

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.**

5. Kapitel § 12 Nr. 1 Satz 2 VerfO

Wirkstoff: Pegvaliase

Dossierbewertung vom 1. Oktober 2019

Datum des Amendments: 19. Dezember 2019

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abbildungsverzeichnis.....	4
Abkürzungsverzeichnis	5
1 Hintergrund.....	6
2 Fragestellung.....	6
3 Liste der verwendeten Quellen.....	7
4 Ergebnisse.....	7
4.1 Informationen zur natürlichen Proteinzufuhr	7
4.2 Charakterisierung der Endpunkte.....	11
4.3 Verzerrungspotential auf Studien- und Endpunktebene	14
4.4 Statistische Methoden	15
4.5 Ergebnisse der Endpunkte „PKU-POMS“ und „ADHD-RS-IV“	16
5 Zusammenfassung.....	20
Referenzen	21

Tabellenverzeichnis

Tabelle 1:	„Natürliche Proteinaufnahme pro Tag" in Studie 165-301 (ITT-Population).....	8
Tabelle 2:	Proteinaufnahme der Studienpopulation der Studie 165-302 zu Baseline des Studienabschnitts 2 (ITT-Population).....	9
Tabelle 3:	Beschreibung der Endpunkte der Kategorie Morbidität	11
Tabelle 4:	Verzerrungspotential der Endpunkte „PKU-POMS“ und „ADHD-RS-IV“ während Studienabschnitt 2 der Studie 165-302	15
Tabelle 5:	Ergebnisse zum Endpunkt „PKU-POMS“ in Studienabschnitt 2 der Studie 165-302 (ITT-Population).....	16
Tabelle 6:	Ergebnisse zum Endpunkt „PKU-POMS“ in Studienabschnitt 4 der Studie 165-302 (ITT-Population).....	18
Tabelle 7:	Ergebnisse zum Endpunkt „ADHD-RS-IV" in Studienabschnitt 2 der Studie 165-302 (ITT-Population).....	18
Tabelle 8:	Ergebnisse zum Endpunkt „Unaufmerksamkeitssubskala des ADHD-RS-IV" in Studie 165-301 (ITT-Population)	20

Abbildungsverzeichnis

Abbildung 1: Durchschnittliche Änderung der täglichen Proteinaufnahme in Gramm aus natürlicher Nahrung zwischen Studienabschnitt 4 der Studie 165-302 im Vergleich zu Baseline der Studie 165-301	10
Abbildung 2: Durchschnittliche Änderung der täglichen Proteinaufnahme in Gramm aus medizinischer Nahrung zwischen Studienabschnitt 4 der Studie 165-302 im Vergleich zu Baseline der Studie 165-301	10

Abkürzungsverzeichnis

ADHD-RS	Attention-Deficit Hyperactivity Disorder Rating Scale
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
MID	Minimal Important Difference
mITT	Modifizierte Intention-to-Treat
Phe	Phenylalanin
PKU	Phenylketonurie
PKU-POMS	PKU-specific Profile of Mood States
POMS	Profile of Mood States
pU	pharmazeutischer Unternehmer
RDA	Recommended Dietary Allowance
SAP	Statistischer Analyseplan
SD	Standardabweichung
SGB	Sozialgesetzbuch
TMD	Total Mood Disturbance
VerfO	Verfahrensordnung des G-BA

1 Hintergrund

Pegvaliase 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: Patientinnen und Patienten ab dem Alter von 16 Jahren mit Phenylketonurie (PKU), deren Phenylalaninwerte im Blut trotz vorausgegangener Anwendung verfügbarer Behandlungsoptionen nicht ausreichend eingestellt sind (Phenylalaninwerte im Blut von über 600 µmol/l).

Der pharmazeutische Unternehmer (pU) hatte am 28. Juni 2019 ein Dossier zur Nutzenbewertung gemäß § 35a SGB V beim Gemeinsamen Bundesausschuss (G-BA) eingereicht. Als bewertungsrelevante Studien wurden die Studien 165-301 und 165-302 für die Nutzenbewertung herangezogen. Bei Studie 165-301 handelt es sich um eine randomisierte, multizentrische Studie der Phase III, in der verschiedene Dosierungen des Wirkstoffs Pegvaliase während einer Induktions-, einer Titrations- und einer Erhaltungsphase miteinander verglichen werden. In beiden Studienarmen erhalten Patientinnen und Patienten Pegvaliase, es unterscheidet sich lediglich die Zieldosierung. Bei Studie 165-302 handelt es sich um eine Studie der Phase III, die sich aus vier verschiedenen Studienabschnitten zusammensetzt. In die Studie wurden Patientinnen und Patienten eingeschlossen, die zuvor eine von mehreren Studien zu Pegvaliase abgeschlossen haben und für mindestens 14 Tage eine stabile Dosis Pegvaliase erhalten haben.

Die Nutzenbewertung wurde am 1. Oktober 2019 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Eine mündliche Anhörung fand am 11. November 2019 statt, in der offene Fragen zu den in den schriftlichen Stellungnahmen dargelegten Daten und Argumentationen erörtert wurden.

In der Nutzenbewertung vom 1. Oktober 2019 wurden die Endpunkte „PKU-POMS“ und „ADHD-RS-IV“ nicht berücksichtigt, da erhebliche Unsicherheiten bezüglich der Validität der Endpunkte bestehen. Im Rahmen des schriftlichen Stellungnahmeverfahrens wurden für den „ADHD-RS-IV“ das Manual des Instruments sowie Schulungsunterlagen nachgereicht. Der Endpunkt „natürliche Proteinzufuhr“ wurde in der Nutzenbewertung nicht per se als patientenrelevant erachtet, da sich Änderungen der Proteinzufuhr in patientenrelevanten Endpunkten der Endpunktkategorien Morbidität, Lebensqualität oder Sicherheit widerspiegeln sollten.

Der Unterausschuss Arzneimittel beauftragte die Darstellung der Ergebnisse zu den Endpunkten „PKU-POMS“ und „ADHD-RS-IV“ für die Studienabschnitte 2 und 4 der Studie 165-302. Ebenso beauftragte der Unterausschuss Arzneimittel die Darstellung der „natürlichen Proteinzufuhr“ als Kriterium zur Beurteilung der Qualität der Studie.

2 Fragestellung

In diesem Amendment werden die vom Unterausschuss Arzneimittel beauftragten Ergebnis- und Datendarstellungen präsentiert.

3 Liste der verwendeten Quellen

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

- Schriftliche Stellungnahme des pU vom 22.10.2019 [2]
- Herstellerdossier zu Pegvaliase [3]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie 165-301 [5]
- Studienbericht inkl. Studienprotokoll und SAP der Studie 165-302 [4]

4 Ergebnisse

Wie in der Nutzenbewertung vom 01.10.2019 beziehen sich alle Angaben zur Studie 165-302 auf den Studienbericht mit Datenschnitt vom 23.09.2016.

4.1 Informationen zur natürlichen Proteinzufuhr

In den Studien 165-301 und 165-302 war gemäß Studienprotokoll vorgegeben, dass Patientinnen und Patienten die Proteinaufnahme im Studienverlauf möglichst beibehalten sollten. Dies galt als erfüllt, wenn sich die Proteinaufnahme sowohl aus natürlichen als auch medizinischen Nahrungsquellen um weniger als 10 % änderte. Bei größeren Abweichungen ($\geq 25\%$) griff die verantwortliche Diätassistentin / der verantwortliche Diätassistent beratend ein und konnte nach Absprache mit dem Monitor ggf. weitere Maßnahmen einleiten. Fiel der Phenylalanin (Phe)-Spiegel auf $30\ \mu\text{mol/l}$ oder niedriger ab, änderte die Diätassistentin / der Diätassistent die Diätvorgaben in Abhängigkeit davon, ob die Proteinaufnahme der Patientin / des Patienten im Normbereich war oder nicht.

Sofern der Phe-Spiegel im Blut auf $\leq 30\ \mu\text{mol/l}$ sank und festgestellt wurde, dass eine Person weniger als die gemäß RDA („recommended dietary allowance“) empfohlene Menge an natürlichem Protein zu sich genommen hat (RDA für Protein aus natürlicher Nahrung: Männer mit einem Alter von 18 Jahren: 52 g/Tag, Männer ≥ 19 Jahre: 56 g/Tag; Frauen mit einem Alter von 18 Jahren: 46 g/Tag, Frauen ≥ 19 Jahre: 46 g/Tag), wird die Diätassistentin / der Diätassistent die Person anweisen, die Zufuhr von Protein aus natürlicher Nahrung um 10 g zu erhöhen und die Zufuhr von Protein aus medizinischer Nahrung um 5 g zu verringern. Diese erhöhte Zufuhr stellte anschließend die neue Referenzgrundlage für die Aufnahme von Nahrungsproteinen dar.

Sofern der Phe-Spiegel im Blut auf $\leq 30\ \mu\text{mol/l}$ sank und festgestellt wurde, dass eine Person mehr als die gemäß RDA empfohlene Menge, aber weniger als das Doppelte der RDA, an Protein aus natürlicher Nahrung zu sich genommen hat, wird die Diätassistentin / der Diätassistent die Person anweisen, die Zufuhr von Protein aus natürlicher Nahrung um 10 % zu erhöhen und die Zufuhr an Protein aus medizinischer Nahrung um 5 g zu verringern. Diese erhöhte Zufuhr stellte anschließend die neue Referenzgrundlage für die Aufnahme von Nahrungsproteinen dar.

Sofern der Phe-Spiegel im Blut auf $\leq 30\ \mu\text{mol/l}$ sank und festgestellt wurde, dass eine Person mehr als das Doppelte der gemäß RDA empfohlenen Menge an Protein aus natürlicher Nahrung zu sich genommen hat, wird die Diätassistentin / der Diätassistent die Person anweisen, die aktuelle Zufuhr von Protein aus natürlicher Nahrung und aus medizinischer Nahrung beizubehalten. Die gleiche Referenzgrundlage sollte für den weiteren Verlauf herangezogen werden.

Die Patientinnen und Patienten erhielten ein Diättagebuch, in dem die Nahrungszufuhr (Nahrung, Getränke, Spezialnahrung mit wenig Protein und medizinische Nahrung) für 3 aufeinanderfolgende Tage vor Visite aufgezeichnet werden sollte. Mithilfe eines webbasierten Softwareprogramms zur Nährstoffanalyse (Metabolic Pro®) wurde auf Basis dieser Angaben u. a. die Proteinaufnahme berechnet.

Studie 165-301

In den Studienunterlagen der Studie 165-301 wurden keine Informationen dazu identifiziert, bei wie vielen Personen eine Anpassung der Proteinzufuhr aufgrund eines geringen Phe-Spiegels ($\leq 30 \mu\text{mol/l}$) vorgenommen wurde oder für wie viele Personen insgesamt eine Anpassung der Proteinzufuhr, d. h. möglicherweise auch aus anderen Gründen, vorgenommen wurde.

Zu Baseline der Studie betrug die tägliche Proteinaufnahme aus natürlicher Nahrung im Durchschnitt 38,5 g (Standardabweichung (SD): 27,7; Median: 29,9 g [min; max: 4; 156]). Die Änderungen der Proteinaufnahme aus natürlicher Nahrung können Tabelle 1 entnommen werden. Nach Woche 16 lagen diesbezügliche Angaben nur noch für $< 70 \%$ der Studienpopulation vor. Die tägliche Proteinaufnahme aus medizinischer Nahrung betrug zu Baseline im Durchschnitt 26,3 g (SD: 28,5; Median: 16,8 g [min; max: 0; 120]). Für die Proteinaufnahme aus medizinischer Nahrung wurden keine Angaben zur Veränderung im Vergleich zu Baseline identifiziert. 41 Personen (15,7 %) erhielten zu Baseline eine proteinrestriktive Diät mit $\geq 75 \%$ der täglichen Proteinaufnahme aus medizinischer Nahrung.

Tabelle 1: „Natürliche Proteinaufnahme pro Tag“ in Studie 165-301 (ITT-Population)

Studie 165-301	Gesamt N = 261
Baseline	
n	250
Mittelwert (SD) [g]	38,5 (27,7)
Median (min; max) [g]	29,9 (k. A.; k. A.)
Veränderung Woche 4 zu Baseline¹⁾	
n	232
Mittelwert (SD) [g]	1,5 (15,9)
Median (min; max) [g]	0,6 (k. A.; k. A.)
Veränderung Woche 8 zu Baseline¹⁾	
n	232
Mittelwert (SD) [g]	4,1 (18,9)
Median (min; max) [g]	2,0 (k. A.; k. A.)
Veränderung Woche 12 zu Baseline¹⁾	
n	225
Mittelwert (SD) [g]	1,8 (18,1)
Median (min; max) [g]	0,5 (k. A.; k. A.)
Veränderung Woche 16 zu Baseline¹⁾	
n	195
Mittelwert (SD) [g]	2,4 (17,2)
Median (min; max) [g]	1,4 (k. A.; k. A.)

¹⁾ Die Veränderung gegenüber Baseline basierte auf Personen mit verfügbaren Messungen zu beiden Zeitpunkten.

Abkürzungen: ITT: Intention-to-Treat; k. A.: keine Angabe; SD: Standardabweichung.

Studie 165-302

Die Proteinaufnahme aus natürlicher und aus medizinischer Nahrung zu Baseline des Studienabschnitts 2 der Studie 165-302 kann Tabelle 2 entnommen werden.

Tabelle 2: Proteinaufnahme der Studienpopulation der Studie 165-302 zu Baseline des Studienabschnitts 2 (ITT-Population)

Studie 165-302	Pegvaliase 20 mg/Tag N = 34	Pegvaliase 40 mg/Tag N = 32	Low dose Placebo N = 15	High dose Placebo N = 14
Proteinrestriktive Diät ¹⁾ , n (%)	1 (2,9)	1 (3,1)	1 (6,7)	3 (21,4)
Tägliche Proteinaufnahme aus medizinischer Nahrung (g)				
n	34	32	15	14
Mittelwert (SD)	10,4 (19,5)	17,3 (21,9)	19,4 (26,0)	28,4 (26,5)
Median (min; max)	0,0 (0,0; 75,0)	1,7 (0,0; 63,3)	15,8 (0,0; 90,0)	25,0 (0,0; 73,2)
Tägliche Proteinaufnahme aus natürlichen Quellen (g)				
n	34	32	15	14
Mittelwert (SD)	48,2 (20,0)	52,4 (32,3)	40,1 (26,6)	39,4 (22,7)
Median (min; max)	47,1 (14,7; 96,2)	42,8 (6,1; 141,4)	25,7 (12,4; 94,7)	36,1 (8,1; 77,4)

¹⁾ Eine proteinrestriktive Diät lag vor, wenn eine Person mehr als 75 % der täglichen Proteinaufnahme aus medizinischer Nahrung aufgenommen hatte. Die gesamte Proteinaufnahme berechnet sich als Summe aus der Proteinaufnahme aus medizinischer und aus natürlicher Nahrung.

Abkürzungen: ITT: Intention-to-Treat; SD: Standardabweichung.

Gemäß klinischem Review traten wichtige Protokollverletzungen („important deviations“) in der mITT-Population während Studienabschnitt 2 der Studie 165-302 bei 18 von 86 in diesen Studienabschnitt eingeschlossenen und zur mITT-Population zählenden Personen aufgrund von Abweichungen von der vorgegebenen Proteinaufnahme auf. Angaben zur Änderung der Proteinaufnahme wurden für die ITT-Population während Studienabschnitt 2 nicht identifiziert. Auf Basis von Untersuchungen der mITT-Population scheinen die sich zwischen den Studienarmen unterscheidenden Änderungen keinen bedeutenden Einfluss auf die primären Wirksamkeitsergebnisse (Phe-Konzentration im Blut) der Studie gehabt zu haben.

Insgesamt wiesen 94 von 215 (43,7 %) in die gesamte Studie 165-302 eingeschlossene Personen zu mindestens zwei aufeinanderfolgenden Messungen Phe-Spiegel $\leq 30 \mu\text{mol/l}$ auf, sodass gemäß Studienprotokoll eine Anpassung der Proteinzufuhr möglich wurde. Bei wie vielen dieser Personen eine Anpassung der Proteinzufuhr vorgenommen wurde, wurde in den Studienunterlagen nicht identifiziert. Ebenfalls wurde nicht identifiziert, für wie viele Personen insgesamt eine Anpassung der Proteinzufuhr, d. h. möglicherweise auch aus anderen Gründen, vorgenommen wurde.

Die durchschnittliche Änderung der täglichen Proteinaufnahme aus natürlicher bzw. medizinischer Nahrung zwischen Studienabschnitt 4 der Studie 165-302 im Vergleich zu Baseline der Studie 165-301 kann Abbildung 1 bzw. 2 entnommen werden. Es ist allerdings zu beachten, dass zu keinem Zeitpunkt Daten für ≥ 70 % der in Studie 165-301 eingeschlossenen Personen vorliegen.

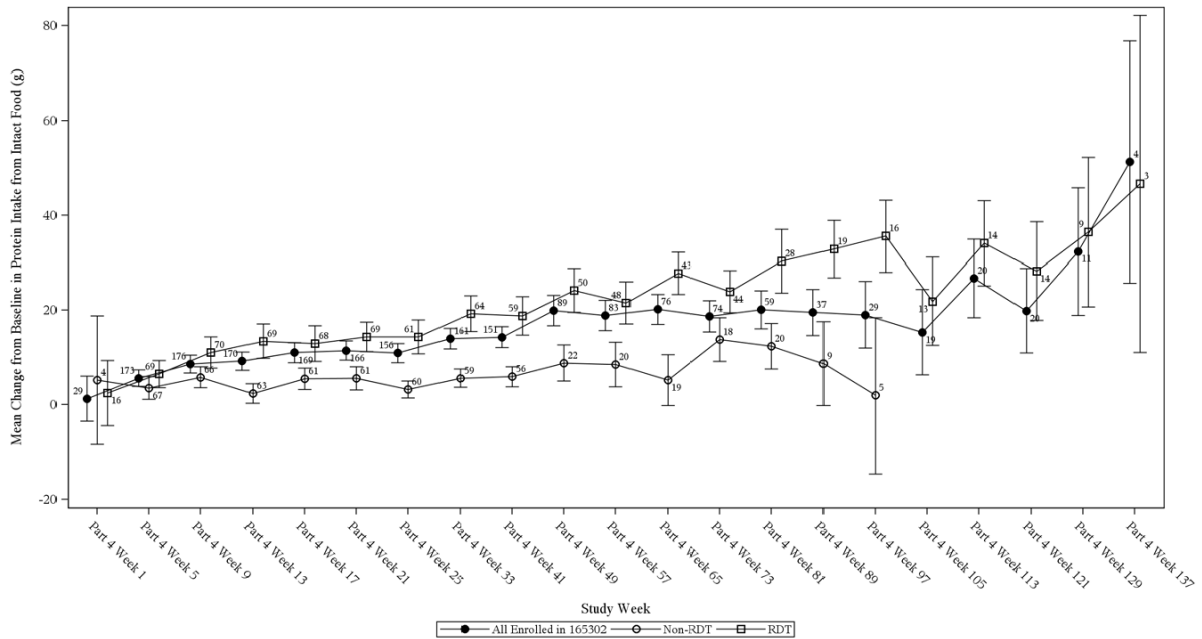


Abbildung 1: Durchschnittliche Änderung der täglichen Proteinaufnahme in Gramm aus natürlicher Nahrung zwischen Studienabschnitt 4 der Studie 165-302 im Vergleich zu Baseline der Studie 165-301

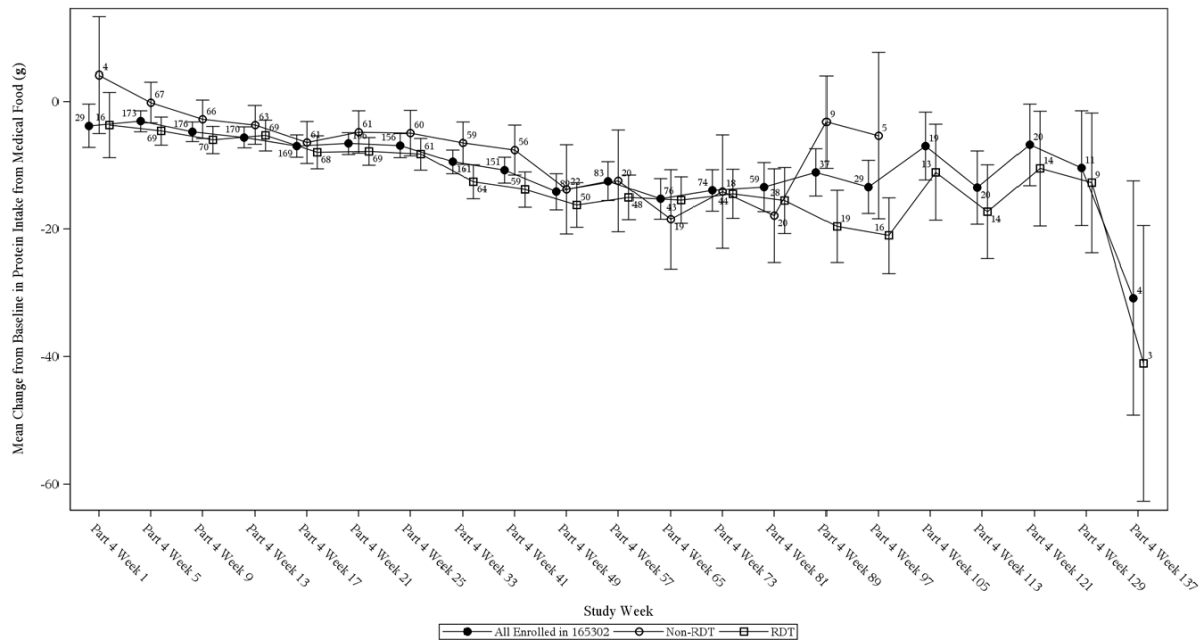


Abbildung 2: Durchschnittliche Änderung der täglichen Proteinaufnahme in Gramm aus medizinischer Nahrung zwischen Studienabschnitt 4 der Studie 165-302 im Vergleich zu Baseline der Studie 165-301

4.2 Charakterisierung der Endpunkte

Tabelle 3: Beschreibung der Endpunkte der Kategorie Morbidität

Endpunktkategorie Morbidität	
Endpunkt	PKU-POMS
Operationalisierung	<p>Der PKU-POMS wurde vom pU zur Beurteilung von Stimmungsschwankungen bei Erwachsenen mit PKU entwickelt [1]. Der Fragebogen besteht aus 20 Adjektiven, die 6 Subskalen zugeordnet sind:</p> <ul style="list-style-type: none"> • Angst (4 Items), • Depression (4 Items), • Wut (3 Items), • Aktivität (3 Items), • Müdigkeit (3 Items) und • Verwirrung (3 Items). <p>Laut Modul 4 des Herstellerdossiers wird der PKU-POMS-Fragebogen mit der gleichen 5-Punkte-Skala bewertet wie das ursprüngliche POMS-Tool. Ein Gesamtwert für den PKU-POMS (Total Mood Disturbance (TMD) Score) kann als Summe der 5 Subskalen-Scores Angst, Depression, Wut, Müdigkeit und Verwirrung und unter Abzug des Scores der Subskala Aktivität berechnet werden. Der PKU-POMS-TMD umfasst einen Score-Bereich von -12 bis 58 Punkten und die Verwirrtheits-Subskala einen Bereich von 0 bis 12 Punkten. Höhere Werte gehen mit einer stärkeren Beeinträchtigung einher. Es wurde in den Studien der ursprüngliche POMS-Fragebogen eingesetzt und eine Teilmenge der beantworteten Items wurde genutzt, um Werte für den PKU-POMS zu berechnen.</p> <p>Erhebungszeitpunkte</p> <p>Da der gesamte POMS erhoben wurde und eine Teilmenge der beantworteten Items genutzt wurde, um Werte für den PKU-POMS zu berechnen, entsprechen die Erhebungszeitpunkte denen des POMS.</p> <p>In Studie 165-301 wurde der POMS gemäß originalem Studienprotokoll nicht eingesetzt. Mit Amendment 1 des Studienprotokolls wurde das Instrument an Tag 1 der Induktion sowie zur letzten Visite bzw. bei vorzeitigem Abbruch eingesetzt.</p> <p>In Studie 165-302 wurde das Instrument während Studienabschnitt 2 an Tag 1, zu Woche 4 und zu Woche 8 eingesetzt. Während Studienabschnitt 4 wurde es in Woche 1, 9, 17, alle 8 Wochen ab Woche 25 sowie zum Abschluss der Studie eingesetzt.</p> <p>Auswertungspopulation</p> <p>Für die Nutzenbewertung ist die ITT-Population die relevante Auswertungspopulation.</p>
Bewertung	<p>Operationalisierung</p> <p>In Modul 4 des Herstellerdossiers ist angegeben, dass der PKU-POMS-Fragebogen mit der gleichen 5-Punkte-Skala bewertet wird wie der POMS. In der vom pU eingereichten Publikation zur Validierungsstudie ist allerdings angegeben, dass für einzelne Items das Antwortformat angepasst wurde, indem Antwortoptionen zusammengelegt wurden [1]. Das Instrument wurde in Studie 165-301 erst ab Amendment 1 des Studienprotokolls eingesetzt. Aus diesem Grund liegen für 91 der 261 (35 %) in die Studie eingeschlossenen Personen für dieses Instrument keine Baseline-Werte vor.</p> <p>Validität</p> <p>Der PKU-POMS wurde in einem mehrstufigen Prozess und unter Beteiligung von Betroffenen entwickelt. Ausgangspunkt der Entwicklung war die Originalversion des POMS.</p> <p>Die 58 Items, die den 6 Subskalen des POMS zugeordnet waren, wurden zunächst in qualitativen Interviews mit 15 an PKU erkrankten Personen hinsichtlich ihrer Verständlichkeit, Akzeptanz und Relevanz im Erkrankungsbild untersucht. Die Erkrankten wurden durch Verteilung von Flyern auf einer nationalen PKU-Konferenz im Jahr 2014 und durch Posten eines Flyers auf der Homepage der Studie rekrutiert. Es ist nicht</p>

Endpunktkategorie Morbidität	
	<p>beschrieben, wie viele der Erkrankten an den Studien 165-301 und 165-302 teilnahmen. 17 Items wurden in diesem Schritt entfernt. Anschließend quantitative Untersuchungen der verbleibenden Items mit 115 Teilnehmenden der Studie 165-301 wurden mittels konfirmatorischer Faktorenanalyse und probabilistischer Verfahren durchgeführt. 21 weitere Items wurden in diesem Schritt entfernt. Die finalen 20 Items zeigten die besten Eigenschaften hinsichtlich Unidimensionalität und Güte der Passung der Daten an das Modell. Bei zwei der Items wurde dazu das Antwortformat angepasst, indem Antwortoptionen zusammengelegt wurden [1].</p> <p>Die so ausgewählten 20 Items bilden den PKU-POMS und wurden mit den Daten der Studie 165-301 hinsichtlich ihrer psychometrischen Eigenschaften untersucht. In Studie 165-301 wurden diese 20 Items nicht als einzelnes Instrument eingesetzt, sondern als Bestandteil des 65 Items umfassenden POMS. Schritte der Fragebogenentwicklung und die Auswertung der psychometrischen Eigenschaften fanden zum einen im gleichen Studiensample statt. Zum anderen wurde der Fragebogen anhand derjenigen Patientinnen und Patienten validiert, die an den Studien 165-301 und 165-302 teilnahmen. Somit liegt für den Fragebogen keine Validierung vor, die an einer externen Population vorgenommen wurde. Die Daten der Studie 165-301 wurden gleichzeitig zur Validierung des PKU-POMS und als Ergebnisse der Studie herangezogen. Zudem handelt es sich beim PKU-POMS, wie er in den Studien 165-301 und 165-302 eingesetzt wurde, nicht um ein separates Instrument, das 20 Items umfasst. In beiden Studien wurde der gesamte POMS-Fragebogen eingesetzt und eine Teilmenge der beantworteten Items wurde genutzt, um Werte für den PKU-POMS zu berechnen. Somit wurden die Antworten auf die 20 im PKU-POMS enthaltenen Items zur Berechnung von zwei Endpunkten verwendet.</p> <p>Der pU gibt in Modul 4 des Herstellerdossiers sowie in der schriftlichen Stellungnahme an, MID für den TMD des PKU-POMS sowie für die Subskalen des PKU-POMS ermittelt zu haben. Zur Einschätzung der Relevanz der Veränderungen zwischen Studienabschnitt 4 der Studie 165-302 und Baseline der Studie 165-301 zieht er einen Wert heran, der verteilungsbasiert anhand der Patientinnen und Patienten berechnet wurde, die in die Studie 165-301 und zum Großteil auch in Studie 165-302 eingeschlossen wurden. Sowohl die Herleitung einer MID anhand der Patientinnen und Patienten der relevanten Zulassungsstudien als auch die Herleitung einer MID anhand von verteilungsbasierten Verfahren wird als nicht adäquat eingeschätzt. Daher werden die auf diese Weise ermittelten Werte für die Ergebnisinterpretation nicht herangezogen.</p> <p>Patientenrelevanz</p> <p>Grundsätzlich werden die mit dem Instrument erfassten Stimmungszustände und -veränderungen als patientenrelevant eingeschätzt.</p> <p>Die Ergebnisse der Studie 165-301 wurden gleichzeitig zur Entwicklung und Validierung des PKU-POMS herangezogen. Somit wurde die Untersuchung der psychometrischen Eigenschaften nicht anhand von Personen vorgenommen, die nicht an den Studien 165-301 und 165-302 teilnahmen. Zudem handelt es sich bei PKU-POMS nicht um einen Item-reduzierten Fragebogen. Stattdessen wurde der originale POMS-Fragebogen mit 65 Items in den Studien 165-301 und 165-302 eingesetzt und eine Teilmenge der beantworteten Items wurde genutzt, um Werte für den PKU-POMS zu berechnen. Aus diesen Gründen ist die Validität der Ergebnisse fraglich.</p>
Endpunkt	ADHD-RS-IV
Operationalisierung	<p>Beim ADHD-RS-IV handelt es um ein Instrument zur Fremdbeurteilung von Unaufmerksamkeits- und Hyperaktivitätssymptomen, das primär zur Diagnose von ADHS bei Kindern und Jugendlichen entwickelt wurde. Das Instrument umfasst zwei Subskalen – Unaufmerksamkeit und Hyperaktivität. Jede dieser Subskalen enthält 9 Items, die auf einer 4-Punkte-Likert-Skala bewertet werden (0 = nie oder selten / 1 = manchmal / 2 = oftmals / 3 = sehr häufig). Die Subskalen können Werte im Bereich von 0 (keine Beeinträchtigung) bis 27 (größtmögliche Beeinträchtigung) annehmen.</p> <p>Der ADHD-RS-IV wurde als semistrukturiertes Interview entwickelt. Die interviewenden Personen, in den relevanten Studien entsprechen diese dem Prüfpersonal, fragen die</p>

Endpunktkategorie Morbidität	
	<p>Betroffene / den Betroffenen bezüglich jedes Items. Anschließend treffen die interviewenden Personen eine klinische Bewertung bezüglich jeden Items, die nicht vollumfänglich dem Bericht der/des Betroffenen entsprechen muss [8]. In Modul 4 des Herstellerdossiers ist angegeben, dass der vorangegangene Monat zur Bewertung herangezogen wurde. Das Manual des Instruments und Schulungsunterlagen für das Prüfpersonal wurden im Rahmen des schriftlichen Stellungsverfahren nachgereicht.</p> <p>Erhebungszeitpunkte In Studie 165-301 sollte der ADHD-RS-IV gemäß originalem Studienprotokoll einmalig zu Baseline eingesetzt werden. Mit Amendment 1 des Studienprotokolls wurde das Instrument an Tag 1 der Induktion sowie zur letzten Visite bzw. bei vorzeitigem Abbruch eingesetzt. In Studie 165-302 wurde das Instrument während Studienabschnitt 2 an Tag 1, zu Woche 4 und zu Woche 8 eingesetzt. Während Studienabschnitt 4 wurde es in Woche 1, 9, 17, alle 8 Wochen ab Woche 25 sowie zum Abschluss der Studie eingesetzt.</p> <p>Auswertungspopulation Für die Nutzenbewertung ist die ITT-Population die relevante Auswertungspopulation.</p>
Bewertung	<p>Operationalisierung Die Operationalisierung ist nachvollziehbar.</p> <p>Validität Für den ADHD-RS-IV liegt eine Studie zur Untersuchung der psychometrischen Eigenschaften anhand von Patientinnen und Patienten mit PKU vor, die nicht denen der Zulassungsstudien von Pegvaliase entsprechen. Das Instrument wurde in dieser Studie allerdings nur bei Kindern eingesetzt. Bei erwachsenen Patientinnen und Patienten mit PKU wurde in dem RCT, der dieser psychometrischen Studie zugrunde liegt, die Adult ADHD Self-Report Scale eingesetzt [7]; ein Instrument, das von den Betroffenen selbst ausgefüllt wird. Es wurden in den vom pU eingereichten Unterlagen keine Untersuchungen der psychometrischen Eigenschaften des ADHD-RS-IV anhand von erwachsenen Patientinnen und Patienten mit PKU identifiziert. Beim ADHD-RS-IV handelt es sich um ein Instrument zur Fremdbeurteilung. Grundsätzlich sind Instrumente zur Selbsteinschätzung zu bevorzugen. Unter Expertinnen und Experten besteht gemäß einer Studie Dissens bezüglich der Frage, wie gut Patientinnen und Patienten mit PKU ihre Symptomatik selbst wahrnehmen. Es wird allerdings konstatiert, dass die Fähigkeit zur Selbsteinschätzung mit stärkerer kognitiver Einschränkung, mehr Symptomen und höheren Phe-Werten abzunehmen scheint [8]. Auf Basis der vorhandenen Baseline-Charakteristika der Studienpopulation von Studie 165-301 ist nicht eindeutig ersichtlich, dass diese Personen grundsätzlich nicht zur Selbsteinschätzung in der Lage wären. Es wurden keine Charakteristika identifiziert, die Rückschlüsse zur kognitiven Einschränkung zulassen. Bezüglich der Unaufmerksamkeit erfasst mittels ADHD-RS-IV erreichen die Patientinnen und Patienten im Mittel ungefähr einen Wert im unteren Drittel der Skala (MW: 9,6 Punkte; Median: 8,0 Punkte; Skala: 0–27 Punkte). Die Phe-Konzentration im Blut ist zu Baseline deutlich erhöht und beträgt im Mittel 1.232,7 µmol/l (Median: 1.221,0 µmol/l). In der bereits erwähnten Validierungsstudie wurde ein von den erwachsenen Betroffenen selbst ausgefülltes Instrument eingesetzt [7]. Die erwachsene Population dieser Studie wies im Mittel eine niedrigere Phe-Konzentration auf als die Population in Studie 165-301 (MW: 1.038,0 µmol/l; Median: 961,0 µmol/l), die insgesamt aber dennoch als deutlich erhöht einzuschätzen ist. Bezüglich der Unaufmerksamkeit erfasst mittels ADHD Self-Report Scale erreichten die Patientinnen und Patienten ebenfalls ungefähr einen Wert im unteren Drittel der Skala (MW: 13,5 Punkte; Median: 13,1 Punkte; Skala: 0–36 Punkte). Auch auf Basis dieser Daten ist nicht eindeutig ersichtlich, dass Patientinnen und Patienten in der Studie 165-301 nicht in der Lage wären, eine Selbsteinschätzung durchzuführen. Untersuchungen der Inter-Rater-Reliabilität wurden nicht identifiziert. Anhand von insgesamt 501 erwachsenen Personen mit ADHS, die an einem RCT teilnahmen, in dem Atomoxetin (ein Wirkstoff zur Behandlung der ADHS) mit Placebo ver-</p>

Endpunktkategorie Morbidität	
	<p>glichen wurde, wurden MID, d. h. individuell relevante Veränderungen, des ADHD-RS-IV sowie dessen Subskalen ermittelt. Es wurde eine ankerbasierte Herleitung angewendet. Als Anker wurde die globale Einschätzung des Schweregrads der Erkrankung durch das klinische Personal verwendet (Clinical Global Impression of Severity [CGI-S]). Das klinische Personal schätzte den Schweregrad unter Berücksichtigung der eigenen Erfahrung mit ADHS-Erkrankten ein. Der CGI-S umfasste 7 Ausprägungen (1 = normal bis 7 = zählt zu den am schwersten betroffenen Personen). Der CGI-S wurde zu jeder Visite eingesetzt; Visiten fanden eine Woche vor erster Einnahme der Studienmedikation, zu Baseline sowie nach 2, 4, 6, 10, 14, 22 und 26 Behandlungswochen statt [6]. Für jede Studienteilnehmende / jeden Studienteilnehmenden wurden Veränderungen seit Baseline unter Verwendung des Last-Observation-Carried-Forward-Ansatzes ermittelt. Personen mit einer Veränderung von 1 Punkt auf dem CGI-S wurden als Anker für die Berechnung der MID für den ADHD-RS-IV und dessen Subskalen verwendet. Die Mittelwerte auf diesen Skalen von Personen, die gemäß CGI-S eine Veränderung von 1 Punkt aufwiesen, wurden als MID bezeichnet. Für den ADHD-RS-IV ergab sich so eine MID von 12,5 Punkten, für die Unaufmerksamkeitsskala eine MID von 6,8 Punkten und für die Hyperaktivitätsskala eine MID von 5,6 Punkten [6].</p> <p>Patientenrelevanz</p> <p>Grundsätzlich werden die mit dem Instrument erfassten Unaufmerksamkeits- und Hyperaktivitätssymptome als patientenrelevant eingeschätzt. Die Validität der Ergebnisse ist aufgrund der beschriebenen Unsicherheiten inklusive der ausschließlichen Fremdbeurteilung allerdings fraglich.</p>

Abkürzungen: ADHD-RS: Attention-Deficit Hyperactivity Disorder Rating Scale; ADHS: Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung; ITT: Intention-to-Treat; MID: Minimal important Difference; MW: Mittelwert; Phe: Phenylalanin; PKU: Phenylketonurie; PKU-POMS: PKU-specific Profile of Mood States; POMS: Profile of Mood States; pU: pharmazeutischer Unternehmer; RCT: Randomisiert kontrollierte Studie; TMD: Total Mood Disturbance.

4.3 Verzerrungspotential auf Studien- und Endpunktebene

Bei Studienabschnitt 4 der Studie 165-302 handelt es sich um eine Langzeitextension ohne Kontrollarm. Insofern wird das Verzerrungspotential für diesen Studienabschnitt und die während dieses Studienabschnitts untersuchten Endpunkte als hoch eingeschätzt. Bei Studienabschnitt 2 handelt es sich um ein randomisiertes, doppelblindes, placebokontrolliertes, vierarmiges Studiendesign, in dem die zuvor verabreichte Medikation in 2 der 4 Behandlungsarmen abgesetzt und durch Placebo ersetzt wird. Das Verzerrungspotential auf Studienebene wird für diesen Studienabschnitt als niedrig eingeschätzt (siehe Nutzenbewertung vom 01.10.2019). Details der Einschätzung des Verzerrungspotentials der Endpunkte „PKU-POMS“ und „ADHD-RS-IV“ kann Tabelle 4 entnommen werden.

Tabelle 4: Verzerrungspotential der Endpunkte „PKU-POMS“ und „ADHD-RS-IV“ während Studienabschnitt 2 der Studie 165-302

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
PKU-POMS	Ja ¹⁾	Ja	Ja ²⁾	Ja ³⁾	Niedrig
ADHD-RS-IV	Ja ¹⁾	Ja	Ja ²⁾	Ja ³⁾	Niedrig

¹⁾ Die subkutan zu verabreichende Placebo-Lösung entspricht bezüglich Aussehen und Konsistenz dem Prüfpräparat. Patientinnen und Patienten, Studienpersonal, privates Pflegepersonal sowie der pU bleiben verblindet, bis alle Teilnehmenden Studienabschnitt 2 abgeschlossen haben.

²⁾ Der SAP ist auf den 24.02.2016 datiert. Zu diesem Zeitpunkt war die Rekrutierung für die Studie bereits abgeschlossen. Im Studienbericht ist angegeben, dass der SAP finalisiert wurde, bevor die Entblindung für Studienabschnitt 2 stattfand. Es wird nicht von einem bedeutenden Einfluss auf das Verzerrungspotential ausgegangen.

³⁾ Ein Carryover-Effekt in den Placebogruppen, der daraus resultiert, dass alle Patientinnen und Patienten bis zum Beginn von Studienabschnitt 2 Pegvaliase erhalten haben, ist nicht auszuschließen. Dies könnte potentiell zu einer Unterschätzung des Effekts von Pegvaliase im Vergleich zu Placebo führen. Es wird nicht von einem bedeutenden Einfluss auf das Verzerrungspotential ausgegangen.

Abkürzungen: ADHD-RS: Attention-Deficit Hyperactivity Disorder Rating Scale; ITT: Intention-to-Treat; PKU-POMS: PKU-specific Profile of Mood States; pU: pharmazeutischer Unternehmer; SAP: Statistischer Analyseplan.

4.4 Statistische Methoden

Statistische Analysen

Studie 165-302 (Studienabschnitt 2)

Die Veränderung der sekundären Endpunkte seit Baseline wird mittels eines gemischten Modells für wiederholte Messungen (Mixed Model Repeated Measures [MMRM]) analysiert. Dieses Modell beinhaltet die Behandlungsgruppe, die Visite, die Interaktion aus Behandlung und Visite und den gemessenen Wert zu Baseline als Faktoren.

Sensitivitätsanalysen

Studie 165-302

Für die sekundären Endpunkte werden die gleichen Analysen durchgeführt wie für den primären Endpunkt (siehe Nutzenbewertung vom 01.10.2019).

Subgruppen

Studie 165-302

Zur Analyse von Subgruppeneffekten wird das Modell der primären Analyse um einen Interaktionsterm aus Subgruppenkovariate und Behandlung erweitert. Für die sekundären Endpunkte handelt es sich um folgende Subgruppen:

- Unaufmerksamkeit zu Baseline, gemessen auf der ADHD-RS-IV (≤ 9 oder > 9 Punkte) für POMS-bezogene Endpunkte
- Geschlecht: weiblich oder männlich
- BMI: < 25 oder ≥ 25 und < 30 oder ≥ 30

Diese Analysen wurden für die ITT-Population nicht identifiziert.

Umgang mit fehlenden Werten

Studie 165-302

PKU-POMS

Falls für ein Item ein Wert fehlen sollte, wird der Mittelwert der beantworteten Items der gleichen Subskala verwendet, um den fehlenden Wert zu ersetzen. Falls in einer Subskala Werte für mehr als ein Item fehlen sollten oder falls Werte für mehr als 2 Messungen des gesamten Fragebogens fehlen sollten, wird kein Wert imputiert, sondern die Werte für die Subskala und den Total Mood Disturbance (TMD) als fehlend gewertet.

4.5 Ergebnisse der Endpunkte „PKU-POMS“ und „ADHD-RS-IV“

PKU-POMS

Die Ergebnisse während Studienabschnitt 2 der Studie 165-302 für den PKU-POMS-TMD und die Verwirrtheitssubskala des PKU-POMS können Tabelle 5 entnommen werden. Sowohl für den PKU-POMS-TMD als auch für die Verwirrtheitssubskala zeigten sich nach 8 Wochen zwischen Pegvaliase und dem entsprechenden Placebo jeweils keine statistisch signifikanten Unterschiede in der Veränderung seit Baseline des Studienabschnitts 2. Für die Subskalen „Angst“, „Depression“, „Wut“, „Aktivität“ und „Müdigkeit“ wurden keine Analysen für Studienabschnitt 2 auf Basis der ITT-Population identifiziert.

Tabelle 5: Ergebnisse zum Endpunkt „PKU-POMS“ in Studienabschnitt 2 der Studie 165-302 (ITT-Population, $N_{ges} = 95$)

Studie 165-302	20 mg Pegvaliase N = 34	Low dose Placebo N = 15	40 mg Pegvaliase N = 32	High dose Placebo N = 14
PKU-POMS-TMD				
Baseline Studienabschnitt 2				
Mittelwert (SD)	7,8 (14,4)	9,9 (11,5)	9,2 (12,5)	5,0 (8,6)
Woche 8 Studienabschnitt 2				
Mittelwert (SD)	12,5 (18,5)	13,3 (11,4)	7,2 (11,9)	6,8 (13,0)
Veränderung von Baseline zu Woche 8				
Mittelwert (SD)	5,0 (12,3)	3,9 (11,6)	-1,0 (10,7)	0,0 (14,8)
LS Mean [95%-KI] ¹⁾	4,94 [0,61; 9,28]	4,90 [-1,49; 11,29]	0,23 [-4,73; 5,18]	1,83 [-5,35; 9,01]
Differenz der LS Mean [95%-KI] ¹⁾	0,04 [-7,68; 7,77]		-1,61 [-10,34; 7,12]	
p-Wert ²⁾	0,9914		0,7149	

Studie 165-302	20 mg Pegvaliase N = 34	Low dose Placebo N = 15	40 mg Pegvaliase N = 32	High dose Placebo N = 14
PKU-POMS-Verwirrtheitsskala				
Baseline Studienabschnitt 2				
Mittelwert (SD)	2,0 (2,0)	2,5 (2,1)	2,3 (2,0)	1,2 (1,5)
Woche 8 Studienabschnitt 2				
Mittelwert (SD)	2,8 (3,0)	3,7 (3,0)	2,3 (2,4)	2,0 (2,0)
Veränderung von Baseline zu Woche 8				
Mittelwert (SD)	0,9 (2,7)	1,4 (2,4)	0,0 (2,2)	0,5 (2,2)
LS Mean [95%-KI] ³⁾	0,98 [0,10; 1,85]	1,50 [0,21; 2,79]	0,49 [-0,51; 1,50]	0,59 [-0,88; 2,07]
Differenz der LS Mean [95%-KI] ³⁾	-0,51 [-2,08; 1,03]		-0,10 [-1,89; 1,69]	
p-Wert ²⁾	0,5053		0,9118	

¹⁾ Gemischtes Modells für wiederholte Messungen (MMRM) mit Berücksichtigung folgender Faktoren: Behandlungsgruppe, Visite, Interaktion aus Behandlung und Visite sowie dem Wert des PKU-POMS-TMD zu Baseline.

²⁾ Es wurde keine Angabe dazu identifiziert, welcher Test durchgeführt wurde.

³⁾ Gemischtes Modell für wiederholte Messungen (MMRM) mit Berücksichtigung folgender Faktoren: Behandlungsgruppe, Visite, Interaktion aus Behandlung und Visite sowie dem Wert der PKU-POMS-Verwirrtheitssubskala zu Baseline.

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model Repeated Measures; PKU-POMS: PKU-specific Profile of Mood States; SD: Standardabweichung; TMD: Total Mood Disturbance.

Für Studienabschnitt 4 liegen Ergebnisse für den PKU-POMS-TMD und für die Verwirrtheitssubskala des PKU-POMS vor. Für die Subskalen „Angst“, „Depression“, „Wut“, „Aktivität“ und „Müdigkeit“ wurden keine Analysen identifiziert.

Zu Woche 25 des Studienabschnitts 4 lag die Rücklaufquote bei 72,4 %, zu späteren Messzeitpunkten fiel sie auf < 70 %. Da die Veränderung seit Baseline in Studie 165-301 gemessen wird, bezieht sich die Rücklaufquote auf die Population der Studie 165-301, die zu Baseline den Fragebogen PKU-POMS ausfüllen konnte. Da der Fragebogen erst mit Amendment 1 des Studienprotokolls der Studie 165-301 eingesetzt werden sollte, konnten Personen, die unter der ersten Version des Studienprotokolls in die Studie eingeschlossen wurden, diesen Fragebogen zu Baseline nicht ausfüllen. Es wird davon ausgegangen, dass die Werte dieser Personen daher rein zufällig fehlen (missing completely at random [MCAR]). Unter Version 1 des Studienprotokolls der Studie 165-301 wurden 91 der insgesamt 261 Studienteilnehmenden in die Studie eingeschlossen. Daher wurden 170 Personen als Ausgangswert zur Berechnung der Rücklaufquote herangezogen.

Zu Woche 25 des Studienabschnitts 4 könnten Personen auf Basis der vorhandenen Informationen eine Gesamtexposition mit Pegvaliase zwischen 39 und 112 Wochen aufgewiesen haben. Dies setzt sich aus einer 14- bis 37-wöchigen Exposition in Studie 165-301, einer 0- bis 36-wöchigen Exposition in Studienabschnitt 1 der Studie 165-302, einer 0- bis 9-wöchigen Exposition in Studienabschnitt 2 der Studie 165-302, einer 0- bis 5-wöchigen Exposition in Studienabschnitt 3 der Studie 165-302 sowie der 25-wöchigen Exposition in Studienabschnitt 4 der Studie 165-302 zusammen. Angaben zur tatsächlichen Spanne der Gesamtexposition mit Pegvaliase wurden nicht identifiziert.

Die mittleren und medianen Veränderungen zu Woche 25 des Studienabschnitts 4 seit Baseline der Studie 165-301 können Tabelle 6 entnommen werden. Es wurden keine statistischen Tests durchgeführt und es liegen keine adäquaten Untersuchungen vor, aus denen sich die Höhe einer relevanten Veränderung ableiten ließe.

Tabelle 6: Ergebnisse zum Endpunkt „PKU-POMS“ in Studienabschnitt 4 der Studie 165-302 (ITT-Population)

Studie 165-302	Jegliche Dosis N = 170
PKU-POMS-TMD	
Woche 25 Studienabschnitt 4	
n	165
Mittelwert (SD)	6,4 (12,2)
Median (min; max)	4,0 (-11; 55)
Veränderung vom naiven Baseline¹⁾ zu Woche 25 von Studienabschnitt 4	
n	123
Mittelwert (SD)	-9,7 (14,4)
Median (min; max)	-9,0 (-48; 26)
PKU-POMS-Verwirrtheitssubskala	
Woche 25 Studienabschnitt 4	
n	165
Mittelwert (SD)	2,0 (1,92)
Median (min; max)	2,0 (0; 9)
Veränderung vom naiven Baseline¹⁾ zu Woche 25 von Studienabschnitt 4	
n	123
Mittelwert (SD)	-2,0 (2,7)
Median (min; max)	-2,0 (-9; 5)

¹⁾ Der naive Baseline-Zeitpunkt ist der Zeitpunkt vor Verabreichung der ersten Dosis der Studienmedikation in Studie 165-301.
Abkürzungen: ITT: Intention-to-Treat; PKU-POMS: PKU-specific Profile of Mood States; SD: Standardabweichung; TMD: Total Mood Disturbance.

ADHD-RS-IV

Die Ergebnisse während Studienabschnitt 2 der Studie 165-302 für die Unaufmerksamkeitsskala des ADHD-RS-IV können Tabelle 7 entnommen werden. Nach 8 Wochen zeigten sich zwischen Pegvaliase und dem entsprechenden Placebo jeweils keine statistisch signifikanten Unterschiede in der Veränderung seit Baseline des Studienabschnitts 2. Für die Subskala „Hyperaktivitätssymptome“ wurden keine Analysen für Studienabschnitt 2 auf Basis der ITT-Population identifiziert.

Tabelle 7: Ergebnisse zum Endpunkt „ADHD-RS-IV“ in Studienabschnitt 2 der Studie 165-302 (ITT-Population, N_{ges} = 95)

Studie 165-302	20 mg Pegvaliase N = 34	Low dose Placebo N = 15	40 mg Pegvaliase N = 32	High dose Placebo N = 14
Unaufmerksamkeitsskala der ADHD-RS-IV				
Baseline Studienabschnitt 2				
Mittelwert (SD)	5,7 (4,5)	5,5 (4,5)	6,0 (6,4)	2,9 (3,7)
Woche 8 Studienabschnitt 2				
Mittelwert (SD)	6,4 (5,3)	6,7 (5,2)	6,7 (6,7)	3,2 (2,9)

Studie 165-302	20 mg Pegvaliase N = 34	Low dose Placebo N = 15	40 mg Pegvaliase N = 32	High dose Placebo N = 14
Veränderung von Baseline zu Woche 8				
Mittelwert (SD)	1,0 (3,7)	1,4 (3,0)	0,6 (5,1)	-0,4 (3,4)
LS Mean [95%-KI] ¹⁾	1,09 [-0,40; 2,58]	1,06 [-1,06; 3,18]	1,17 [-0,47; 2,81]	-0,36 [-2,82; 2,09]
Differenz der LS Mean [95%-KI] ¹⁾	0,03 [-2,56; 2,63]		1,53 [-1,42; 4,49]	
p-Wert ²⁾	0,9815		0,3054	

¹⁾ Gemischtes Modell für wiederholte Messungen (MMRM) mit Berücksichtigung folgender Faktoren: Behandlungsgruppe, Visite, Interaktion aus Behandlung und Visite sowie dem Wert der Unaufmerksamkeitsskala der ADHD-RS-IV zu Baseline.

²⁾ Es wurde keine Angabe dazu identifiziert, welcher Test durchgeführt wurde.

Abkürzungen: ADHD-RS: Attention-Deficit Hyperactivity Disorder Rating Scale; ITT: Intention-to-Treat; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model Repeated Measures; SD: Standardabweichung.

Die ADHD-RS-IV sollte ab Version 1 des Studienprotokolls in Studie 165-301 eingesetzt werden. Somit sollte dieses Instrument bei allen 261 in die Studie eingeschlossenen Personen zu Baseline eingesetzt werden. 203 dieser 261 in Studie 165-301 eingeschlossenen Personen wurden anschließend in Studie 165-302 eingeschlossen. Gemäß Studienprotokoll der Studie 165-302 sollten ungefähr 250 Personen in Studie 165-302 eingeschlossen werden, die zuvor eine andere Studie zu Pegvaliase abgeschlossen und 14 Tage vor Screening für die Studie eine stabile Dosis Pegvaliase erhalten haben. Die Studie 165-301 haben 213 Personen abgeschlossen und es wurden anschließend 203 Personen in Studie 165-302 eingeschlossen. Es ist unklar, weswegen 10 Personen, die die Studie 165-301 abgeschlossen haben, nicht in Studie 165-302 eingeschlossen wurden. Insgesamt wurden 58 Personen, die an Studie 165-301 teilnahmen, nicht in Studie 165-302 eingeschlossen. Aus verschiedenen Gründen (siehe Nutzenbewertung vom 01.10.2019) haben Personen Studie 165-301 abgebrochen, hauptsächlich aufgrund unerwünschter Ereignisse (n = 17) oder auf eigenen Wunsch (n = 17). Insgesamt wurden 215 Personen in Studie 165-302 eingeschlossen; 12 dieser Personen nahmen zuvor nicht an Studie 165-301 sondern an Phase-II-Studien teil.

Unter Berücksichtigung der Gründe für den Abbruch der Teilnahme an Studie 165-301 ist nicht davon auszugehen, dass 58 Personen rein zufällig nicht in Studie 165-302 eingeschlossen wurden. Daher werden alle 261 Personen, für die zu Baseline in Studie 165-301 der ADHD-RS-IV eingesetzt werden sollte, zur Berechnung der Rücklaufquote herangezogen. Die so berechnete Rücklaufquote liegt zu allen Messzeitpunkten während Studienabschnitt 4 der Studie 165-302 bei < 70 % (maximale Rücklaufquote zu Woche 9 während Studienabschnitt 4: 65,5 %). Aus diesem Grund werden die Veränderungen dieses Endpunkts zwischen Baseline der Studie 165-301 und Studienabschnitt 4 der Studie 165-302 nicht dargestellt.

Stattdessen wird die Veränderung des ADHD-RS-IV für Studie 165-301 dargestellt. In Studie 165-301 wurde der ADHD-RS-IV nur zu Baseline und bei Studienende (letzte Studienvisite oder vorzeitiger Abbruch der Studie) eingesetzt. Die Rücklaufquote zwischen Baseline und Studienende betrug für die Unaufmerksamkeitsskala knapp unter 70 % (181 von 261 Personen); die Veränderung kann Tabelle 8 entnommen werden. Die mediane und mittlere Veränderung zwischen Baseline und Studienende (letzte Studienvisite oder vorzeitiger Abbruch) der Studie 165-301 liegen unterhalb der MID (Minimal Important Difference) für individuelle Veränderungen. Es ist allerdings zu beachten, dass einzelne Personen unterschiedlich lange behandelt wurden. Die mediane

Behandlungsdauer betrug 25,9 Wochen (min; max: 0,1; 36,9). Für die Hyperaktivitätsskala lag die Rücklaufquote bei deutlich < 70 %.

Tabelle 8: Ergebnisse zum Endpunkt „Unaufmerksamkeitssubskala des ADHD-RS-IV“ in Studie 165-301 (ITT-Population)

Studie 165-301	Gesamt (N = 261)
Baseline	
n	253
Mittelwert (SD)	9,8 (6,1)
Median (min; max)	9,0 (0,0; 26,0)
Studienende¹⁾	
n	189
Mittelwert (SD)	6,6 (5,1)
Median (min; max)	6,0 (0,0; 24,0)
Veränderung seit Baseline²⁾	
n	181
Mittelwert (SD)	-3,5 (5,2)
Median (min; max)	-3,0 (-23,0; 10,0)

¹⁾ Erhebung bei der letzten Studienvisite bzw. bei vorzeitigem Abbruch.

²⁾ Die Veränderung gegenüber Baseline basierte auf Personen mit verfügbaren Messungen zu beiden Zeitpunkten.

Abkürzungen: ADHD-RS: Attention-Deficit Hyperactivity Disorder Rating Scale; ITT: Intention-to-Treat; SD: Standardabweichung.

5 Zusammenfassung

Es bestehen erhebliche Unsicherheiten bezüglich der Validität der Endpunkte „PKU-POMS“ und „ADHD-RS-IV“. Das Verzerrungspotential wird für beide Endpunkte während des 8-wöchigen Studienabschnitts 2 der Studie 165-302 als niedrig eingeschätzt. Während Studienabschnitt 4 wird das Verzerrungspotential als hoch eingeschätzt.

Während Studienabschnitt 2 zeigten sich für beide Endpunkte keine statistisch signifikanten Unterschiede zwischen Pegvaliase und Placebo bezüglich der Veränderung seit Baseline des Studienabschnitts 2.

Für den Endpunkt „PKU-POMS“ werden die mittleren und medianen Veränderungen zwischen Woche 25 des Studienabschnitts 4 der Studie 165-302 im Vergleich zu Baseline der Studie 165-301 dargestellt. Zu Woche 25 des Studienabschnitts 4 könnten Personen auf Basis der vorhandenen Information eine Gesamtexposition mit Pegvaliase zwischen 39 und 112 Wochen aufgewiesen haben. Angaben zur tatsächlichen Spanne der Gesamtexposition mit Pegvaliase wurden nicht identifiziert. Es wurden keine statistischen Tests durchgeführt und es liegen keine adäquaten Untersuchungen vor, aus denen sich die Höhe einer relevanten Veränderung ableiten ließe.

Für den Endpunkt „ADHD-RS-IV“ liegt die Rücklaufquote bezogen auf alle Personen, für die zu Baseline der Studie 165-301 ein Wert vorliegen könnte, zu allen Messzeitpunkten während Studienabschnitt 4 der Studie 165-302 bei < 70 %. Aus diesem Grund werden die Veränderungen dieses Endpunkts zwischen Baseline der Studie 165-301 und Studienabschnitt 4 der Studie 165-302 nicht dargestellt. Die mediane und mittlere Veränderung zwischen Baseline und Studienende (letzte Studienvisite oder vorzeitiger Abbruch) der Studie 165-301 liegen unterhalb der MID für individuelle Veränderungen.

Referenzen

1. **Bacci ED, Wyrwich KW, Gries KS, Chen Y, Jain R, Konkol L, et al.** An adaptation of the profile of mood states for use in adults with phenylketonuria. *J Inborn Errors Metab Screen* 2016;4:1-11.
2. **BioMarin International.** Schriftliche Stellungnahme zur Nutzenbewertung nach § 35a SGB V und Kosten-Nutzen-Bewertung nach § 35b SGB V: Pegvaliase [unveröffentlicht]. 22.10.2019.
3. **BioMarin Pharmaceutical.** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Pegvaliase (Palynziq) [online]. Berlin (GER): Gemeinsamer Bundesausschuss (G-BA); 28.06.2019. [Zugriff: 19.11.2019]. URL: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/471/#dossier>.
4. **BioMarin Pharmaceutical.** A four-part, phase 3, randomized, double-blind, placebo-controlled, four-arm, discontinuation study to evaluate the efficacy and safety of subcutaneous injections of bmn 165 self administered by adults with phenylketonuria; study 165-302; clinical study report. 2017.
5. **BioMarin Pharmaceutical.** A phase 3, open-label, randomized, multi-center study to assess the safety and tolerability of an induction, titration, and maintenance dose regimen of BMN 165 self-administered by adults with phenylketonuria not previously treated with BMN 165; study 165-301; clinical study report. 2017.
6. **Spencer TJ, Adler LA, Meihua Q, Saylor KE, Brown TE, Holdnack JA, et al.** Validation of the adult ADHD investigator symptom rating scale (AISRS). *J Atten Disord* 2010;14(1):57-68.
7. **Wyrwich KW, Auguste P, Yu R, Zhang C, Dewees B, Winslow B, et al.** Evaluation of neuropsychiatric function in phenylketonuria: psychometric properties of the ADHD rating scale-IV and adult ADHD self-report scale inattention subscale in phenylketonuria. *Value Health* 2015;18(4):404-412.
8. **Wyrwich KW, Gries KS, Al-Jassar G, Bacci ED, Chen Y, Jain R, et al.** Assessing the content validity of the investigator-rated ADHD rating scale version IV instrument inattention subscale for use in adults with phenylketonuria. 2016;4:1-7.