Arzt/Ärztin die entsprechenden Dosisanpassungen für die Infusionen übermittelt, anstatt der Rohwerte des Plasma-Arginins. Um die Verblindung der Studie aufrechtzuerhalten, werden dem ärztlichen Prüfpersonal die Ergebnisse der Arginin-, Ornithin- oder Guanidinverbindungen aus der doppelblinden oder verblindeten Nachbeobachtungsphase erst nach Abschluss der Studie mitgeteilt. Der Baselinewert ist der Mittelwert (MW) aller gemessenen Werte während der Screening-Phase. Bei Durchlaufen

mehrerer Screening-Phasen wurde nur der MW der finalen Screening-Phase berücksichtigt. Der finale Arginin-Wert ist der MW der letzten 4 Messwerte der RCT-Phase, wenn folgende Kriterien für die Messungen gelten:

- Die Messung wurde nach dem geplanten Termin für die 20. Dosis der Studienmedikation durchgeführt.
- Die vorangegangenen 2 Dosen der Studienmedikation wurden wie geplant verabreicht, d. h. 7 bzw. 14 Tage vor der Messung.

Sofern nur mindestens 1 Messwert, aber weniger als 4 Messwerte diese Kriterien erfüllten, gingen nur diese Werte bzw. dieser Wert in die Berechnung des MW ein. Wenn keiner der letzten 4 Messwerte die Kriterien erfüllte, wurde der letzte nach Baseline erhobene Messwert verwendet. Falls keine Post-Baseline-Werte vorlagen, wurde die Veränderung zu Baseline mit 0 ersetzt.

Ob das Therapieziel von ($< 200 \mu\text{mol/L}$) noch der aktuellen Behandlungsstrategie entspricht, ist unklar, da aktuelle und gültige Leitlinien fehlen.

Bewertung

Die Operationalisierung ist mit Einschränkungen nachvollziehbar. Die angelegten Kriterien des finalen Arginin-Wertes sind bei Fehlen aller 4 Werte hinfällig. In diesem Fall können Werte in die Analyse eingehen, auch wenn es sich nicht um Werte nach Woche 20 handelt oder ggf. die vorangegangenen 2 Dosen der Studienmedikation nicht wie geplant verabreicht wurden oder auch der Baselinewert selbst. Es ist unklar, wie viele Personen in den jeweiligen Armen davon betroffen sind, wodurch eine abschließende Beurteilung der Relevanz nicht erfolgen kann.

Patientenrelevanz

Als Laborparameter wird die Arginin-Konzentration als nicht unmittelbar patientenrelevant bewertet. Die Arginin-Konzentration im Plasma kann jedoch der Therapiesteuerung dienen [18]. Studien zur Surrogat-Validierung konnten nicht identifiziert werden.

Validität

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

Ornithin, Guanidino-Verbindungen, Aminosäuren und Ammoniak

Der Endpunkt Ornithin, Guanidino-Verbindungen, Aminosäuren und Ammoniak wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Patientenrelevanz.

Operationalisierung

Beschreibung

Die Bestimmung erfolgte standardisiert in einem Zentrallabor. Im Modul 4 werden folgende Untersuchungen unter den GC-Verbindungen angeführt: 5-Guanidino-2-Hydroxy-Pentansäure (Argininic Acid, ArgA), Guanidinoessigsäure (Guanidinoacetic Acid, GAA), Alpha-Keto-Delta-Guanidino-Valeriansäure (Alpha-Keto-Delta-Guanidinovaleric Acid, GVA) und N-Acetyl-L-Arginin (NAA). In den Studienunterlagen finden sich zusätzlich noch Untersuchungen zu folgenden GC-Verbindungen: Kreatin, Homoarginin oder andere. Um die Verblindung der Studie aufrechtzuerhalten, werden den Prüfarzten/Prüfärztinnen die Ergebnisse der Arginin-, Ornithin- oder Guanidinverbindungen aus der doppelblinden oder verblindeten Nachbeobachtungsphase erst nach Abschluss der Studie mitgeteilt. Die Probennahme für Aminosäuren erfolgt nach einem etwa 4-stündigen Fasten von Nahrung und essentiellen Aminosäuren, es sei denn, dies ist operativ und/oder klinisch unpraktisch. Die Ergebnisse der Aminosäuren (außer Arginin und Ornithin) werden den Prüfarzten/Prüfärztinnen zur Verfügung gestellt. Das Aminosäure-Panel umfasst: Taurine, Aspartic Acid, Hydroxyproline,

Threonine, Serine, Asparagine, Glutamic Acid, Glutamine, Proline, Glycine, Alanine, Citrulline, Valine, Methionine, Cystine, Isoleucine, Leucine, Tyrosine, Phenylalanine, a-Amino-n-butyric acid, Tryptophan, Lysine und Histidine.

Ammoniak wird durch das lokale Labor erhoben, da die Proben instabil sind. Die einrichtungsspezifischen Verfahren zur Gewinnung und Verarbeitung von Proben für Ammoniak-Tests müssen sorgfältig befolgt werden, um ein genaues Ergebnis zu gewährleisten. Bei erhöhten Werten wird ein Wiederholungstest in Betracht gezogen. Die Abnahme erfolgt ebenfalls nach einer 4-stündigen Nahrungskarenz.

Ornithin ist neben Harnstoff der zweite Metabolit, der aus der Katalyse der ARG1-D hervorgeht.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Es ist nicht weiter beschrieben, welche GC-Verbindungen unter den Begriff „Andere“ fallen.

Patientenrelevanz

Bei den vorliegenden Endpunkten handelt es sich um Laborparameter, die nicht als unmittelbar patientenrelevant eingeschätzt werden. Studien zur Surrogatvalidierung legt der pU nicht vor.

Validität

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

Mobilitätsendpunkte

2 Minutes Walking Test (2MWT)

Der Endpunkt 2MWT wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Zur Ermittlung der Mobilität (Ausdauer) wurde in der vorliegenden PEACE-Studie der 2-Minuten-Gehtest (2MWT) erhoben. Hierbei handelt es sich um eine kürzere Modifikation der 6-Minuten-Version (NIH Toolbox motor domain - 2MWT [10]). Der Test wurde auf Grundlage der Richtlinie der ATS durchgeführt (American Thoracic Society 2002) [1]. Die Ergebnisse des 2MWT werden als zurückgelegte Strecke in Metern erhoben. Die Probandinnen und Probanden werden angewiesen, Schuhe zu tragen, die den gesamten Fuß sicher bedecken (z. B. Turnschuhe usw.), und keine Sandalen oder Hausschuhe usw. Wenn eine Gehhilfe zum Laufen benötigt wurde, konnte diese laut Angaben des pU während des Tests benutzt werden (Stock, Gehhilfe usw.), wie auch in der ATS Guideline beschrieben ist [1]. Der 2MWT-Test sollte gemäß Studienprotokoll erst ab einem Alter von 3 Jahren eingesetzt werden.

Bewertung

Die Operationalisierung ist nur eingeschränkt nachvollziehbar.

Die der Durchführung dienliche Guideline der ATS bezieht sich auf den 6MWT bei Patientinnen und Patienten mit mittelschweren bis schweren Herz- oder Lungenerkrankungen. Es findet sich darin keine Information zum Umgang oder zur Testdurchführung bei Kindern bzw. Personen jeglichen Alters in der vorliegenden Indikation. Es wird vorliegend jedoch von der Sicherstellung einer Studienzentren-übergreifenden standardisierten Durchführung ausgegangen. Die Nutzung einer Gehhilfe bei der Durchführung des 2MWT wurde erlaubt. Anhand des 2MWT können jedoch Gruppen mit ggü. ohne Gehhilfe unterschieden werden (siehe unten zur Known-groups-Validität). Kritisch wird insbesondere gesehen, dass ein

Umgang mit einem möglichen Wechsel einer Gehhilfe (z. B. von Stock zu Rollstuhl oder umgekehrt) innerhalb der Studie bei Erhebung des 2MWT nicht beschrieben ist. Dies könnte ggf. zu Abweichungen der Messwerte führen, die nicht bzw. nicht ausschließlich auf eine Veränderung der Morbidität der Patientin / des Patienten zurückzuführen sind. Es ist anzunehmen, dass u. a. das Alter, die Größe und die damit verbundene Schrittlänge eines Kindes auch Einfluss auf die zurücklegbare Distanz in 2 Minuten hat. Es kann außerdem nicht ausgeschlossen werden, dass – neben Alter und Größe – das entwicklungsbedingte Erlernen vom Laufen für einen Teil der Patientinnen und Patienten zu einer Verbesserung der Gehstrecke unabhängig von der Studienmedikation führen kann. Eine Standardisierung (bspw. hinsichtlich des Alters) anhand einer geeigneten Referenzpopulation war in der vorgelegten PEACE-Studie nicht geplant. Da der 2MWT erst ab 3 Jahren durchgeführt werden sollte und der Einschluss von Personen ab 2 Jahren möglich war, ist eine Erhebung der gesamten Studienpopulation ggf. nicht abgedeckt.

Patientenrelevanz

Die körperliche Belastbarkeit bzw. Ausdauerleistungsfähigkeit werden als patientenrelevanter Morbiditätsendpunkt angesehen. Spastische Lähmungen, welche ein relevantes Symptom im vorliegenden Anwendungsgebiet darstellen, können zu Bewegungseinschränkungen führen. Inwiefern diese nur das Gangbild oder auch die zurücklegbare Gehstrecke in Metern beeinflussen, ist vorliegend unklar.

Validität

Der Endpunkt 2MWT wurde anhand von gesunden Patientinnen und Patienten validiert [37]. Die Auswahl des 2-Minuten-Gehtests (2MWT) zur Messung der Ausdauer im motorischen Bereich der NIH-Toolbox basierte auf der einfachen Durchführung und Aufgabenerfüllung, der hohen Test-Retest-Zuverlässigkeit (ICC: 0,89 (0,82–0,93)) und der hohen Korrelation mit dem 6-Minuten-Gehtest (6MWT) ($r = 0,96$) unter Verwendung von Daten gesunder Personen im Alter von 3–85 Jahren ($n = 328$). In der Literatur werden Normwerte für jedes Jahr in der Altersgruppe 3–17 Jahre und für die Altersgruppen 18–29, 30–39, 40–49, 50–59, 60–69 und 70–85 Jahre, auf der Grundlage einer repräsentativen US-Population von 4.859 Teilnehmenden ermittelt, beschrieben, die vorliegend nicht berücksichtigt wurden [11,12]. Grundsätzlich werden Normwerte in diesem Anwendungsgebiet als relevant eingeschätzt, um altersbedingte Einflussfaktoren in der Auswertung zu berücksichtigen. Untersuchungen zur Validität und Übertragbarkeit auf Personen im vorliegenden Anwendungsgebiet konnten nicht identifiziert werden. Eine Studie zur Untersuchung der psychometrischen Güte bei Kindern (6–12 Jahre) mit neuromuskulären Erkrankungen (davon 63 % mit Zerebralparese) zeigte eine hohe Inter- und Intra-Rater-Reliabilität (ICC > 0,8). Es zeigte sich eine Known-Groups-Validität, sodass der Test zwischen gesunden Kindern und solchen mit neuromuskulären Erkrankungen und zusätzlich noch zwischen solchen mit Gehhilfen unterscheiden kann. Untersuchungen zur Änderungssensitivität und Korrelationen zu anderen Instrumenten waren nicht Teil dieser Untersuchung [36]. Grundsätzlich erscheint eine Übertragbarkeit der Validitätsbewertung zur Zerebralparese aufgrund ähnlicher Symptomatik möglich.

Gross Motor Function Measure (GMFM) Dimensionen D (Stehen) und E (Gehen, Laufen, Springen)

Der Endpunkt GMFM (Dimension D und E) wird in der Nutzenbewertung berücksichtigt. Die Dimension D wird jedoch aufgrund von Unklarheiten in den Auswertungen im Anhang dargestellt.

Operationalisierung

Beschreibung

Der Gross Motor Function Measure (GMFM) ist ein Beobachtungsinstrument zur Bestimmung der grobmotorischen Fähigkeiten von Kindern mit Zerebralparese [19].

Aufbau des Instruments

Der GMFM-88 umfasst insgesamt 88 Testaufgaben, die sich an einem gesunden, motorisch normal entwickelten 5-jährigen Kind orientieren und von diesem gewöhnlich problemlos ausgeführt werden können. Die Aufgaben sind, entsprechend der kindlichen motorischen Entwicklungsschritte, in 5 Dimensionen unterteilt:

- Dimension A: Liegen und Drehen (17 Aufgaben)
- Dimension B: Sitzen (20 Aufgaben)
- Dimension C: Krabbeln und Knien (14 Aufgaben)
- Dimension D: Stehen (13 Aufgaben)
- Dimension E: Gehen, Laufen und Springen (24 Aufgaben)

Im Vorfeld können individuelle Zieldimensionen (A–E) ausgewählt werden, bei denen die größten Veränderungen erwartet werden. In der Studie PEACE wurden zu Studienbeginn alle Dimensionen des GMFM eingesetzt. Ab Protokollversion 3 vom 25.07.2019 wurde die Anwendung auf die Dimensionen D und E beschränkt. Gemäß pU erfolgte dies aufgrund von Deckeneffekten (Aufgaben waren alle erfüllbar) (in Studie 101A) in den Dimensionen A, B und C und da diese Dimensionen aufgrund der spastischen Lähmungen der unteren Extremitäten im vorliegenden Anwendungsgebiet besonders relevant sind [13].

Durchführung und Auswertung des Tests

Das Testhandbuch definiert allgemeine Richtlinien zur standardisierten Durchführung des Tests. Für jedes Item sind maximal 3 Versuche zulässig, wovon der beste Versuch aufgezeichnet wird. Verbale Ermunterungen, der Einsatz von Spielsachen zum Anreiz, die Demonstration der Aufgaben und Hilfestellung beim Einnehmen der Ausgangsstellung, nicht aber bei der Durchführung, sind erlaubt. Es wird empfohlen, die Testung zunächst ohne Hilfsmittel und Orthesen durchzuführen. Der pU gibt an, dass die Durchführung zunächst ohne Hilfsmittel zur Fortbewegung erfolgte. Laut Testhandbuch [38] sollten die Testungen ausschließlich von in der Anwendung der GMFM trainierten Personen durchgeführt werden.

Die Bewertung jeder einzelnen Testaufgabe erfolgt auf Grundlage von Beobachtungen mit einer 4-Punkte-Ordinalskala (0–3 Punkte):

- 0 = initiiert nicht
- 1 = initiiert (Durchführung von < 10 %)
- 2 = führt Aufgabe teilweise aus (Durchführung von 10 bis 100 %)
- 3 = führt gestellte Aufgabe aus (Durchführung 100 %)

Für den Großteil der Aufgaben sind genaue Bewertungsrichtlinien formuliert, die für die Bewertung herangezogen werden müssen.

Die dimensionsspezifische Auswertung erfolgt gemäß Testhandbuch in Prozent (0–100 %) für den Anteil der erreichten Punktzahl einer Person an der maximal erreichbaren Punktzahl. Ein höherer Wert entspricht dabei stärker ausgeprägten grobmotorischen Funktionen.

Aus den Angaben des pU lässt sich folgern, dass dieser die Scores der Dimension D (0–39) und E (0–72) als Rohwerte verwendet hat. Des Weiteren hat der pU präspezifiziert die Patientinnen und Patienten je nach motorischen Fähigkeiten mittels Gross Motor Function Classification System (GMFCS I–V) zu Baseline klassifiziert, ein Instrument für die Klassifikation der motorischen Fähigkeiten von Personen mit Zerebralparese. Dieses System wurde speziell für Kinder mit Zerebralparese entwickelt und validiert [30,31,32,45]. Es umfasst 5 Klassen,

wobei höhere Klassen einer schlechteren motorischen Fähigkeit entsprechen. Zu den GMFCS I–III wurden unterschiedliche MID zugeordnet [29]. Keine dieser MID lag bei $\geq 15\%$. Personen mit einem GMFCS ≥ 4 wurden von der Analyse ausgeschlossen, da keine MID identifiziert werden konnte.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Es fehlen Angaben, ob die Tests gemäß der im Testmanual definierten Richtlinien standardisiert durchgeführt wurden. Das Manual selbst wurde vom pU nicht eingereicht. Lediglich, dass zunächst keine Hilfsmittel eingesetzt werden sollten, ist erwähnt. Ein weiterer Umgang damit und der Einfluss auf die Bewertung ist jedoch nicht weiter beschrieben. Ob die Testung, wie im Manual beschrieben, ausschließlich von in der Anwendung der GMFM trainierten Personen durchgeführt wurde, ist ebenfalls nicht beschrieben. Auch kann bspw. nicht ausgeschlossen werden, dass das entwicklungsbedingte Erlernen vom Laufen für einen Teil der Patientinnen und Patienten zu einer Verbesserung unabhängig von der Studienmedikation führen kann. Ab wann ein Fragebogen als unvollständig angesehen wurde, ist nicht beschrieben.

Patientenrelevanz

Patienten und Patientinnen mit ARG1-D leiden meist unter einer spastischen Lähmung in der Muskulatur, welche alle Extremitäten (spastische Tetraparese) oder beide Beine betreffen kann (spastische Paraparese). Eine Verbesserung der Grobmotorik stellt daher einen patientenrelevanten Endpunkt dar. Die vom GMFM erfassten Aspekte der motorischen Entwicklung werden als patientenrelevant bewertet.

Validität

Der GMFM wurde 1989 anhand von Kindern mit Zerebralparese im Alter von 5 Monaten bis 16 Jahren entwickelt und validiert [19,40]. Die Entwicklung des Instruments orientierte sich an den motorischen Fähigkeiten eines motorisch normal entwickelten 5-jährigen Kindes [40]. Die Items der Dimensionen D und E sind zufolge der Item-Map von Russel schwieriger als die der anderen Dimensionen [39]. Für Kinder mit weniger gravierenden motorischen Einschränkungen (die z. B. selbstständig gehen können) erscheint das alleinige Testen der Dimensionen D und E plausibel. Auch in anderen Interventionsstudien wurden nur die Dimensionen (C,) D und E erhoben und ausgewertet [24,26,29,33]. Allerdings umfasst das erwartete Anwendungsgebiet Kinder ab 2 Jahren. Es ist daher fraglich, ob die Auswahl der Dimensionen D und E sinnvoll ist, da diese komplexere Aufgaben beinhalten und sich das Instrument an den motorischen Fähigkeiten eines motorisch normal entwickelten 5-jährigen Kindes orientiert [40]. So ist z. B. unklar, wie ein Kind, welches entwicklungs-technisch noch nicht auf einem Bein stehen oder hüpfen kann, in die Bewertung einbezogen wurde. Dies scheint auch vor dem Hintergrund relevant, da laut Modul 3, die Kinder erst verhältnismäßig spät ihre ersten Schritte erlernen [13]. Es ist außerdem unklar, ob sich das Instrument auch für Erwachsene eignet, da an den Validierungsstudien nur Kinder teilgenommen haben.

Insgesamt zeigten sich für die Dimensionen D und E des GMFM über alle untersuchten Gütekriterien (interne Konsistenz, Konstruktvalidität, Test-Retest-Reliabilität) gute bis sehr gute Werte, insbesondere für Kinder mit Zerebralparese. Es ist jedoch darauf hinzuweisen, dass keine Studien zur Validierung der GMFM in der Indikation Hyperargininämie vorgelegt wurden. Die Übertragbarkeit der GMFM-Testergebnisse von Studien mit Kindern mit Zerebralparese auf die Indikation Hyperargininämie erscheint aufgrund der ähnlichen Symptomatik grundsätzlich plausibel, wird jedoch im Dossier nicht weiter diskutiert.

In verschiedenen Anwendungsstudien zeigten sich beim Gesamtscore Boden- und Deckeneffekte. Eine Beurteilung von sehr leicht oder sehr schwer betroffenen Kindern ist damit aufgrund der geringen Differenzierungsfähigkeit eingeschränkt [7]. Vos-Vromans

konnte in einer Studie mit 55 Kindern mit Zerebralparese jedoch zeigen, dass die Deckeneffekte in diesem Anwendungsgebiet insbesondere bei den Dimensionen A bis C vorliegen, der Score der Dimension D lag zwischen 18 und 97 (Mittelwert: 82, SD: 17,9) und bei der Dimension E zwischen 17 und 99 (Mittelwert: 70,5, SD: 21,3) [43].

Functional Mobility Scale (FMS)

Der Endpunkt FMS wird in der Nutzenbewertung im Anhang berichtet aufgrund von Unklarheiten hinsichtlich Patientenrelevanz und Validität. Die Abhängigkeit von Hilfsmitteln wird außerdem bereits über den CaGI-C erhoben.

Operationalisierung

Beschreibung

Die funktionelle Mobilität sollte in der PEACE-Studie mittels der Functional Mobility Scale (FMS) ermittelt werden. Die FMS erfasst jeweils auf einer 6-stufigen Ordinalskala die hinsichtlich der alltäglichen Fortbewegung genutzten Hilfsmittel bzw. Hilfestellungen in 3 unterschiedlichen Umgebungen: 1) zu Hause über kurze Distanzen (5 m), 2) zwischen Unterrichtsstunden in der Schule (50 m) und 3) über lange Distanzen in der Öffentlichkeit, wie zum Beispiel auf dem Schulweg oder bei einem Spaziergang (500 m).

Die angegebenen Distanzen in Klammern sind Richtwerte und der Fokus der Fragen liegt auf der jeweiligen Umgebung. Ein niedriger Wert auf der Skala entspricht einer schlechteren funktionellen Mobilitätsleistung (1 = „Benutzt Rollstuhl, Kinderwagen, Roller, Einkaufswagen, Wagen oder wird getragen ODER geht nur mit hochspezialisierter/unterstützender Gehhilfe ODER macht begrenzte Schritte mit erheblicher Unterstützung durch eine andere Person“ bis 6 = „Selbstständiges Gehen und Laufen auf allen Untergründen ohne Hilfsmittel oder Hilfe einer anderen Person“). Die FMS wurde in der PEACE-Studie mittels Fremdeinschätzung durch die Eltern/Betreuungsperson erhoben und hinsichtlich jeder Umgebung für sich ausgewertet.

Bewertung

Die Operationalisierung ist nur eingeschränkt nachvollziehbar. Das Instrument wurde vom pU nicht vorgelegt, wodurch auch die einzelnen Skalen nicht in Gänze nachvollzogen werden können. Auch kann bspw. nicht ausgeschlossen werden, dass das entwicklungsbedingte Erlernen vom Laufen für einen Teil der Patientinnen und Patienten zu einer Verbesserung unabhängig von der Studienmedikation führen kann.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung nicht als unmittelbar patientenrelevant eingeschätzt. Patienten und Patientinnen mit ARG1-D leiden meist unter einer spastischen Lähmung in der Muskulatur, welche alle Extremitäten (spastische Tetraparese) oder beide Beine betreffen kann (spastische Paraparese). Patientinnen und Patienten können aufgrund möglicher Bewegungseinschränkungen auf Hilfsmittel zur Fortbewegung angewiesen sein. Die krankheitsspezifische Angewiesenheit auf Hilfsmittel zur Fortbewegung kann in der vorliegenden Indikation als patientenrelevant angesehen werden. Gemäß Fragebogeninstruktionen erfasst der FMS jedoch die generelle Alltagsleistung hinsichtlich der Fortbewegungsmittel bzw. Hilfestellungen – ohne Einbezug anderer direkter Morbiditäts- und Lebensqualitätsparameter. Die Patientenrelevanz dieser Alltagsleistung wird aufgrund weiterer potentieller Einflussgrößen (siehe Validität) bei fehlender Originalfragebogenversion als unklar bewertet.

Die Erhebung erfolgte durch die Fremdbeurteilung der Eltern/Betreuungsperson. Diese wird jedoch aufgrund des jungen Alters der Patientinnen und Patienten sowie möglicher kognitiver Einschränkungen in der PEACE-Studie als adäquat erachtet.

Validität

In einer Validierungsstudie, welche 310 Patientinnen und Patienten zwischen 4–18 Jahren mit Zerebralparese untersuchte, weist das Instrument sowohl eine sehr gute Intrarater-Reliabilität (ICC: 0,88–0,99) als auch interne Konsistenz (Cronbach's α : 0,86–0,99) auf [17]. Die Änderungssensitivität konnte mittels einer stetigen Verbesserung des Scores zu jeweils 3, 6, 9 und 12 Monaten nach einem orthopädischen Eingriff, von 35 Patientinnen und Patienten mit spastischer Diplegie nachgewiesen werden. Die in der PEACE-Studie definierte Verbesserung um eine Skalenstufe, (MID) deckt sich mit den klinisch bedeutsamen Veränderungen, welche an 64 Patientinnen und Patienten mit motorischen Störungen untersucht wurden [8]. Es liegen keine Validierungsstudien für das vorliegende Anwendungsgebiet vor. Der pU begründet dies mit der Seltenheit des ARG1-D. Eine Übertragbarkeit scheint aufgrund ähnlicher Symptomatik der Zerebralparese möglich.

Der Fragebogen wurde für die Fremdeinschätzung von Kindern ab 4 Jahren entwickelt. Aus den Ausführungen des pU wird jedoch nicht ersichtlich, inwiefern Altersunterschiede der Patientinnen und Patienten bei den Antworten berücksichtigt werden. So könnten intraindividuelle Entwicklungsunterschiede bzw. -fortschritte von Probandinnen und Probanden das alltägliche Nutzungsverhalten von Hilfsmitteln (bspw. „Kinderwagen“) bzw. die Zuhilfenahme von Hilfestellungen (bspw. „macht begrenzte Schritte mit erheblicher Unterstützung durch eine andere Person“) während der Studie beeinflussen. Da ein allgemeiner Wechsel des alltäglichen Nutzungsverhaltens unabhängig von der vorliegenden Erkrankung erfolgen kann, z. B. nach einem Sturz, nach Präferenz oder auch in Abhängigkeit des Selbstvertrauens, der Verfügbarkeit von Gehhilfen und dem Vorliegen anderer Erkrankungen bei Studienteilnehmenden, wird die Validität der Erfassung der funktionellen Mobilität anhand des FMS insgesamt als eingeschränkt eingeschätzt. Die Abhängigkeit von Hilfsmitteln wird außerdem bereits über den CaGI-C erhoben.

Gillette Functional Assessment Questionnaire (GFAQ)

Der Endpunkt GFAQ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Der Gillette Functional Assessment Questionnaire dient der Erfassung der selbstständigen, maximalen Funktionen unter Verwendung von Hilfsmitteln und Orthesen. Dieser setzt sich aus einer „Walking Scale“ sowie aus einem Fragenkatalog bestehend aus 22 Items zusammen. In der PEACE-Studie wird lediglich die 10-stufige Walking Scale erhoben, welche die übliche Fähigkeit zur Fortbewegung mit Hilfe der normalerweise verwendeten Hilfsmittel der Patienten und Patientinnen erfasst. Ein niedrigerer Wert auf der Skala entspricht einer schlechteren Gehfähigkeit (1 = „Macht gar keine Schritte“ bis 10 = „Geht, rennt und klettert auf ebenem und unebenem Gelände und bewältigt Treppen ohne Schwierigkeiten und ohne Hilfestellung. Ist normalerweise fähig, mit Gleichaltrigen Schritt zu halten“). Erhoben wurde die Skala durch die Befragung der Erziehungsberechtigten. Als Responder wurden Personen gewertet, die eine Verbesserung ≥ 2 Skalenstufen von Baseline zu Woche 24 hatten, was $> 15\%$ der Skalenspannweite entspricht.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Es werden keine genauen Angaben bezüglich eines Bezugszeitraums gegeben, in welcher die übliche Fähigkeit zu Fortbewegung erfasst wird.

Patientenrelevanz

Patienten und Patientinnen mit ARG1-D leiden meist unter einer spastischen Lähmung in der Muskulatur, welche alle Extremitäten (spastische Tetraparese) oder beide Beine betreffen kann (spastische Paraparese). Eine Verbesserung der funktionellen Mobilität stellt daher einen patientenrelevanten Endpunkt dar. Die Erhebung erfolgte durch die Fremdbeurteilung der Eltern, diese wird jedoch aufgrund des jungen Alters der Patientinnen und Patienten sowie möglicher kognitiver Einschränkungen als adäquat erachtet.

Validität

Eine separate Erhebung der GFAQ-Walking-Scale ist zulässig. Es kann jedoch nicht ausgeschlossen werden, dass das entwicklungsbedingte Erlernen vom Laufen für einen Teil der Patientinnen und Patienten zu einer Verbesserung unabhängig von der Studienmedikation führen kann. Die Validität der GFAQ-Walking-Scale wurde in einer Studie mit 41 Patienten und Patientinnen zwischen 2–17 Jahren mit verschiedenen neuromotorischen Erkrankungen untersucht [28]. Das Instrument weist sowohl eine sehr gute Test-Retest-Reliabilität (ICC = 0,92) als auch Intrarater-Reliabilität (ICC = 0,92) auf. Zudem korrelierte die GFAQ-Walking-Scale signifikant mit anderen Messinstrumenten zur Erfassung der Gehfähigkeit (POSNA, NI, O2, WeeFIM). Der GFAQ wurde lediglich in einer Population validiert, die zu Baseline Werte von 6–10 aufwies. Die Validität des GFAQ-Scores für Personen < 6 ist unklar. Ein GFAQ-Score zwischen 2 und 9 war ein Einschlusskriterium für die PEACE-Studie. Angaben zum Umfang der Personen mit einem GFAQ-Score < 6 zu Studienbeginn konnten nicht identifiziert werden. Laut einer weiteren Untersuchung hinsichtlich klinisch bedeutsamer Veränderungen wird eine Veränderung um mindestens 2 Skalenpunkte in der GFAQ-Walking-Scale als klinisch relevant gewertet [8]. Dies entspricht dem für die PEACE-Studie definierten MID. Es liegen keine Validierungsstudien hinsichtlich Patientinnen und Patienten mit ARG1-D vor. Eine Übertragbarkeit scheint aufgrund ähnlicher Symptomatik möglich.

Adaptives Verhalten

Vineland Adaptive Behavior Scales, Second Edition (VABS-II)

Der Endpunkt VABS-II wird in der Nutzenbewertung nicht berücksichtigt aufgrund von Unklarheiten in der Operationalisierung.

Operationalisierung

Beschreibung

Das Instrument VABS-II dient der Fremdbewertung des adaptiven Verhaltens, d. h. der Fähigkeit, die Herausforderungen des täglichen Lebens zu bewältigen, bis zu einem Alter von 90 Jahren. Ziel der Vineland-II ist es, die persönlichen und sozialen Fähigkeiten von Kindern oder Erwachsenen in alltäglichen Situationen zu erfassen, im Besonderen bei Personen, die kognitive Einschränkungen aufweisen oder Schwierigkeiten haben, in Testsituationen zu bestehen. Die Beurteilung des adaptiven Verhaltens erfolgt durch die Betreuungsperson. Es gibt verschiedene Versionen des Instruments. Die vorliegende Betreuer-Version umfasst 5 Domänen, die wiederum 2–3 Subdomänen (insgesamt 11 Subdomänen) enthalten. Alle Items sind zur Beantwortung an spezifische Altersspannen geknüpft, weshalb nur bestimmte Items für eine Patientin / einen Patienten beantwortet werden müssen.

Adaptive Verhaltensdomänen:

- Kommunikation (99 Items)
 - Rezeptiv (≥ 0 Jahre)
 - Expressiv (≥ 0 Jahre)
 - Geschriebene Sprache (≥ 3 Jahre)

- Fähigkeiten des alltäglichen Lebens (109 Items)
 - Persönlich (≥ 0 Jahre)
 - Häuslich (≥ 1 Jahre)
 - Gemeinschaft (≥ 1 Jahre)
- Soziale Fähigkeiten (99 Items)
 - Zwischenmenschliche Beziehungen (≥ 0 Jahre)
 - Spielen und Freizeit (≥ 0 Jahre)
 - Bewältigungsstrategien (≥ 1 Jahre)
- Motorische Fähigkeiten (76 Items)
 - Grobmotorik (≥ 0 bis < 7 Jahre)
 - Feinmotorik (≥ 0 bis < 7 Jahre)

Maladaptives Verhalten (optional)

- Index für maladaptives Verhalten (36 Items)
 - Teil A (≥ 3 Jahre)
 - Teil B (≥ 3 Jahre)
 - Teil C (≥ 3 Jahre)
- Maladaptive verhaltenskritische Elemente (Teil D, 14 Items, ≥ 3 Jahre)

Die Antwortoptionen entsprechen einer Häufigkeitsabfrage des jeweiligen Items und reichen von 2 (= üblicherweise), 1 (= manchmal), 0 (= niemals) bis DK (= weiß nicht). Bei einigen Items existiert als weitere Antwortoption N/O (= nicht möglich), sofern bislang keine entsprechende Situation aufgetreten ist. Beispielsweise die Arbeitsleistung, sofern die betroffene Person nicht arbeitstätig ist. Abgefragt wird dabei jeweils der Zustand zum aktuellen Zeitpunkt. Höhere Werte weisen auf ein besseres adaptives Verhalten hin.

Die Erhebung der optionalen Domäne „maladaptives Verhalten“ war laut SAP nicht vorgesehen und wurde für die PEACE-Studie nicht ausgewertet. Die Fragen der VABS-II wurden in der PEACE-Studie durch den Patienten / die Patientin selbst beantwortet oder, falls dieser/diese dazu nicht in der Lage war, durch die Eltern oder die Betreuungsperson.

Die Werte der einzelnen Subdomänen wurden für jede Domäne zum jeweiligen Domain Composite Score (DCS) addiert. Die DCS wurden zum Adaptive Behaviour Composite Score (ABCS) zusammengefasst, wobei die Domäne „motorische Fähigkeiten“ für Personen ≥ 7 Jahre nicht berücksichtigt wurde. ABCS und DCS haben laut Modul 4 eine Skalenspannweite von 120 Punkten. Als Responder wurden Personen gewertet, die eine Verbesserung $\geq 7,5$ Punkte von Baseline zu Woche 24 hatten. Diese MID entspricht nicht dem 15 % Responsekriterium und wird daher nicht berücksichtigt.

Bewertung

Die Operationalisierung ist unzureichend beschrieben. Der pU legt keine Version des VABS-II, welcher in der Studie PEACE angewendet wurde, mit den Studienunterlagen vor. Laut Protokoll wurde die Caregiver-reported-Version verwendet, laut Angaben des pU sollte diese jedoch, wenn möglich, primär von den Patientinnen und Patienten selbst ausgefüllt werden. In welchem Umfang das erfolgte und ob innerhalb der Studienlaufzeit zwischen Eigen- und Fremdbewertung gewechselt werden durfte, ist nicht beschrieben. Des Weiteren ist nicht beschrieben, ob die Fragen selbst ausgefüllt oder interviewbasiert erhoben werden. Normwertskalen, die das Alter der Patientinnen und Patienten einbeziehen, werden nicht berechnet.

Das Instrument besteht aus altersspezifischen Items. Es ist nicht beschrieben, wie damit umgegangen wurde, wenn Patientinnen und Patienten die Altersgrenze im Verlauf der 6-monatigen Studienphase überschreiten. Zudem ist in den Studienunterlagen weder das

Scoring der DCS ausreichend beschrieben noch finden sich Angaben zur konkreten Berechnung des im SAP erwähnten ABCS. Angegeben ist, dass die motorische Domäne lediglich für Patientinnen und Patienten < 7 Jahre berücksichtigt werden sollte. Ab welcher Anzahl an fehlenden Items eine Subdomäne, deren Domäne und das adaptive Verhalten nicht bewertet werden kann, ist ebenfalls nicht erläutert.

Patientenrelevanz

Der ARG1-Mangel ist mit kognitiven Einschränkungen assoziiert und geht mit einer Vielzahl von Verhaltensauffälligkeiten einher [13]. Die Veränderung des adaptiven Verhaltens wird als ein patientenrelevanter Endpunkt bewertet.

Validität

Der pU führt keine Quellen zur Validierung des Endpunkts oder zum Einsatz des Instruments in der vorliegenden Indikation an und begründet dies mit der fehlenden Durchführbarkeit durch die Seltenheit der Erkrankung. In mehreren Studien wurden die psychometrischen Gütekriterien bei unterschiedlichen Patientenpopulationen in jeglichen Altersgruppen untersucht. Es zeigten sich u. a. eine gute Reliabilität (Test-Retest- und Interrater-Reliabilität), interne Konsistenz, Inhalts-, Kriteriums- und Konstruktvalidität. Die Vineland-II wurde zudem bereits bei Patientinnen und Patienten mit Epilepsie im Kindesalter eingesetzt [9,14,27,34,35] und auch bei Kindern mit Zerebralparese [13]. Aufgrund der generischen Art des Fragebogens und der vielfältigen Validierung der Vineland-II in unterschiedlichen Patientenpopulationen, darunter auch Personen mit Entwicklungsstörungen, wird die Validität des Instruments als hinreichend übertragbar angesehen. Unklarheiten ergeben sich aufgrund der vorliegenden Operationalisierungseinschränkungen, insbesondere bezüglich der Anwendung und der Auswertung des Fragebogens.

Die Erhebung des Endpunkts kann durch eine Fremdeinschätzung des altersabhängigen, adäquaten, adaptiven Verhaltens erfolgen. Dies erscheint in der vorliegenden Indikation akzeptabel; zum einen aufgrund der Krankheitssymptomatik, zum anderen, weil davon auszugehen ist, dass die betreuende Person die Patientin / den Patienten täglich betreut und deren/dessen Funktionalität entsprechend gut einschätzen kann. Angaben über den Umfang der Fremdeinschätzung bzw. separate Auswertungen zu den Eigeneinschätzungen finden sich jedoch nicht.

Symptomatik

Caregiver Global Impressions of Severity (CaGI-S) und Clinician Global Impressions of Severity (CGI-S)

Der Endpunkt CaGI-S wird in der Nutzenbewertung nicht dargestellt, da wesentliche Unsicherheiten zur Validität bestehen. Zusätzlich liegen keine adäquaten Analysen mit Rücklaufquoten > 70 % vor.

Operationalisierung

Beschreibung

Der CaGI-S und der CGI-S sind Fragebögen, die vorliegend jeweils im Rahmen einer Fremdbewertung den Eindruck hinsichtlich der aktuellen (d. h. zu diesem Zeitpunkt) Defizite des/der Studienteilnehmenden in Bezug auf Mobilitätsaspekte, Alltagskompetenzen, soziale Fähigkeiten und adaptives Verhalten im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel erfassen. Der CaGI-S wird durch die Betreuungsperson des/der Studienteilnehmenden beantwortet, der CGI-S durch den den Prüfarzt / die Prüfarztin. Aus Konsistenzgründen wurden die Fragen, sofern möglich, während der gesamten Studie jeweils durch dieselben Personen beantwortet.

Im Rahmen der PEACE-Studie wurden die folgenden Fragen im Rahmen beider Fragebögen eingesetzt:

1. Unter Berücksichtigung der Qualität und des Selbstvertrauens des Studienteilnehmenden beim Gehen, wie ist Ihr Eindruck von der Gehfähigkeit des Studienteilnehmenden, im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?
2. Wie beurteilen Sie in Anbetracht der Abhängigkeit des Studienteilnehmenden von Hilfsmitteln (Gehhilfen) die Fähigkeit des Studienteilnehmenden, sich von einem Ort zum anderen zu bewegen, im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?

Die folgenden Fragen waren ausschließlich für den CaGI-S vorgesehen:

3.
 - a) Welchen Eindruck haben Sie von den grobmotorischen Fähigkeiten (Bewegungen des ganzen Körpers, Arme und Beine: z. B. einen Ball fangen, balancieren, klettern, Trampolin springen, Fangen spielen, Rennen laufen) im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?
 - b) Welchen Eindruck haben Sie von den feinmotorischen Fähigkeiten (Bewegungen der Hände und Finger: z. B. Malen mit Buntstiften, Schneiden mit der Schere, Spielen mit kleinen Gegenständen wie LEGO oder Perlen, mit Messer und Gabel essen, Schnürsenkel binden, einen Bleistift halten, zeichnen oder schreiben) im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?
 - c) Welchen Eindruck haben Sie von den Alltagskompetenzen (Aktivitäten für eine unabhängige Lebensführung: z. B. sich selbst ernähren, baden, auf die Toilette gehen, sich anziehen, sich pflegen) im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?
 - d) Welchen Eindruck haben Sie von der Kommunikationsfähigkeit (z. B. Anweisungen befolgen, Geschichten erzählen, Kommentare zu Ereignissen abgeben) im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?
 - e) Welchen Eindruck haben Sie von der Sozialisationsfähigkeit (z. B. Teilen von Spielzeug, Spielen mit Gleichaltrigen, Rücksichtnahme auf andere) im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?
 - f) Unter Berücksichtigung Ihrer Antworten auf die Fragen a–e: Was ist Ihr Eindruck zur Anpassungsfähigkeit des Studienteilnehmenden an seine/ihre Umwelt im Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel (z. B. andere Familienmitglieder, Freunde, Kinder, Schulkameraden)?

Die möglichen Antworten für jede Frage sind: sehr viel schlechter als andere, viel schlechter als andere, schlechter als andere, ein wenig schlechter als andere, kein deutlicher Unterschied zu anderen. Laut Modul 4 erfolgte die Bewertung auf einer 5-stufigen Skala, wobei ein niedriger Wert einer geringeren Symptomschwere entsprach (1 = „kein klarer Unterschied zu anderen“; 2 = „etwas schlechter als andere“; 3 = „schlechter als andere“; 4 = „viel schlechter als andere“; 5 = „sehr viel schlechter als andere“). Da präspezifiziert lediglich deskriptive Auswertungen geplant waren, geschah dies für weiterführende Post-hoc-Analysen zu Gruppenunterschieden. Eine Auswertung erfolgte auf Itemebene.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar.

Patientenrelevanz

In der vorliegenden Therapiesituation soll die Betreuungsperson (CaGI-S) bzw. der Prüfarzt / die Prüfarztin (CGI-S) Mobilitätsaspekte, soziale Fähigkeiten und das adaptive Verhalten (im Vergleich zu Personen derselben Altersgruppe ohne ARG1-Mangel) zum aktuellen Zeitpunkt einschätzen. Unabhängig von der Vergleichskomponente wird insbesondere die Verbesserung von Mobilitätsaspekten im vorliegenden Anwendungsgebiet als patientenrelevant bewertet. Grundsätzlich wird die Selbsteinschätzung für die Nutzenbewertung bevorzugt. Aufgrund des mehrheitlich sehr jungen Alters der Studienteilnehmenden und der möglichen kognitiven Einschränkungen gestaltet sich die Erhebung in Form einer Selbsteinschätzung zumindest für einen Teil der Studienpopulation schwierig, sodass vorliegend eine Einschätzung durch eine dritte Person im Rahmen der Nutzenbewertung als adäquat erachtet würde.

Validität

Bei einem Global-Impression-Fragebogen (wie der CaGI-S und der CGI-S) handelt es sich nicht um ein studienübergreifendes, beständiges Instrument mit den gleichen Fragen, sondern um verschiedene Fragebögen, die häufig ähnlich aufgebaut sind, sodass diese bspw. häufig krankheitsspezifisch auf Itemebene einen Eindruck zur Morbidität erfassen. Im Rahmen der PEACE-Studie bestehen einige Unklarheiten zur Validität der eingesetzten Global-Impression-Fragebögen, insbesondere zu CaGI-S und CGI-S, welche nicht den unmittelbaren Eindruck von den Defiziten der Studienteilnehmenden vorsehen, sondern diese in einen Vergleich zu einer subjektiv variierenden (da vorgestellten) Referenzpopulation setzen (siehe unten). Der vorliegende CaGI-S scheint zudem nicht die Schwere von krankheitsspezifischen Symptomen auf globaler Ebene oder einzelne Symptome direkt zu erfragen, sondern besteht aus einer Fragensammlung zu teils komplexen Fähigkeiten und Kompetenzen. Zur diesbezüglichen Auswahl liegt keine Begründung vor.

Insgesamt ist zu berücksichtigen, dass keine Selbsteinschätzung durch die Patientin / den Patienten, sondern eine Fremdeinschätzung durch die jeweilige Betreuungsperson (CaGI-S) bzw. das ärztliche Prüfpersonal (CGI-S) erfolgt. Im Kontext der Hyperargininämie, bei der auch Kleinkinder ab 2 Jahren betroffen sind und welche mit kognitiven Einschränkungen einhergehen kann, erscheint eine Selbsteinschätzung durch alle Patientinnen und Patienten allerdings schwer umsetzbar, weshalb die Erhebung des Endpunkts durch eine dritte Person als angemessen angesehen wird. Fremdbewertungen liegen sowohl durch die betreuende Person als auch das ärztliche Prüfpersonal vor. Da die betreuenden Personen vorrangig Eltern der Testpersonen sind, wird von einem engen Patienten-Betreuer-Verhältnis ausgegangen, weshalb dieser dem CGI-S vorzuziehen wäre.

Wie eingangs erwähnt, erfragen CaGI-S und CGI-S jedoch nicht den unmittelbaren Eindruck einer/eines Dritten zu den Defiziten einer Patientin / eines Patienten, sondern erheben Defizite im imaginierten Vergleich zu anderen Personen derselben Altersgruppe ohne ARG1-Mangel. Dies erfolgt zudem durch vorgegebene Antwortoptionen, welche mit Ausnahme einer Antwort („kein klarer Unterschied“) nur defizitäre Vergleiche zu anderen Personen zulassen. Der beschriebene Vergleich kann von verschiedenen Faktoren beeinflusst sein, die außerhalb der Morbidität der jeweiligen betroffenen Person liegen – insbesondere von der dem Beurteilenden jeweils bekannten Referenzpopulation. Die Anforderungen an die fremdbeurteilende Person werden entsprechend als deutlich höher (und als interindividuell variabler) eingeschätzt als bei einer alleinigen Befragung zum Eindruck von bspw. den Symptomen einer Patientin / eines Patienten. Ob die beschriebenen Anforderungen in

vergleichbarer Weise durch alle Beurteilenden erbracht werden können und eine valide Aussage zu etwaigen Defiziten der Patientinnen und Patienten zulassen, erscheint in Ermangelung vorliegender Validierungsstudien als fraglich.

Caregiver Global Impressions of Change (CaGI-C) und Clinician Global Impressions of Change (CGI-C)

Die Erhebung durch den CaGI-C erscheint valider als jene durch den CaGI-S. Es liegen jedoch keine Auswertungen mit adäquaten Effektschätzern vor. Deskriptive Ergebnisse werden in der vorliegenden Nutzenbewertung im Anhang berichtet.

Der Endpunkt CGI-C wird in der Nutzenbewertung nicht berücksichtigt, da die Validität als nachrangig gegenüber dem CaGI-C angesehen wird.

Operationalisierung

Beschreibung

Der CaGI-C und der CGI-C sind Fragebögen, die unabhängig voneinander der Betreuungsperson und dem Prüfarzt / der Prüfarztin des Probanden gestellt werden, um ihren Eindruck von der Veränderung in der Qualität von Mobilitätsaspekten, Alltagskompetenzen, sozialen Fähigkeiten und vom adaptiven Verhalten während des doppelblinden Teils der Studie zu erfassen. Der CaGI-C wird durch die Betreuungsperson des Studienteilnehmenden beantwortet, der CGI-C durch den Prüfarzt / die Prüfarztin. Aus Konsistenzgründen wurden die Fragen, sofern möglich, während der gesamten Studie durch dieselben Personen beantwortet.

Im Rahmen der PEACE-Studie waren die folgenden Fragen für beide Betreuungsperson (CaGI-C) und Prüfarzt/Prüfarztin (CGI-C) vorgesehen:

1. Wie beurteilen Sie die Qualität und das Selbstvertrauen der Person beim Gehen, was ist Ihr Eindruck, wie sich die Gehfähigkeit des Studienteilnehmenden während der Studie verändert hat?
2. Welchen Eindruck haben Sie in Anbetracht der Abhängigkeit des Studienteilnehmenden von Hilfsmitteln (Gehhilfen), was ist Ihr Eindruck, wie sich die Fähigkeit des Studienteilnehmenden, sich von einem Ort zum anderen zu bewegen, während der Studie verändert hat?

Die folgenden Fragen waren ausschließlich für die Betreuungsperson vorgesehen:

3.
 - a) Welchen Eindruck haben Sie davon, wie sich die grobmotorischen Fähigkeiten (Bewegungen des ganzen Körpers, Arme und Beine: z. B. einen Ball fangen, balancieren, klettern, Trampolin springen, Fangen spielen, Rennen laufen) seit Beginn der Studie verändert haben?
 - b) Welchen Eindruck haben Sie davon, wie sich die feinmotorischen Fähigkeiten (Bewegungen der Hände und Finger: z. B. Malen mit Buntstiften, Schneiden mit der Schere, Spielen mit kleinen Gegenständen wie LEGO oder Perlen, mit Messer und Gabel essen, Schnürsenkel binden, einen Bleistift halten, zeichnen oder schreiben) seit Beginn der Studie verändert haben?
 - c) Welchen Eindruck haben Sie davon, wie sich die Alltagskompetenzen (Aktivitäten für eine unabhängige Lebensführung: z. B. sich selbst ernähren, baden, auf die Toilette gehen, sich anziehen, sich pflegen) seit Beginn der Studie verändert haben?
 - d) Welchen Eindruck haben Sie davon, wie sich die Kommunikationsfähigkeit (z. B. Anweisungen befolgen, Geschichten erzählen, Kommentare zu Ereignissen abgeben) seit Beginn der Studie verändert haben?

- e) Welchen Eindruck haben Sie davon, wie sich die Sozialisationsfähigkeit (z. B. Teilen von Spielzeug, Spielen mit Gleichaltrigen, Rücksichtnahme auf andere) seit Beginn der Studie verändert hat?
- f) Unter Berücksichtigung Ihrer Antworten auf die Fragen a–e: Wie hat sich die Fähigkeit zur Anpassung an die Umwelt des Patienten / der Patientin seit Beginn der Studie verändert?

Eine Auswertung erfolgte auf Itemebene. Der Bezugszeitraum ist seit Beginn der Studie.

Die möglichen Antworten für jede Frage erfolgten auf einer 5-stufigen Ordinalskala: viel schlechter, schlechter, keine eindeutige Veränderung, besser, viel besser. Für das Dossier erfolgte eine Kodierung der Antwortmöglichkeiten von 1 bis 5: 1 = „viel besser“; 2 = „besser“; 3 = „keine eindeutige Veränderung“; 4 = „schlechter“; 5 = „viel schlechter“. Da präspezifiziert lediglich deskriptive Auswertungen geplant waren, geschah dies für weiterführende Post-hoc-Analysen zum Gruppenunterschied.

Bewertung

Die Operationalisierung ist nur in Teilen nachvollziehbar. Es ist unklar, ob die Fragen interviewbasiert erhoben oder selbst ausgefüllt werden. Es wurden keine geeigneten Effektschätzer vorgelegt (siehe Kapitel 2.4.), eine deskriptive Darstellung erfolgt im Anhang.

Patientenrelevanz

In der vorliegenden Therapiesituation soll die Betreuungsperson (CaGI-C) bzw. der Prüfarzt / die Prüfarztin (CGI-C) die Verbesserung von Mobilitätsaspekten, Alltagskompetenzen, sozialen Fähigkeiten und vom adaptiven Verhalten im Vergleich zum Studienbeginn einschätzen. Insbesondere die Verbesserung von Mobilitätsaspekten wird im vorliegenden Anwendungsgebiet als patientenrelevant bewertet. Grundsätzlich wird die Selbsteinschätzung für die Nutzenbewertung bevorzugt. Aufgrund des mehrheitlich sehr jungen Alters der Studienteilnehmenden und der möglichen kognitiven Einschränkungen gestaltet sich die Erhebung in Form einer Selbsteinschätzung zumindest für einen Teil der Studienpopulation schwierig, sodass eine Einschätzung durch eine dritte Person im Rahmen der Nutzenbewertung als adäquat erachtet wird.

Validität

Wie auch bei CaGI-S (und CGI-S) handelt es sich beim CaGI-C (und CGI-C) nicht um ein studienübergreifendes, beständiges Instrument mit den gleichen Fragen, sondern um verschiedene Fragebögen, die häufig ähnlich aufgebaut sind, sodass diese bspw. häufig krankheitsspezifisch auf Itemebene einen Eindruck zur Morbidität erfassen. Es konnten weder aus den Unterlagen des pU noch aus einer orientierenden Recherche weiterführende Informationen zur Validität hinsichtlich der beiden Fragebögen identifiziert werden, sodass vorliegend eine abschließende Bewertung nicht getroffen werden kann. Ähnlich wie der CaGI-S scheint der vorliegende CaGI-C nicht die Veränderung von krankheitsspezifischen Symptomen auf globaler Ebene oder einzelne Symptome direkt zu erfragen, sondern besteht aus einer Fragenauswahl zu teils komplexen Fähigkeiten und Kompetenzen. Zur diesbezüglichen Auswahl liegt keine Begründung vor. Die Bevorzugung des CaGI-S ggü. dem CGI-S hinsichtlich der Einschätzung anhand einer Betreuungsperson ggü. einer Einschätzung vom Prüfpersonal ist direkt übertragbar, weshalb der CaGI-C dem CGI-C gegenüber vorgezogen wird.

Eine Verzerrung durch Recall-Bias ist beim CaGI-C zu einem späteren Zeitpunkt der Erhebung (24 Wochen) möglich, weshalb wiederholte zeitnahe Erhebungen der Symptomschwere (inklusive einer Erhebung zu Baseline) grundsätzlich präferiert werden. In der vorliegenden PEACE-Studie wird der CaGI-S (mit wiederholter zeitnahe Erhebung) jedoch nicht als valide

eingeschätzt. Dieses begründet sich insbesondere darin, dass der CaGI-S im Rahmen der Fragestellungen einen Vergleich mit einer subjektiv variablen Referenzpopulation vorsieht, weshalb der CaGI-C mit einer patientenzentrierteren Erhebung (ohne Vergleich) demgegenüber als validere Erhebungsmethode eingeschätzt wird.

2.3.3 Lebensqualität

Pediatric Quality of Life Inventory (PedsQL)

Die Erhebungen anhand des PedsQL werden im vorliegenden Anwendungsgebiet als patientenrelevant und valide erachtet. Für die Nutzenbewertung liegen jedoch keine validen Auswertungen für den PedsQL vor, weshalb keine Ergebnisdarstellung erfolgt.

Operationalisierung

Beschreibung

Der Pediatric Quality of Life (PedsQL) dient zur Messung der allgemeinen gesundheitsbezogenen Lebensqualität bei Kindern und Jugendlichen. In der PEACE-Studie wurde der PedsQL für Patientinnen und Patienten im Alter von 2 bis 18 Jahren erhoben. Krankheitsspezifische Module des PedsQL wurden in der Studie PEACE nicht erfasst.

Die Datenerhebung erfolgte mittels Selbst- beziehungsweise durch Fremdbewertung der Eltern. Bei Kindern von 2 bis 4 Jahren erfolgte eine obligate Bewertung durch die Eltern. Es kamen die altersgerechten, selbst- und fremdberichteten Versionen für die Altersgruppen 5–7 Jahre, 8–12 Jahre und 13–18 Jahre zum Einsatz. Die Fähigkeit zur Selbstbewertung des PedsQL wurde zu Baseline durch das Prüfpersonal bzw. den testenden Psychologen / die testende Psychologin festgelegt. Falls sowohl eine Selbstbewertung als auch eine Bewertung durch die Eltern vorlag, wurde jeweils die Selbstbewertung verwendet. Der Bezugszeitraum beträgt einen Monat.

Die Erfassung der gesundheitsbezogenen Lebensqualität von Patientinnen und Patienten ≥ 19 Jahre war in der Studie durch den SF-36 geplant.

Der PedsQL besteht aus 4 Skalen mit insgesamt 23 Items:

- Physische Funktion (8 Items)
- Emotionale Funktion (5 Items)
- Soziale Funktion (5 Items)
- Schulische Funktion (5 Items)

Die Items wurden auf einer 5-Punkte-Likert-Skala von 0 (nie) bis 4 (fast immer) bewertet. Die Werte wurden zur Analyse wie folgt in eine Skala von 0 bis 100 transformiert: 0 = 100; 1 = 75; 2 = 50; 3 = 25; 4 = 0. Höhere Werte entsprechen einer besseren Funktion. Skalenwerte wurden nur berechnet, sofern mindestens 50 % der zugehörigen Items beantwortet wurden.

Der Gesamtscore bzw. Summenwert ergibt sich dann aus dem Mittelwert der beantworteten Items der einbezogenen Skalen.

Es können die 3 Summenwerte gebildet werden:

- Gesamtscore (23 Items)
- Physische Gesundheit (8 Items) (\cong Skalen „Emotionale Funktion“, „Soziale Funktion“ und „Schulische Funktion“)
- Summenwert der psychosozialen Gesundheit (15 Items)

Eine Auswertung innerhalb der verschiedenen Alterskategorien des PedsQL war nicht geplant.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Der pU gibt an, dass die Beurteilung der Fähigkeit zur Selbstbewertung durch das Prüfungspersonal bzw. einen Psychologen / eine Psychologin erfolgte, jedoch nicht, welche konkreten Kriterien für die Beurteilung herangezogen wurden. Es wird davon ausgegangen, dass Personen, denen die Fähigkeit zur Selbsteinschätzung abgesprochen wurde, auch keine Selbsteinschätzung vornehmen durften, sodass bei der vorrangigen Betrachtung der Selbsteinschätzung von einem Kollektiv ausgegangen wird, welches diese Fähigkeit besitzt. In der PEACE-Studie gibt es Personen, für die sowohl eine Fremd- als auch eine Selbsteinschätzung vorlag. Separate Auswertungen für die Selbsteinschätzungen liegen nicht vor.

Patientenrelevanz

Die durch den PedsQL erhobene gesundheitsbezogene Lebensqualität ist als patientenrelevant zu erachten. Eine Selbstbewertung, sofern diese möglich ist, wird gegenüber einer Fremdbewertung bevorzugt. Die Erhebung erfolgte teilweise durch die Fremdbeurteilung der Eltern, wenn keine Fähigkeit zur Selbstbeurteilung vorlag. Im vorliegenden Anwendungsgebiet wird dies für die Altersgruppe ab 5 Jahren durch das mögliche Vorhandensein von kognitiven Einschränkungen für eine Teilpopulation als adäquat erachtet. Für die Altersgruppe 2–4 Jahre wird die fremdberichtete Elternversion berücksichtigt.

Validität

Der PedsQL ist ein etabliertes Messinstrument, dessen Validität in vorherigen Nutzenbewertungsverfahren bereits bewertet und als valide eingeschätzt wurde. Das Instrument wurde unter anderen an Patientinnen und Patienten mit Zerebralparese validiert [42]. In der Untersuchung konnten für die meisten generischen Kernskalen akzeptable bis gute interne Konsistenzen (Cronbach's α : 0,72–0,89) nachgewiesen werden. Lediglich die durch die Kinder selbstberichtete soziale Funktion wies eine schlechte interne Konsistenz (Cronbach's α : 0,54) auf und sollte laut der Autoren lediglich zur deskriptiven Analyse herangezogen werden. Es konnte zudem gezeigt werden, dass Kinder mit Zerebralparese konsistent eine niedrigere Lebensqualität berichten als gesunde Kinder. Untersuchungen zur psychometrischen Güte des PedsQL für Patientinnen und Patienten mit ARG1-D konnten nicht identifiziert werden. Eine Übertragbarkeit scheint aufgrund ähnlicher Symptomatik zwischen der Zerebralparese und ARG1-D möglich.

Eine gemeinsame Auswertung der Fremd- und Eigenbewertung des PedsQL wird aufgrund der Unterschiede in den Instrumenten und einer fehlenden Validierung der gemeinsamen Auswertungen als nicht sachgerecht eingeschätzt. Separate Auswertungen für alle Personen, welche gemäß Einschätzung des Prüfungspersonals bzw. der Psychologin / des Psychologen die Fähigkeit zur Selbstbewertung besitzen, und Auswertungen der Fremdbeurteilungen ausschließlich für die Personen, denen die Fähigkeit zur Selbsteinschätzung aberkannt wurde, konnten nicht identifiziert werden. Zudem liegen Auswertungen von PedsQL-Fremdbeurteilungen zu einer größeren Anzahl an Personen vor, als im Dossier berichtet sind, weshalb unklar ist, ob darunter Personen sind, die gemäß der Einschätzung des Prüfungspersonals bzw. der Psychologin / des Psychologen die Fähigkeit zur Selbstbewertung besitzen.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt Unerwünschte Ereignisse (UE) wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Als UE wurde in der PEACE-Studie jedes unerwünschte medizinische Ereignis (einschließlich abnormaler Laborwerte) definiert, welches in einem Zusammenhang mit der Anwendung eines Arzneimittels auftrat, unabhängig davon, ob ein kausaler Zusammenhang mit der Studienmedikation vermutet wurde. Unter UE wurden ebenfalls Krankheiten oder Verletzungen verstanden, welche eine Verschlechterung bereits bestehender Erkrankungen bzw. Verletzungen darstellten. Wann immer möglich, sollte vorzugsweise eine Diagnose als UE erfasst werden als eine Reihe von zugehörigen Symptomen, die zu dieser Diagnose gehört. Es wurden jene UE und SUE erfasst, welche ab dem Tag der Studieneinwilligung bis zur letzten Studienvsitedate auftraten. Innerhalb von COVID-bedingten Pausen wurden UE weiterhin erfasst. Die Kodierung der UE erfolgte gemäß „Medical Dictionary for Regulatory Activities“ (MedDRA) Version 24 unter Verwendung der Systemorganklassen (SOC) und Preferred Terms (PT).

Schweregradeinteilung

Die Einstufung des Schweregrades wurde wie folgt dokumentiert:

- **Mild:** Ein normalerweise vorübergehendes Ereignis, das nur minimale oder keine Behandlung erfordert und die täglichen Aktivitäten des Patienten / der Patientin im Allgemeinen nicht beeinträchtigt.
- **Moderat:** Ein Ereignis, das normalerweise durch zusätzliche spezifische therapeutische Maßnahmen gelindert werden kann. Das Ereignis beeinträchtigt die üblichen Aktivitäten des täglichen Lebens und verursacht Unbehagen, stellt aber kein signifikantes oder dauerhaftes Risiko für den Patienten / die Patientin dar und führt zu einem geringen Maß an Unannehmlichkeiten oder Bedenken hinsichtlich der therapeutischen Maßnahmen.
- **Schwer:** Ein Ereignis, das die gewohnten täglichen Aktivitäten des Patienten / der Patientin beeinträchtigt und möglicherweise intensive therapeutische Intervention benötigt.

Schwerwiegende unerwünschte Ereignisse (SUE)

Als SUE wurden jene UE definiert, die:

- zum Tod führten
- lebensbedrohlich waren. Diese Definition schloss keine UE ein, die im Falle eines schwereren Verlaufes zum Tod der Person hätten führen können.
- einen stationären Krankenhausaufenthalt (≥ 24 Stunden) oder die Verlängerung eines bestehenden Krankenhausaufenthaltes erforderten
- zu einer dauerhaften oder erheblichen Behinderung oder Invalidität führten, d. h. die Fähigkeit, übliche Lebensfunktionen auszuführen, war erheblich gestört
- zu angeborenen Anomalien oder Geburtsfehlern führten
- ein wichtiges medizinisches Ereignis oder ein schwerwiegender medizinischer Zustand waren, der keines der oben genannten Kriterien erfüllte, aber die Person gefährden konnte oder eine medizinische/chirurgische Intervention erforderte, um eines der oben genannten Ergebnisse zu verhindern

UE von besonderem Interesse (AESI)

Es werden folgende UE von besonderem Interesse mittels MedDRA PT definiert:

- Reaktionen an der Injektionsstelle
- Überempfindlichkeitsreaktion: umfasst alle UE des MedDRA PT „Überempfindlichkeit“ sowie UE, die im Ermessen des ärztlichen Prüfpersonals eine Überempfindlichkeitsreaktion im Zusammenhang mit der Studienmedikation darstellten und während bzw. bis zu 4 Stunden nach der Infusionsverabreichung auftraten. Diese lassen sich wie folgt unterteilen:
 - Mild: Selbstlimitierende, sich spontan zurückbildende Reaktionen, die mit einer vorübergehenden Unterbrechung der Gabe der Studienmedikation oder einer Verringerung der Infusionsrate behandelt werden können
 - Moderat: Reaktionen, die nicht mit einfachen Maßnahmen abklingen und eine längere Beobachtung und Therapieunterbrechung erfordern
 - Schwer: Reaktionen, die ein Eingreifen erfordern, um schwerwiegende Folgen zu vermeiden.
- Hyperammonämische Episoden
 - Hyperammonämie
 - Hyperammonämische Krise
 - Hyperammonämische Enzephalopathie

Post hoc wurden UE des HLT „Leberfunktionsanalysen“ als UE „anomaler Leberfunktionstest“ als andere wichtige UE erfasst, da Einschränkungen der Leberfunktion für Personen mit ARG1-D von besonderer Bedeutung sind.

Für das Modul 4 wurden folgende UE als erkrankungsbezogene UE definiert: „Ammoniak erhöht“, „Alaninaminotransferase erhöht“, „Aspartataminotransferase erhöht“, „Hyperammonämie“, „Appetit vermindert“, „Hyperammonämische Enzephalopathie“, „Krampfanfall“ und „Tonische Konvulsion“.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Es ist jedoch unklar, weshalb in der PEACE-Studie mehr SUE als schwere UE gewertet wurden. Ein SUE wird demnach nicht auch als ein schweres UE gewertet, was aus der beschriebenen Operationalisierung für ein schweres UE nicht hervorgeht. Die Auswahl an erkrankungsbezogenen Ereignissen ist nur mit Einschränkungen nachvollziehbar. Es finden sich keine Erläuterungen zu deren Herleitungen. Ob z. B. das PT „Appetit vermindert“ inhaltlich plausibel ist, kann nicht nachvollzogen werden, da Angaben diesbezüglich fehlen.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen. Die Patientenrelevanz von Laborparametern ist unklar.

Validität

Bewertung

Die Erhebung des Endpunkts wird als valide erachtet. Bei der Ergebnisinterpretation ist zu beachten, dass die berichteten UE auch Ereignisse der Grunderkrankung umfassen können. Der pU berichtet mit dem Modul 4 post hoc die Gesamtraten der UE unter Herausrechnung der erkrankungsbezogenen Ereignisse.

2.3.5 Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 6.

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

Studienvisite (Woche) Endpunkt ¹⁾	Screening ²⁾	Baseline	Woche 12	Woche 24
Todesfälle ³⁾	kontinuierlich			
Arginin ⁴⁾	x ⁵⁾	x ⁶⁾		
2MWT ⁷⁾	x	- ⁸⁾	x	x
GMFM-E/(-D ⁴⁾)	x	- ⁸⁾	x	x
FMS ⁴⁾	x	- ⁸⁾	x	x
GFAQ	x	- ⁸⁾	x	x
CaGI-C ⁴⁾ /CGI-C	- ⁹⁾	- ⁸⁾	x	x
PedsQL ¹⁰⁾	x	- ⁸⁾	x	x
Sicherheit	x ³⁾	kontinuierlich		

¹⁾ Erhebungen können sich über mehr als einen Tag erstrecken. Dabei ist unklar, ob eine Erhebung zu einem Endpunkt mehrere Tage andauern kann.

²⁾ Es gab 3 Screeningtermine innerhalb etwa eines Monats vor der ersten Studienmedikationsgabe (1: Tag -28 bis -21; 2: Tag -14; 3: Tag -7 jeweils ± 1 Tag).

³⁾ Im Rahmen der Sicherheit erfasst.

⁴⁾ Der Endpunkt wird im Anhang dargestellt.

⁵⁾ Erhoben zu allen 3 Screeningterminen.

⁶⁾ Wöchentlich; 168 Stunden nach der Gabe der letzten Dosis der Studienmedikation und unmittelbar vor der Gabe der nächsten Dosis.

⁷⁾ 2MWT wird als erster der neuromotorischen Endpunkte erhoben (FMS, GMFM, CaGI-S/C, 9HPT).

⁸⁾ In den Studienunterlagen wird für diese Endpunkte jeweils eine Veränderung zu Baseline beschrieben, was jedoch nach den Angaben im Protokoll der Zeitpunkt zum Screening ist. Unklar ist, an welchem Screening-Tag die Erhebung durchgeführt wurde (1: Tag -28 bis -21; 2: Tag -14; 3: Tag -7).

⁹⁾ Es wird angenommen, dass keine Erhebung im Screening erfolgte. Die Angaben sind jedoch uneindeutig.

¹⁰⁾ Ergebnisse zum PedsQL werden aufgrund fehlender adäquater Auswertungen nicht dargestellt.

Abkürzungen: 2MWT: 2-Minuten-Gehtest (2-Minute Walking Test); 9HPT: 9-Hole Pegboard Test; CaGI-C/CGI-C: Caregiver/Clinician Global Impressions of Change; CaGI-S/CGI-S: Caregiver/Clinician Global Impressions of Severity; FMS: Functional Mobility Scale; GFAQ: Gillette Functional Assessment Questionnaire; GMFM: Gross Motor Function Measure-88; PedsQL: Pediatric Quality of Life Inventory

2.4 Statistische Methoden

Der für die vorliegende Nutzenbewertung herangezogene Datenschnitt erfolgte am 14. Oktober 2021 nach Abschluss der RCT-Phase. Die aktuellste vorliegende SAP-Version (3.0) wurde am 11. November 2021 vor dem Datenbankschluss am 19. November 2021 finalisiert.

Analysepopulationen

FAS-Population: Die FAS umfasst alle Personen, die randomisiert wurden und die mindestens eine Dosis der verblindeten Studienbehandlung erhalten haben. Alle Sicherheits- und Wirksamkeitsanalysen werden mit dieser Gruppe durchgeführt.

Die FAS-Population entspricht der ITT-Population, da alle randomisierten Teilnehmenden auch eine Studienmedikation erhalten haben.

Datenschnitte

Es liegt eine Interimsanalyse zum präspezifizierten Datenschnitt vom 14. Oktober 2021 vor (Ende der 24-wöchigen RCT-Phase). Dieser ist Gegenstand der vorliegenden Nutzenbewertung.

Ein weiterer Datenschnitt erfolgte laut Dossier zu Woche 144 der LTE-Phase am 24. Januar 2023. Diesbezügliche Daten liegen zur Nutzenbewertung nicht vor (siehe Kapitel 1).

Präspezifizierte Subgruppenanalysen

Subgruppenanalysen waren geplant für die primäre Analyse der Veränderung der Plasma-Arginin-Werte gegenüber dem Ausgangswert nach 24 Wochen und für die wichtigsten sekundären Analysen der Veränderung der 2MWT gegenüber dem Ausgangswert in Woche 24 und der Veränderung der GMFM-E-Werte gegenüber dem Ausgangswert in Woche 24. Die Anzahl der Personen in den Subgruppen musste ≥ 4 Personen in jeder Gruppe enthalten.

- Alter (< 18 Jahre bei der Untersuchung, ≥ 18 Jahre bei der Untersuchung)
- Geschlecht (männlich, weiblich)
- Region (US, nicht-US)
- GMFCS-Einstufung zu Studienbeginn (I, \geq II)

Präspezifizierung geplanter und durchgeführter Analysen

Die Randomisierung erfolgte stratifiziert, die Analysen jedoch unstratifiziert. Dies ist aufgrund der geringen Studiengröße nachvollziehbar.

Für die Sensitivitätsanalyse des GMFM-D wurde 1 Person ausgeschlossen, die keine Erhebung zu Baseline hatte. Diese ging in der Hauptanalyse irrtümlicherweise mit 0 zu Baseline ein. Zusätzlich wurde festgestellt, dass die ursprünglich geplante Auswertung aufgrund großer Varianzunterschiede zwischen den Gruppen nicht angemessen ist und dem wurde mit einem angepassten MMRM-Modell Rechnung getragen, indem die Varianzen separat für die Gruppen geschätzt wurden. Im Vergleich zur Hauptanalyse führt die Entfernung dieser einen Person aus der Berechnung in der Sensitivitätsanalyse zu einer deutlichen Änderung des Baselinemittelwerts sowie der Differenzen zu Woche 24 und zu einem signifikanten Unterschied zwischen den Gruppen. Die Robustheit des Effekts, z. B. mit Hilfe einer Imputation des Baselinewerts, wurde nicht weiter geprüft.

Auswertungen zum CaGI-C waren ausschließlich deskriptiv geplant. Mit dem Dossier werden post hoc durchgeführte MMRM-Analysen für den CaGI-C vorgelegt. In das Modell gehen vermutlich 2 Messungen ein (Woche 12 und Woche 24). Gemäß Operationalisierung überschneiden sich dabei die Recall-Perioden vollständig (erfragt wird jeweils eine Veränderung zu Studienbeginn). Das durchgeführte MMRM-Modell wird folglich als nicht adäquat erachtet.

Die post hoc berechneten Responderanalysen zu den Endpunkten GMFM und GFAQ, welche ausschließlich eine Verbesserung und keine Verschlechterungen erfassen, erscheinen im vorliegenden Anwendungsgebiet, welches zum gegenwärtigen Zeitpunkt eine irreversible und progressiv degenerative Erkrankung darstellt, nicht als alleinig geeignete Auswertungsmethode. Für die Nutzenbewertung werden daher vorrangig die kontinuierlichen Auswertungen berichtet. Eine Darstellung der Responderanalysen zur Verbesserung erfolgt ergänzend.

Fehlende Werte und Daten-Imputation

Für den primären Endpunkt Arginin im Plasma sollte beim Fehlen aller 4 finalen Messwerte der letzte Post-Baselinerwert verwendet werden. Sollte kein Post-Baselinerwert vorhanden sein, wurde die Veränderung von Baseline als 0 imputiert (siehe Kapitel 2.3.2).

Imputationen waren für andere Endpunkte nicht vorgesehen.

Für Responderanalysen sollte beim Fehlen der interessierenden Daten die Person als Non-Responder gewertet werden.

Einschätzung der statistischen Auswertungen

Der pU folgt seinen präspezifisch geplanten Auswertungen für den primären und für sekundäre Endpunkte mit Abweichungen, wie weiter oben beschrieben.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 7: Verzerrungspotential der RCT-Phase der Studie PEACE

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
PEACE	Ja ¹⁾	Ja ²⁾	Ja	Ja	Nein	Ja ³⁾⁴⁾	Hoch

¹⁾ Ein computergenerierter Randomisierungsplan wird von einem Beauftragten des Sponsors erstellt und verwaltet. Es wird von einer unabhängigen Verwaltung von Sponsoren/Untersuchern ausgegangen, jedoch liegen dazu keine eindeutigen Informationen vor.

²⁾ Zuteilung zur Behandlungsgruppe erfolgte mittels Interactive Web/Voice Response System (IXRS).

³⁾ Es zeigt sich eine Ungleichverteilung in zahlreichen Baselinecharakteristika (z. B. Enzymaktivität, Level der Spastik, kognitive Verzögerung, Lebertestanomalien) und auch in den Baselinewerten der erhobenen Endpunkte (Arginin, Gehfähigkeit, benötigte Hilfsmittel, Lebensqualität) jeweils zugunsten des Interventionsarms. Diese starke Ungleichverteilung weist auf eine Randomisierung hin, die nicht erfolgreich war.

⁴⁾ Die verschriebenen Ernährungsvorgaben (bspw. zur Proteinrestriktion und zu Kalorien) werden im Pegzilarginasearm im Mittel unterschritten. Im Placeboarm hingegen werden die verschriebenen Ernährungsvorgaben im Mittel überschritten (siehe Tabelle 25). Da ein ernährungsbedingter Einfluss bei ARG1-D ein relevanter Faktor ist, der die Symptomatik der Patientinnen und Patienten beeinflusst, kann eine Verzerrung über alle Endpunkte hinweg nicht ausgeschlossen werden.

Abkürzungen: ARG1-D: Arginase-1-Mangel; RCT: Randomisierte kontrollierte Studie

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

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

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
2MWT	Ja	Unklar ¹⁾	Unklar ²⁾	Ja ³⁾⁴⁾⁵⁾⁶⁾	Hoch
GMFM-E	Ja	Ja	Unklar ²⁾	Ja ³⁾⁴⁾⁵⁾⁶⁾⁷⁾	Hoch
GFAQ	Ja	Ja	Unklar ²⁾	Ja ⁴⁾⁵⁾⁶⁾	Hoch
Sicherheit	Ja	Ja	Nein	Ja ⁵⁾⁶⁾	Hoch

¹⁾ Der 2MWT sollte erst ab 3 Jahren durchgeführt werden. Es gibt im Pegzilarginasearm mindestens 1 Person, die zu Studienbeginn 2 Jahre alt ist. Es ist unklar, weshalb für 1 Person keine Berechnung für die Veränderung von Baseline zu Woche 24 vorliegt. Die Rückläufe können nicht in Gänze nachvollzogen werden.

²⁾ In den Studienunterlagen wird für diese Endpunkte jeweils eine Veränderung zu Baseline beschrieben, was jedoch nach den Angaben im Protokoll der Zeitpunkt zum Screening ist. Unklar ist, an welchem Screening-Tag bzw. -Tagen (bis zu drei) die Erhebung durchgeführt wurde und ob bei etwaiger Mehrfacherhebung Kriterien zur Auswahl des Erhebungszeitpunktes für die Auswertung vorlagen.

³⁾ Ein Einfluss auf die Motorik und manuelle Geschicklichkeit unabhängig von der neuromuskulären Funktion z. B. durch kognitive Einschränkungen, welche zu Baseline deutlich unterschiedlich in den Studienarmen verteilt waren, kann nicht ausgeschlossen werden (29 % der Personen im Pegzilarginasearm ggü. 46 % im Placeboarm hatten zu Baseline eine kognitive Störung und 57 % im Pegzilarginasearm ggü. 91 % im Placeboarm eine kognitive Beeinträchtigung). Ausgeschlossen von der PEACE-Studie waren Personen, deren Gesundheitszustand oder Komorbiditäten nach Ansicht des Prüfpersonals die Einhaltung der Studie bzw. die Interpretation der Studiendaten beeinträchtigen (z. B. schwere geistige Beeinträchtigungen). Eine Verzerrung durch die starke Ungleichverteilung kann jedoch nicht gänzlich ausgeschlossen werden.

⁴⁾ Das entwicklungsbedingte Erlernen vom Laufen, Hüpfen, auf einem Bein Stehen und der Feinmotorik kann für einen Teil der Patientinnen und Patienten zu einer Verbesserung der Motorik / manuellen Geschicklichkeit unabhängig von der Studienmedikation führen. Der Pegzilarginasearm ist im Median (min; max) zu Baseline jünger (8 Jahre (2; 28)) als der Placeboarm (12 Jahre (5; 29)), und 5 Personen (23,8 %) im Pegzilarginasearm sind zu Baseline im Alter von 2 bis 6 Jahren, während im Placeboarm nur 1 Person (9,1 %) im Alter von 5 Jahren in dieser Altersgruppe vorliegt. Ein entwicklungsbedingtes Erlernen bzw. Verbessern von motorischen Fähigkeiten während der Studie könnte den Effekt zugunsten von Pegzilarginase modifizieren.

⁵⁾ Es zeigt sich eine Ungleichverteilung in zahlreichen Baselinecharakteristika (z. B. Enzymaktivität, Level der Spastik, kognitive Verzögerung, Lebertestanomalien) und auch in den Baselinewerten der erhobenen Endpunkte (Arginin, Gehfähigkeit, benötigte Hilfsmittel, Lebensqualität) jeweils zugunsten des Interventionsarms. Diese starke Ungleichverteilung weist auf eine Randomisierung hin, die nicht erfolgreich war.

⁶⁾ Die verschriebenen Ernährungsvorgaben (bspw. zur Proteinrestriktion und zu Kalorien) werden im Pegzilarginasearm im Mittel unterschritten. Im Placeboarm hingegen werden die verschriebenen Ernährungsvorgaben im Mittel überschritten (siehe Tabelle 25). Da ein ernährungsbedingter Einfluss bei ARG1-D ein relevanter Faktor ist, der die Symptomatik der Patientinnen und Patienten beeinflusst, kann eine Verzerrung über alle Endpunkte hinweg nicht ausgeschlossen werden.

⁷⁾ Es ist unklar, wie viele Personen im Pegzilarginase-Arm < 5 Jahre waren und ob das Instrument für diese Kinder valide eingesetzt werden konnte.

Abkürzungen: ARG1-D: Arginase-1-Mangel; 2MWT: 2-Minuten-Gehtest (2-Minute Walking Test); GFAQ: Gillette Functional Assessment Questionnaire; GMFM: Gross Motor Function Measure-88; ITT: Intention to Treat

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

Tabelle 9: Allgemeine Angaben; Studie PEACE

Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
ITT-Population, n (%)	21 (100)	11 (100)
FAS ¹⁾ , n (%)	21 (100)	11 (100)
Abbruch der Studie, n (%)	1 (4,8)	0
Aufgrund von: Entzug der Einverständniserklärung	1 (4,8)	-
Abbruch der Studienmedikation, n (%)	k. A.	k. A.
Mediane Behandlungsdauer der RCT-Phase in Wochen (min; max) ²⁾	24 (6; 25)	24 (22; 25)
Mediane Beobachtungsdauer der RCT-Phase in Wochen (min; max) ³⁾	28,1 (12; 65)	29,0 (27; 65)
COVID-19 Pause, n (%) ⁴⁾		
Screening-Phase	2 (9,5)	3 (27,3)
RCT-Phase	2 (9,5)	1 (9,1)
Mediane Dauer der COVID-19-bedingten Pause in Wochen (min; max)		
Screening-Phase	k. A.	k. A.
RCT-Phase	k. A.	k. A.
Übergang in die Long-Term-Extension-Phase n (%)	20 (95,2)	11 (100)

¹⁾ Definition, siehe Kapitel 2.4 Statistische Methoden

²⁾ Pausierungszeiten wurden von der Behandlungsdauer abgezogen. Die Probanden durften die Studie wieder aufnehmen, wenn sie wegen der COVID-19-Pandemie eine vorübergehende Pause eingelegt hatten.

³⁾ Zeit von erster Visite bis zur letzten Visite in der 24-wöchigen RCT-Phase. Unklar ist, ob die erste Visite in der Screening-Phase liegen konnte oder ob es sich dabei um die Baselinevisite der RCT-Phase handelt.

⁴⁾ Das Alter der Personen mit COVID-19-bedingter Pausen ist nicht berichtet.

Abkürzungen: COVID-19: Coronavirus SARS-CoV-2; FAS: Full Analysis Set; IDM: Individuelles Disease Management; ITT: Intention to Treat; k. A.: keine Angabe; RCT: Randomisierte kontrollierte Studie

In der PEACE-Studie zeigen sich deskriptive Baselineunterschiede in zahlreichen Charakteristika. Dies betrifft z. B. das Alter. Es befinden sich im Median jüngere Kinder im Pegzilarginasearm (8 Jahre (2; 28)) als im Placeboarm (12 Jahre (5; 29)) und 5 Personen (23,8 %) im Pegzilarginasearm sind zu Baseline im Alter von 2 bis 6 Jahren, während im Placeboarm nur 1 Person (9,1 %) im Alter von 5 Jahren in dieser Altersgruppe vorliegt. Ein deskriptives Ungleichgewicht findet sich u. a. auch in Parametern der Krankheitsschwere (Spastik, kognitive Einschränkungen, Gehfähigkeit, Krampfanfälle im letzten Jahr), zu welcher deskriptiv höhere Werte berichtet werden als im Placeboarm. Es finden sich keine Angaben zur Verteilung des Stratifikationsfaktors der Anzahl hyperammonämischer Episoden in den vergangenen 90 Tagen und dem vergangenen Jahr (siehe Tabelle 2) zu Baseline.

Tabelle 10: Charakterisierung der Studienpopulation der Studie PEACE

Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
<i>Alter (Jahre)²⁾</i> Mittelwert (SD) Median (min; max)	9,6 (6,16) 8 (2; 28)	12,9 (6,77) 12 (5; 29)
<i>Altersgruppe (Jahre), n (%)</i> 2 bis < 6 6 bis < 12 12 bis < 18 ≥ 18	5 (23,8) 8 (38,1) 7 (33,3) 1 (4,8)	1 (9,1) 4 (36,4) 4 (36,4) 2 (18,2)
<i>Geschlecht, n (%)</i> Männlich Weiblich	12 (57,1) 9 (42,9)	7 (63,6) 4 (36,4)
<i>Abstammung (genetisch), n (%)</i> Kaukasisch/weiß Asiatisch Afroamerikanisch Andere Mehrere Ethnien Fehlender Wert	10 (47,6) 3 (14,3) 0 6 (28,6) 1 (4,8) 1 (4,8)	4 (36,4) 3 (27,3) 2 (18,2) 0 1 (9,1) 1 (9,1)
<i>Region, n (%)</i> USA Nicht-USA ⁵⁾	8 (38,1) 13 (61,9)	6 (54,5) 5 (45,5)
<i>BMI (kg/m²)</i> Mittelwert (SD) Median (min; max)	18,0 (3,04) 17,6 (13,7; 24,6)	20,3 (4,0) 20,4 (15,0; 27,7)
<i>Alter bei Auftreten erster Symptome (Jahre)</i> Mittelwert (SD) Median (min; max)	1,6 (2,52) 1 (0; 10)	2,5 (2,02) 2 (0; 7)
<i>Alter bei Diagnose des ARG1-Mangels (Jahre)</i> Mittelwert (SD) Median (min; max)	2,8 (4,06) 0,7 (0; 15)	4,2 (3,07) 4,6 (0; 11)
<i>Methode der ARG1-D-Diagnose⁷⁾, n (%)</i> Argininlevel Enzymaktivität in den RBC DNA Testung Neugeborenenenscreening Andere	18 (85,7) 1 (4,8) 17 (81,0) 6 (28,6) 3 (14,3)	10 (90,9) 4 (36,4) 9 (81,8) 2 (18,2) 0
<i>Medizinische Vorgeschichte, n (%)³⁾</i> Entwicklungsverzögerung Alanin-Aminotransferase erhöht Hyperammonämie ⁴⁾ Fußdeformität Kognitive Störung Botulinumtoxin-Injektion	3 (14,3) 12 (57,1) 10 (47,6) 0 6 (28,6) 1 (4,8)	4 (36,4) 7 (63,6) 4 (36,4) 2 (18,2) 5 (45,5) 3 (27,3)

Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
<i>Historische Arginin-Level (μM)</i> n (%) Mittelwert (SD) Median (min; max)	19 (90,5) 409,4 (114,8) 409 (173; 724)	11 (100) 476,1 (124,1) 454 (277; 664)
<i>Enzymaktivität in den RBC (units)</i> n (%) Mittelwert (SD) Median (min; max)	19 (90,5) 0,06 (0,157) 0 (0,0; 0,5)	11 (100) 0,02 (0,060) 0 (0,0; 0,2)
<i>Hyperammonämie in der Krankheitsgeschichte, n (%)³⁾⁴⁾</i>	12 (57,1)	6 (54,5)
<i>Gehfähigkeit kurze Distanz²⁾, n (%)</i> Normal Minimal oder moderat beeinträchtigt Schwere Beeinträchtigung Nicht in der Lage	12 (57,1) 7 (33,3) 2 (9,5) 0	2 (18,2) 8 (72,7) 1 (9,1) 0
<i>Gehfähigkeit moderate Distanz²⁾, n (%)</i> Normal Minimal oder moderat beeinträchtigt Schwere Beeinträchtigung Nicht in der Lage	9 (42,9) 8 (38,1) 2 (9,5) 2 (9,5)	2 (18,2) 6 (54,5) 2 (18,2) 1 (9,1)
<i>Gehfähigkeit längere anhaltende Distanz²⁾, n (%)</i> Normal Minimal oder moderat beeinträchtigt Schwere Beeinträchtigung Nicht in der Lage	6 (28,6) 9 (42,9) 4 (19,0) 2 (9,5)	1 (9,1) 5 (45,5) 1 (9,1) 4 (36,4)
<i>Fähigkeit Treppe zu steigen, n (%)⁶⁾</i> Normal Minimal oder moderat beeinträchtigt Schwere Beeinträchtigung Nicht in der Lage	9 (42,9) 6 (28,6) 1 (4,8) 4 (19,0)	2 (18,2) 5 (45,5) 1 (9,1) 3 (27,3)
<i>Personen, die den Einsatz von Hilfsmitteln benötigen, n (%)⁷⁾</i> <i>Art des Hilfsmittels</i> Knöchel-Fuß-Orthese Einspitziger Gehstock Gehstock mit 4 Spitzen Unterarmgehstützen Standard-Gehgestell Gehgestell mit 2 Rädern Gehgestell mit 4 Rädern Rollstuhl Andere Hilfsmittel	7 (33,3) 4 (19,0) 1 (4,8) - - 1 (4,8) - - 3 (14,3) 1 (4,8)	5 (45,5) 3 (27,3) - - - 1 (9,1) - - 3 (27,3) 1 (9,1)
<i>Distanz, für die das Hilfsmittel benötigt wird, n (%)⁷⁾</i> Kurze Distanzen Moderate Distanzen Längere Distanzen	6 (28,6) 5 (23,8) 6 (28,6)	3 (27,3) 5 (45,5) 4 (36,4)

Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
<i>GMFCS-Level zu Baseline, n (%)⁸⁾</i>		
I	9 (42,9)	5 (45,5)
II	9 (42,9)	4 (36,4)
III	0	0
IV	3 (14,3)	2 (18,2)
V	0	0
<i>Level der Spastik, n (%)</i>		
Keine	8 (38,1)	3 (27,3)
Mild	7 (33,3)	2 (18,2)
Moderat	5 (23,8)	4 (36,4)
Schwer	1 (4,8)	2 (18,2)
<i>Lokalisation der Spastik, n (%)⁷⁾</i>		
Obere Extremität rechts	1 (4,8)	3 (27,3)
Obere Extremität links	1 (4,8)	3 (27,3)
Untere Extremität rechts	13 (61,9)	8 (72,7)
Untere Extremität links	13 (61,9)	8 (72,7)
<i>Personen mit epileptischen Krampfanfällen vor Studieneinschluss, n (%)</i>	7 (33,3)	4 (36,4)
<i>Personen mit kognitiver Verzögerung, n (%)</i>	12 (57,1)	10 (90,9)
<i>Personen mit sprachlicher Verzögerung, n (%)</i>	13 (61,9)	7 (63,6)
<i>Personen mit Muskelkrämpfen, n (%)</i>	3 (14,3)	4 (36,4)
<i>Personen mit Lebertestanomalien, n (%)</i>	14 (66,7)	9 (81,8)
ALT/AST-Erhöhung	13 (61,9)	9 (81,8)
PT/INR-Erhöhung	8 (38,1)	3 (27,3)
Fibrinogen abnormal	0	0
GGT erhöht	2 (9,5)	0
Andere	1 (4,8)	0

¹⁾ Alter = (Datum der Einwilligung nach Inkenntnissetzung - Geburtsdatum) / 365,25, als ganze Zahl trunziert.

²⁾ Kurz: zu Hause (5 m); moderat: Schule (50 m); länger anhaltend: Einkaufen (500 m)

³⁾ Angaben aus der medizinischen Vorgeschichte nach SOC und PT ≥ 20 % oder Unterschied von mind. 15 % in den Studienarmen.

⁴⁾ Abweichungen bei Hyperammonämie zwischen medizinischer Vorgeschichte und Krankheitsgeschichte sind nicht plausibel.

⁵⁾ Genauere Angaben liegen nicht vor. Neben den USA (n = 9) nahmen Studienzentren in Kanada (n = 1) und in Europa (n = 9) an der Studie teil (siehe Tabelle 2).

⁶⁾ Angaben fehlen für 1 Person.

⁷⁾ Mehrfachangaben sind möglich.

⁸⁾ Der GMFCS umfasst 5 Klassen, wobei höhere Klassen einer schlechteren motorischen Fähigkeit entsprechen.

Abkürzungen: ALT: Alanin-Aminotransferase; ARG1-D: Arginase-1-Mangel; AST: Aspartat-Aminotransferase; GGT: Gamma-Glutamyltransferase; GMFCS: Gross Motor Function Classification System; IDM: Individuelles Disease Management; μ M: Mikromolar; PT/INR: Prothrombin Time/International Normalized Ratio; PT: Preferred Terms nach MedDRA; RBC: Red Blood Cells; SD: Standardabweichung; SOC: Systemorganklasse

Protokollverletzungen

Anhand der für die PEACE-Studie berichteten wesentlichen Protokollverletzungen sind eindeutig verzerrende Aspekte nicht ersichtlich. Es ist jedoch anzumerken, dass bei 76 % im Pegzilarginasearm und bei 91 % im Kontrollarm mindestens eine wesentliche Protokollverletzung über die gesamte Studienlaufzeit auftrat. Es lagen keine gesonderten

Auswertungen für Protokollverletzungen in der RCT-Phase vor. Bei einzelnen Unterkategorien fehlen eingehende Angaben (z. B. zu study procedure criteria).

Begleitbehandlung

In Tabelle 11 sind die Begleitbehandlungen für die PEACE-Studie berichtet. Während der Studie erhielten 13 % Personen mehr eine Begleitmedikation mit Stickstofffängern im Pegzilarginasearm. Die verschriebenen Ernährungsvorgaben (bspw. zu Protein und Kalorien) während des Studienverlaufs werden im Pegzilarginasearm im Mittel unterschritten. Im Placeboarm werden die verschriebenen Ernährungsvorgaben im Mittel überschritten. Aus den kategoriellen Aminosäureeinteilungen geht nicht eindeutig hervor, wie viele Personen je Behandlungsarm mit Aminosäuren behandelt wurden. Eine detaillierte Dokumentation zu den Ernährungsvorgaben und dem tatsächlichen Konsum findet sich in Tabelle 25 im Anhang. Einschränkend ist an dieser Stelle zu nennen, dass die Validität der Ernährungstagebücher durch Unsicherheiten in der Compliance bei der Dokumentation der Ernährungsdaten vorliegend als eingeschränkt bewertet wird, weshalb die Daten lediglich im Anhang berichtet werden.

Tabelle 11: Angaben zur Begleitbehandlung mit Stickstofffängern, Aminosäurehaltige Ergänzungsmittel und Ernährungsanpassungen; Studie PEACE

Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
<i>Begleitende Therapie zu Studienbeginn mit Stickstofffängern, n (%)</i>	20 (95,2)	9 (81,8)
Aminosäurehaltige Ergänzungsmittel ¹⁾ , n (%)		
ATC-Klasse		
Bevorzugter Arzneimittelname		
Aminosäuren	2 (9,5)	0 (0)
Aminosäuren und Derivate	3 (14,3)	1 (9,1)
Aminosäuren/Kohlenhydrate/Mineralien/Vitamine	14 (66,7)	7 (63,6)
EAA	6 (28,6)	2 (18,2)
Fett/Kohlenhydrate/Proteine/Mineralien/Vitamine, Kombinationen	1 (4,8)	2 (18,2)
Protein Supplements	1 (4,8)	0
<i>Durchschnittliche Überschreitung der Ernährungsvorgaben während des Studienverlaufs, n (%)²⁾</i>		
Protein	k. A.	k. A.
Kalorien	k. A.	k. A.
u. a.	k. A.	k. A.

¹⁾ Es geht nicht eindeutig aus den Studienunterlagen hervor, ob mit den vorliegenden Extraktionen alle EAA abgedeckt sind, die als Begleitmedikation innerhalb des IDM verabreicht werden sollen.

²⁾ Anzahl Personen, die im Mittel über den Zeitraum von 24 Wochen ihre Ernährungsvorgaben nicht eingehalten haben.

Abkürzungen: ATC: Anatomisch-therapeutisch-chemische Klassifikation; EAA: Essentielle Aminosäuren; IDM: Individuelles Disease Management; k. A.: keine Angabe; SD: Standardabweichung

Anhand der in der Studie PEACE dokumentierten Begleitmedikation ergeben sich neben den in Tabelle 11 genannten keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

3.2 Mortalität

Es traten keine Todesfälle während der RCT-Phase der Studie PEACE auf.

3.3 Morbidität

2MWT

In Tabelle 12 sind die Ergebnisse des 2MWT berichtet. Bei der Interpretation sind die Unterschiede zu Baseline sowie die Abweichungen zwischen Mittelwert und Median zu Baseline (Pegzilarginasearm) und zu Woche 24 (beide Arme) zu berücksichtigen. Im Pegzilarginasearm befindet sich 1 Person weniger in der Auswertung.

Tabelle 12: Veränderung der Gehstrecke in Metern zu Studienwoche 24 (RCT) in der Studie PEACE, FAS

2MWT	Studie PEACE	Pegzilarginase + IDM N = 21 ¹⁾	Placebo + IDM N = 11
Baseline ²⁾			
n (%)		20 (95,2)	11 (100)
Mittelwert (SD)		109,0 (55,76)	99,9 (49,00)
Median (min; max)		122,0 (2; 202)	102,0 (0; 171)
Woche 24			
n (%)		20 (95,2)	10 (90,9)
Mittelwert (SD)		115,9 (51,81)	102,3 (51,1)
Median (min; max)		124,5 (1; 176)	117,0 (0; 163)
Veränderung Woche 24 zu Baseline ⁴⁾			
n (%)		19 ³⁾ (90,5)	10 (90,9)
LS Mean [95%-KI]		7,4 [k. A.]	1,9 [k. A.]
Gruppenunterschied der Veränderungen ⁴⁾			
LS Mean [95%-KI]			5,5 [-15,6; 26,7]
p-Wert			0,5961 ⁵⁾

¹⁾ Der 2MWT sollte erst ab 3 Jahren durchgeführt werden. Es gibt im Pegzilarginasearm mindestens 1 Person, die zu Studienbeginn 2 Jahre alt ist.

²⁾ Der Baselinewert wurde während der Screeningphase (Tag -28 bis -7 vor Baseline) erhoben. Zu welchem Zeitpunkt genau die Messung erfolgte, ist unklar.

³⁾ Es ist unklar, weshalb für 1 Person keine Berechnung für die Veränderung von Baseline zu Woche 24 vorliegt. Die Rückläufe können nicht in Gänze nachvollzogen werden.

⁴⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

⁵⁾ Auch im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich kein signifikanter Unterschied ($p = 0,66$). Effektschätzer liegen jedoch nicht vor.

Abkürzungen: 2MWT: 2 Minutes Walking Test; FAS: Full Analysis Set; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

GMFM (Dimension E)

Ergebnisse zur Dimension D werden im Anhang berichtet, da wesentliche Einschränkungen in den durchgeführten Analysen und den berechneten Effektschätzern bestehen (Tabelle 21 und Tabelle 22).

In Tabelle 13 werden die kontinuierlichen Auswertungen des GMFM der Dimension E berichtet. Bei der Interpretation sind die Abweichungen zwischen Mittelwert und Median

insbesondere im Placeboarm zu berücksichtigen. Die post hoc berechneten Responderanalysen mit einer Responseschwelle von $\geq 15\%$ zeigten keinen signifikanten Unterschied zwischen den Studienarmen.

Tabelle 13: Veränderung im GMFM Dimension E zu Studienwoche 24 (RCT) in der Studie PEACE, FAS

GMFM E¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Baseline ²⁾			
n (%)		21 (100)	11 (100)
Mittelwert (SD)		48,3 (19,93)	46,5 (24,56)
Median (min; max)		53,0 (5; 71)	56 (0; 72)
Woche 24			
n (%)		20 (95,2)	11 (100)
Mittelwert (SD)		52,0 (21,27)	46,1 (25,71)
Median (min; max)		57 (2; 72)	59 (0; 71)
Veränderung Woche 24 zu Baseline ³⁾			
n (%)		20 (95,2)	11 (100)
LS Mean [95%-KI]		4,2 [k. A.]	-0,4 [k. A.]
Gruppenunterschied der Veränderungen ³⁾			
LS Mean [95%-KI]			4,6 [-1,1; 10,2]
p-Wert			0,1087 ⁴⁾

¹⁾ Wertebereich 0–72. Höhere Werte entsprechen einer höheren grobmotorischen Funktion.

²⁾ Der Baselinewert wurde während der Screeningphase (Tag -28 bis -7 vor Baseline) erhoben. Zu welchem Zeitpunkt genau die Messung erfolgte, ist unklar.

³⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

⁴⁾ Auch im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich kein signifikanter Unterschied ($p = 0,32$). Effektschätzer liegen jedoch nicht vor.

Abkürzungen: FAS: Full Analysis Set; GMFM: Gross Motor Function Measure; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Gillette Functional Assessment Questionnaire (GFAQ)

In Tabelle 14 sind die Ergebnisse der kontinuierlichen Auswertung für den Endpunkt GFAQ dargestellt. Keine Person konnte die prädefinierte Responseschwelle von 2 Punkten erreichen. Bei der Interpretation sind die Abweichungen zwischen Mittelwert und Median insbesondere im Placeboarm zu Woche 24 zu berücksichtigen, wobei sich im Mittelwert keine Verbesserung zeigt, im Median hingegen schon.

Tabelle 14: Ergebnisse für den GFAQ zu Studienwoche 24 (RCT) in der Studie PEACE, FAS

GFAQ¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Baseline ²⁾			
n (%)		21 (100)	11 (100)
Mittelwert (SD)		7,9 (2,05)	7,5 (2,62)
Median (min; max)		8,0 (2; 10)	8,0 (2; 10)
Woche 24			
n (%)		20 (95,2)	11 (100)
Mittelwert (SD)		8,0 (2,26)	7,2 (2,48)
Median (min; max)		9,0 (2; 10)	9,0 (2; 9)

GFAQ ¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Veränderung Woche 24 zu Baseline ³⁾			
n (%)		20 (95,2)	11 (100)
LS Mean [95%-KI]		0,1 [k. A.]	-0,3 [k. A.]
Gruppenunterschied der Veränderungen ³⁾			
LS Mean [95%-KI]			0,4 [-0,3; 1,0]
p-Wert			0,23 ⁴⁾

¹⁾ Wertebereich von 1–10. Höhere Werte entsprechen einer besseren Gehfähigkeit.

²⁾ Der Baselinewert wurde während der Screeningphase (Tag -28 bis -7 vor Baseline) erhoben. Zu welchem Zeitpunkt genau die Messung erfolgte, ist unklar.

³⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

⁴⁾ Auch im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich kein signifikanter Unterschied ($p = 0,27$). Effektschätzer liegen jedoch nicht vor.

Abkürzungen: FAS: Full Analysis Set; GFAQ: Gillette Functional Assessment Questionnaire; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

3.4 Lebensqualität

Es liegen keine adäquaten Analysen für die Lebensqualität vor.

3.5 Sicherheit

Im Folgenden sind die Ergebnisse der Sicherheitsauswertungen für die PEACE-Studie dargestellt. Der pU erfasst unter TEAE alle UE, welche ab dem Tag der ersten Studienmedikation bis zum Behandlungsende ausgewertet wurden. Die mediane Behandlungsdauer (min; max) beträgt in beiden Armen 24 Wochen (Pegzilarginasearm: 24 (6; 25); Placeboarm: 24 (22; 25)). Die Beobachtungsdauer (min; max) beträgt im Pegzilarginasearm 28 Wochen (12; 65) und im Placeboarm 29 Wochen (27; 65). Es traten bei mehr Personen SUE als schwere UE in beiden Studienarmen auf. Effektschätzer zu Gruppenunterschieden wurden post hoc für das Modul 4 vorgelegt.

Tabelle 15: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der RCT-Phase der Studie PEACE, FAS

Studie PEACE Zusammenfassung der UE (Personen mit ≥ 1 Ereignis)	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)	RR [95%-KI] ³⁾ p-Wert
UE (ergänzend dargestellt) ⁶⁾	18 (85,7)	11 (100)	– ¹⁾
Schwere UE ²⁾⁶⁾	1 (4,8)	0 (0)	1,64 [0,07; 37,15] 1,0000 ⁵⁾
SUE ⁶⁾	4 (19,0)	4 (36,4)	0,52 [0,16; 1,70] 0,3970
UE, welches zum Abbruch der Studienmedikation führte	0 (0) ⁴⁾	0 (0)	–

¹⁾ Patientenrelevanz von Laborparametern unklar.

²⁾ Es wurden studieneigene Klassifikationskriterien verwendet (siehe Kapitel 2.3.4). Ein schweres Ereignis war definiert als ein Ereignis, das die gewohnten täglichen Aktivitäten des Patienten / der Patientin beeinträchtigt und möglicherweise intensive therapeutische Intervention benötigt.

- ³⁾ Effektschätzer wurden post hoc für das Dossier berechnet. Dafür wurde das RR unstratifiziert aus einer Vierfeldertafel berechnet. 95%-KI wurden mittels Normalverteilungsapproximation berechnet. Der p-Wert wurde mittels exaktem Fisher-Test berechnet.
- ⁴⁾ 1 Person brach die Studie aufgrund von eigenem Wunsch nach der 6. Dosis der Studienmedikation ab. Im 150-Tage-Bericht der EMA [16] wird angegeben, dass dieser Abbruch als Studienabbruch aufgrund von UE (Muskelkrämpfe und Abdominalschmerz) gewertet werden kann.
- ⁵⁾ Nullzellen-Korrektur um Wert 0,5 vorgenommen.
- ⁶⁾ Unter Herausrechnung der erkrankungsbezogenen Ereignisse (siehe Kapitel 2.3.4) ergeben sich post hoc für die Gesamtrate der UE gemäß pU 1 Person weniger im Placeboarm als in Tabelle 15 dargestellt. SUE ohne Erkrankungsbezug werden im Dossier bei 1 Person (ggü. vorliegend 4 Personen) im Pegzilarginasearm berichtet und bei keiner Person (ggü. vorliegend 4 Personen) im Placeboarm. Die Anzahl von Personen mit schweren UE stimmt zwischen den beiden Auswertungsvorgehen überein.

Abkürzungen: FAS: Full Analysis Set; EMA: European Medicines Agency; IDM: Individuelles Disease Management; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RCT: Randomisierte kontrollierte Studie; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; - : nicht berechenbar.

Unerwünschte Ereignisse

Tabelle 16: UE mit Inzidenz $\geq 10\%$ nach Systemorganklasse und Preferred Term in der RCT-Phase der PEACE-Studie, FAS

Studie PEACE MedDRA-Systemorganklasse Preferred Term	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)	RR [95%-KI] ¹⁾ p-Wert
Erkrankungen des Gastrointestinaltrakts	10 (47,6)	6 (54,5)	0,87 [0,43; 1,76] 1,0000
Erbrechen	6 (28,6)	3 (27,3)	1,05 [0,32; 3,40] 1,0000
Obstipation	3 (14,3)	1 (9,1)	1,57 [0,18; 13,39] 1,0000
Übelkeit	1 (4,8)	4 (36,4)	0,13 [0,02; 1,03] 0,0367
Abdominalschmerz	1 (4,8)	3 (27,3)	0,17 [0,02; 1,49] 0,1055
Infektionen und parasitäre Erkrankungen	8 (38,1)	0	9,27 [0,58; 147,11] 0,0292
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	7 (33,3)	0 (0)	8,18 [0,51; 131,22] 0,0664
Fieber	4 (19,0)	0 (0)	4,91 [0,29; 83,67] 0,2720
Untersuchungen	7 (33,3)	3 (27,3)	1,22 [0,39; 3,82] 1,0000
Ammoniak erhöht ³⁾	3 (14,3)	2 (18,2)	0,79 [0,15; 4,03] 1,0000
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	5 (23,8)	0 (0)	6,00 [0,36; 99,49] 0,1378
Erkrankungen des Nervensystems	5 (23,8)	3 (27,3)	0,87 [0,25; 2,99] 1,0000
Erkrankungen der Atemwege, des Brustraums und Mediastinums	5 (23,8)	2 (18,2)	1,31 [0,30; 5,69] 1,0000
Husten	4 (19,0)	1 (9,1)	2,10 [0,27; 16,54]

Studie PEACE MedDRA-Systemorganklasse Preferred Term	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)	RR [95%-KI] ¹⁾ p-Wert
			0,6367
Augenerkrankungen	4 (19,0)	0 (0)	4,91 [0,29; 83,67] 0,2720
Stoffwechsel- und Ernährungsstörungen	3 (14,3)	4 (36,4)	0,39 [0,11; 1,45] 0,1967
Hyperammonämie ²⁾	2 (9,5)	3 (27,3)	0,35 [0,07; 1,79] 0,3098
Appetit vermindert	0	2 (18,2)	0,11 [0,01; 2,09] 0,1109
Erkrankungen der Haut und des Unterhautgewebes	3 (14,3)	2 (18,2)	0,79 [0,15; 4,03] 1,0000
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	1 (4,8)	2 (18,2)	0,26 [0,03; 2,58] 0,2661
Psychiatrische Erkrankungen	2 (9,5)	2 (18,2)	0,52 [0,08; 3,23] 0,5932

¹⁾ Effektschätzer wurden post hoc für das Dossier berechnet. Dafür wurde das RR unstratifiziert aus einer Vierfeldertafel berechnet. 95%-KI wurden mittels Normalverteilungsapproximation berechnet. Der p-Wert wurde mittels exaktem Fisher-Test berechnet.

²⁾ Als UE von besonderem Interesse definiert.

³⁾ Gehört zu dem post hoc definierten anderen wichtigen UE (siehe Kapitel 2.3.4)

Abkürzungen: FAS: Full Analysis Set; IDM: Individuelles Disease Management; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RCT: Randomisierte kontrollierte Studie; RR: Relatives Risiko; UE: Unerwünschtes Ereignis

Schwere unerwünschte Ereignisse

Es traten im Pegzilarginasearm bei 1 Person schwere UE in den Kategorien: „Aspartataminotransferase erhöht“, „Hyperammonämie“ und „Panikattacke“ auf.

Schwerwiegende unerwünschte Ereignisse

Tabelle 17: Schwerwiegende UE, die bei $\geq 5\%$ der Patientinnen und Patienten in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term in der RCT-Phase der PEACE-Studie, FAS

Studie PEACE MedDRA-Systemorganklasse ¹⁾ Preferred Term	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)	RR [95%-KI] ²⁾ p-Wert
Hyperammonämie ³⁾	2 (9,5)	3 (27,3)	0,35 [0,07; 1,79] 0,3098
Hyperammonämische Enzephalopathie ³⁾	1 (4,8)	1 (9,1)	0,52 [0,04; 7,59] 1,0000

¹⁾ Schwerwiegende UE $> 5\%$ traten in den beiden berichteten PT auf, die zu den SOC Stoffwechsel- und Ernährungsstörungen und Erkrankungen des Nervensystems gehören. In diesen SOC traten keine weiteren SUE auf.

²⁾ Effektschätzer wurden post hoc für das Dossier berechnet. Dafür wurde das RR unstratifiziert aus einer Vierfeldertafel berechnet. 95%-KI wurden mittels Normalverteilungsapproximation berechnet. Der p-Wert wurde mittels exaktem Fisher-Test berechnet.

³⁾ Als UE von besonderem Interesse definiert.

Abkürzungen: FAS: Full Analysis Set; IDM: Individuelles Disease Management; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; PT: Preferred Term; RCT: Randomisierte kontrollierte Studie; RR: Relatives Risiko; SOC: Systemorganklasse; (S)UE: (Schwerwiegendes) unerwünschtes Ereignis

Unerwünschte Ereignisse von besonderem Interesse

In Tabelle 18 werden die UE von besonderem Interesse gelistet. UE von besonderem Interesse, die bereits unter den UE mit einer Inzidenz $\geq 10\%$ oder SUE gelistet sind, werden hier nicht erneut aufgeführt. Es handelt sich um die AESI: „Hyperammonämie“ und „Hyperammonämische Enzephalopathie“ und um das PT „Ammoniak erhöht“, was laut pU zu den post hoc definierten „anomaler Leberfunktionstest“ gehört und als andere wichtige UE geführt wird.

Tabelle 18: UE von besonderem Interesse (AESI) in der RCT-Phase der PEACE-Studie, FAS

Studie PEACE UE von besonderem Interesse (Personen mit ≥ 1 Ereignis)	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)	RR [95%-KI] ¹⁾ p-Wert
Überempfindlichkeit	2 (9,52)	0 (0)	2,73 [0,14; 52,30] 0,5343
Reaktion an der Injektionsstelle	0 (0)	0 (0)	-
Hyperammonämische Episoden ²⁾	3 (14,3)	4 (36,4)	0,39 [0,11; 1,45] 0,1967
Anomaler Leberfunktionstest ³⁾	4 (19,0)	2 (18,2)	1,05 [0,23; 4,85] 1,0000
Alaninaminotransferase erhöht	2 (9,5)	0	n. v.
Aspartataminotransferase erhöht	2 (9,5)	0	n. v.

¹⁾ Effektschätzer wurden post hoc für das Dossier berechnet. Dafür wurde das RR unstratifiziert aus einer Vierfeldertafel berechnet. 95%-KI wurden mittels Normalverteilungsapproximation berechnet. Der p-Wert wurde mittels exaktem Fisher-Test berechnet.

²⁾ Beinhaltet laut des pU die folgenden Preferred Terms: Hyperammonämie, Hyperammonämische Enzephalopathie und Hyperammonämische Krise.

³⁾ Post hoc definiert: andere wichtige UE (siehe Kapitel 2.3.4)

Abkürzungen: FAS: Full Analysis Set; IDM: Individuelles Disease Management; KI: Konfidenzintervall; n. v.: nicht vorgelegt; pU: pharmazeutischer Unternehmer; RCT: Randomisierte kontrollierte Studie; RR: Relatives Risiko; UE: Unerwünschtes Ereignis; - : nicht berechenbar.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Pegzilarginase

Gemäß Fachinformation ist Pegzilarginase (Loargys®) zur Behandlung von Arginase-1-Mangel (ARG1-D), auch bekannt als Hyperargininämie, bei Erwachsenen, Jugendlichen und Kindern im Alter ab 2 Jahren zugelassen. Pegzilarginase ist bestimmt für die chronische Behandlung von Patientinnen und Patienten mit ARG1-D in Verbindung mit einer individualisierten Behandlung der Erkrankung, wie z. B. der Proteinrestriktion, Aminosäure-haltigen Ergänzungsmitteln und pharmakologischer Behandlung (z. B. Stickstofffängern). Die Therapie sieht eine wöchentliche Injektion vor. Die empfohlene Anfangsdosis beträgt 0,1 mg/kg pro Woche, welche auf Grundlage der Arginin-Konzentrationen im Plasma angepasst werden kann, um die Plasmakonzentration von Arginin innerhalb des Normbereichs zu halten. Anpassungen erfolgen in Schritten von 0,05 mg/kg. Dosen über 0,2 mg/kg/Woche wurden in klinischen Studien zu ARG1-D nicht untersucht. Pegzilarginase sollte durch intravenöse Infusion oder subkutane Injektion gegeben werden, wobei die Dosis gleich ist [20].

Das Einschlusskriterium der Studie PEACE waren Kinder ≥ 2 Jahre, Jugendliche und Erwachsene, sodass die Anforderungen aus der Zulassungspopulation erfüllt sind. Die älteste teilnehmende Person war zu Baseline 29 Jahre alt. Inwiefern dies die obere Altersspanne der Zulassungspopulation darstellen könnte oder für das obere Altersspektrum der Zulassungspopulation keine Daten gewonnen werden konnten, wird seitens des pU nicht diskutiert und bleibt daher offen. Die Reversibilität der Symptome bei erwachsenen Patientinnen und Patienten mit bereits seit langem bestehenden neurologischen Schädigungen scheint eingeschränkt zu sein. Patientinnen und Patienten mit einer moderaten Hyperargininämie (< 250 uM) wurden von der Studie ausgeschlossen. Nach dem Wortlaut der Indikation kommen diese Personen jedoch für eine Behandlung in Frage.

Zum Einschluss einer Person musste diese in der Lage sein, die Untersuchungen durchzuführen und zu Baseline Defizite im 2MWT oder GMFM-D oder -E aufweisen (siehe Tabelle 10/Abbildung 1). Inwiefern die Cut-Offs zum Ausschluss von Personen mit geringeren Defiziten begründet sind und wie die ausgeschlossenen Personen in Abgrenzung zur Studienpopulation hinsichtlich ihrer Krankheitslast definiert werden können, konnte aus den eingereichten Unterlagen nicht nachvollzogen werden. Laut EMA würde für Personen mit geringeren Defiziten, die ausgeschlossen werden sollten, auch die Erhaltung einer normalen motorischen Funktion als legitimes Behandlungsziel angesehen werden. Personen mit extremen Mobilitätsdefiziten, definiert als jene Personen, welche ungeeignet für eine Beurteilung mittels des Gillette Functional Assessment Questionnaire (GFAQ) sind bzw. welche einen GFAQ-Gesamtwert von 1 aufweisen, wurden ebenfalls von der Studie ausgeschlossen. Entsprechend waren in Abgrenzung zur Zulassungspopulation weder Personen mit keinen bzw. bis zu den jeweiligen Cut-Offs vorliegenden Mobilitäts-einschränkungen noch solche mit starken Einschränkungen Teil der Studienpopulation.

An der PEACE-Studie nahmen überwiegend Studienzentren aus den USA teil, jedoch auch ein Studienzentrum aus Deutschland. Zumindest ein Teil der Studienteilnehmenden wurde mit Hilfe des Neugeborenen Screenings diagnostiziert (siehe Tabelle 10). Aus welchen Ländern diese Diagnosen stammen, ist nicht berichtet. Im Pegzilarginasearm wurden mehr Kinder mittels Neugeborenen Screening diagnostiziert (29 % vs. 18 %). Die Kinder waren im Mittel jünger bei Auftreten der ersten Symptome (1,6 vs. 2,5 Jahre) und bei Diagnosestellung (2,8 vs. 4,2 Jahre). Die Personen der PEACE-Studie erkrankten im Pegzilarginasearm demnach früher, wurden jedoch auch früher diagnostiziert und konnten einer möglichen Therapie zugeführt werden. Ob dies auch an eine bessere Prognose geknüpft ist, kann an dieser Stelle klinisch

nicht beurteilt werden. In Deutschland gibt es zur Zeit kein Neugeborenencreening für den ARG1-D. Es konnten jedoch auch keine Informationen identifiziert werden, dass in anderen Ländern regelhaft ein Neugeborenencreening auf ARG1-D durchgeführt wird. Inwieweit die Studienpopulation den deutschen Versorgungskontext abbildet und die Ergebnisse übertragbar sind, ist unklar.

4.2 Design und Methodik der Studie

Bei der Studie PEACE handelt es sich um eine randomisierte, doppelblinde, Placebo-kontrollierte, multizentrische Studie der Phase III im Parallelgruppendesign (2:1) mit 2 Behandlungsarmen. In der Studie wurde die Wirksamkeit und Sicherheit von Pegzilarginase im Vergleich zu Placebo, jeweils in Kombination mit individuellem Krankheitsmanagement bei Kindern und Jugendlichen und Erwachsenen ab 2 Jahren erfasst. Eine ARG1-D-Diagnose musste in der Krankenakte dokumentiert sein, wobei eine der folgenden Bedingungen erfüllt sein musste: entweder erhöhte Arginin-Konzentration im Plasma oder Mutationsanalyse, mit Diagnose einer pathogenen Mutation oder reduzierte Arginase-Aktivität in den Erythrozyten. Anhand der Baselinecharakteristika zeigt sich, dass für alle Personen bestätigende Mutationsanalysen vorlagen. Primäres Zielkriterium war die Veränderung der Arginin-Konzentration im Plasma von Baseline zu Woche 24.

Nach Abschluss der 24-wöchigen RCT-Phase sollten die Studienteilnehmenden in die einarmige Langzeit-Extensionsphase (LTE) von bis zu 150 Wochen übergehen. Die ersten 8 Wochen der LTE-Phase sollten weiterhin verblindet sein, um eine Entblindung von noch laufenden Personen in der RCT-Phase zu vermeiden. Es gingen im Placeboarm alle Personen und im Pegzilarginasearm 20 von 21 Personen in die LTE-Phase über. 1 Person brach die Studie auf eigenen Wunsch zu Woche 6 der RCT-Phase ab. Die vorgelegten Ergebnisse entstammen der geplanten Interimsanalyse vom 21.02.2022 mit einem Datenschnitt vom 14.10.2021 für die PEACE-Studie (Ende der RCT-Phase). Daten der laut Modul 4 am 01.02.2023 abgeschlossenen PEACE-Studie inklusive LTE-Phase wurde für die Erstellung der Nutzenbewertung nicht eingereicht. Die Daten wurden beim pU angefragt, jedoch mit der Begründung eines fehlenden Studienberichtes nicht nachgereicht. Eine ergebnisgesteuerte Berichterstattung kann diesbezüglich nicht gänzlich ausgeschlossen werden. Welche Ergebnisse und Bewertungen aus den einarmigen und somit nicht-vergleichenden Langzeitdaten hätten abgeleitet werden können, kann nicht weitergehend beurteilt werden. Die EMA fordert die finalen Daten der PEACE-Studie zum 31.03.2024. Eine Zulassung erfolgte durch eine Genehmigung für das Inverkehrbringen unter außergewöhnlichen Umständen.

Es wurden in den Placeboarm der Studie 11 Personen, in den Pegzilarginasearm 21 Personen randomisiert. Die Randomisierung in die Studienarme erfolgte stratifiziert nach der Schwere der Hyperammonämie in der Vorgeschichte (≥ 1 hyperammonämische Episode innerhalb von 90 Tagen vor der Einwilligung oder ≥ 2 hyperammonämische Episoden innerhalb von 365 Tagen vor der Einwilligung und keine hyperammonämische Episode innerhalb von 365 Tagen vor der Einwilligung oder 1 hyperammonämische Episode > 90 Tage vor der Einwilligung). Angaben zur Verteilung des Stratifikationsfaktors finden sich in den Studienunterlagen nicht. Es wird lediglich berichtet, dass 57 % im Pegzilarginasearm und 54 % im Placeboarm eine Hyperammonämie in der Krankheitsgeschichte aufwiesen. Weshalb innerhalb der Medical History unter dem SOC „Stoffwechsel- und Ernährungsstörungen“ im PT Hyperammonämie weniger Ereignisse berichtet sind (Pegzilarginasearm: 48 % und Placeboarm 36 %), ist unklar.

Auffällig ist eine unterschiedliche Verteilung vieler Baselinemerkmale z. B. Alter, Level der Spastik, kognitive Verzögerungen, sprachliche Verzögerungen und Muskelkrämpfe, sowie

Lebertestanomalien. Im Pegzilarginasearm betrug das mediane Alter 8 Jahre, im Placeboarm lag es mit 12 Jahren höher. Ebenso gab es im Pegzilarginasearm mehr Kinder im Alter von 2 bis < 6 Jahren (24 %) als im Placeboarm (9 %). 18 % der Personen im Placeboarm waren \geq 18 Jahre, hingegen nur 5 % im Pegzilarginasearm. Dies ist insbesondere wegen des natürlichen Krankheitsverlaufes relevant. Dieser ist progredient degenerativ, was eine zunehmende und zumeist irreversible Verschlechterung mit zunehmendem Alter bewirkt [13,41]. Wenig Unterschied in den Baselinemerkmalen zwischen den Studienarmen zeigt sich in der Anzahl der Personen mit einem Alter von 6 bis < 12 und 12 bis < 18 Jahren (33 % bis 38 %). Während für 55 % der Studienteilnehmenden im Placeboarm eine moderate oder schwere Spastik zu Baseline berichtet wird, betrifft dies nur 29 % im Pegzilarginasearm. Im Placeboarm werden Spastiken in den oberen Extremitäten (27 %) und unteren Extremitäten (73 %) häufiger beschrieben als im Pegzilarginasearm (5 % bzw. 62 %). Eine kognitive Verzögerung wird bei 91 % im Placeboarm und bei 57 % im Pegzilarginasearm beschrieben. Lebertestanomalien liegen bei 67 % der Personen im Pegzilarginasearm und bei 82 % im Placeboarm vor. Auch verschiedene Baselinewerte der erhobenen Endpunkte weisen eine unterschiedliche Verteilung in den Studienarmen auf. So hat der Placeboarm beispielsweise einen höheren historischen Arginin-Wert (409 μ M vs. 476 μ M) und eine schlechtere Gehfähigkeit über moderate und längere Distanzen. 12 % mehr Personen haben im Placeboarm schwere Beeinträchtigungen beim Treppe steigen oder sind dazu nicht in der Lage. Auf Hilfsmittel sind ca. 13 % mehr Personen im Placeboarm angewiesen, wobei über 20 % der Studienteilnehmenden im Placeboarm diese Hilfsmittel bereits für moderate Distanzen (50 m) benötigen. Insgesamt scheint eine ältere Population mit einer schwereren Krankheitslast im Placeboarm vorzuliegen. Eine Verzerrung aufgrund eines Ungleichgewichts in den Ausgangseigenschaften ist wahrscheinlich.

Ca. 95 % der Studienteilnehmenden im Pegzilarginasearm waren zu Baseline stabil auf mindestens einen Stickstofffänger eingestellt. Die häufigsten waren dabei aus der Gruppe der Glycerolphénylbutyraten und der Natriumbenzoaten. Im Placeboarm erhielten 82 % der Personen eine Begleittherapie mit Stickstofffängern. Anpassungen waren während der doppelblinden Periode nicht erlaubt.

Insgesamt scheint der Krankheitsverlauf von verschiedenen Faktoren beeinflusst zu sein. Besonders relevant sind dabei die Arginase-Restaktivität und auch die Einhaltung der Ernährung. Die Restaktivität gemessen in den Erythrozyten ist im Pegzilarginasearm mit 0,06 Units höher als im Placeboarm mit 0,02 Units. Eine Relevanz dessen kann an dieser Stelle klinisch nicht beurteilt werden. Anhand der Ernährungsdaten (siehe Tabelle 25 im Anhang) zeigt sich, dass die Ernährungsvorgaben im Pegzilarginasearm (Gesamtkalorien, EAA und natürliche Proteine/Tag) weitestgehend eingehalten oder sogar unterschritten wurden, im Placeboarm hingegen lagen die Einnahmen über weite Teile der Studie über den Ernährungsvorgaben. Da die Compliance in der Ernährung von Personen mit ARG1-D ein relevanter Faktor sein kann, der zur Verschlechterung der Symptomatik beitragen kann, kann ein Einfluss dieser Unterschiede auf die Ergebnisse der Studie PEACE nicht ausgeschlossen werden und trägt bei der Ergebnisinterpretation des Gruppenunterschiedes bei den Endpunkten zur Morbidität zu Unsicherheiten bei.

Die Ergebnissicherheit ist aufgrund der geringen Studiengröße zudem grundsätzlich limitiert. Auch ist fraglich, ob eine Verbesserung in der Motorik bei einer langsam progressiv fortschreitenden Erkrankung überhaupt erwartbar ist, weshalb für die RCT-Phase besonders die kontinuierlichen Auswertungen betrachtet wurden, um ggf. auch das Auftreten keiner Verschlechterung vs. einer Verschlechterung abbilden zu können.

Das Verzerrungspotential der 24-wöchigen RCT-Phase der Studie PEACE wird bedingt durch die große Ungleichverteilung der Baselinecharakteristika und einer unterschiedlichen Compliance bei der Einhaltung der Ernährungsvorgaben insgesamt als hoch bewertet.

4.3 Mortalität

Todesfälle wurden im Rahmen der Sicherheitserfassung dokumentiert. Im Verlauf der RCT-Phase der Studie PEACE ist keine Person verstorben.

4.4 Morbidität

Allgemein zu den Endpunkten der Mobilität

Ein etwaiger Einfluss der Studienmedikation auf die Motorik zeigt sich möglicherweise erst über einen langen Beobachtungszeitraum. Eine Studienlaufzeit von 24 Wochen ist daher unter Umständen nicht ausreichend, um etwaige Veränderungen abzubilden.

Bei der Interpretation der Ergebnisse sind die generellen Ungleichgewichte der Baselinecharakteristika, welche in Kapitel 4.2 ausgeführt wurden, zu berücksichtigen. Im Folgenden wird zusätzlich auf Ungleichgewichte von Baselinecharakteristika eingegangen, die endpunktspezifisch als besonders relevant eingeschätzt werden.

2MWT

Die Gehfähigkeit als Indikator für die Leistungsfähigkeit (Ausdauer) der Patientinnen und Patienten wurde über den 2MWT erhoben. Der Test wurde auf Grundlage der Richtlinie der ATS durchgeführt (American Thoracic Society 2002), wobei anzumerken ist, dass diese keine Durchführungsvorgaben für Kinder enthalten [1]. Die Nutzung einer Gehhilfe bei Durchführung des 2MWT wurde erlaubt. Es liegen jedoch keine Informationen dazu vor, inwiefern Patientinnen und Patienten zu unterschiedlichen Zeitpunkten beim 2MWT unterschiedliche Gehhilfen in Anspruch nehmen konnten bzw. genommen haben. Gleichzeitig weisen die Ergebnisse zum CaGI-C (siehe Item „Hilfsmittel für die Fortbewegung“ in Tabelle 24) darauf hin, dass es im Studienverlauf Veränderungen bei der Abhängigkeit von Fortbewegungshilfsmitteln bei einzelnen Patientinnen und Patienten in beiden Studienarmen gab. Diese Unklarheit hinsichtlich der Anwendung von Hilfsmitteln bei Anwendung des 2MWT wird kritisch gesehen.

Der Test sollte validierungskonform ab einem Alter von 3 Jahren eingesetzt werden. Aus den Baselinecharakteristika geht nicht hervor, wie viele Personen zwischen 2 und 3 Jahren im Pegzilarginasearm vorlagen (mindestens 1 Person). Im Placeboarm war keine Person unter 3 Jahren. Insgesamt können die Rückläufe für den 2MWT jedoch nicht in Gänze nachvollzogen werden, da eindeutige Angaben bzw. Gründe für fehlende Werte fehlen. Unklar ist zudem, weshalb eine Rücklaufquote von 20 Personen zu Baseline und nach Woche 24 berichtet werden, eine Auswertung der Veränderung jedoch nur für 19 Personen erfolgte.

Es zeigte sich kein statistisch signifikanter Unterschied in den Veränderungen der absoluten Gehstrecke von Baseline zu Woche 24 zwischen den Studienarmen. Bei der Interpretation sind die Abweichungen zwischen Mittelwert und Median zu Baseline (Pegzilarginasearm) und zu Woche 24 (beide Arme) zu berücksichtigen. Abweichungen (zwischen Mittelwert und Median) können auch aufgrund von kleinen Fallzahlen zustande kommen. Eine altersspezifische Standardisierung anhand einer geeigneten Referenzpopulation war nicht geplant, wird jedoch in der vorliegenden Studie mit Kindern ab 2 Jahren und zwischen den Studienarmen differierenden Altersverteilungen als relevant erachtet, um etwaige entwicklungsbedingte

Unterschiede zu berücksichtigen. Es kann nicht ausgeschlossen werden, dass das entwicklungsbedingte Erlernen bzw. Verbessern vom Laufen für einen Teil der Patientinnen und Patienten zu einer Verbesserung der Gehstrecke unabhängig von der Studienmedikation führen kann. Das jüngere Ausgangsalter im Pegzilarginasearm (im Median 8 Jahre) ggü. dem Placeboarm (im Median 12 Jahre) könnte in der Auswertung zu einer Verzerrung zugunsten des Prüfpräparats führen. Weitere Unterschiede ergeben sich zur Gehfähigkeit zu Baseline. Alle Personen wurden zu Baseline auf kurze Distanzen (von etwa 5 m) als gehfähig beschrieben, jedoch wurden im Pegzilarginasearm bei weniger Personen Beeinträchtigungen berichtet, sodass zu 57 % ggü. zu 18 % im Placeboarm eine normale Gehfähigkeit berichtet wurde. Ein ähnliches Bild ergibt sich bei einer moderaten Distanz (von etwa 50 m) zu Baseline, zu welcher ein ähnlich hoher Anteil an Personen (9 % ggü. 9,5 % in Placebo) als nicht gehfähig beschrieben wurde, aber 43 % im Pegzilarginasearm ggü. 19 % im Placeboarm als normal gehfähig berichtet wurden. Zu länger anhaltenden Distanzen (von etwa 500 m) wurden im Pegzilarginasearm 9,5 % der Personen und im Placeboarm 36 % der Personen als nicht gehfähig beschrieben, während die diesbezügliche Gehfähigkeit von 29 % der Personen im Pegzilarginasearm ggü. von 9 % im Placeboarm als normal beschrieben wurde. Die dargelegten Unterschiede sowie solche zu weiteren Charakteristika (Fähigkeit Treppe zu steigen, Schwere und Lokalisation der Spastik etc.) deuten insgesamt auf eine höhere mobilitätsspezifische Krankheitslast im Placeboarm ggü. dem Pegzilarginasearm zu Baseline hin und sollten in Form der Verzerrung bei der Ergebnisinterpretation berücksichtigt werden (s. u.).

Unabhängig von der neuromuskulären Funktion kann auch ein Einfluss durch kognitive Einschränkungen auf die Gehstrecke bzw. die einheitliche Testdurchführbarkeit nicht ausgeschlossen werden. Diese waren zu Baseline deutlich unterschiedlich in den Studienarmen verteilt. 29 % der Personen im Pegzilarginasearm ggü. 46 % im Placeboarm hatten zu Baseline eine kognitive Störung und 57 % im Pegzilarginasearm ggü. 91 % im Placeboarm eine kognitive Beeinträchtigung. Ausgeschlossen von der PEACE-Studie waren Personen, deren Gesundheitszustand oder Komorbiditäten nach Ansicht des ärztlichen Prüfpersonals die Einhaltung der Studie bzw. die Interpretation der Studiendaten beeinträchtigen (z. B. schwere geistige Beeinträchtigungen). Eine Verzerrung durch die starke Ungleichverteilung kann jedoch nicht gänzlich ausgeschlossen werden.

Insgesamt zeigte sich bei hohem Verzerrungspotential kein Unterschied in der zurückgelegten Gehstrecke in 2 Minuten zwischen den Studienarmen.

GMFM-E

Die Gross Motor Function Measure (GMFM) ist ein Beobachtungsinstrument zur Bestimmung der grobmotorischen Fähigkeiten von Kindern mit Zerebralparese [19]. Ergebnisse zur Dimension D werden im Anhang berichtet. In der Dimension E wird das Gehen, Laufen und Springen erhoben. Es zeigte sich kein statistisch signifikanter Unterschied in den Veränderungen des GMFM Dimension E von Baseline zu Woche 24 zwischen den Studienarmen.

In Ermangelung ausreichender Informationen und der Einreichung eines Testmanuals kann nicht final beurteilt werden, ob eine standardisierte Durchführung anhand eines Testmanuals erfolgte. Ähnlich wie beim 2MWT sind mobilitätsspezifische (Alter, Krankheitslast etc.) und kognitive Unterschiede zwischen den Studienarmen zu berücksichtigen, welche den Effekt zugunsten von Pegzilarginase modifizieren könnten. Hinsichtlich einer potentiell entwicklungsbedingten Effektmodifikation muss zudem einschränkend berücksichtigt werden, dass sich die Aufgaben im Rahmen des GMFM an den Fähigkeiten von einem gesunden, motorisch normal entwickelten 5-jährigen Kind orientieren. In den

Pegzilarginasearm wurden allerdings auch Kinder mit ARG1-D ab 2 Jahren (und unter 5 Jahren) randomisiert, weshalb für einen Teil der Studienpopulation unklar ist, ob sie die enthaltenen Aufgaben (wie z. B. Hüpfen, auf einem Bein Stehen oder Rückwärtslaufen) altersgemäß überhaupt durchführen können. Dies betrifft mindestens 1 Person und maximal 5 Personen im Pegzilarginasearm (24 %). Genauere Angaben hierzu werden vom pU nicht gemacht. Sensitivitätsanalysen unter Ausschluss von Kindern unter 5 Jahren wären diesbezüglich wünschenswert. Die dargelegten Baselineunterschiede wirken sich auch beim GMFM E negativ auf das Verzerrungspotential aus.

Insgesamt zeigte sich bei hohem Verzerrungspotential kein Unterschied in der Veränderung der grobmotorischen Fähigkeit von Baseline zu Woche 24 zwischen den Studienarmen.

GFAQ

Der Gillette Functional Assessment Questionnaire dient der Erfassung der selbstständigen, maximalen Funktionen. Die eingesetzte „Walking Scale“ erfasst dabei die übliche Fähigkeit zur Fortbewegung mit Hilfe der normalerweise verwendeten Hilfsmittel der Patienten und Patientinnen. Auch beim GFAQ sind mobilitätsspezifische (Alter, Krankheitslast etc.) Unterschiede zwischen den Studienarmen zu berücksichtigen, welche den Effekt zugunsten von Pegzilarginase modifizieren könnten. Es zeigte sich kein statistisch signifikanter Unterschied in den Veränderungen des GFAQ von Baseline zu Woche 24 zwischen den Studienarmen bei hohem Verzerrungspotential.

4.5 Lebensqualität

PedsQL

Es lagen keine adäquaten Analysen für die Fremd- und Selbstbewertung der Studienpopulation vor, weshalb die Ergebnisse zum PedsQL nicht herangezogen werden konnten. Der Einfluss von Pegzilarginase auf die Lebensqualität kann anhand der vorgelegten Daten nicht beurteilt werden.

4.6 Sicherheit

Es wurden für die Auswertungen der Endpunkte der Kategorie „Sicherheit“ UE berücksichtigt, die ab der 1. Dosis der Studienmedikation bis zum Behandlungsende der RCT-Phase zu Woche 24 auftraten. In jedem der Studienarme betrug die Dauer der Behandlung im Median 24 Wochen. Da COVID-bedingte Dosisunterbrechungen nicht in die Dauer der Behandlung zählen, ergibt sich eine Diskrepanz zur Beobachtungszeit, welche im Median im Pegzilarginasearm bei 28 Wochen und im Placeboarm bei 29 Wochen lag. Die Sicherheitspopulation entspricht der FAS-Population.

Der pU legt der Einteilung in die Schweregrade der UE eine eigene Klassifikation zugrunde und stellt in den mit dem Nutzendossier eingereichten Unterlagen die UE mit Grad 3 („schwer“) separat dar. Innerhalb der Behandlungsperiode bis Woche 24 trat bei 18 Personen (85,7 %) im Pegzilarginasearm und bei allen Personen im Placeboarm mindestens ein UE auf. Nur 1 Person im Pegzilarginasearm hatte ein UE mit dem Schweregrad „schwer“. Mit 4 Ereignissen in beiden Armen traten im Placeboarm im Verhältnis (36 % vs. 19 %) mehr SUE auf, der Gruppenunterschied war jedoch nicht signifikant. Kein UE führte laut Studienunterlagen zu einem Abbruch der Studienmedikation. Im 150-Tage-Bericht der EMA [16] wird angegeben, dass 1 Person die Studie aufgrund von eigenem Wunsch nach der 6. Dosis der Studienmedikation abbrach und dass dieser Abbruch auch als Studienabbruch aufgrund von UE (Muskelkrämpfe und Abdominalschmerz) gewertet werden kann.

Es ist nicht abschätzbar, ob die Angaben, dass SUE in beiden Studienarmen numerisch häufiger auftraten als schwere UE (Pegzilarginasearm: 19 vs. 5 %, Placeboarm: 36 vs. 0 %) plausibel sind. Einschränkungen könnten sich diesbezüglich aus der eingangs genannten studienspezifischen Schweregradierung ergeben.

Für den Endpunkt „UE“ zeigte sich auf SOC-Ebene zwischen den beiden Studienarmen jeweils ein statistisch signifikanter Unterschied zum Vorteil von Pegzilarginase für den PT „Übelkeit“ (RR: 0,13 (95%-KI: [0,02; 1,03]); p = 0,0367). Für die SOC „Infektionen und parasitäre Erkrankungen“ zeigt sich hingegen ein statistisch signifikanter Nachteil für Pegzilarginase (RR: 9,27 (95%-KI: [0,58; 147,11]); p = 0,0292).

Für den Endpunkt „Schwere UE“ zeigten sich lediglich bei 1 Person im Pegzilarginasearm Ereignisse in den PT „Aspartataminotransferase erhöht“, „Hyperammonämie“ und „Panikattacke“.

Für den Endpunkt „SUE“ zeigte sich auf PT-Ebene bei der Beurteilung des Gruppenunterschieds kein statistisch signifikanter Unterschied zwischen den Studienarmen. Mehr als 5 % SUE traten in den als AESI gekennzeichneten Kategorien „Hyperammonämie“ und „Hyperammonämische Enzephalopathie“ auf, jeweils mit höheren Inzidenzen im Placeboarm.

Post hoc für das Dossier definiert und berichtet der pU UE, die als Ereignisse der Grunderkrankung gewertet werden. Es finden sich jedoch keine Erläuterungen oder Angaben zu deren Herleitung. So ist z. B. unklar, ob der PT „Appetit vermindert“ inhaltlich plausibel ist. Unter Berücksichtigung und Herausrechnung ergeben sich jedoch keine signifikanten Effekte.

Beim Anteil der Personen mit einem „UE von besonderem Interesse“ zeigten sich zwischen den Studienarmen, in den durch den pU definierten Kategorien „Überempfindlichkeit“, „Reaktion an der Injektionsstelle“, „Hyperammonämische Episoden“ und in dem post hoc definierten anderen wichtigen UE „Anomaler Leberfunktionstest“, keine signifikanten Unterschiede. Eine Überempfindlichkeit zeigte sich bei 2 Personen im Pegzilarginasearm und bei keiner Person im Placeboarm. Reaktionen an der Einstichstelle wurden während der RCT-Phase nicht verzeichnet.

Die EMA stuft das Sicherheitsprofil von Pegzilarginase insgesamt als akzeptabel ein. Es wird jedoch darauf hingewiesen, dass die Sicherheitsdatenbank sehr klein ist und eine zuverlässige Bewertung der unerwünschten Arzneimittelwirkungen noch nicht möglich scheint. Die festgestellten Sicherheitsrisiken der Überempfindlichkeit und der Reaktionen an der Injektionsstelle werden seitens der EMA für die klinische Praxis als akzeptabel und beherrschbar angesehen [15].

Folgende Auflagen wurden durch die EMA an das Inverkehrbringen geknüpft: jährliche Berichte einer nichtinterventionellen Unbedenklichkeitsprüfung (Post-Authorisation Safety Study) mit einem ersten aktualisierten Bericht über die Unbedenklichkeit innerhalb von 6 Monaten nach der Zulassung [15].

Insgesamt wurden nur in dem PT „Übelkeit“ und der SOC „Infektionen und parasitäre Erkrankungen“ Unterschiede zwischen den Studienarmen festgestellt. Es wurden nur sehr wenige schwere UE und wenige SUE beobachtet, fast alle gehörten zu den erkrankungsbezogenen Ereignissen. Das Verzerrungspotential wurde in der Gesamtschau mit hoch eingestuft, da auch hier die Ungleichverteilungen in den Baselinecharakteristika und Einhaltung der Ernährungsvorgaben bestehen, welche den Effekt zugunsten von Pegzilarginase modifizieren könnten.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Pegzilarginase ist zugelassen für die Behandlung von Arginase-1-Mangel (ARG1-D), auch bekannt als Hyperargininämie, bei Erwachsenen, Jugendlichen und Kindern im Alter ab 2 Jahren. Die Nutzenbewertung von Pegzilarginase basiert auf der zulassungs begründenden Studie PEACE. Es handelt sich bei der Studie PEACE um eine parallele, doppelblinde, multizentrische, Placebo-kontrollierte Phase-III-Studie zur Untersuchung der Wirksamkeit und Verträglichkeit von Pegzilarginase im Vergleich zu Placebo, jeweils in Kombination mit individuellem Krankheitsmanagement, bei Kindern (≥ 2 Jahre), Jugendlichen und Erwachsenen mit Arginase-1-Mangel (ARG1-D).

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

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

Tabelle 19: Zusammenfassende Darstellung der Ergebnisse der Studie PEACE

Studie PEACE Darstellung der Ergebnisse	Pegzilarginase + IDM N = 21			Placebo + IDM N = 11			Pegzilarginase vs. Placebo	Effekt
Mortalität								
Todesfälle¹⁾	N	Personen mit Ereignis, n (%)		N	Personen mit Ereignis, n (%)		Effektschätzer	
	21 ²⁾	0 (0)		11 ²⁾	0 (0)		-	\leftrightarrow
Morbidität								
2MWT (Gehstrecke in Metern)	N	Baseline MW (SD)	Änderung Woche 24 zu Baseline LS Mean [95%-KI]	N	Baseline MW (SD)	Änderung Woche 24 zu Baseline LS Mean [95%-KI]	Effektschätzer ⁴⁾ LS-Mean- Differenz [95%-KI]; p-Wert	
	21 ²⁾³⁾	109,0 (55,76)	7,4 [k. A.]	11 ²⁾³⁾	99,9 (49,00)	1,9 [k. A.]	5,5 [-15,6; 26,7]; 0,60	\leftrightarrow
GMFM-E⁵⁾ (Gehen, Laufen, Springen)	N	Baseline MW (SD)	Änderung Woche 24 zu Baseline LS Mean [95%-KI]	N	Baseline MW (SD)	Änderung Woche 24 zu Baseline LS Mean [95%-KI]	Effektschätzer ⁴⁾ LS-Mean- Differenz [95%-KI]; p-Wert	
	21 ²⁾⁶⁾	48,3 (19,93)	4,2 [k. A.]	11 ²⁾	46,5 (24,56)	-0,4 [k. A.]	4,6 [-1,1; 10,2]; 0,11	\leftrightarrow

Studie PEACE Darstellung der Ergebnisse	Pegzilarginase + IDM N = 21			Placebo + IDM N = 11			Pegzilarginase vs. Placebo	Effekt
	N	Baseline MW (SD)	Änderung Woche 24 zu Baseline LS Mean [95%-KI]	N	Baseline MW (SD)	Änderung Woche 24 zu Baseline LS Mean [95%-KI]		
GFAQ ⁷⁾ (Walking Scale)	21 ²⁾ 6)	7,9 (2,05)	0,1 [k. A.]	11 ²⁾	7,5 (2,62)	-0,3 [k. A.]	0,4 [-0,3; 1,0]; 0,23	↔
Sicherheit⁸⁾								
	N	Personen mit Ereignis, n (%)	N	Personen mit Ereignis, n (%)	RR ¹¹⁾ [95%-KI]; p-Wert			
Schwere UE ⁹⁾ 10)	21 ²⁾ 6)	1 (4,8)	11 ²⁾	0 (0)	1,64 [0,07; 37,15]; 1,00			↔
SUE ¹⁰⁾	21 ²⁾ 6)	4 (19,0)	11 ²⁾	4 (36,4)	0,52 [0,16; 1,70]; 0,40			↔
UE, das zum Abbruch der Studien- medikation führte	21 ²⁾ 6)	0 (0)	11 ²⁾	0 (0)	-			↔

¹⁾ Todesfälle wurden über die Sicherheit erhoben.

²⁾ Full Analysis Set (FAS; siehe Kapitel 2.4)

³⁾ Der 2MWT sollte erst ab 3 Jahren durchgeführt werden. Es gibt im Pegzilarginasearm mindestens 1 Person, die zu Studienbeginn 2 Jahre alt ist. Es ist unklar, weshalb für 1 Person keine Berechnung für die Veränderung von Baseline zu Woche 24 vorliegt. Die Rückläufe können nicht in Gänze nachvollzogen werden.

⁴⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

⁵⁾ Wertebereich 0–72. Höhere Werte entsprechen einer höheren grobmotorischen Funktion.

⁶⁾ 1 Person brach die Studie zu Woche 6 ab, weshalb zu Woche 24 im Pegzilarginasearm 20 Personen in die Berechnung eingingen.

⁷⁾ Wertebereich von 1–10. Höhere Werte entsprechen einer besseren Gehfähigkeit.

⁸⁾ Weitere Ergebnisse zur Sicherheit (UE nach Organsystemen und Einzelereignisse) sind im Ergebniskapitel 3.5 Sicherheit dargestellt.

⁹⁾ Es wurden studieneigene Kriterien zur Schweregradeinteilung (mild, moderat, schwer) verwendet (siehe Kapitel 2.3.4).

¹⁰⁾ Post hoc definierte SUE ohne Erkrankungsbezug werden im Dossier bei 1 Person (ggü. vorliegend 4 Personen) im Pegzilarginasearm berichtet und bei keiner Person (ggü. vorliegend 4 Personen) im Placeboarm. Die Anzahl von Personen mit schweren UE stimmt zwischen den beiden Auswertungsvorgehen überein.

¹¹⁾ RR wurde unstratifiziert aus einer Vierfeldertafel berechnet. 95%-KI wurden mittels Normalverteilungsapproximation berechnet. Der p-Wert wurde mittels exaktem Fisher-Test berechnet.

Abkürzungen: 2MWT: 2-Minuten-Gehtest; FAS: Full Analysis Set; GFAQ: Gillette Functional Assessment Questionnaire; GMFM: Gross Motor Function Measure-88; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; MMRM: Mixed Models for Repeated Measure; MW: Mittelwert; RCT: Randomisierte kontrollierte Studie; RR: Relatives Risiko; SD: Standardabweichung; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis

Referenzen

1. **A. T. S. Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories.** ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002;166(1):111-117.
2. **Aeglea Biotherapeutics.** An open-label, multicenter extension study to evaluate the long-term safety, tolerability and effects of intravenous AEB1102 in patients with arginase 1 deficiency who previously received treatment in study CAEB1102-101A; interim clinical study report [unveröffentlicht]. 28.10.2021.
3. **Aeglea Biotherapeutics.** PEACE (pegzilarginase effect on arginase 1 deficiency clinical endpoints): a randomized, double-blind, placebo-controlled phase 3 study of the efficacy and safety of pegzilarginase in children and adults with arginase 1 deficiency; clinical study protocol version 1.0 - version 6.0 [unveröffentlicht]. 11.12.2020.
4. **Aeglea Biotherapeutics.** PEACE (pegzilarginase effect on arginase 1 deficiency clinical endpoints): a randomized, double-blind, placebo-controlled phase 3 study of the efficacy and safety of pegzilarginase in children and adults with arginase 1 deficiency; interim clinical study report [unveröffentlicht]. 21.02.2022.
5. **Aeglea Biotherapeutics.** PEACE (pegzilarginase effect on arginase 1 deficiency clinical endpoints): a randomized, double-blind, placebo-controlled phase 3 study of the efficacy and safety of pegzilarginase in children and adults with arginase 1 deficiency; statistical analysis plan, version 3.0 [unveröffentlicht]. 11.11.2021.
6. **Aeglea Biotherapeutics.** A phase 1/2 open-label study in patients with arginase 1 deficiency to investigate the safety, pharmacokinetics, and pharmacodynamics of intravenous AEB1102: clinical study report [unveröffentlicht]. 29.07.2019.
7. **Alotaibi M, Long T, Kennedy E, Bavishi S.** The efficacy of GMFM-88 and GMFM-66 to detect changes in gross motor function in children with cerebral palsy (CP): a literature review. *Disabil Rehabil* 2014;36(8):617-627.
8. **Ammann-Reiffer C, Bastiaenen CHG, Van Hedel HJA.** Measuring change in gait performance of children with motor disorders: assessing the Functional Mobility Scale and the Gillette Functional Assessment Questionnaire walking scale. *Dev Med Child Neurol* 2019;61(6):717-724.
9. **Berg AT, Caplan R, Baca CB, Vickrey BG.** Adaptive behavior and later school achievement in children with early-onset epilepsy. *Dev Med Child Neurol* 2013;55(7):661-667.
10. **Bohannon RW, Bubela D, Magasi S, McCreath H, Wang YC, Reuben D, et al.** Comparison of walking performance over the first 2 minutes and the full 6 minutes of the Six-Minute Walk Test. *BMC Res Notes* 2014;7:269.
11. **Bohannon RW, Wang YC, Bubela D, Gershon RC.** Normative two-minute walk test distances for boys and girls 3 to 17 years of age. *Phys Occup Ther Pediatr* 2018;38(1):39-45.
12. **Bohannon RW, Wang YC, Gershon RC.** Two-minute walk test performance by adults 18 to 85 years: normative values, reliability, and responsiveness. *Arch Phys Med Rehabil* 2015;96(3):472-477.

13. **Carvalho DR, Brum JM, Speck-Martins CE, Ventura FD, Navarro MM, Coelho KE, et al.** Clinical features and neurologic progression of hyperargininemia. *Pediatr Neurol* 2012;46(6):369-374.
14. **De Bildt A, Kraijer D, Sytema S, Minderaa R.** The psychometric properties of the Vineland Adaptive Behavior Scales in children and adolescents with mental retardation. *J Autism Dev Disord* 2005;35(1):53-62.
15. **European Medicines Agency (EMA).** Loargys (international non-proprietary name pegzilarginase): European public assessment report EMEA/H/C/005484/0000 [online]. 12.10.2023. Amsterdam (NED): EMA. [Zugriff: 01.02.2024]. URL: https://www.ema.europa.eu/en/documents/assessment-report/loargys-epar-public-assessment-report_en.pdf.
16. **European Medicines Agency (EMA).** Rapporteurs day 150 joint assessment report of the responses to the list of questions - clinical: Loargys; Pegzilarginase, EMEA/H/C/005484/0000. Amsterdam (NED): EMA; 26.06.2023.
17. **Graham HK, Harvey A, Rodda J, Natrass GR, Pirpiris M.** The Functional Mobility Scale (FMS). *J Pediatr Orthop* 2004;24(5):514-520.
18. **Häberle J, Burlina A, Chakrapani A, Dixon M, Karall D, Lindner M, et al.** Suggested guidelines for the diagnosis and management of urea cycle disorders: first revision. *J Inher Metab Dis* 2019;42(6):1192-1230.
19. **Harvey AR.** The Gross Motor Function Measure (GMFM). *J Physiother* 2017;63(3):187.
20. **Immedica Pharma.** Loargys 5 mg/ml Injektions-/Infusionslösung [online]. 15.12.2023. Berlin (GER). [Zugriff: 01.02.2024]. URL: <http://www.fachinfo.de>.
21. **Immedica Pharma Germany.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4A Anhang 4-G: Pegzilarginase (Loargys), Arginase-1-Mangel, auch bekannt als Hyperargininämie, bei Erwachsenen, Jugendlichen und Kindern ab 2 Jahren; Ergänzende Informationen zur Studie PEACE [unveröffentlicht]. 12.01.2024.
22. **Immedica Pharma Germany.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4A: Pegzilarginase (Loargys), Arginase-1-Mangel, auch bekannt als Hyperargininämie, bei Erwachsenen, Jugendlichen und Kindern ab 2 Jahren; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 12.01.2024.
23. **Kallen M, Slotkin J, Griffith J, Magasi S, Salsman J, Nowinski C, et al.** NIH toolbox: technical manual; domain motor, subdomain dexterity [online]. Evanston (USA): Healthmeasures Northwestern University; 2012. [Zugriff: 22.03.2024]. URL: https://www.healthmeasures.net/images/nihtoolbox/Technical_Manuals/Motor/Toolbox_Grip_Strength_Test_Technical_Manual.pdf.
24. **Lee JH, Sung IY, Yoo JY.** Therapeutic effects of strengthening exercise on gait function of cerebral palsy. *Disabil Rehabil* 2008;30(19):1439-1444.
25. **Mendoza-Sanchez S, Molina-Rueda F, Florencio LL, Carratala-Tejada M, Cuesta-Gomez A.** Reliability and agreement of the Nine Hole Peg Test in patients with unilateral spastic cerebral palsy. *Eur J Pediatr* 2022;181(6):2283-2290.
26. **Meyer-Heim A, Ammann-Reiffer C, Schmartz A, Schafer J, Sennhauser FH, Heinen F, et al.** Improvement of walking abilities after robotic-assisted locomotion training in children with cerebral palsy. *Arch Dis Child* 2009;94(8):615-620.

27. **Middleton HA, Keene RG, Brown GW.** Convergent and discriminant validities of the scales of independent behavior and the revised Vineland Adaptive Behavior Scales. *Am J Ment Retard* 1990;94(6):669-673.
28. **Novacheck TF, Stout JL, Tervo R.** Reliability and validity of the Gillette Functional Assessment Questionnaire as an outcome measure in children with walking disabilities. *J Pediatr Orthop* 2000;20(1):75-81.
29. **Oeffinger D, Bagley A, Rogers S, Gorton G, Kryscio R, Abel M, et al.** Outcome tools used for ambulatory children with cerebral palsy: responsiveness and minimum clinically important differences. *Dev Med Child Neurol* 2008;50(12):918-925.
30. **Palisano R, Rosenbaum P, Bartlett D, Livingston M.** Gross motor function classification system: expanded and revised [online]. Hamilton (CAN): CanChild Mac Master University; 2007. [Zugriff: 22.03.2024]. URL: https://www.canchild.ca/system/tenon/assets/attachments/000/000/058/original/GMFCS-ER_English.pdf.
31. **Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B.** Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol* 1997;39(4):214-223.
32. **Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH.** Content validity of the expanded and revised Gross Motor Function Classification System. *Dev Med Child Neurol* 2008;50(10):744-750.
33. **Park ES, Rha DW, Yoo JK, Kim SM, Chang WH, Song SH.** Short-term effects of combined serial casting and botulinum toxin injection for spastic equinus in ambulatory children with cerebral palsy. *Yonsei Med J* 2010;51(4):579-584.
34. **Pearson Education.** Vineland Adaptive Behavior Scales, second edition (Vineland™-II) [online]. London (GBR): Pearson. [Zugriff: 26.02.2019]. URL: <https://www.pearsonclinical.com/psychology/products/100000668/vineland-adaptivebehavior-scales-second-edition-vinelandiiivineland-ii.html#tab-details>.
35. **Perry A, Factor DC.** Psychometric validity and clinical usefulness of the Vineland Adaptive Behavior Scales and the AAMD Adaptive Behavior Scale for an autistic sample. *J Autism Dev Disord* 1989;19(1):41-55.
36. **Pin TW.** Psychometric properties of 2-minute walk test: a systematic review. *Arch Phys Med Rehabil* 2014;95(9):1759-1775.
37. **Reuben DB, Magasi S, McCreath HE, Bohannon RW, Wang YC, Bubela DJ, et al.** Motor assessment using the NIH toolbox. *Neurology* 2013;80(11 Suppl 3):S65-75.
38. **Russell D, Rosenbaum P, Avery LM, Lane M.** Gross Motor Function Measure (GMFM-66 and GMFM-88) user's manual. London: Mac Keith Press; 2002.
39. **Russell DJ, Avery LM, Rosenbaum PL, Raina PS, Walter SD, Palisano RJ.** Improved scaling of the gross motor function measure for children with cerebral palsy: evidence of reliability and validity. *Physical Therapy* 2000;80(9):873-885.
40. **Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S.** The Gross Motor Function Measure: a means to evaluate the effects of physical therapy. *Dev Med Child Neurol* 1989;31(3):341-352.

41. **Schlune A, Vom Dahl S, Haussinger D, Ensenuer R, Mayatepek E.** Hyperargininemia due to arginase I deficiency: the original patients and their natural history, and a review of the literature. *Amino Acids* 2015;47(9):1751-1762.
42. **Varni JW, Burwinkle TM, Berrin SJ, Sherman SA, Artavia K, Malcarne VL, et al.** The PedsQL in pediatric cerebral palsy: reliability, validity, and sensitivity of the Generic Core Scales and Cerebral Palsy Module. *Dev Med Child Neurol* 2006;48(6):442-449.
43. **Vos-Vromans DC, Ketelaar M, Gorter JW.** Responsiveness of evaluative measures for children with cerebral palsy: the gross motor function measure and the pediatric evaluation of disability inventory. *Disabil Rehabil* 2005;27(20):1245-1252.
44. **Wang YC, Magasi SR, Bohannon RW, Reuben DB, McCreath HE, Bubela DJ, et al.** Assessing dexterity function: a comparison of two alternatives for the NIH Toolbox. *J Hand Ther* 2011;24(4):313-320; quiz 321.
45. **Wood E, Rosenbaum P.** The gross motor function classification system for cerebral palsy: a study of reliability and stability over time. *Dev Med Child Neurol* 2000;42(5):292-296.

Anhang

Arginin

Im Folgenden werden die Ergebnisse der kontinuierlichen Auswertung der Arginin-Konzentration berichtet (Tabelle 20). Anzumerken ist, dass bereits zu Baseline ein Unterschied in der Argininkonzentration im Plasma zwischen den Studienarmen vorliegt, wobei diese deutlich höher liegt im Placeboarm. Die präspezifizierte Analyse für den primären Endpunkt sah eine ANCOVA vor. Der pU weicht davon ab und verwendet ein MMRM für die primäre Analyse mit der Begründung, dass während der Auswertung festgestellt wurde, dass ein MMRM bezüglich des Fehlers 1. Art und zur Biasreduktion geeigneter sei. Woran (ob an neuen wissenschaftlichen Erkenntnissen oder anhand der tatsächlichen Daten) dies festgemacht wird, kann nicht nachvollzogen werden. Die Ergebnisse der ursprünglich vorgesehenen Analyse zeigen jedoch ein konsistentes Ergebnis, sodass der Darstellung der primären Analyse des pU und der EMA gefolgt wird.

Tabelle 20: Veränderung der Argininkonzentration im Plasma von Baseline zu Studienwoche 24 (RCT) der log-transformierten Arginin-Daten in der Studie PEACE, FAS

Arginin	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Baseline Argininkonzentration (μM) ¹⁾			
Mittelwert (Log-Skala SD)		354,0 (1,30)	464,7 (1,21)
Median (min; max)		368,3 (202; 572)	483,8 (294; 573)
Woche 24 Argininkonzentration (μM) ²⁾			
Mittelwert (Log-Skala SD)		86,4 (1,60)	426,5 (1,31)
Median (min; max)		83,7 (48; 376)	478,6 (213; 555)
Veränderung Woche 24 zu Baseline ³⁾			
GLS Mean [95%-KI]		0,23 [k. A.]	1,0 [k. A.]
Gruppenunterschied der Veränderungen ³⁾			
GLS Mean [95%-KI]			0,23 [0,17, 0,33]
p-Wert			< 0,0001

¹⁾ Der Baselinewert wird als das arithmetische Mittel aller Plasmargininwerte definiert, die während des Screeningzeitraums und vor Verabreichung der ersten Dosis der Studienmedikation ermittelt wurden.

²⁾ Der finale Argininspiegel wird definiert als der Mittelwert der letzten 4 verfügbaren logarithmierten Werte vor der Verabreichung (Werte, die 168 Stunden nach der Verabreichung erhoben wurden), und die folgenden Kriterien erfüllen: (1) das Datum der Probennahme lag nach dem geplanten Datum für die Dosis, (2) die beiden vorangegangenen Dosen wurden wie geplant verabreicht, d. h. etwa 7 Tage und 14 Tage vor der Entnahme der Argininprobe. Wenn mindestens 1, aber weniger als 4 Werte die Kriterien erfüllen, dann werden nur diese Werte in den Mittelwert einbezogen. Erfüllt keiner der Werte die Kriterien, so wird der letzte einzelne Argininwert nach dem Baselinewert berücksichtigt.

³⁾ MMRM mit log-transformierten Daten mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem logarithmierten Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

Abkürzungen: FAS: Full Analysis Set; GLS: Geometric Least Square; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; log: logarithmiert; μM : Mikromolar; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

GMFM Dimension D

Im Folgenden (Tabelle 21) sind die Ergebnisse der Hauptanalyse für den GMFM der Dimension D (Stehen) dargestellt. Die post hoc durchgeführte Sensitivitätsanalyse (Tabelle 22) mit Ausschluss einer Person, für die kein Baselinewert erhoben wurde (siehe Kapitel 2.4), zeigte einen Gruppenunterschied von 2,3 mit einem 95%-KI [0,38; 4,22] und einem p-Wert von

0,0208. Im Gegensatz dazu sind die Ergebnisse der Hauptanalyse nicht signifikant. Im EPAR wird die Sensitivitätsanalyse herangezogen, jedoch zeigt sich zwischen den Angaben im Modul 4 und den Angaben im EPAR eine Diskrepanz bezüglich des KI, welches die 0 miteinschließt, und somit kann kein signifikanter Unterschied abgeleitet werden (Effektschätzer und 95%-KI im EPAR: 2,3 [-0,4; 4,9]).

Bei der Interpretation der Ergebnisse sollte berücksichtigt werden, dass der Effektschätzer der Hauptanalyse (wegen des Messwertfehlers) nicht als adäquat erachtet wird und Unsicherheiten bezüglich des Effektschätzers der Sensitivitätsanalyse bestehen bezüglich der Diskrepanzen zu den Daten der EMA und nachträglicher Modellanpassung mit starkem Einfluss einer Person auf die Ergebnisse (siehe Kapitel 2.4.).

Das Verzerrungspotential der Dimension D entspricht dem der Dimension E. Eine Verzerrung wegen einer möglichen ergebnisgesteuerten Berichterstattung kann jedoch für die Dimension D zusätzlich vorliegen.

Tabelle 21: Veränderung im GMFM Dimension D zu Studienwoche 24 (RCT-Phase) in der Studie PEACE, FAS

GMFM D¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Baseline			
n (%)		20 (95,2)	11 (100)
Mittelwert (SD)		28,0 (9,61)	26,8 (14,76)
Median (min; max)		30 (1; 38)	32 (0; 39)
Woche 24			
n (%)		20 (95,2)	11 (100)
Mittelwert (SD)		30,5 (10,09)	28,2 (13,28)
Median (min; max)		33,5 (0; 39)	33 (0; 39)
Veränderung Woche 24 zu Baseline ²⁾			
n (%)		20 (95,2)	11 (100)
LS Mean [95%-KI]		2,7 [k. A.]	1,3 [k. A.]
Gruppenunterschied der Veränderungen ²⁾			
LS Mean [95%-KI]			1,4 [-1,4; 4,2]
p-Wert			0,3037 ³⁾

¹⁾ Wertebereich 0–39. Höhere Werte entsprechen einer höheren grobmotorischen Funktion.

²⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

³⁾ Auch im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich kein signifikanter Unterschied ($p = 0,25$). Effektschätzer liegen jedoch nicht vor.

Abkürzungen: FAS: Full Analysis Set; GMFM: Gross Motor Function Measure; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Tabelle 22: Sensitivitätsanalyse der Veränderung im GMFM Dimension D zu Studienwoche 24 (RCT-Phase) in der Studie PEACE, FAS

GMFM D¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Baseline			
n (%)		20 (95,2)	10 (90)
Mittelwert (SD)		28,0 (9,61)	29,50 (12,421)
Median (min; max)		30 (1; 38)	k. A.
Woche 24			
n (%)		20 (95,2)	10 (90)
Mittelwert (SD)		30,5 (10,09)	k. A.
Median (min; max)		33,5 (0; 39)	k. A.
Veränderung Woche 24 zu Baseline ²⁾			
n (%)		20 (95,2)	10 (90)
LS Mean [95%-KI]		2,7 [k. A.]	0,40 [k. A.]
Gruppenunterschied der Veränderungen ²⁾			
LS Mean [95%-KI]			2,30 [0,38; 4,22]
p-Wert			0,0208

¹⁾ Wertebereich 0–39. Höhere Werte entsprechen einer höheren grobmotorischen Funktion.

²⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt. Das MMRM-Modell wurde angepasst, da die Analyse mittels Levene-Test ungleiche Varianzen zwischen den Studienarmen zeigte ($p = 0,0217$).

Abkürzungen: FAS: Full Analysis Set; GMFM: Gross Motor Function Measure; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

FMS

Tabelle 23: Ergebnisse für die FMS – der RCT-Phase der PEACE-Studie, FAS

FMS¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
FMS 5			
Baseline			
n (%)		20 (95,2)	11 (100)
Mittelwert (SD)		5,2 (1,39)	4,8 (1,72)
Woche 24			
n		20 (95,2)	11 (100)
Mittelwert (SD)		5,2 (1,40)	4,7 (1,85)
Veränderung Woche 24 zu Baseline ²⁾			
LS Mean [95%-KI]		0,1 [k. A.]	-0,1 [k. A.]
Gruppenunterschied der Veränderungen ²⁾			
LS-Mean-Differenz [95%-KI]			0,2 [-0,3; 0;6]
p-Wert ⁴⁾			0,44
FMS 50			
Baseline			
n (%)		19 (90,5)	11 (100)
Mittelwert (SD)		5,0 (1,49)	4,5 (1,81)

FMS¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Woche 24			
n		20 (95,2)	11 (100)
Mittelwert (SD)		5,2 (1,53)	4,6 (1,80)
Veränderung Woche 24 zu Baseline ²⁾			
LS Mean [95%-KI]		0,2 [k. A.]	0,1 [k. A.]
Gruppenunterschied der Veränderungen ²⁾			
LS-Mean-Differenz [95%-KI]			0,1 [-0,3; 0,6]
p-Wert ⁴⁾			0,53
FMS 500			
Baseline			
n (%)		19 (90,5)	11 (100)
Mittelwert (SD)		4,6 (1,68)	3,7 (2,24)
Woche 24			
n		18 (85,7) ³⁾	11 (100)
Mittelwert (SD)		4,4 (1,89)	3,9 (2,34)
Veränderung Woche 24 zu Baseline ²⁾			
LS Mean [95%-KI]		0,0 [k. A.]	0,2 [k. A.]
Gruppenunterschied der Veränderungen ²⁾			
LS-Mean-Differenz [95%-KI]			0,1 [-0,4; 0,1]
p-Wert ⁴⁾			0,25

¹⁾ Werte 1–6. Höhere Werte sprechen für eine bessere funktionelle Mobilitätsleistung.

²⁾ MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als feste Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert. Es wurde keine stratifizierte Analyse vorgelegt.

³⁾ Es ist nicht ersichtlich, warum 1 Person weniger für die Auswertung der Veränderung von Baseline zu Woche 24 herangezogen wurde.

⁴⁾ Auch im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigten sich keine signifikanten Unterschiede (FMS 5: p = 0,87; FMS 50: p > 0,99; FMS 500: p = 0,54). Effektschätzer liegen jedoch nicht vor.

Abkürzungen: FAS: Full Analysis Set; FMS: Functional Mobility Scale; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Models for Repeated Measure; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Caregiver Global Impression of Change

Tabelle 24: Veränderung im CaGI-C zu Studienwoche 24, deskriptive Darstellung der Caregiver-Bewertung in der Studie PEACE, FAS

CaGI-C¹⁾²⁾	Studie PEACE	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)
Gefähigkeit			
Viel schlechter		0	0
Schlechter		1 (4,8)	1 (9,1)
Keine eindeutige Veränderung		11 (52,4)	7 (63,6)
Besser		6 (28,6)	1 (9,1)
Viel besser		1 (4,8)	2 (18,2)
Keine Antwortmöglichkeit ausgewählt		0	0

Studie PEACE	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)
CaGI-C¹⁾²⁾		
Hilfsmittel für die Fortbewegung		
Viel schlechter	0	0
Schlechter	0	0
Keine eindeutige Veränderung	13 (61,9)	6 (54,5)
Besser	4 (19,0)	2 (18,2)
Viel besser	1 (4,8)	2 (18,2)
Keine Antwortmöglichkeit ausgewählt	1 (4,8)	1 (9,1)
Feinmotorik		
Viel schlechter	0	0
Schlechter	0	0
Keine eindeutige Veränderung	11 (52,4)	5 (45,5)
Besser	8 (38,1)	5 (45,5)
Viel besser	0	1 (9,1)
Keine Antwortmöglichkeit ausgewählt	0	0
Grobmotorik		
Viel schlechter	0	0
Schlechter	0	0
Keine eindeutige Veränderung	14 (66,7)	7 (63,6)
Besser	4 (19,0)	3 (27,3)
Viel besser	1 (4,8)	1 (9,1)
Keine Antwortmöglichkeit ausgewählt	0	0
Alltagskompetenz		
Viel schlechter	0	0
Schlechter	0	0
Keine eindeutige Veränderung	14 (66,7)	6 (54,5)
Besser	4 (19,0)	4 (36,4)
Viel besser	1 (4,8)	1 (9,1)
Keine Antwortmöglichkeit ausgewählt	0	0
Kommunikation		
Viel schlechter	0	0
Schlechter	0	0
Keine eindeutige Veränderung	13 (61,9)	4 (36,4)
Besser	4 (19,0)	5 (45,5)
Viel besser	2 (9,5)	2 (18,2)
Keine Antwortmöglichkeit ausgewählt	0	0
Sozialisierung		
Viel schlechter	0	0
Schlechter	0	0
Keine eindeutige Veränderung	15 (71,4)	7 (63,6)
Besser	3 (14,3)	3 (27,3)
Viel besser	1 (4,8)	1 (9,1)
Keine Antwortmöglichkeit ausgewählt	0	0

CaGI-C ¹⁾²⁾	Studie PEACE	Pegzilarginase + IDM N = 21 n (%)	Placebo + IDM N = 11 n (%)
Anpassungsfähigkeit an das Umfeld			
Viel schlechter		0	0
Schlechter		1 (4,8)	0
Keine eindeutige Veränderung		12 (57,1)	5 (45,5)
Besser		5 (23,8)	5 (45,5)
Viel besser		1 (4,8)	1 (9,1)
Keine Antwortmöglichkeit ausgewählt		0	0

¹⁾ 5-stufige Skala, wobei ein niedriger Wert einer größeren Verbesserung entspricht (1 = „viel besser“; 2 = „besser“; 3 = „keine eindeutige Veränderung“; 4 = „schlechter“; 5 = „viel schlechter“)

²⁾ Angaben stammen aus Modul 4.

Abkürzungen: CaGI-C: Caregiver Global Impression of Change; FAS: Full Analysis Set; IDM: Individuelles Disease Management

Ernährungstagebuch

Tabelle 25: Ergebnisse für das Diättagebuch – der RCT-Phase der PEACE-Studie, FAS

3-Tage-Diättagebuch ¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Verschriebene Kalorienmenge (kcal/Tag)			
Baseline			
n (%)		18 (85,71)	9 (81,82)
MW (SD)		1603,8 (367,55)	1710,6 (411,75)
Median (min; max)		1515 (1033; 2425)	1800 (985; 2250)
Woche 24			
n (%)		17 (80,95)	8 (72,73)
MW (SD)		1621,1 (363,29)	1801,3 (330,39)
Median (min; max)		1637 (1033; 2368)	1800 (1200; 2250)
Tatsächlich konsumierte Kalorienmenge (kcal/Tag)			
Baseline			
n (%)		19 (90,48)	10 (90,91)
MW (SD)		1402,5 (474,52)	1936,2 (773,06)
Median (min; max)		1450 (571; 2070)	1884 (1083; 3692)
Woche 24			
n (%)		18 (85,71)	9 (81,82)
MW (SD)		1478,1 (619,17)	1922,7 (686,04)
Median (min; max)		1393,5 (344; 2613)	1724 (1018; 3020)
Verschriebene natürliche Proteinmenge (g/Tag)			
Baseline			
n (%)		18 (85,71)	9 (81,82)
MW (SD)		16,8 (9,45)	22,2 (10,24)
Median (min; max)		16,5 (5; 40)	18 (11; 40)
Woche 24			
n (%)		17 (80,95)	8 (72,73)
MW (SD)		17,5 (9,75)	23,1 (10,51)
Median (min; max)		18 (5; 40)	21,5 (11; 40)

3-Tage-Diättagebuch¹⁾	Studie PEACE	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Tatsächlich konsumierte natürliche Proteinmenge (g/Tag)			
Baseline			
n (%)		19 (90,48)	10 (90,91)
MW (SD)		15,0 (9,01)	24,1 (15,91)
Median (min; max)		15,2 (1; 34)	22,6 (6; 61)
Woche 24			
n (%)		18 (85,71)	9 (81,82)
MW (SD)		16,0 (7,70)	23,6 (16,29)
Median (min; max)		17,5 (0; 27)	23,3 (7; 61)
Verschriebene EAA-Proteinmenge (g/Tag)			
Baseline			
n (%)		19 (90,48)	10 (90,91)
MW (SD)		13,3 (6,93)	13,2 (6,89)
Median (min; max)		13,5 (0; 28)	12,1 (3; 25)
Woche 24			
n (%)		18 (85,71)	9 (81,82)
MW (SD)		13,0 (6,85)	12,9 (6,98)
Median (min; max)		13 (0; 28)	12,2 (3; 25)
Tatsächlich konsumierte EAA-Proteinmenge (g/Tag)			
Baseline			
n (%)		19 (90,48)	10 (90,91)
MW (SD)		10,1 (7,86)	14,1 (7,85)
Median (min; max)		9 (0; 23)	15,5 (3; 25)
Woche 24			
n (%)		18 (85,71)	9 (81,82)
MW (SD)		13,1 (10,54)	13,4 (7,20)
Median (min; max)		11,8 (0; 50)	12,2 (3; 25)

¹⁾ Aufzeichnung der gesamten Nahrungsaufnahme des Studienteilnehmenden an 3 aufeinanderfolgenden Tagen vor der jeweiligen Visite zu Baseline und in der Woche 24. In den Wochen 6, 12, 18 ergibt sich ein ähnliches Bild, weshalb die Daten lediglich zu Woche 24 berichtet werden.

Abkürzungen: EAA: Essentielle Aminosäuren; FAS: Full Analysis Set; IDM: Individuelles Disease Management; kcal: Kilokalorie; MW: Mittelwert; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung