

# Amendment



Gemeinsamer  
Bundesausschuss

**zur Nutzenbewertung 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**

Dossierbewertung vom 15. April 2024

Datum des Amendments: 14. Juni 2024

## **Inhaltsverzeichnis**

Tabellenverzeichnis .....	3
Abkürzungsverzeichnis .....	4
1 Hintergrund .....	5
2 Fragestellung .....	5
3 Liste der verwendeten Quellen .....	5
4 Ergebnisse.....	6
4.1 Morbidität.....	6
4.1.1 GMFM-D.....	6
4.2 Lebensqualität .....	10
4.2.1 PedsQL .....	10
4.3 Verzerrungspotential GMFM-D .....	10
5 Zusammenfassung.....	12
Anhang .....	14

## **Tabellenverzeichnis**

Tabelle 1: Veränderung im GMFM Dimension D zu Studienwoche 24 (RCT) in der Studie PEACE, FAS .....	7
Tabelle 2: Verzerrungspotential des in der Nutzenbewertung nachträglich berücksichtigten Endpunktes GMFM-D der Studie PEACE .....	10
Tabelle 3: Veränderung im PedsQL (Selbstbewertung) zu Studienwoche 24 (RCT) in der Studie PEACE, FAS.....	14

## Abkürzungsverzeichnis

2MWT	2 Minutes Walking Test
ARG1-D	Arginase-1-Mangel
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FAS	Full Analysis Set
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
GMFCS	Gross Motor Function Classification System
GMFM	Gross Motor Function Measure
IDM	Individuelles Disease Management
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to Treat
KI	Konfidenzintervall
LS	Least Square
MCAR	Missing completely at random
MMRM	Mixed Models for Repeated Measure
MW	Mittelwert
MWD	Mittelwertsdifferenz
N	Anzahl
PedsQL	Pediatric Quality of Life Inventory
pU	Pharmazeutischer Unternehmer
RCT	Randomisierte kontrollierte Studie/n (randomized controlled trial/s)
SGB	Sozialgesetzbuch
VerfO	Verfahrensordnung
WRS	Wilcoxon-Rangsummentest

## 1 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.

In die Bewertung werden gemäß Zulassungsstatus Patientinnen und Patienten in folgender Indikation eingeschlossen: Arginase-1-Mangel (ARG1-D), auch bekannt als Hyperargininämie, bei Erwachsenen, Jugendlichen und Kindern im Alter ab 2 Jahren.

Der pharmazeutische Unternehmer (pU) hatte am 12. Januar 2024 ein Dossier zur Nutzenbewertung gemäß § 35a SGB V beim Gemeinsamen Bundesausschuss (G-BA) eingereicht. Als bewertungsrelevante Studie wurde die pivotale Studie PEACE für die Nutzenbewertung herangezogen. Dabei handelt es sich 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 ( $\geq 2$  Jahre), Jugendlichen und Erwachsenen mit Arginase-1-Mangel (ARG1-D).

Die Nutzenbewertung wurde 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. Eine mündliche Anhörung fand am 27. Mai 2024 statt, in der offene Fragen zu den in den schriftlichen Stellungnahmen dargelegten Daten und Argumentationen erörtert wurden.

Mit der Stellungnahme reicht der pU Sensitivitätsanalysen für den in der Nutzenbewertung im Anhang dargestellten Morbiditätsendpunkt GMFM-D nach sowie eine getrennte Auswertung für den bislang in der Nutzenbewertung nicht dargestellten Lebensqualitätsfragebogen PedsQL.

Vor diesem Hintergrund bat der Unterausschuss Arzneimittel um eine Nachbewertung der mit der Stellungnahme nachgereichten ergänzenden Analysen und Auswertungen zu den Endpunkten GMFM-D und PedsQL.

## 2 Fragestellung

Mit diesem Amendment werden die am 06. Mai 2024 und 31. Mai 2024 durch den pU im Stellungnahmeverfahren zur Nutzenbewertung nachgereichten Analysen zu den Endpunkten GMFM-D (Morbidität) sowie PedsQL (Lebensqualität) bewertet.

## 3 Liste der verwendeten Quellen

Zur Bestimmung des Ausmaßes des Zusatznutzens für Pegzilarginase wurden folgende Quellen für dieses Amendment herangezogen:

- schriftliche Stellungnahme des pU vom 06. Mai 2024 und 31. Mai 2024 [4,5]
- mündliche Anhörung [2]
- Herstellerdossier [3]
- finaler Studienbericht [1]
- Testmanual GMFM [6]

## 4 Ergebnisse

### 4.1 Morbidität

#### 4.1.1 GMFM-D

Zum GMFM-D wurden u. a. folgende Auswertungen eingereicht:

- Präspezifizierte Analyse (enthält Person mit 0 zu Baseline; siehe Tabelle 21 in der Nutzenbewertung)
- EPAR-Auswertung (präspezifiziertes Modell ohne Person mit fehlendem Baselinewert) (siehe Tabelle 1)
- Sensitivitätsanalyse 1 (post hoc; ohne Person mit fehlendem Baselinewert und mit angepasstem MMRM-Modell aufgrund von Varianzheterogenität; siehe Tabelle 1)
- Sensitivitätsanalyse 2 (post hoc; Imputation des fehlenden Baselinewerts mit dem durchschnittlichen Baselinewert aller anderen Personen mit Growth Motor Function Classification System Level IV (7,25 Punkte); siehe Tabelle 1)
- Sensitivitätsanalyse 3 (post hoc; Imputation des fehlenden Baselinewerts mit dem durchschnittlichen Baselinewert aller anderen Personen (28,5 Punkte); siehe Tabelle 1)
- p-Wert des präspezifizierten nicht-parametrischen Wilcoxon-Rangsummentests (Angabe in Fußnoten der jeweiligen Tabelle)

Post hoc durchgeführte und durch den pU nachgereichte Outputs von Responderanalysen waren aufgrund relevanter fehlender Angaben nicht für das Amendment verwertbar.

Bei der Auswertung des GMFM-D stellte sich für den pU heraus, dass für einen Patienten / eine Patientin keine Messung des Baselinewerts vorhanden war und stattdessen wurde der Baselinewert mit 0 ersetzt. Für das Dossier wurde eine Auswertung gemäß der präspezifizierten Analyse einschließlich der Person mit 0 zu Baseline dargestellt (präspezifizierte Analyse) als auch eine Sensitivitätsanalyse (1) ohne diese Person und mit post hoc modifiziertem MMRM-Modell. Im EPAR wurde die Auswertung ohne diese Person stattdessen mit dem präspezifizierten Modell dargestellt (EPAR-Auswertung). Der pU begründet das post hoc abweichende Vorgehen damit, dass das präspezifizierte Modell aufgrund von signifikanter Varianzheterogenität nicht mehr adäquat erschien. Die EMA folgte gemäß pU dem Vorgehen nicht und zog zwar eine Auswertung ohne die Person mit fehlendem Baselinewert, jedoch auf Grundlage des präspezifizierten Modells vor (EPAR-Auswertung). Vorliegend wurden vom pU u. a. 2 weitere Sensitivitätsanalysen (2 und 3), jeweils mit singulärer Imputation des fehlenden Baselinewerts nachgereicht. Die Person mit fehlendem Baselinewert weist ein Growth Motor Function Classification System (GMFCS) Level von IV auf. Das GMFCS umfasst 5 Level (I bis V), wobei höhere Klassen einer schlechteren motorischen Fähigkeit entsprechen (siehe Kapitel 2.3.4 in der Nutzenbewertung). Die Sensitivitätsanalyse 2 imputiert entsprechend den fehlenden Baselinewert anhand des Mittelwerts der Baselinewerte aller anderen Personen mit GMFCS-Level IV. Die Sensitivitätsanalyse 3 imputiert den fehlenden Baselinewert anhand des Mittelwerts der Baselinewerte aller anderen Personen (aus beiden Armen).

Für alle Modelle wurde zudem der präspezifizierte nicht-parametrische Wilcoxon-Rangsummentest (WRS) durchgeführt (Angabe in Fußnoten der jeweiligen Tabelle). Die genannten Auswertungen können Tabelle 1 entnommen werden.

**Tabelle 1: Veränderung im GMFM Dimension D zu Studienwoche 24 (RCT) in der Studie PEACE, FAS**

<b>Studie PEACE Veränderung im GMFM-D<sup>1)</sup></b>	<b>Pegzilarginase + IDM N = 21</b>	<b>Placebo + IDM N = 11</b>
<b>Präspezifizierte Analyse</b> (enthält Person mit 0 zu Baseline)		
	<i>siehe Tabelle 21 in der Nutzenbewertung</i>	
<b>EPAR-Auswertung</b> (präspezifiziertes Modell ohne Person mit fehlendem Baselinewert)		
Baseline		
n (%)	21 (100)	10 (90,9) <sup>2)</sup>
Mittelwert (SD)	28,05 (9,61)	29,5 (12,42) <sup>2)</sup>
Median (min; max)	30 (1; 38)	k. A.
<b>Woche 24</b>		
n (%)	20 (95,2)	10 (90,9)
Mittelwert (SD)	30,45 (10,09)	29,9 (12,64)
Median (min; max)	33,5 (0; 39)	k. A.
<b>Veränderung Woche 24 zu Baseline<sup>3)</sup></b>		
n (%)	20 (95,24)	10 (90,9) <sup>2)</sup>
LS Mean [95%-KI]	2,67 [k. A.]	0,42 [k. A.]
<b>Gruppenunterschied der Veränderungen<sup>3)</sup></b>		
LS Mean [95%-KI]	2,25 [-0,37; 4,87]	
p-Wert	0,090 <sup>4)</sup>	
<b>Sensitivitätsanalyse 1</b> (post hoc; ohne Person mit fehlendem Baselinewert und mit angepasstem MMRM-Modell aufgrund von Varianzheterogenität)		
Baseline		
n (%)	<i>siehe EPAR-Auswertung</i>	
Mittelwert (SD)		
Median (min; max)		
<b>Woche 24</b>		
n (%)	<i>siehe EPAR-Auswertung</i>	
Mittelwert (SD)		
Median (min; max)		
<b>Veränderung Woche 24 zu Baseline<sup>5)</sup></b>		
n (%)	20 (95,2)	10 (90,9) <sup>2)</sup>
LS Mean [95%-KI]	2,70 [0,9; 4,5]	0,40 [-0,3; 1,1]
<b>Gruppenunterschied der Veränderungen<sup>5)6)</sup></b>		
LS Mean [95%-KI]	2,30 [0,38; 4,22]	
p-Wert	0,021 <sup>4)</sup>	
Hedges' g [95%-KI]	0,66 [-0,12; 1,44] <sup>7)</sup>	
<b>Sensitivitätsanalyse 2</b> (post hoc; Imputation des fehlenden Baselinewerts mit dem durchschnittlichen Baselinewert aller anderen Personen mit einem GMFCS-Level IV (7,25 Punkte))		
Baseline		
n (%)	21 (100)	11 (100) <sup>8)</sup>
Mittelwert (SD)	28,05 (9,61)	27,48 (13,56) <sup>8)</sup>
Median (min; max)	k. A.	k. A.
<b>Woche 24</b>		
n (%)	20 (95,2)	11 (100)
Mittelwert (SD)	30,45 (10,09)	28,18 (13,28)
Median (min; max)	k. A.	k. A.



Studie PEACE Veränderung im GMFM-D <sup>1)</sup>	Pegzilarginase + IDM N = 21	Placebo + IDM N = 11
Veränderung Woche 24 zu Baseline <sup>3)</sup> n (%) LS Mean [95%-KI]	20 (95,2) 2,69 [k. A.]	11 (100) <sup>8)</sup> 0,68 [k. A.]
Gruppenunterschied der Veränderungen <sup>3)</sup> LS Mean [95%-KI] p-Wert	2,01 [-0,51; 4,52]] 0,11 <sup>9)</sup>	
<b>Sensitivitätsanalyse 3</b> (post hoc; Imputation des fehlenden Baselinewerts mit dem durchschnittlichen Baselinewert aller anderen Personen (28,5 Punkte))		
Baseline n (%) Mittelwert (SD) Median (min; max)	21 (100) 28,05 (9,61) k. A.	11 (100) <sup>10)</sup> 29,41 (11,79) <sup>10)</sup> k. A.
Woche 24 n (%) Mittelwert (SD) Median (min; max)	20 (95,2) 30,45 (10,09) k. A.	11 (100) 28,18 (13,28) k. A.
Veränderung Woche 24 zu Baseline <sup>3)</sup> n (%) LS Mean [95%-KI]	20 (95,2) 2,68 [k. A.]	11 (100) <sup>10)</sup> -1,22 [k. A.]
Gruppenunterschied der Veränderungen <sup>3)</sup> LS Mean [95%-KI] p-Wert	3,90 [0,37; 7,43] 0,032 <sup>11)</sup>	
Hedges' g [95%-KI]	0,81 [0,05; 1,58]	

<sup>1)</sup> Wertebereich 0–39. Höhere Werte entsprechen einer höheren grobmotorischen Funktion.

<sup>2)</sup> Ausschluss von 1 Person (mit GMFCS-Level IV) für die kein Baselinewert vorlag.

<sup>3)</sup> 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.

<sup>4)</sup> Im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich kein signifikanter Unterschied ( $p = 0,094$ ).

<sup>5)</sup> MMRM mit Besuch, Studienbehandlung und Interaktion zwischen Besuch und Studienbehandlung als Effekte, und dem Baselinewert als Kovariate. Standard-Kovarianzstrukturtyp = unstrukturiert, jedoch wurde gemäß Stellungnahme hier eine Compound-Symmetry-Struktur angewendet. Es wurde keine stratifizierte Analyse vorgelegt.

<sup>6)</sup> Unter Verwendung des Mancl-DeRouen-Kovarianzschätzers, welcher auf Basis der vorliegenden eingeschränkten Informationen als gleichwertig eingeschätzt wird, zeigt sich ein signifikanter Effekt ( $p = 0,028$ ) mit einem Effektschätzer von 2,25 [0,27; 4,23] und einem Hedges' g von 0,64 [-0,14; 1,41]. Aus Übersichtsgründen werden die Ergebnisse nicht zusätzlich tabellarisch dargestellt.

<sup>7)</sup> Die Bewertbarkeit des Hedges' g wird bei ungleicher Varianz vom pU in Frage gestellt [7].

<sup>8)</sup> Für 1 Person (mit GMFCS-Level IV) lag kein Baselinewert vor. Der fehlende Baselinewert wurde imputiert mit 7,25, was dem durchschnittlichen Baselinewert aller anderen Personen mit einem GMFCS-Level von IV entspricht.

<sup>9)</sup> Im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich kein signifikanter Unterschied ( $p = 0,15$ ).

<sup>10)</sup> Für 1 Person (mit GMFCS-Level IV) lag kein Baselinewert vor. Der fehlende Baselinewert wurde imputiert mit 28,5, was dem durchschnittlichen Baselinewert aller anderen Personen (unabhängig von Behandlungsarm und GMFCS-Level) entspricht.

<sup>11)</sup> Im präspezifizierten nicht parametrischen Wilcoxon-Rangsummentest zeigte sich ein signifikanter Unterschied ( $p = 0,045$ ).

Abkürzungen: EPAR: European Public Assessment Report; FAS: Full Analysis Set; GMFCS: Gross Motor Function Classification System; 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



Die Effektschätzer der EPAR-Auswertung und der vom pU im Dossier präferierten Sensitivitätsanalyse 1 sind konsistent, nicht jedoch die Signifikanz des p-Werts (nur in der Sensitivitätsanalyse 1) und das zugehörige 95%-KI. Die singuläre Imputation des Baseline-MW aller Personen mit GMFCS-Level IV (Sensitivitätsanalyse 2) ergab bezogen auf den Gruppenunterschied ein ähnliches Ergebnis (MWD: 2,01 [95%-KI: -0,51; 4,52], p-Wert: 0,11). Die größte Differenz wurde über die singuläre Imputation des Baseline-MW anhand aller Personen (Sensitivitätsanalyse 3) erreicht (MWD: 3,9 [95%-KI: 0,37; 7,43], p-Wert: 0,035). In diesem Fall lag ein Hedges' g von 0,81 vor. Das 95%-KI schließt die Irrelevanzgrenze von 0,2 mit ein (95%-KI der SMD: 0,05; 1,51). Die kleinste, weiterhin positive Differenz wurde in der Präspezifizierten Analyse mit einem Baseline-Wert von 0 für den fehlenden Baselinewert erreicht (MWD: 1,4 [95%-KI: -1,4; 4,2], p-Wert: 0,30, siehe Tabelle 21 in der Nutzenbewertung).

Es zeigt sich in den Auswertungen konsistent eine positive Effektrichtung zugunsten von Pegzilarginase, wobei in 2 der 5 diskutierten MMRM-Auswertungen eine Signifikanz des Effektes vorliegt. Die beiden signifikanten Effekte (Sensitivitätsanalyse 1 und Sensitivitätsanalyse 3) würden anhand von Hedges' g jeweils nicht als klinisch relevant beurteilt. Zudem zeigt sich, dass gemäß pU nur unter vollständigem Ausschluss der Person mit fehlendem Baselinewert eine Varianzheterogenität vorlag. Darauf basierend wird keine der beiden Analysen unter Ausschluss der Person mit fehlendem Baselinewert (EPAR-Auswertung mit präspezifiziertem Modell ohne Varianzadjustierung versus Sensitivitätsanalyse 1 mit Varianzadjustierung allerdings post hoc) ggü. der jeweils anderen vorgezogen. Zugleich wird deutlich, dass insgesamt die Berücksichtigung bzw. Nichtberücksichtigung, die Art der Berücksichtigung dieser einen Person und das gewählte Auswertungsmodell einen entscheidenden Einfluss auf die Signifikanz der Ergebnisse hat. Dies lässt nicht auf einen robusten Effekt schließen. Es erscheint nachvollziehbar, dass eine Imputation von 0 für einen fehlenden Baselinewert im Placeboarm zu einer Verzerrung zugunsten des Placeboarms führen könnte und die präspezifizierte Analyse (siehe Tabelle 21 in der Nutzenbewertung) demnach als die konservativste der eingereichten Auswertungen und als nicht adäquat betrachtet werden kann. Entsprechend sind die weiteren Sensitivitätsanalysen mit anderweitigen Imputationen zu begrüßen. Da zu Baseline verschiedene Merkmale, wie bspw. Alter, Level der Spastik u. a., auf eine ältere Population mit einer schwereren Krankheitslast im Placeboarm als im Pegzilarginasearm hindeuten (siehe Nutzenbewertung) und die Person mit fehlendem Baselinewert deutliche Mobilitätseinschränkungen (mittels GMFCS festgestellt) aufweist, erscheint eine singuläre Imputation des fehlenden Baselinewerts einer Person im Placeboarm anhand vom Mittelwert aller Personen (Sensitivitätsanalyse 3) jedoch als weniger sachgerecht und die Analyse als zu anti-konservativ. Entsprechend wird die Sensitivitätsanalyse 2 mit Imputation des fehlenden Baselinewerts mit dem durchschnittlichen Baselinewert aller anderen Personen mit einem GMFCS-Level IV vorliegend als verlässlichere Auswertung mit Imputation eingeschätzt.

Insgesamt ist hinsichtlich der validen Erhebung für beide Arme gleichermaßen zu berücksichtigen, dass in Ermangelung ausreichender Informationen zur Nutzung des Testmanuals nicht final beurteilt werden kann, ob eine standardisierte Durchführung anhand eines Testmanuals erfolgte. Zudem sind hinsichtlich des Verzerrungspotentials (siehe Tabelle 2) ähnlich wie beim 2MWT und dem GMFM-E mobilitätsspezifische (Alter, Krankheitslast etc.) und kognitive Unterschiede zwischen den Studienarmen zu berücksichtigen, welche den Effekt zugunsten von Pegzilarginase modifizieren könnten. Die dargelegten Baselineunterschiede und die in der Nutzenbewertung bereits thematisierte durchschnittliche Überschreitung der verschriebenen Ernährungsvorgaben im Placeboarm wirken sich auch beim GMFM-D negativ auf das Verzerrungspotential aus. Nachgereichte post hoc durchgeführte Sensitivitätsanalysen unter Herausrechnung der < 5-Jährigen für den GMFM-D deuten jedoch nicht auf verzerrende Aspekte bezüglich der abweichenden Altersverteilungen der < 5-Jährigen zwischen den Studienarmen hin.

In der Gesamtschau wird aufgrund der Darlegungen nicht von einem robusten und klinisch relevanten Effekt zugunsten von Pegzilarginase im Endpunkt GMFM-D ausgegangen.

## 4.2 Lebensqualität

### 4.2.1 PedsQL

Die Fremdbewertung des PedsQL wird wegen geringer Rückläufe und begründeten Zweifeln an der Annahme Missing completely at random (MCAR) nicht berichtet. Die Selbstbewertung des PedsQL wird aufgrund von Validitätseinschränkungen (siehe Anhang) nicht für die Nutzenbewertung herangezogen. Aus Transparenzgründen wird die Selbstbewertung jedoch im Anhang dargestellt.

#### Fremdbewertung

Die Rücklaufquoten sind in beiden Armen gering und unterschreiten im Placebo-Arm ab Woche 12 die 70 % (bezogen auf alle Personen < 18 Jahren). Aufgrund der studienspezifischen Vorgaben zur Anwendung des PedsQL (siehe Kapitel 2.3.3 in der Nutzenbewertung), insbesondere in Abhängigkeit von der psychologischen Beurteilung der Fähigkeit zur Selbstbewertung, kann zudem bei den fehlenden Erhebungen nicht von einer Missing-completely-at-random-Annahme der fehlenden Rückläufe ausgegangen werden. Die Fremdbewertung des PedsQL wird daher nicht zur Nutzenbewertung herangezogen.

## 4.3 Verzerrungspotential GMFM-D

Tabelle 2: Verzerrungspotential des in der Nutzenbewertung nachträglich berücksichtigten Endpunktes GMFM-D der Studie PEACE

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
GMFM-D	Ja	Ja	Nein <sup>1)</sup>	Ja <sup>2)3)4)5)</sup>	Hoch

<sup>1)</sup> Bei dieser sehr kleinen Patientenpopulation (N = 32), insbesondere im Placeboarm (n = 11), zeigt sich durch die Herausrechnung bzw. den Umgang mit fehlenden Werten einer Person ein deutlicher Einfluss auf die Ergebnisse. Die Vorlegung umfangreicher Sensitivitätsanalysen zeigt ein differenziertes Bild, weshalb in der Gesamtschau keine eindeutigen Hinweise für eine ergebnisgesteuerte Berichterstattung vorliegen.

<sup>2)</sup> 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.

<sup>3)</sup> Es zeigt sich eine Ungleichverteilung in zahlreichen Baselinecharakteristika (z. B. Enzymaktivität, Level der Spastik, kognitive Verzögerung, Lebertestabnormalien) 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.

- <sup>4)</sup> 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 Nutzenbewertung Tabelle 25). Da ein ernährungsbedingter Einfluss bei ARG1-D ein relevanter Faktor ist, der die Symptomatik der Patientinnen und Patienten beeinflusst, kann eine Verzerrung nicht ausgeschlossen werden.
- <sup>5)</sup> Bei der Entwicklung des GMFM wurde sich an einem gesunden, motorisch normal entwickelten 5-jährigen Kind orientiert. 5 Kinder im Pegzilarginase-Arm waren < 5 Jahre. Sensitivitätsanalysen unter Ausschluss der Personen unter 5 Jahren wurden vorgelegt. Eine Verzerrung wird nicht angenommen.

Abkürzungen: ARG1-D: Arginase-1-Mangel; GMFM: Gross Motor Function Measure-88; ITT: Intention to Treat

## 5 Zusammenfassung

Im Rahmen der schriftlichen Stellungnahme reichte der pU Sensitivitätsanalysen für den in der Nutzenbewertung ergänzend dargestellten Morbiditätsendpunkt GMFM-D nach sowie hinsichtlich Fremd- und Selbstbewertung getrennte Auswertungen für den in der Nutzenbewertung nicht dargestellten Lebensqualitätsfragebogen PedsQL.

Für die Endpunktkategorie Morbidität zeigte sich für den Endpunkt GMFM-D anhand der dargestellten Analysen, dass sich ein konsistenter positiver Effekt zugunsten von Pegzilarginase abzeichnet. Dieser weist jedoch nicht über alle Analysen hinweg einen signifikanten Unterschied auf, insbesondere nicht in einer als verlässlicher bewerteten Auswertung mit Imputation des fehlenden Baselinewerts anhand des durchschnittlichen Baselinewerts von Personen mit gleichem GMFCS-Level IV (Sensitivitätsanalyse 2; siehe Tabelle 1). Auf Basis der jeweiligen Werte von Hedges'  $g$  (einem möglichen Hilfskriterium zur Einschätzung klinischer Relevanz) lässt sich auch bei den beiden signifikanten Auswertungsergebnissen nicht ableiten, dass ein klinisch relevanter Effekt für den GMFM-D vorliegt. Bei hohem Verzerrungspotential (siehe Tabelle 2) wird in der Gesamtschau nicht von einem robusten und klinisch relevanten Effekt zugunsten von Pegzilarginase im Endpunkt GMFM-D ausgegangen.

Die Nachbewertung des PedsQL in der Endpunktkategorie Lebensqualität ergab, dass die Fremdbewertung, welche grundsätzlich von allen Studienteilnehmenden hätte erhoben werden können, aufgrund von geringen Rückläufen und begründeten Zweifeln an der Missing-Completely-at-Random-Annahme nicht herangezogen werden konnte. Für insgesamt 7 Personen (6 Personen im Pegzilarginase-Arm und 1 Person im Placebo-Arm) lagen neben den Selbsteinschätzungen auch Fremdbewertungen vor. Ob diese Personen in der Lage waren, eine valide Selbsteinschätzung abzugeben, ist unklar und konnte auch vom pU im Stellungnahmeverfahren nicht aufgeklärt werden. Ergebnisse zur Selbstbewertung, welche grundsätzlich bevorzugt werden, werden durch diese Validitätseinschränkungen nicht zur Nutzenbewertung herangezogen, jedoch aus Transparenzgründen im Anhang berichtet.

## Referenzen

1. **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; final clinical study report [unveröffentlicht]. 15.03.2024.
2. **Gemeinsamer Bundesausschuss (G-BA).** Mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses: Pegzilarginase D-1022; stenografisches Wortprotokoll [online]. Berlin (GER): G-BA; 27.05.2024. [Zugriff: 04.06.2024]. URL: [https://www.g-ba.de/downloads/91-1031-1052/2024-05-27\\_Wortprotokoll\\_Pegzilarginase\\_D-1022.pdf](https://www.g-ba.de/downloads/91-1031-1052/2024-05-27_Wortprotokoll_Pegzilarginase_D-1022.pdf).
3. **Immedica Pharma Germany.** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Pegzilarginase (Loargys) [online]. Berlin (GER): Gemeinsamer Bundesausschuss (G-BA); 12.01.2024. [Zugriff: 04.06.2024]. URL: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/1052/#dossier>.
4. **Immedica Pharma Germany.** Schriftliche Stellungnahme zur Nutzenbewertung nach § 35a SGB V und Kosten-Nutzen-Bewertung nach § 35b SGB V: Pegzilarginase (Loargys); Nachreichung [unveröffentlicht]. 06.05.2024.
5. **Immedica Pharma Germany.** Schriftliche Stellungnahme zur Nutzenbewertung nach § 35a SGB V und Kosten-Nutzen-Bewertung nach § 35b SGB V: Pegzilarginase (Loargys) [unveröffentlicht]. 30.05.2024.
6. **Russell D, Wright M, Rosenbaum P, Avery LM.** Gross Motor Function Measure (GMFM-66 and GMFM-88) user's manual; 3rd ed. London (GBR): Mac Keith Press; 2021.
7. **Shieh G.** Confidence intervals and sample size calculations for the standardized mean difference effect size between two normal populations under heteroscedasticity. Behav Res Methods 2013;45(4):955-967.

## Anhang

### PedsQL (Selbstbewertung)

Es lagen weder formalisierte Kriterien noch dokumentierte Gründe für die psychologische Beurteilung der Fähigkeit zur Selbstbewertung vor. Die Einteilung erfolgte in individueller Abwägung auf Basis der fachlichen Expertise des Prüfpersonals bzw. des Psychologen / der Psychologin. Eine Begründung zur Beurteilung der Fähigkeit zur Selbstbewertung wurde nicht systematisch erfasst, weshalb für einen Teil der Studienpopulation ggf. keine validen Erhebungen vorliegen. Insbesondere trifft dies auf Personen zu, für die sowohl eine Selbst- als auch eine Fremdbeurteilung vorlag. Solche überschneidenden Erhebungen liegen für 6 Personen im Pegzilarginase-Arm und für 1 Person im Placebo-Arm vor. Die vorliegenden Rücklaufquoten beziehen sich daher auf alle Studienteilnehmenden im Alter von 5 bis < 18 Jahren, da verlässliche Angaben zur Anzahl an Personen mit der Fähigkeit zur validen Selbstbeurteilung nicht vorliegen. Die diesbezüglichen Rücklaufquoten liegen unter 70 %.

Im Folgenden wird der Gesamtscore des selbstberichteten PedsQL berichtet. Die einzelnen Skalen (physische Funktion, emotionale Funktion, soziale Funktion, schulische Funktion) weisen übereinstimmende Rücklaufzahlen mit dem Gesamtscore auf und werden vorliegend nicht zusätzlich dargestellt. Es zeigt sich in keiner Skala und in keinem weiteren daraus gebildeten Summenwert (physische Gesundheit, psychosoziale Gesundheit) ein signifikanter Effekt.

*Tabelle 3: Veränderung im PedsQL (Selbstbewertung) zu Studienwoche 24 (RCT) in der Studie PEACE, FAS*

<b>Studie PEACE Veränderung im PedsQL<sup>1)</sup> (Selbstbewertung)</b>	<b>Pegzilarginase + IDM N = 15<sup>2)</sup></b>	<b>Placebo + IDM N = 9<sup>2)</sup></b>
Baseline n (%) <sup>2)</sup> Mittelwert (SD)	10 (66,7) 74,9 (12,8)	5 (55,6) 70,2 (20,9)
Woche 24 n (%) <sup>2)</sup> Mittelwert (SD)	9 (60,0) 72,4 (12,9)	4 (44,5) 68,2 (21,0)
Veränderung Woche 24 zu Baseline <sup>3)</sup> n (%) <sup>2)</sup> LS Mean [95%-KI]	9 (60,0) -4,4 [k. A.]	4 (44,5) -7,6 [k. A.]
Gruppenunterschied der Veränderungen <sup>3)</sup> LS Mean [95%-KI] p-Wert <sup>4)</sup>		3,2 [-11,6; 17,9] 0,65

<sup>1)</sup> Wertebereich zwischen 0 bis 100 Punkten, wobei höhere Werte eine bessere Lebensqualität repräsentieren.

<sup>2)</sup> An Studientag 1 waren 5 Kinder aus dem Pegzilarginase-Arm < 5 Jahre alt und 1 Person ≥ 18 Jahre alt. Im Placebo-Arm war keine Person < 5 Jahre und 2 Personen ≥ 18 Jahre alt. Prozentuale Angaben beziehen sich demnach auf 15 (von insgesamt 21) Personen im Pegzilarginasearm und 9 (von insgesamt 11) Personen im Placeboarm, da verlässliche Angaben zur Anzahl an Personen mit der Fähigkeit zur validen Selbstbeurteilung nicht vorliegen.

<sup>3)</sup> Weiterführende Angaben zum Modell und der Berücksichtigung von Effektmodifikatoren fehlen.

<sup>4)</sup> Angaben zum Test fehlen.

Abkürzungen: FAS: Full Analysis Set; IDM: Individuelles Disease Management; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; PedsQL: Pediatric Quality of Life Inventory; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung