

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.**

5. Kapitel § 12 Nr. 1 Satz 2 Verfo

Wirkstoff: Glycerolphenylbutyrat

Datum der Veröffentlichung: 1. Juni 2018

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abbildungsverzeichnis.....	4
Abkürzungsverzeichnis	5
Hintergrund.....	7
1 Einführung	8
2 Nutzenbewertung	10
2.1 Fragestellung.....	10
2.2 Zulassungsbegründende Studien.....	10
2.3 Liste der verwendeten Quellen.....	15
2.4 Studiencharakteristika	16
2.5 Endpunkte und statistische Methoden	19
2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte	19
2.5.2 Verzerrungspotential auf Studien- und Endpunktebene.....	22
2.5.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz	24
2.5.4 Statistische Methoden.....	30
2.6 Ergebnisse zum Zusatznutzen	32
2.6.1 Charakterisierung der Studienpopulation	32
2.6.2 Mortalität.....	42
2.6.3 Morbidität.....	42
2.6.4 Lebensqualität	44
2.6.5 Sicherheit	45
2.6.6 Subgruppenanalysen	52
3 Methodische Anmerkungen und Bewertung der Unterlagen.....	53
3.1 Zulassungsstatus und Zulassungspopulation von Glycerolphenylbutyrat.....	53
3.2 Design und Methodik der berücksichtigten Studien.....	54
3.3 Wirksamkeit	57
3.4 Lebensqualität	59
3.5 Sicherheit.....	59
4 Anforderung an eine qualitätsgesicherte Anwendung	61
5 Zusammenfassung der Bewertung	62
Referenzen	64
Anhang	69

Tabellenverzeichnis

Tabelle 1: Übersicht berücksichtigter supportiver Studien mit dem Wirkstoff Glycerolphenylbutyrat.....	11
Tabelle 2: Charakterisierung der Studie HPN-100-006.....	16
Tabelle 3: Charakterisierung der Intervention in der Studie HPN-100-006.....	18
Tabelle 4: In der Nutzenbewertung berücksichtigte Endpunkte.....	19
Tabelle 5: Verzerrungspotential der Studie HPN-100-006.....	22
Tabelle 6: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte in der Studie HPN-100-006.....	23
Tabelle 7: Beschreibung der Endpunkte der Kategorie Mortalität	24
Tabelle 8: Beschreibung der Endpunkte der Kategorie Morbidität	24
Tabelle 9: Beschreibung der Endpunkte der Kategorie Sicherheit.....	27
Tabelle 10: Allgemeine Angaben zu der Studienpopulation in der Studie HPN-100-006.....	33
Tabelle 11: Demographische Eigenschaften der ITT-Population in der Studie HPN-100-006....	34
Tabelle 12: Charakterisierung der Safety-Population in der Studie HPN-100-006.....	34
Tabelle 13: Allgemeine Angaben zur Medikation, Therapietreue und Proteinzufuhr und in der Studie HPN-100-006.....	36
Tabelle 14: Charakterisierung der Studienpopulation der supportiven einarmigen Beobachtungsstudien.....	39
Tabelle 15: Ergebnisse zum Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut in der Studie HPN-100-006.....	43
Tabelle 16: Ergebnisse zum Endpunkt Anzahl und Schweregrad hyperammonämischer Krisen in der Studie HPN-100-006 – Safety-Population	44
Tabelle 17: Zusammenfassung der Ergebnisse für unerwünschte Ereignisse in der Studie HPN-100-006 – Safety-Population	45
Tabelle 18: Zusammenfassung der Ergebnisse für unerwünschte Ereignisse gemäß MedDRA-Systemorganklasse und Preferred Term mit einer Inzidenz $\geq 5\%$ in der Studie HPN-100-006 in einem der Studienarme der Safety-Population	46
Tabelle 19: Kurzübersicht der Ergebnisse für unerwünschte Ereignisse und hyperammonämische Krisen in den supportiven Studien auf Patientenebene.....	47
Tabelle 20: Kurzübersicht über eine Auswahl von Ergebnissen für unerwünschte Ereignisse, die entsprechend dem Risk Management Plan des EPAR als spezifisch für Glycerolphenylbutyrat eingeschätzt sind, gemäß MedDRA-Systemorganklasse und Preferred Term mit einer Inzidenz $\geq 10\%$ in den supportiven Studien	48
Tabelle 21: Kurzübersicht der Ergebnisse für unerwünschte Ereignisse CTCAE-Grad ≥ 3 gemäß MedDRA-Systemorganklasse und Preferred Term in den supportiven Studien	50
Tabelle 22: Zusammenfassende Darstellung der Studienergebnisse.....	62
Tabelle 23: Ergänzende Beschreibung der Endpunkte der Kategorie Morbidität.....	69
Tabelle 24: Ergänzende Beschreibung der Endpunkte der Kategorie Lebensqualität.....	81
Tabelle 25: Ergebnisse zum Endpunkt SF-36 in der Studie HPN-100-007 für erwachsene Studienteilnehmer – Safety-Population.....	82

Abbildungsverzeichnis

Abbildung 1: Schematische Übersicht über die berücksichtigten Studien nach Studiendauer, Altersgruppe und erhobenen Endpunkten (eigene Darstellung)	14
Abbildung 2: Patientenfluss der Studie HPN-100-011 [27]	38

Abkürzungsverzeichnis

AL	Argininosuccinat-Lyase
ALT	Alaninaminotransferase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANOVA	Analysis of Variance
ARG	Arginase 1
ASS	Argininosuccinat-Synthetase
AST	Aspartataminotransferase
AUC	area under the curve
CPS	Carbamoylphosphat-Synthetase 1
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	Data and Safety Monitoring Board
EMA	European Medicines Agency
EPAR	European Public Assessment Report
ESPED	Erhebungseinheit für Seltene Pädiatrische Erkrankungen in Deutschland
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
HAC	hyperammonämische Krise/n
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
k.A.	keine Angabe/n
KI	Konfidenzintervall
MedDRA	Medical Dictionary for Regulatory Activities
MITT	Modified Intention-to-Treat
MW	Mittelwert
N/n	Anzahl
NAGS	N-Acetyl-Glutamat-Synthetase
NaPBA	Natriumphenylbutyrat
n.b.	nicht berechnet
PAGN	Phenylacetylglutamin
PP	Per-Protokoll
OTC	Ornithin-Transcarbamylase
PT	Preferred Term/s
pU	pharmazeutischer Unternehmer
RCT	randomisierte kontrollierte Studie/n (randomized controlled trial/s)
RR	Relatives Risiko

SAP	Statistischer Analyseplan
SF-15	15-Item Short Form Health Survey
SF-36	36-Item Short Form Health Survey
SGB	Sozialgesetzbuch
SD	Standardabweichung
SUE	schwerwiegende/s unerwünschte/s Ereignis/se
UCD	Harnstoffzyklusstörungen (Urea Cycle Disorders)
UE	unerwünschte/s Ereignis/se
VerfO	Verfahrensordnung des G-BA
WHO-DD	World Health Organization Drug Dictionary

Hintergrund

Glycerolphenylbutyrat (Ravicti®) 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.

Dementsprechend hat der G-BA in seiner Sitzung am 15. März 2012 den mit Beschluss vom 1. August 2011 erteilten Auftrag an das IQWiG zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen gemäß § 35a Absatz 2 SGB V in der Weise abgeändert, dass bei Orphan Drugs eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann erfolgt, wenn der Umsatz des betreffenden Arzneimittels die gesetzliche Grenze von 50 Millionen Euro überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt (vgl. § 35a Absatz 1 Satz 12 SGB V).

Der G-BA bestimmt gemäß 5. Kapitel § 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 Glycerolphenylbutyrat 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 heran und bewertet die darin enthaltenen Daten nach Maßgabe der im 5. Kapitel § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Glycerolphenylbutyrat in seiner Sitzung am 22. Mai 2018 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 28. Februar 2018 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pharmazeutischen Unternehmers in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung
- Anzahl der GKV-Patienten in der Zielpopulation

Die Nutzenbewertung wird am 1. Juni 2018 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 Einführung

Störungen des Harnstoffzyklus

Unter Harnstoffzyklusstörungen werden angeborene Enzymdefekte subsummiert, die zu einer potentiell lebensbedrohlichen Anreicherung von toxisch wirkendem Ammoniak führen. Im Harnstoffzyklus wird im Rahmen des Stickstoffstoffwechsels Ammoniak enzymatisch zu Harnstoff entgiftet und im Urin ausgeschieden. Beim Menschen findet der Harnstoffzyklus in der Leber statt. Störungen des Harnstoffzyklus können durch verschiedene Enzymdefekte entstehen:

- Mangel an Carbamoylphosphat-Synthetase 1 (CPS)
- Mangel an Ornithin-Transcarbamylase (OTC)
- Mangel an Argininosuccinat-Synthetase (ASS)
- Mangel an Argininosuccinat-Lyase (AL)
- Mangel an Arginase 1 (ARG)
- Mangel an (CPS aktivierender) N-Acetyl-Glutamat-Synthetase (NAGS) bzw. Mangel an Ornithin-Translokase

Jeder dieser Enzymdefekte kann zu einer Anreicherung von Ammoniak führen (Hyperammonämie). Das Manifestationsalter ist variabel, teilweise liegen bereits bei Neugeborenen Symptome vor, teilweise tritt die Symptomatik erst im Erwachsenenalter auf. Als Auslöser für Spätmanifestationen werden u. a. katabolische Ereignisse (bspw. durch stark erhöhten Proteinabbau infolge einer Nekrose), chirurgische Eingriffe, bestimmte Arzneimittel oder Proteinüberladung genannt [23]. Störungen des Harnstoffzyklus sind nicht Bestandteil des Neugeborenen Screenings.

Enzymdefekte im Harnstoffzyklus können überwiegend auf autosomal-rezessiv vererbte Mutationen zurückgeführt werden. Die kumulative Inzidenz von Harnstoffzyklusstörungen bei Kindern und Jugendlichen unter 16 Jahren wird entsprechend einer aktuellen Erhebung des ESPED-Projekts (<http://www.esped.uni-duesseldorf.de/esped/home>) mit etwa 1:49.000 Neugeborenen pro Jahr angegeben.

Bei Neugeborenen treten unspezifische neurologische Symptome nach Einsetzen der Proteinzufuhr auf. Diese umfassen Nahrungsverweigerung, Erbrechen, Tachypnoe, Konvulsionen, Lethargie und Koma. Bei Kleinkindern und Jugendlichen stehen neurologische Symptome wie Erbrechen, Ataxie, Verwirrung, Agitiertheit, Reizbarkeit, Lethargie, Somnolenz und Koma im Vordergrund. Bei Erwachsenen können akut Zeichen einer Enzephalopathie mit psychiatrischen Symptomen, Erbrechen, Krampfanfällen, Bewusstseinstörung sowie Organversagen auftreten. Zu den chronischen Manifestationen zählen Entwicklungs- und Gedeihstörungen und mentale Retardiertheit, neurologisch-psychiatrische Symptome, Bauchschmerzen und Erbrechen, Hepatomegalie, ggf. Proteinaversion. Symptome können auch episodisch auftreten und ohne Behandlung wieder abklingen.

In der akuten Hyperammonämie wird die Proteinzufuhr bei gleichzeitiger Glukosezufuhr gestoppt und es muss eine medikamentöse Therapie eingeleitet werden, die vom Alter und dem Ammoniakspiegel abhängig ist. Zu den medikamentösen Therapieoptionen gehört die Gabe von L-Arginin, L-Citrullin, Natriumbenzoat, Natriumphenylacetat, Natriumphenylbutyrat (NaPBA), Carbamylglutamat, Carnitin, Vitamin B12 und Biotin. Bei hohem Ammoniakspiegel kann eine Hämodialyse erforderlich sein. Prognostisch relevante Faktoren sind die Dauer des Komas und der höchste Ammoniaklevel.

Das Ziel der Langzeittherapie besteht darin, eine normale geistige und körperliche Entwicklung zu gewährleisten und Hyperammonämien zu vermeiden. Zu den therapeutischen Maßnahmen gehören eine Diät mit individuell titrierter Proteinzufuhr, Supplementation mit essentiellen Aminosäuren, Vitaminen und Mineralstoffen, die Gabe von Medikamenten, die die Stickstoffausscheidung erhöhen (u. a. Natriumphenylbutyrat) sowie als einzige kurative Therapieoption die Lebertransplantation (indiziert bei neonataler Manifestation).

Glycerolphanylbutyrat

Glycerolphanylbutyrat ist ein Arzneimittel zur medikamentösen Zusatztherapie bei erwachsenen und pädiatrischen Patienten im Alter von ≥ 2 Monaten mit Harnstoffzyklusstörungen (Urea Cycle Disorders, UCD), die durch diätetische Eiweißrestriktion und/oder Aminosäuresubstitution allein nicht behandelt werden können. Es handelt sich um eine Flüssigkeit zum Einnehmen.

Es handelt sich um ein Prodrug, dessen Metabolit Phenylacetat Stickstoff über Glutamin bindet. Es entsteht Phenylacetylglutamin, das über die Nieren ausgeschieden wird. Glycerolphanylbutyrat ist seit 2013 in den USA für die Behandlung von Harnstoffzyklusstörungen zugelassen. Der Wirkmechanismus ist mit dem des auf dem Markt verfügbaren Natriumphenylbutyrat vergleichbar.

2 Nutzenbewertung

2.1 Fragestellung

Glycerolphenylbutyrat (Ravicti®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 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 50 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. In die Bewertung werden gemäß Zulassungsstatus Patienten in folgender Indikation eingeschlossen:

- Erwachsene und pädiatrische Patienten im Alter von ≥ 2 Monaten mit Harnstoffzyklusstörungen, die durch diätetische Eiweißrestriktion und/oder Aminosäuresubstitution allein nicht behandelt werden können.

2.2 Zulassungsbegründende Studien

Pivotal Studie

- Studie HPN-100-006: Randomisierte, aktiv kontrollierte, doppelt-verblindete Cross-over-Studie der Phase III zur Ermittlung der Wirksamkeit und Sicherheit von Glycerolphenylbutyrat zur Behandlung von Erwachsenen mit Harnstoffzyklusstörungen [32].

Supportive Studien

- Studien HPN-100-005 und HPN-100-005SE: „A Phase 2, Fixed-Sequence, Open-Label, Switch-Over Study of the Safety and Tolerability of HPN-100 Compared to Sodium Phenylbutyrate in Children 6-17 Years of Age with Urea Cycle Disorders, with a Long-Term Safety Extension“ [28,29].
- Studie HPN-100-007: „A Phase 3, Open-Label Study of the Safety of HPN-100 for the Long-Term Treatment of Urea Cycle Disorders“ [Sicherheits-Follow-up von Patienten aus Studie HPN-100-007 und weiteren Patienten] [31].
- Studien HPN-100-012 und HPN-100-012SE: „A Switch-Over, Open-Label Study of the Safety, Pharmacokinetics, and Efficacy of HPN-100, Followed by Long-Term Treatment With HPN-100, in Pediatric Subjects under 6 Years of Age with Urea Cycle Disorders (UCDs)“ [33,34].
- Studie HPN-100-009: „An Open Label Study of the Safety, Efficacy and Pharmacokinetics of Glycerol Phenylbutyrate (GPB; RAVICTI®) in Pediatric Subjects Under Two Years of Age with Urea Cycle Disorders (UCDs)“ [26].
- Studie HPN-100-011: „Long Term Use of HPN-100 in Urea Cycle Disorders“ [Sicherheits-Follow-up von Patienten aus Studien HPN-100-005/HPN-100-005SE, HPN-100-006/HPN-100-007 und HPN-100-012/HPN-100-012SE] [27].

Neben der pivotalen Studie HPN-100-006 (Cross-over-RCT) konnten für die Nutzenbewertung von Glycerolphenylbutyrat sieben supportive Studien (einarmige Beobachtungsstudien) aus dem Studienprogramm des pharmazeutischen Unternehmers (pU) berücksichtigt werden. Für die vorliegende Nutzenbewertung konnten hingegen fünf vom pU vorgelegte supportive Studien nicht berücksichtigt werden: Eine schrittweise Dosisescalationsstudie (UP-1204-003) [30], eine abgebrochene Studie (HPN-100-004) und drei laufende Studien, von denen noch keine Studienberichte

vorlagen (HPN-100-014, HPN-100-021, HZNP-RAV-401). Eine tabellarische Übersicht mit dem PICO-Schema der berücksichtigten supportiven Studien findet sich in Tabelle 1.

Tabelle 1: Übersicht berücksichtigter supportiver Studien mit dem Wirkstoff Glycerolphenylbutyrat

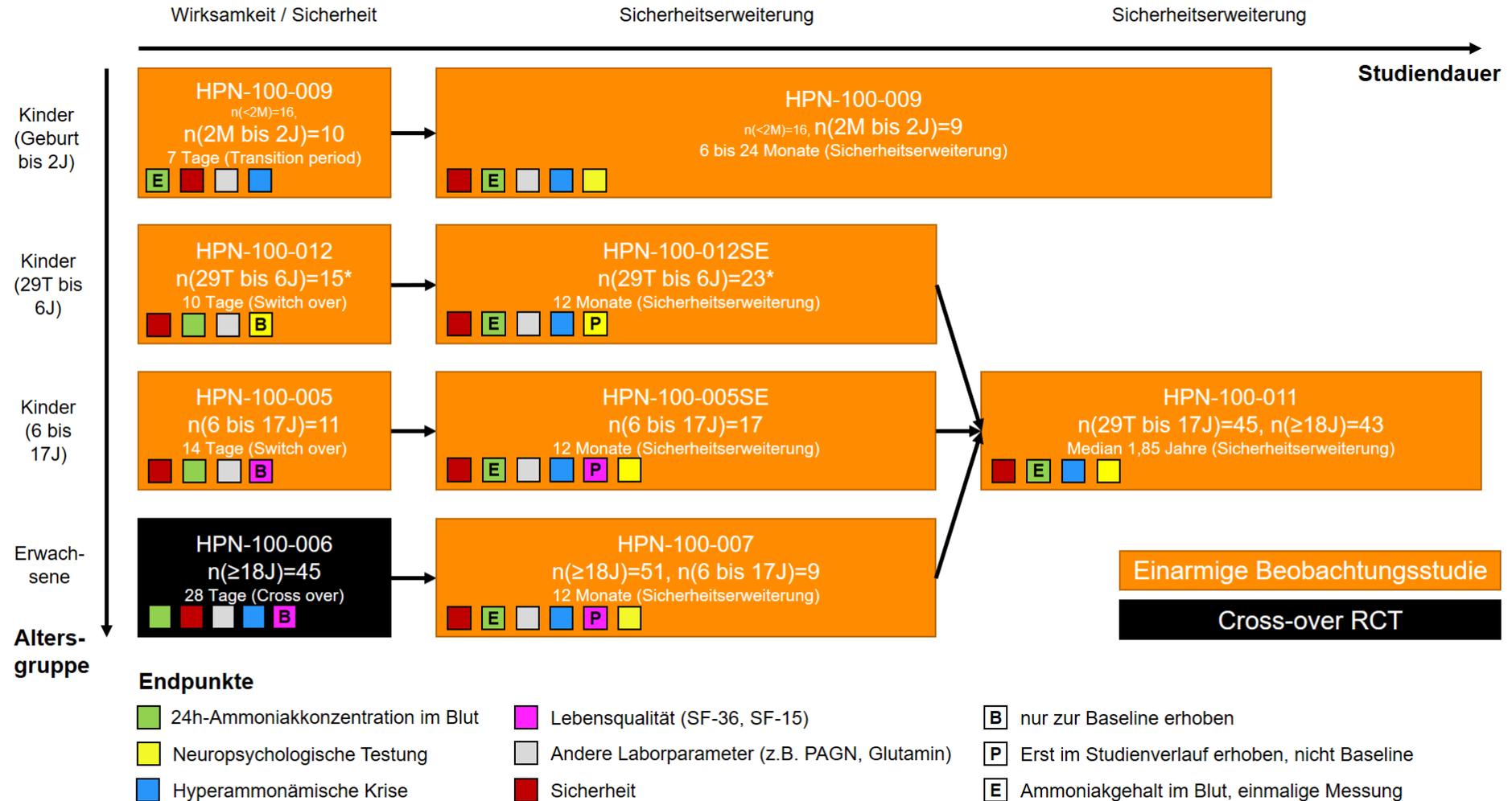
Studie	Studiencharakteristika	Kommentar
Kinder zwischen 6 und 17 Jahren		
HPN-100-005 und HPN-100-005SE	<p><u>Design:</u> Offene, multizentrische, einarmige Phase-II-Beobachtungsstudie.</p> <p><u>Einschluss:</u> UCD-Patienten (keine Einschränkung bei Subtypen); Alter: 6 bis 17 Jahre, stabil eingestellt auf NaPBA; HPN-100-005: N=11, HPN-100-005SE: N=17.</p> <p><u>Intervention:</u> Tägliche Dosis Glycerolphenylbutyrat (ml) = NaPBA-Dosis (g) × 0,95 / 1,1 (und individuelle diätetische Eiweißrestriktion und Aminosäuresubstitution). Dosierungsanpassungen waren während der Sicherheitserweiterung erlaubt.</p> <p><u>Begleitmedikation:</u> Begleitmedikation, die schon vor Studienteilnahme vorlag, wurde fortgesetzt und aufgezeichnet. Bestimmte Medikamente, die z. B. HAC auslösen können oder den Proteinstoffwechsel beeinflussen, waren nicht erlaubt (Ausschlusskriterium).</p> <p><u>Primärer Endpunkt:</u> Auftreten von UE.</p> <p><u>Studiendauer:</u> HPN-100-005: Switch-over-Phase von NaPBA zu Glycerolphenylbutyrat 14 Tage (Woche 1: NaPBA, Woche 2 Glycerolphenylbutyrat), N=11. HPN-100-005SE: Sicherheitserweiterung für 12 Monate, Rekrutierung von 6 weiteren Patienten, N=17.</p>	<p><u>Limitation:</u> Einarmige Beobachtungsstudie, die keine Kontrollgruppe einschließt</p>
Erwachsene und Kinder ab 6 Jahren		
HPN-100-007 (Sicherheitserweiterung der pivotalen Studie HPN-100-006)	<p><u>Design:</u> Offene, multizentrische, einarmige Phase-III-Beobachtungsstudie.</p> <p><u>Einschluss:</u> Studienteilnehmer der Studie HPN-100-006 (40 von 45 Patienten) sowie weitere UCD-Patienten ab 6 Jahren oder Studienteilnehmer, die die Einschlusskriterien der Studie HPN-100-006 nicht erfüllten (z. B. andere UCD-Subtypen oder keine stabile Einstellung auf NaPBA); neu eingeschlossene Patienten (9 Kinder zwischen 6 und 17 Jahren sowie 11 Erwachsene) wurden vor Studienbeginn auf Glycerolphenylbutyrat eingestellt; N=60.</p> <p><u>Intervention:</u> Tägliche Dosis Glycerolphenylbutyrat (ml) = NaPBA-Dosis (g) × 0,95 / 1,1 (und individuell angepasste diätetische Eiweißrestriktion und Aminosäuresubstitution). Dosierungsanpassungen waren erlaubt.</p> <p><u>Begleitmedikation:</u> Begleitmedikation, die schon vor Studienteilnahme vorlag, wurde fortgesetzt und aufgezeichnet. Bestimmte Medikamente, die z. B. HAC auslösen können oder den Proteinstoffwechsel beeinflussen, waren nicht erlaubt (Ausschlusskriterium).</p> <p><u>Primärer Endpunkt:</u> Auftreten von UE.</p> <p><u>Studiendauer:</u> 12 Monate.</p>	<p><u>Limitation:</u> Einarmige Beobachtungsstudie, die keine Kontrollgruppe einschließt</p>

Studie	Studiencharakteristika	Kommentar
Kinder zwischen 29 Tagen und 6 Jahren		
HPN-100-012 und HPN-100-012SE	<p><u>Design:</u> Offene, multizentrische, einarmige Phase-III-Beobachtungsstudie.</p> <p><u>Einschluss:</u> UCD-Patienten außer mit dem Subtyp NAGS-Mangel; Alter: 29 Tage bis 6 Jahre, stabil eingestellt auf NaPBA (min. 5 Tage); HPN-100-012: N=15, HPN-100-012SE: N=23.</p> <p><u>Intervention:</u> Initiale Dosis Glycerolphenylbutyrat (ml) = NaPBA-Dosis (g) × 0,87 (und individuell angepasste diätetische Eiweißrestriktion und Aminosäuresubstitution). Dosisadjustierungen waren im Rahmen der Sicherheitserweiterung basierend auf Ammoniakkonzentration im Blut, PAGN-Werten und klinischer Einschätzung erlaubt.</p> <p>Neu eingeschlossene Patienten in der Sicherheitserweiterung (HPN-100-012SE): Tägliche Dosis Glycerolphenylbutyrat (ml) = NaPBA-Dosis (g) × 0,95 / 1,1 (und individuelle Dosisadjustierung).</p> <p><u>Begleitmedikation:</u> Begleitmedikation, die schon vor Studienteilnahme vorlag, wurde fortgesetzt und aufgezeichnet.</p> <p><u>Primärer Endpunkt:</u> Auftreten von UE.</p> <p><u>Studiendauer:</u> HPN-100-012: Switch-over-Phase von NaPBA zu Glycerolphenylbutyrat 10 Tage (Tag 1: NaPBA, Tag 2–10: Glycerolphenylbutyrat), N=15. HPN-100-012SE: Sichererweiterung für 12 Monate, Rekrutierung von 8 weiteren Patienten, N=23.</p>	<p><u>Limitation:</u> Einarmige Beobachtungsstudie, die keine Kontrollgruppe einschließt</p> <p><u>Alter:</u> Tatsächliche Spannweite Alter: 2 Monate bis 5 Jahre (Zeitpunkt Screening)</p>
Kinder bis 2 Jahre		
HPN-100-009	<p><u>Design:</u> Offene, multizentrische, einarmige Phase-IV-Beobachtungsstudie.</p> <p><u>Einschluss:</u> UCD-Patienten (UCD-Diagnose oder -Verdachtsdiagnose) außer mit dem Subtyp NAGS-Mangel; Alter: bis 2 Jahre; Medikation bei Studieneintritt entweder Glycerolphenylbutyrat oder NaPBA; N < 2 Monate: N=16, ≥ 2 Monate und < 2 Jahre: N=10.</p> <p><u>Intervention:</u></p> <ul style="list-style-type: none"> • Neu diagnostiziert mit HAC: Initiale Glycerolphenylbutyrat-Dosis: 11,2 ml/m²/Tag • Neu diagnostiziert ohne HAC: Initiale Glycerolphenylbutyrat-Dosis: 8,5 ml/m²/Tag • Patienten bereits stabil auf NaPBA eingestellt: Tägliche Dosis Glycerolphenylbutyrat (ml) = NaPBA-Dosis (g) × 0,86 (bei Tablettengabe) • Patienten bereits stabil auf Natriumbenzoat eingestellt: Tägliche Dosis Glycerolphenylbutyrat (ml) = Natriumbenzoat Dosis (g) × 0,5 <p>Zudem wurden bei jedem Studienteilnehmer individuell diätetische Eiweißrestriktion und Aminosäuresubstitution angepasst. Dosierungsanpassungen waren erlaubt.</p> <p><u>Begleitmedikation:</u> Begleitmedikation, die schon vor Studienteilnahme vorlag, wurde fortgesetzt und aufgezeichnet.</p> <p><u>Primärer Endpunkt:</u> Auftreten von UE.</p>	<p><u>Limitation:</u> Einarmige Beobachtungsstudie, die keine Kontrollgruppe einschließt</p> <p><u>Anmerkungen:</u> Die Studie wurde von der FDA gefordert; NaPBA-Medikation war kein Einschlusskriterium</p>

Studie	Studiencharakteristika	Kommentar
	<u>Studiendauer:</u> Transition period (7 Tage) gefolgt von einer Sicherheitserweiterung, die mindestens 6 Monate dauerte und bis zu 2 Jahren dauern konnte.	
Langzeitbeobachtung über alle Altersgruppen		
HPN-100-011 (weitere Sicherheitserweiterung der supportiven Studien HPN-100-005SE, HPN-100-007 und HPN-100-012SE)	<p><u>Design:</u> Offene, multizentrische, einarmige Phase-IV-Beobachtungsstudie; zweite Sicherheitserweiterung.</p> <p><u>Einschluss:</u> Studienteilnehmer, die bereits die Studien HPN-100-005SE, HPN-100-007 oder HPN-100-012SE abgeschlossen haben (88 von 100 Patienten, davon 45 Kinder).</p> <p><u>Intervention:</u> Gleiche Dosis wie in den vorangehenden Studien</p> <ul style="list-style-type: none"> • Patienten < 20 kg: 0,53 ml/kg/Tag Glycerolphenylbutyrat • Patienten > 20 kg: 11,48 ml/m²/Tag Glycerolphenylbutyrat • Maximaldosis: 17,4 ml/Tag Glycerolphenylbutyrat <p>Zudem wurden bei jedem Studienteilnehmer individuell diätetische Eiweißrestriktion und Aminosäuresubstitution angepasst. Eine individuelle Dosisadjustierung war möglich.</p> <p><u>Begleitmedikation:</u> Begleitmedikation, die schon vor Studienteilnahme vorlag, wurde fortgesetzt und aufgezeichnet. Bestimmte Medikamente, die z. B. HAC auslösen können, waren nicht erlaubt.</p> <p><u>Primärer Endpunkt:</u> Auftreten von UE.</p> <p><u>Studiendauer:</u> Unbefristet. Studienteilnehmer aus den USA beendeten mit Markteinführung die Studienteilnahme; Studienteilnehmer aus Kanada haben zum Zeitpunkt des letzten Datenschnitts die Studie noch nicht abgeschlossen; Median 1,85 Jahre.</p>	<p><u>Limitation:</u> Einarmige Beobachtungsstudie, die nicht auf einen Vergleich zur Baseline-NaPBA-Medikation (jeweils Beginn der Studien HPN-100-005, HPN-100-006 und HPN-100-012) ausgelegt war</p> <p><u>Anmerkung:</u> Zum Zeitpunkt des letzten Datenschnitts (29.10.2015) war die Studie noch nicht abgeschlossen, 6 von 82 Patienten (Studienteilnehmer aus Kanada) wurden noch beobachtet</p>

Abkürzungen: FDA: U.S. Food and Drug Administration; HAC: hyperammonämische Krise/n; NaPBA: Natriumphenylbutyrat; PAGN: Phenylacetylglutamin; UCD: Harnstoffzyklusstörungen; UE: unerwünschte/s Ereignis/se.

Die in der Nutzbewertung berücksichtigten supportiven Studien umfassen mehrere Altersgruppen, unterscheiden sich hinsichtlich der Dauer und sind alle einarmige Beobachtungsstudien (siehe Tabelle 1). Eine schematische Übersicht über die pivotale und die supportiven berücksichtigten Studien hinsichtlich der Dimensionen Altersgruppen und Studiendauer findet sich in Abbildung 1 (pivotales Cross-over-RCT in schwarz dargestellt, supportive einarmige Beobachtungsstudien in orange dargestellt).



* der jüngste Patient in den Studien HPN-100-012 bzw. HPN-100-012SE war zum Zeitpunkt des Screenings min. 2 Monate alt.

Abbildung 1: Schematische Übersicht über die berücksichtigten Studien nach Studiendauer, Altersgruppe und erhobenen Endpunkten (eigene Darstellung).

Mit dem Nutzendossier legte der pU Unterlagen für eigens berechnete metaanalytisch zusammengefasste Schätzer über mehrere Studien vor. Für den Einschluss einzelner Studien in die Metaanalysen setzte der pU voraus, dass ein Vergleich zwischen einer Intervention und einer Vergleichsgruppe vorlag. Dies sah der pU als gegeben, wenn Patienten zum Studienbeginn auf die Vergleichstherapie (NaPBA) eingestellt waren, und dementsprechend wurde als Kontrollbedingung in den einarmigen Beobachtungsstudien die Baseline (NaPBA-Medikation) herangezogen. Dieser Argumentation des pU wird nicht gefolgt. Einarmige Beobachtungsstudien bilden in diesem Zusammenhang den Krankheitsverlauf während der Glycerolphenylbutyrat-Intervention ab und liefern keinen validen Vergleich gegen eine Kontrollgruppe.

Die einzige randomisierte kontrollierte Studie (RCT) im vorgelegten Nutzendossier wurde bei erwachsenen Patienten mit Harnstoffzyklusstörungen über einen Zeitraum von vier Wochen durchgeführt. Prinzipiell ist es bei der vorliegenden Fragestellung möglich, RCT mit einer längeren Studiendauer sowie bei Patienten mit Harnstoffzyklusstörungen im Alter zwischen 2 Monaten und 18 Jahren durchzuführen. Da der pU als einzige verfügbare Evidenz für Langzeiteffekte sowie für Patienten mit Harnstoffzyklusstörungen unter 18 Jahren einarmige Beobachtungsstudien vorgelegt hat, werden diese in der Nutzenbewertung lediglich als supportive Studien berücksichtigt. Die Endpunkte Lebensqualität (ergänzend im Anhang dargestellt), neuropsychologische Tests (Charakterisierung des Endpunkts im Anhang dargestellt) und Sicherheit werden im Verlauf dargestellt, da die Studien keinen validen Vergleich gegen eine Kontrollgruppe im Parallelgruppendesign liefern.

2.3 Liste der verwendeten Quellen

Zur Bestimmung des Ausmaßes des Zusatznutzens für Glycerolphenylbutyrat wurden folgende Studien und Daten herangezogen:

- Herstellerdossier zu Glycerolphenylbutyrat [49]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [17]
- Unterlagen der U.S. Food and Drug Administration (FDA) [53,54]
- Unterlagen der Canadian Agency for Drugs and Technologies in Health (CADTH) [9]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der pivotalen Studie HPN-100-006 [32]
- Studienberichte der supportiven Studien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE [26,27,28,29,31,33,34]

2.4 Studiencharakteristika

Tabelle 2: Charakterisierung der Studie HPN-100-006

Charakteristikum	Beschreibung
Design	<p>Es handelt sich um eine randomisierte, aktiv kontrollierte, doppelt-verblindete Cross-over-Studie der Phase III zur Ermittlung der Wirksamkeit und Sicherheit von Glycerolphenylbutyrat im Vergleich zu NaPBA zur Behandlung von Erwachsenen mit Harnstoffzyklusstörungen.</p> <p><u>Studienziel</u> Ziel der Studie war es, die Nichtunterlegenheit von Glycerolphenylbutyrat gegenüber NaPBA anhand des Ammoniakblutspiegels zu testen.</p> <p>Erwachsene Patienten mit bestätigter Diagnose einer Harnstoffzyklusstörung, die zu Studienbeginn für mindestens eine Woche auf eine stabile NaPBA-Dosis eingestellt waren, die keine Krankheitssymptome aufwiesen und deren Ammoniakspiegel unter 100 µmol/l lag, wurden im Verhältnis 1:1 in einen von zwei Behandlungsarmen randomisiert. In Behandlungsarm A erhielten die Patienten NaPBA plus Placebo für 2 Wochen, danach Glycerolphenylbutyrat plus Placebo für 2 Wochen. In Behandlungsarm B erhielten die Patienten zuerst Glycerolphenylbutyrat plus Placebo für 2 Wochen, danach NaPBA plus Placebo für 2 Wochen. Eine Wash-out-Phase war aus Sicherheitsgründen nicht vorgesehen. Bei der Screeningvisite wurde die NaPBA-Dosis ermittelt und daraus die erforderliche Dosis Glycerolphenylbutyrat abgeleitet. Die Gesamtbehandlungsdauer betrug demnach jeweils 4 Wochen. Während der Studiendauer waren die Patienten angehalten, eine definierte Diät einzuhalten.</p> <p>Ein DSMB wurde eingerichtet, das die Aufgabe hatte, die Daten zur Sicherheit zu beurteilen, nachdem ca. 50 % der Patienten die Prüfintervention erhalten hatten. Hierzu wurden Abbruchkriterien festgelegt.</p> <p>Nach Abschluss der Behandlungsphase konnten die Patienten in eine Open-label-Extensionsstudie mit einer Beobachtungsdauer von 12 Monaten wechseln (Studie HPN-100-007).</p>
Population	<p>Wesentliche Einschlusskriterien Erwachsene Patienten (≥ 18 Jahre), die folgende Kriterien erfüllen:</p> <ul style="list-style-type: none"> • Enzymatisch, biochemisch oder genetisch bestätigte Diagnose einer Harnstoffzyklusstörung aufgrund eines Defekts der CPS, OTC oder ASS • Zu Studienbeginn seit mindestens einer Woche auf eine stabile NaPBA-Dosis eingestellt. • Keine Zeichen einer Hyperammonämie (Ammoniakspiegel < 100 µmol/l) während der 2 Wochen vor der Screeningvisite. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Zeichen einer Hyperammonämie bis zu 2 Wochen vor der Screeningvisite oder Ammoniakspiegel ≥ 100 µmol/l zur Screeningvisite oder zu Baseline. • Infektion oder andere Erkrankung, die mit einer Erhöhung des Ammoniumspiegels einhergeht. • Klinischer oder labordiagnostischer Befund Grad 3 oder höher gemäß CTCAE (mit Ausnahme von Leberenzym erhöhungen). • Medikamentöse Therapie mit Wirkstoffen, die zu Änderung der renalen Clearance oder einer Erhöhung des Proteinkatabolismus führen oder den Ammoniumspiegel erhöhen bis 24 Stunden vor Studienbeginn und während der Studie. • Einnahme von Natriumbenzoat bis eine Woche vor Studienbeginn. • Verlängerung der QTc-Intervalls in der Anamnese oder bei der Screeningvisite oder zu Baseline auf > 450 msec. • Bekannte Überempfindlichkeit für Phenylacetate oder Phenylbutyrate. • Zustand nach Lebertransplantation einschließlich hepatozelluläres Transplantat.

Charakteristikum	Beschreibung									
	<p>Folgende Stopping Rules wurden festgelegt</p> <p>a) Individuelle Studienteilnehmer</p> <ul style="list-style-type: none"> ○ Auftreten einer HAC (klinische Symptome [Erbrechen, Proteinintoleranz, Lethargie, Psychose, auffällige Befunde in der neurologischen Untersuchung, Gehirnödem, Kopfschmerzen] in Verbindung mit einem Ammoniakspiegel $\geq 100 \mu\text{mol/l}$). ○ Auftreten eines Ereignisses des CTCAE-Grads 4 oder ein lebensbedrohendes Ereignis, das nicht nach CTCAE klassifiziert ist. ○ Klinisch relevante Allergie oder Überempfindlichkeit gegenüber Glycerolphenylbutyrat. ○ QTc-Intervall $> 500 \text{ msec.}$ oder ein Anstieg im QTc-Intervall um $> 60 \text{ msec.}$ von Baseline ○ Lebertransplantation ○ Schwangerschaft <p>b) Auf Studienebene nach Vorgaben des DSMB, wenn</p> <ul style="list-style-type: none"> ○ zwei Patienten das gleiche Ereignis des CTCAE-Grad 4 erleiden (außer Hyperammonämie) oder ○ bei 25 % von je 10 eingeschlossenen Patienten oder bei 5 Patienten (je nachdem, was zuerst zutrifft) eine Hyperammonämie eintritt. 									
Intervention und Zahl der Patienten	<p>Insgesamt 46 Patienten wurden im Verhältnis 1:1 entsprechend eines computergenerierten zentralen Schemas in einen von zwei Behandlungsarmen randomisiert:</p> <table border="1" data-bbox="470 996 1313 1193"> <thead> <tr> <th></th> <th>Periode 1 (2 Wochen)</th> <th>Periode 2 (2 Wochen)</th> </tr> </thead> <tbody> <tr> <td>Behandlungsarm A (N=22)</td> <td>NaPBA + Placebo</td> <td>Glycerolphenylbutyrat + Placebo</td> </tr> <tr> <td>Behandlungsarm B (N=24)</td> <td>Glycerolphenylbutyrat + Placebo</td> <td>NaPBA + Placebo</td> </tr> </tbody> </table> <p>Die Gesamtbehandlungsdauer betrug jeweils 4 Wochen. Nach Ende der jeweils ersten und zweiten Behandlungsperiode wurden die Patienten für jeweils einen Tag stationär aufgenommen und es wurden körperliche Untersuchung sowie Blut- und Urinuntersuchungen durchgeführt.</p> <p>Es handelte sich um eine doppelt-verblindete Studie, d. h. Studienärzte, Studienpersonal einschließlich Pharmazeut sowie Patienten waren verblindet. Bei Notfällen, in denen die Information über die aktuelle Zuordnung zum Behandlungsarm erforderlich wurde, konnte über ein interaktives Voice Response System eine Entblindung durchgeführt werden.</p> <p>44 Patienten beendeten die Studie. Je ein Patient pro Behandlungsarm schied aus der Studie aus: Ein Patient in Behandlungsarm A wurde nach dem ersten Studientag aufgrund von Nebenwirkungen wegen Non-Compliance hinsichtlich der Diät ausgeschlossen; ein Patient in Behandlungsarm B wurde nach der Randomisierung, aber vor der Einnahme der Studienmedikation wegen Verletzung eines Einschlusskriteriums ausgeschlossen.</p> <p>Ein Patient in Behandlungsarm B erhielt die Studienmedikation in der falschen Reihenfolge und wurde daher aus der Per-Protokoll-Analyse ausgeschlossen.</p>		Periode 1 (2 Wochen)	Periode 2 (2 Wochen)	Behandlungsarm A (N=22)	NaPBA + Placebo	Glycerolphenylbutyrat + Placebo	Behandlungsarm B (N=24)	Glycerolphenylbutyrat + Placebo	NaPBA + Placebo
	Periode 1 (2 Wochen)	Periode 2 (2 Wochen)								
Behandlungsarm A (N=22)	NaPBA + Placebo	Glycerolphenylbutyrat + Placebo								
Behandlungsarm B (N=24)	Glycerolphenylbutyrat + Placebo	NaPBA + Placebo								
Ort und Zeitraum der Durchführung	<p>Studienbeginn (erster Patient eingeschlossen): 12.10.2009 Studienende (letzte Visite): 09.09.2010 Die Studie wurde an 19 Zentren in den USA und Kanada durchgeführt.</p>									

Charakteristikum	Beschreibung
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt 24-Stunden-AUC für Ammoniak im Blut (gemessen als $\mu\text{mol h/l}$) an den Tagen 14 und 28 (am Ende der Behandlungsperiode wurde ein „steady state“ für die Studienmedikation vorausgesetzt). Die AUC wurde aus einer Sequenz von 8 Blutentnahmen, davon eine vor und sieben nach 2, 4, 8, 12, 16, 20 und 24 Stunden nach Medikamenteneinnahme kalkuliert. Die Proben wurden in den lokalen Labors der Studienzentren ausgewertet, dabei kamen verschiedene Analyseverfahren zur Anwendung.</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Korrelation zwischen der 24-Stunden-PAGN-Ausscheidung im Urin und der Ammoniakkonzentration im venösen Blut. • Maximale Ammoniakwerte im Vergleich von NaPBA mit Glycerolphenylbutyrat. • Anteil der Ammoniakwerte oberhalb des ULN im Vergleich von NaPBA mit Glycerolphenylbutyrat. • Anzahl und Schweregrad symptomatischer HAC. • Pharmakokinetische Messwerte für relevante Metaboliten von NaPBA und Glycerolphenylbutyrat in Blut und Urin. • Anteil von UE in jeder Behandlungsgruppe.
Subgruppenanalysen	Keine a priori festgelegten Subgruppenanalysen geplant.

Abkürzungen: AUC: area under the curve; ASS: Argininosuccinat-Synthetase; CPS: Carbamylphosphat-Synthetase; CTCAE: Common Terminology Criteria for Adverse Events; DSMB: Data and Safety Monitoring Board; HAC: hyperammonämische Krise/n; OTC: Ornithin-Transcarb-amylose; PAGN: Phenylacetylglutamin; NaPBA: Natriumphenylbutyrat; SAP: Statistischer Analyseplan; UE: unerwünschte/s Ereignis/se; ULN: upper limit of normal.

Tabelle 3: Charakterisierung der Intervention in der Studie HPN-100-006

Intervention	Kontrolle
<p><u>Wirkstoff</u> Glycerolphenylbutyrat</p> <p><u>Dosis</u> Die tägliche Dosis wurde aus der NaPBA-Dosis nach der Formel $\text{NaPBA-Dosis (g)} \times 0,95 / 1,1 = \text{Glycerolphenylbutyrat-Tagesdosis (in ml)}$ abgeleitet, die für die Kontrolle des Ammoniakspiegels unter Berücksichtigung des Enzymdefekts, der persönlichen Diät und der Zufuhr von Aminosäuren oder anderen Supplementen erforderlich war. Die Einnahme erfolgte dreimal täglich zu den Mahlzeiten.</p> <p><u>Darreichungsform</u> Glycerolphenylbutyrat lag als farblos bis gelbliche geruch- und geschmacklose Flüssigkeit vor.</p>	<p><u>Wirkstoff</u> NaPBA</p> <p><u>Dosis</u> Die tägliche Dosis, die für die Kontrolle des Ammoniakspiegels unter Berücksichtigung des Enzymdefekts, der persönlichen Diät und der Zufuhr von Aminosäuren oder anderen Supplementen erforderlich war, wurde individuell festgelegt. Die empfohlene Dosis NaPBA beträgt 9,9 bis 13 g/m^2 Körperoberfläche bei Erwachsenen. Die Einnahme erfolgte dreimal täglich zu den Mahlzeiten.</p> <p><u>Darreichungsform</u> NaPBA lag in Tablettenform mit jeweils 500 mg Wirkstoff vor. Patienten, die vor Studienbeginn NaPBA als Pulver einnahmen, sollten möglichst auf Tabletten umgestellt werden. Lediglich wenn Patienten Probleme beim Schlucken der Tabletten hatten, konnten sie den Wirkstoff auch in Form von Pulver einnehmen.</p>
<p>Die erlaubte maximale Tagesdosis betrug 20 g NaPBA (entsprechend 17,4 ml Glycerolphenylbutyrat). Bei vier Patienten wurde eine höhere Dosis genehmigt, um den Ammoniakspiegel zu kontrollieren. Für jeden Wirkstoff wurden korrespondierende Placebos hergestellt, die in Farbe und Konsistenz dem Verum ähnelten.</p>	

Intervention	Kontrolle
<p>Wenn Patienten bereits vor Beginn der Studie mit NaPBA adäquat behandelt waren, wurde die Dosierung während der Studie beibehalten. Während der Studie war keine Dosisänderung erlaubt. Während der Studiendauer waren die Patienten angehalten, eine definierte Diät einzuhalten. Die Compliance mit der Studienmedikation wurde anhand von study drug compliance diaries und der zurückgegebenen Studienmedikation bestimmt.</p> <p><u>Begleitmedikation</u> Innerhalb von 30 Tagen vor Studienbeginn war die Einnahme einer (anderen) Studienmedikation nicht erlaubt. Arzneimittel, die zu einer Änderung der renalen Clearance (z. B. Probenecid) oder einer Erhöhung des Proteinkatabolismus (z. B. Kortikosteroide) führen oder den Ammoniumspiegel erhöhen (z. B. Valproat) waren ab 24 Stunden vor Studienbeginn und während der Studie nicht erlaubt.</p> <p><u>Ernährung</u> Alle Patienten folgten ihrer individuell verschriebenen proteinarmen Diät sowie Aminosäuresubstitution. Die individuelle Diät orientierte sich am Alter und der residuellen Enzymaktivität und wurde in der Studie nicht geändert.</p>	

Es liegen keine Amendments zum Studienprotokoll vom 30. Juni 2009 (Version 1.0) vor.

2.5 Endpunkte und statistische Methoden

2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte

Tabelle 4: In der Nutzenbewertung berücksichtigte Endpunkte

Endpunkt ¹⁾	Kategorie	Vom pU als patientenrelevant eingestuft	In der Nutzenbewertung berücksichtigt
Todesfälle ²⁾	Mortalität	●	●
24-Stunden-AUC-Ammoniakkonzentration im Blut	Morbidität	●	○
Tägliche durchschnittliche Ammoniakkonzentration im Blut		●	-
Neuropsychologische Tests ³⁾		● ⁴⁾	_ ⁵⁾
Anzahl und Schweregrad von HAC ⁶⁾		●	●
SF-36 ⁷⁾	Lebensqualität	●	● ⁸⁾
SF-15 ⁷⁾		●	_ ⁵⁾
UE	Sicherheit	●	●
SUE		●	●
Studienabbruch aufgrund UE		●	●
UCD-Symptomprofil		●	-

● Endpunkt wird als bewertungsrelevant eingestuft und in der Nutzenbewertung berücksichtigt.

○ Endpunkt wird nicht als bewertungsrelevant eingestuft aber ergänzend in der Nutzenbewertung dargestellt.

- Endpunkt wird nicht als bewertungsrelevant eingestuft.

¹⁾ Dargestellte Endpunkte beziehen sich auf die pivotale Studie HPN-100-006 sowie die vom pU im Nutzendossier als patientenrelevant eingestuften Endpunkte.

²⁾ Todesfälle wurden im Rahmen der Sicherheitsbewertung erhoben und vom pU ergänzend dargestellt.

³⁾ Folgende neuropsychologischen Tests wurden je nach Altersgruppe erhoben: Bayley III, WPPSI-III, Child Behavior Checklist, BRIEF-P, BRIEF, WASI, Grooved Pegboard, California Verbal Learning Test II und Digit Span. Die Tests wurden

ausschließlich in den einarmigen supportiven Studien erhoben. Die Charakterisierung der Endpunkte wurde ergänzend im Anhang dargestellt.

- ⁴⁾ Vom pU wurden im Nutzendossier der BRIEF und BRIEF-P als patientenrelevant eingestuft und dargestellt, die anderen Tests sind den anderen supportiven Studien entnommen.
- ⁵⁾ Aufgrund fehlender Informationen zu verschiedenen Testgütekriterien, wie z. B. Änderungssensitivität, bleibt die Aussagekraft der Ergebnisse unklar. Die Ergebnisse werden nicht dargestellt. Die Charakterisierung der Endpunkte wurden ergänzend im Anhang dargestellt.
- ⁶⁾ Hyperammonämische Krisen wurden zusätzlich im Rahmen der Sicherheitsbewertung erhoben.
- ⁷⁾ Die Fragebögen zur Lebensqualität wurden in den einarmigen supportiven Studien erhoben (Ausnahme ist die Baseline-Erhebung des SF-36 bei erwachsenen Studienteilnehmern; Daten wurden in der pivotalen Studie HPN-100-006 erhoben). Die Charakterisierung der Endpunkte wurde ergänzend im Anhang dargestellt.
- ⁸⁾ Die Aussagekraft der Ergebnisse des SF-36 sind aufgrund der fehlenden Kontrollgruppe im Parallelgruppendesign eingeschränkt und sind ergänzend im Anhang dargestellt.

Abkürzungen: AUC: area under the curve; HAC: hyperammonämische Krise/n; pU: pharmazeutischer Unternehmer; SF-15: 15-Item Short Form Health Survey; SF-36: 36-Item Short Form Health Survey; (S)UE. (schwerwiegende/s) unerwünschte/s Ereignis/se; UCD: Harnstoffzyklusstörungen.

Mit dem Nutzendossier legte der pU Unterlagen für einige Endpunkte vor, die als nicht bewertungsrelevant eingeschätzt werden. Alle Endpunkte werden nachfolgend mit entsprechender Begründung der Einschätzung der Berücksichtigung für die Nutzenbewertung aufgeführt.

Mortalität

Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst und werden als patientenrelevant bewertet.

Morbidität

Bei den Endpunkten 24-Stunden-AUC-Ammoniakkonzentration im Blut und tägliche durchschnittliche Ammoniakkonzentration im Blut handelt es sich um einen Laborparameter. Der vom pU im Nutzendossier ausgeführten Begründung für eine Validierung des Endpunktes 24-Stunden-AUC-Ammoniakkonzentration im Blut als Surrogat für Morbiditäts- und Mortalitätsendpunkt wird nicht gefolgt. Die dargelegte Evidenz ist nicht ausreichend für eine Surrogatvalidierung (siehe

Tabelle 8), daher ist die Relevanz des Endpunktes unklar. Da es sich bei dem Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut um den primären Endpunkt der pivotalen Studie HPN-100-006 handelt, wird der Endpunkt ergänzend dargestellt. Bei dem Endpunkt tägliche durchschnittliche Ammoniakkonzentration im Blut wurden dieselben Blutproben wie beim Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut verwendet (je acht Blutproben an Studientag 14 und 28), allerdings wurden die Daten anders ausgewertet. Dieser Endpunkt wird nicht berücksichtigt und nicht dargestellt, da keine Daten vorliegen, die die Validität des Endpunktes als Surrogatparameter für patientenrelevante Endpunkte belegt.

Der Endpunkt neuropsychologische Tests wurde mit verschiedenen Instrumenten, angepasst an das jeweilige Alter der Studienpopulation, erfasst. Dazu gehörten die Testverfahren Bayley III, WPPSI-III, Child Behavior Checklist, BRIEF-P, BRIEF, WASI, Grooved Pegboard, California Verbal Learning Test II und Digit Span. Der Endpunkt wurde ausschließlich in den supportiven einarmigen Beobachtungsstudien erhoben. Vergleiche aus diesen Studien zwischen Baseline (NaPBA-Medikation) und Verlaufsdaten (Glycerolphenylbutyrat-Medikation) sind ohne eine prospektive parallele Vergleichsgruppe nicht aussagekräftig. Die Charakterisierungen der Endpunkte aus diesen Studien werden ergänzend im Anhang dargestellt. Aufgrund fehlender Informationen zu verschiedenen Testgütekriterien, wie z. B. Änderungssensitivität, bleibt die Aussagekraft der Ergebnisse unklar. Die Ergebnisse werden nicht dargestellt.

Der Endpunkt Anzahl und Schweregrad hyperammonämischer Krisen wurde im Rahmen der Sicherheitsbewertung erfasst und wird als patientenrelevant bewertet. Der Endpunkt wird zusätzlich für die Studie HPN-100-006 auch in der Kategorie Morbidität dargestellt.

Lebensqualität

Für den Endpunkt Lebensqualität wurden die Fragebögen SF-15 (für Kinder und Jugendliche) und SF-36 (für Erwachsene) verwendet. In der pivotalen Studie HPN-100-006, die Erwachsene ab 18 Jahren einschloss, wurde der SF-36 an Studientag 1 erhoben (Baseline-Erhebung). Zu den weiteren Beobachtungszeitpunkten wurde der Fragebogen nicht nochmals erhoben (z. B. an den Studientagen 14 und 28, nach der zweiwöchigen Behandlungsdauer). Somit sind keine validen Vergleiche zwischen den beiden Behandlungsperioden (NaPBA und Glycerolphenylbutyrat) aus der pivotalen Studie HPN-100-006 ableitbar.

Weiterhin wurden der SF-15 und SF-36 in den supportiven einarmigen Beobachtungsstudien verwendet. Die Vergleiche zum Endpunkt Lebensqualität aus den einarmigen Beobachtungsstudien zwischen Baseline (NaPBA-Medikation) und Verlaufsdaten (Glycerolphenylbutyrat-Medikation) sind ohne eine prospektive parallele Vergleichsgruppe nicht aussagekräftig. Die Charakterisierung der Endpunkte sowie die Verlaufsdaten des SF-36 werden ergänzend im Anhang dargestellt. Aufgrund fehlender Informationen zu verschiedenen Testgütekriterien, wie z. B. Änderungssensitivität, bleibt die Aussagekraft der Ergebnisse des SF-15 unklar. Die Ergebnisse des SF-15 werden nicht dargestellt.

Sicherheit

Die Endpunkte zur Sicherheit werden als patientenrelevant erachtet mit Ausnahme des UCD-Symptomprofils. Für die Sicherheitsendpunkte wurden sowohl die pivotalen Studie HPN-100-006 als auch die supportiven Studien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE herangezogen. Bei den unerwünschten Ereignissen (UE) sind zudem Todesfälle und die Anzahl hyperammonämischer Krisen berücksichtigt. Die Sicherheitsdaten für die supportiven einarmigen Beobachtungsstudien können nur im aggregierten intraindividuellen Verlauf dargestellt werden, da eine valide Kontrollgruppe im Parallelgruppendesign nicht erhoben wurde.

Der vom pU als patientenrelevant eingestufte Sicherheitsendpunkt UCD-Symptomprofil kann in der vorliegenden Nutzenbewertung nicht berücksichtigt werden. Unter diesem Endpunkt werden Symptome erfasst, die im Zusammenhang mit der Behandlung der Harnstoffzyklusstörung stehen. Diese Symptome werden prinzipiell auch als UE erfasst. Der pU argumentiert im Dossier, dass es nicht zu einer doppelten Zählung käme, d. h. dass beispielsweise ein Symptom wie Bauchschmerzen nicht im UCD-Symptomprofil und bei den UE erfasst würde. Der Grund dafür liege darin, dass die UE in den Kurzzeitstudien (die pivotalen Studie und die kurzfristigen supportiven einarmigen Beobachtungsstudien) und das UCD-Symptomprofil in den Langzeitstudien (die langfristigen supportiven einarmigen Beobachtungsstudien zur Sicherheitserweiterung) dargestellt werden würden. Dieser Argumentation des pU wird nicht gefolgt. In den Langzeitstudien HPN-100-005SE und HPN-100-007 werden UE bei jeder Untersuchung erfasst (insgesamt 14 Untersuchungen über 12 Monate, letzte Erfassung nach 12 Monaten), während das UCD-Symptomprofil drei Mal erfasst wurde: zu Baseline, Woche 1 und Monat 3. In den übrigen Langzeitstudien HPN-100-009 (Sicherheitserweiterung), HPN-100-011 und HPN-100-012SE wurde das UCD-Symptomprofil nicht als gesonderter Endpunkt erfasst, UE hingegen schon. Die

Symptome aus dem UCD-Symptomprofil sind somit ausreichend unter den UE erfasst (siehe auch Tabelle 9).

2.5.2 Verzerrungspotential auf Studien- und Endpunktebene

Die pivotale Studie HPN-100-006 ist eine randomisierte, aktiv kontrollierte, doppelt-verblindete Cross-over-Studie der Phase III. Die Studienteilnehmer wurden zufällig und unter Einsatz eines computergestützten zentralen Schemas randomisiert und im Verhältnis 1:1 den beiden Studienarmen NaPBA + Placebo → Glycerolphenylbutyrat + Placebo oder Glycerolphenylbutyrat + Placebo → NaPBA + Placebo zugeordnet.

Die Studie wurde doppelt verblindet durchgeführt. Hierbei ist zu beachten, dass das Medikament NaPBA einen aversiven Geruch und Geschmack hat [9,44]. Es ist unklar, inwiefern die gustatorischen und olfaktorischen Eigenschaften der Placebo-Medikamente dem Originalpräparat (NaPBA) gleichen. Sollten die Eigenschaften nicht vergleichbar sein, könnten sich die Studienteilnehmer ihrer Gruppenzugehörigkeit entsprechend klarwerden. Von einem Einfluss auf das Verzerrungspotential auf Studien- und Endpunktebene (Mortalität und Morbidität) ist jedoch nicht auszugehen. Es ist allerdings nicht auszuschließen, dass eine mögliche Entblindung einen Einfluss auf die unerwünschten Ereignisse haben könnte. Daher wird das Verzerrungspotential bei diesem Endpunkt als unklar klassifiziert.

Es ergeben sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Es ergeben sich zudem keine Hinweise auf sonstige, das Verzerrungspotential beeinflussende Aspekte wie beispielsweise Carry-over-Effekte (sowohl NaPBA als auch Glycerolphenylbutyrat haben relativ kurze Halbwertszeiten, siehe auch Abschnitt 3.2).

Das Verzerrungspotential auf Studienebene wird insgesamt mit niedrig bewertet (vgl. Tabelle 5). Das Verzerrungspotential auf Endpunktebene ist in Tabelle 6 dargestellt.

Hinsichtlich der supportiven einarmigen Beobachtungsstudien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE handelt es sich um Studien ohne Kontrollgruppe im Parallelgruppendesign. Es wird daher bei diesen Studien von einem hohen Verzerrungspotential auf Studien- und Endpunktebene ausgegangen.

Tabelle 5: Verzerrungspotential der Studie HPN-100-006

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patient	Verblindung Behandler	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
HPN-100-006	ja	ja	unklar ¹⁾	ja	nein	nein	niedrig

¹⁾ Es ist unklar, inwiefern die gustatorischen und olfaktorischen Eigenschaften der Placebo-Medikamente dem Originalpräparat (NaPBA) gleichen. Die Studienteilnehmer könnten sich daher ihrer Gruppenzugehörigkeit klarwerden.

Tabelle 6: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte in der Studie HPN-100-006

Endpunkte Studie HPN-100-006	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität					
Todesfälle	unklar ¹⁾	ja ³⁾	nein	nein	niedrig
Morbidität					
24-Stunden-AUC-Ammoniakkonzentration im Blut	unklar ¹⁾	ja ³⁾	nein	nein	niedrig
Anzahl und Schweregrad von HAC	unklar ^{1,2)}	ja ³⁾	nein	nein	niedrig
Sicherheit					
UE	unklar ^{1,2)}	ja ³⁾	nein	nein	unklar
SUE	unklar ^{1,2)}	ja ³⁾	nein	nein	unklar
Studienabbruch aufgrund UE	unklar ^{1,2)}	ja ³⁾	nein	nein	unklar

¹⁾ Es ist unklar, inwiefern die gustatorischen und olfaktorischen Eigenschaften der Placebo-Medikamente dem Originalpräparat (NaPBA) gleichen. Die Studienteilnehmer könnten sich daher ihrer Gruppenzugehörigkeit klarwerden.

²⁾ Im Falle eines medizinischen Notfalls gab es während der Studie HPN-100-006 jederzeit die Möglichkeit über ein interaktives Voice Response System die Verblindung aufzuheben. Eine solche Entblindung musste dem Studienleiter mitgeteilt werden. Während der Studie ist der Fall der Entblindung nicht aufgetreten.

³⁾ Die ITT-Population bezieht sich nicht auf die in der gesamten Studie randomisierten Studienpopulation. Es wurde ein Studienteilnehmer – aufgrund der Verletzung eines Einschlusskriteriums – nach der Randomisierung, aber vor der ersten Medikation ausgeschlossen.

Abkürzungen: AUC: area under the curve; HAC: hyperammonämische Krise/n; ITT: Intention-to-Treat; NaPBA: Natriumphénylbutyrat; (S)UE: (schwerwiegende/s) unerwünschte/s Ereignis/se.

2.5.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz

Laut SGB V § 35a Absatz 1 Satz 11 1. Halbs. gilt der medizinische Zusatznutzen durch die Zulassung als belegt. Nachweise gemäß § 5 Absatz 1 bis 6 sind nicht verpflichtend vorzulegen. Zur Beurteilung der Patientenrelevanz der vorgelegten Endpunkte wurden das Dossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Das Ergebnis der Betrachtung der Endpunkte findet sich aufgeteilt nach Endpunktkategorien in den Tabellen 7 bis 9.

Tabelle 7: Beschreibung der Endpunkte der Kategorie Mortalität

Endpunktkategorie Mortalität	
Endpunkt	Todesfälle
Operationalisierung	<p>Todesfälle wurden im Rahmen der Sicherheitsbewertung (siehe Tabelle 9) erfasst. Unerwünschte Ereignisse wurden gemäß CTCAE (Version 3.0) erfasst; Grad 5 entsprach dabei einem UE, das mit dem Todesfall zusammenhing. Todesfälle jeglichen Grundes wurden als SUE erfasst.</p> <p>Erhebungszeitpunkte <u>Studie HPN-100-006:</u> Die Erhebung des Endpunktes schloss den gesamten Studienzeitraum ein (ab dem Zeitpunkt der Einwilligungserklärung) und erstreckte sich bis 30 Tage nach der Einnahme der letzten Studienmedikation (siehe auch Tabelle 9).</p> <p><u>Supportive Studien:</u> Die Erhebung der Todesfälle im Rahmen der Sicherheitsbewertung erstreckte sich über den gesamten jeweiligen Studienzeitraum (siehe auch Tabelle 9).</p> <p>Die Gesamtrate an Todesfällen wurden in Modul 4 des Herstellerdossiers dargestellt.</p>
Bewertung	<p>Operationalisierung Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz Mortalität gilt als valider Endpunkt. Entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“</p>

Tabelle 8: Beschreibung der Endpunkte der Kategorie Morbidität

Endpunktkategorie Morbidität	
Endpunkt	24-Stunden-AUC-Ammoniakkonzentration im Blut
Operationalisierung	<p>Für den Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut wurde der Ammoniakspiegel im Blut über 24 Stunden an Tag 14 (Ende der ersten Studienphase) und an Tag 28 (Ende der zweiten Studienphase) ermittelt. Die beiden Erhebungstage waren jeweils am Ende der zweiwöchigen Behandlungsperiode. Zu diesen Zeitpunkten wurde ein „steady state“ für die Studienmedikation vorausgesetzt. An den beiden Tagen wurden Blutproben (je 2 ml venöses Blut) zu folgenden Zeitpunkten entnommen: vor der ersten Dosis sowie nach 2, 4, 8, 12, 16, 20 und 24 Stunden (relativ zur ersten Dosis des Tages gesehen).</p>

Endpunktkategorie Morbidität	
	<p>Die Blutproben wurden vom Labor des Prüfarztes bzw. Studienzentrums gemäß SOP (Standard Operating Procedures) analysiert. Da es sich um eine multizentrische Studie handelt, wurden die Blutproben in verschiedenen Laboren analysiert (insgesamt 19 Labore). Die ermittelten Ammoniakkonzentrationen können über die verschiedenen Labore variieren. Um diesen Abweichungen Rechnung zu tragen, wurden die ermittelten Werte über die beteiligten Labore standardisiert (für Details siehe Abschnitt 2.5.4).</p> <p>Für die Abschätzung der Ammoniakkonzentration im Blut über 24 Stunden wurde anschließend über die acht gemessenen Werte über 24 Stunden die AUC bestimmt. Bei fehlenden Werten wurden die fehlenden Messwerte imputiert oder als fehlende Messwerte klassifiziert (für Details siehe Abschnitt 2.5.4).</p> <p>Die Analyse der 24-Stunden-AUC-Ammoniakkonzentration im Blut erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung Die Definition und Erhebung des Endpunktes ist nachvollziehbar beschrieben.</p> <p>Validität und Patientenrelevanz Der pU führt den Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut als Surrogatparameter für Mortalität und Morbidität insbesondere für hyperammonämische Krisen auf. Laut pU führen eine Reduktion und eine Kontrolle des Ammoniakspiegels im Blut zu einer Verbesserung der Mortalität und der Morbidität. Grundlage für diese Argumentation sind zum einen Kohortenstudien zum natürlichen Krankheitsverlauf und Interventionsmöglichkeiten zur Kontrolle des Ammoniakhaushalts der Patienten [5,6,42,43]. Zum anderen zitiert der pU eine Studie unter der Kontrolle der Symptomatik mittels Glycerolphenylbutyrat (siehe dazu Ausführungen in Modul 4 des Dossiers).</p> <p>Hinsichtlich der Studien zum Krankheitsverlauf werden im Dossier (Modul 4) vier Studien zitiert [5,6,42,43], die eine Korrelation zwischen Ammoniak-Plasmaspiegel und Morbidität und Mortalität zeigen sollen. Bei den zitierten Quellen handelt es sich um Kohortenstudien, unkontrollierten Beobachtungsstudien oder Fallserien, in denen unter anderem Ammoniakkonzentration im Blut als Endpunkt erhoben wurde. Maße für eine Korrelation zwischen dem Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut (oder einem vergleichbaren Endpunkt zur Ammoniakkonzentration im Blut) und einem Morbiditäts- bzw. Mortalitätsendpunkt wurde vom pU nicht zitiert.</p> <p>Bei der Studie, die unter der Kontrolle der Symptomatik mittels Glycerolphenylbutyrat zitiert wird [38], handelt es sich um eine Beobachtungsstudie, bei der eine Kohorte aus medizierten UCD-Patienten beobachtet wurde. Die Kohorte setzt sich aus Studienteilnehmern verschiedener Studien zusammen, und zwar folgende: UP-1204-003, HPN-100-005, HPN-100-006 und HPN-100-012 und die jeweiligen Sicherheitserweiterungen. Insgesamt wurden 100 Patienten berücksichtigt. Die berücksichtigten Studienteilnehmer in dieser Validierungsstudie sind somit dieselben, die der pU auch im Dossier für die vorliegende Nutzenbewertung eingereicht hat. Der pU gibt an, dass die Studie folgende Zusammenhänge aufzeige (der maximale Beobachtungszeitraum der Studien betrug 12 Monate) [38]:</p> <ul style="list-style-type: none"> • Anzahl der Patienten mit einer HAC stieg proportional zu einem Anstieg des Ammoniakspiegels (gemessen mit ULN¹). • Zeitraum bis zur ersten HAC war kürzer bei Patienten mit einem Nüchternammoniakspiegel² $\geq 1,0$ ULN gegenüber Patienten mit einem Nüchternammoniakspiegel $< 0,5$ ULN. • Risiko einer HAC war größer bei Patienten mit einem Nüchternammoniakspiegel $\geq 1,0$ ULN gegenüber Patienten mit einem Nüchternammoniakspiegel $< 0,5$ ULN (Der Zusammenhang war in einer Sensitivitätsanalyse ohne Patienten < 6 Jahren noch ausgeprägter). • Kumulative Rate von HAC über 12 Monate korrelierte mit dem Nüchternammoniakspiegel. <p>Ein Korrelationskoeffizient oder ein anderes Zusammenhangsmaß zwischen dem Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut (oder einem vergleichbaren</p>

Endpunktkategorie Morbidität	
	<p>Endpunkt zur Ammoniakkonzentration im Blut) und einem Morbiditäts- bzw. Mortalitätsendpunkt wurde vom pU nicht zitiert. Zudem wurden die Zusammenhänge zwischen dem RCT (HPN-100-006) und den einarmigen Beobachtungsstudien nicht dargestellt.</p> <p>Der pU argumentiert weiterhin mit einem Dokument der EMA (CHMP zu HEPARSEC) in der Indikation UCD, in dem beschrieben sei, dass Ammoniak ein wesentliches Toxin sei, das im Blut kumuliert negative Ereignisse induziere. In dem entsprechenden Abschnitt stellt die EMA das Krankheitsbild der UCD dar [18]. Korrelationskoeffizienten zwischen dem Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut und einem Morbiditätsendpunkt sind nicht dargestellt.</p> <p>Die Darstellung von Zusammenhängen zwischen Blut-Ammoniak-Endpunkten und Morbiditäts- bzw. Mortalitätsendpunkten aus den dargestellten Beobachtungsstudien kann nicht mit einer Surrogatvalidierung gleichgesetzt werden. Kriterien für Surrogatendpunkte werden folgendermaßen im IQWiG-Methodenpapier (Version 5.0) dargestellt: „Ein Surrogatendpunkt kann dann als valide gelten, wenn der Effekt auf den zu ersetzenden patientenrelevanten Endpunkt durch den Effekt auf den Surrogatendpunkt in einem ausreichenden Ausmaß erklärt wird. [...] Diese Verfahren bedürfen in der Regel einer Metaanalyse von mehreren randomisierten Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden.“ [35]. Eine Möglichkeit der Validierung basiert u. a. auf der Darstellung von Korrelationen: „Da in der Nutzenbewertung des Instituts Aussagen bezogen auf Gruppen von Patientinnen und Patienten gemacht werden, stützt sich die Bewertung der Validität eines Surrogatendpunkts hier primär auf den Grad der Korrelation auf der Ebene der Behandlungseffekte. Zusätzlich zur Höhe der Korrelation wird für die Bewertung der Validität eines Surrogatendpunkts die Aussagesicherheit der Ergebnisse zur Validierung berücksichtigt.“ [35]. Für eine entsprechende Surrogatvalidierung für den Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut wäre dementsprechend mindestens ein Korrelationsmaß mit einem hohen Grad einer Korrelation zwischen dem Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut und dem Surrogatendpunkt (z. B. HAC) in mehreren unabhängigen RCT mit einer vergleichbaren Intervention und Population nötig, wobei auch die Aussagesicherheit der Ergebnisse entsprechend hoch sein müsste. Eine entsprechende Validierung mit der Darstellung des Grades der Korrelation sowie der Aussagesicherheit der Ergebnisse wurde vom pU nicht vorgelegt.</p> <p>Zusammenfassend liegen keine Daten vor, die die Validität des Endpunktes 24-Stunden-AUC-Ammoniakkonzentration im Blut als Surrogatparameter für patientenrelevante Endpunkte belegen. Damit ist die Relevanz des Endpunktes unklar. Da es sich bei dem Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut um den primären Endpunkt handelt, werden die Ergebnisse ergänzend dargestellt.</p>
Endpunkt	Anzahl und Schweregrad von HAC
Operationalisierung	<p>Der Endpunkt Anzahl und Schweregrad hyperammonämischer Krisen (HAC) wurde vom pU als sekundärer Endpunkt definiert und wurde zudem im Rahmen der Sicherheitsbewertung erfasst (siehe Tabelle 9).</p> <p>Das Auftreten einer HAC führte zum Ausschluss aus der Studie HPN-100-006, außerdem wurde das DSMB informiert.</p> <p>Eine HAC war definiert als das Auftreten klinischer Symptome in Verbindung mit einem Ammoniakspiegel $\geq 100 \mu\text{mol/l}$. Klinische Symptome waren:</p> <ul style="list-style-type: none"> • Erbrechen • Proteinintoleranz, d. h. körperliche Symptome treten wiederholt bei hoher Proteinaufnahme auf, sodass eine selbstinduzierte Reduktion der Proteinzufuhr erfolgt. • Lethargie • Psychose • Auffällige neurologische Untersuchung (Hypotonie, Spastiken, Hyperreflexie und/oder Klonus).

Endpunktkategorie Morbidität	
	<ul style="list-style-type: none"> • Gehirnödem (in MRT oder CT nachgewiesen) • Kopfschmerzen <p>Mögliche auslösende Faktoren (z. B. Infektionen, Änderung der Diät, Non-Compliance), ggf. Gründe für Non-Compliance mit der Studienmedikation, Ammoniaklevel bei Krankenhausaufnahme und Symptomatik wurden dokumentiert. Hospitalisation aufgrund einer HAC wurde als SUE gewertet.</p> <p>HAC wurden pro Behandlungsgruppe gezählt (Anzahl der Patienten mit mindestens einer HAC) ebenso der Schweregrad; der Gruppenvergleich sollte mittels Fisher's exact test erfolgen.</p>
Bewertung	<p>Operationalisierung</p> <p>Die qualitative Operationalisierung des Endpunkts erscheint sachgerecht. Eine Operationalisierung des Schweregrades, wie im Studienprotokoll vorgesehen, wird vom pU nicht vorgenommen; abgesehen davon, dass eine HAC, die zur Hospitalisierung führt, als SUE eingestuft wurde. Für UE (siehe Tabelle 9) wurde im Übrigen die Einteilung nach Schweregrad entsprechend der Vorgaben der CTCAE vorgenommen. Für UE, die nicht mittels CTCAE erfasst wurden, wurde hinsichtlich des Schweregrades folgende Einteilung vorgenommen:</p> <ul style="list-style-type: none"> • mild: keine Beeinträchtigung von Routineaktivitäten • moderat: Beeinträchtigung von Routineaktivitäten • schwer: Routineaktivitäten können nicht mehr ausgeführt werden • lebensbedrohlich: lebensbedrohliches oder zur Behinderung führendes UE • fatal: UE mit dem Endpunkt Tod <p>Validität und Patientenrelevanz</p> <p>Eine HAC ist als klinisches Ereignis mit Patientenrelevanz einzustufen, da es sich um eine potentiell lebensbedrohliche Krise handelt. Die klinische Symptomatik in Verbindung mit einem hohen Ammoniakspiegel sowie bildgebender Diagnostik ist als valide einzustufen.</p> <p>Es ist zu beachten, dass HAC auch als Sicherheitsendpunkt klassifiziert war.</p>

¹⁾ ULN bezieht sich auf den ermittelten Ammoniakspiegel im Blut, der von Labor zu Labor variieren kann. Aus dem Studienbericht der Studie HPN-100-006 ergibt sich, dass die ULN-Werte der 19 Labore zwischen 26 und 64 µmol/l variierten. Die Werte wurden anschließend auf den einheitlichen Wert von 35 µmol/l normalisiert. Unklar bleibt, ob diese Werte auch in der vorliegenden Validierungsstudie verwendet wurden.

²⁾ Der Nüchternammoniakspiegel wurde durch die beiden Messwerte der Prä-Dosis und nach 24 Stunden bei der Baseline-Messung der 24-Stunden-AUC Ammoniakkonzentration im Blut ermittelt.

Tabelle 9: Beschreibung der Endpunkte der Kategorie Sicherheit

Endpunktkategorie Sicherheit	
Endpunkt	UE
Operationalisierung	<p>Ein unerwünschtes Ereignis (UE) ist definiert als ein nachteiliges und ungewolltes Zeichen (einschließlich z. B. abnormale Laborbefunde), ein Symptom oder eine Erkrankung, welche zeitlich mit dem Gebrauch der Studienmedikation assoziiert sein konnten und/oder im Zusammenhang mit der Studienmedikation stehen konnten.</p> <p>Eine Klassifizierung der UE wurde gemäß Medical Dictionary for Regulatory Activities (MedDRA, Version 12.1) vorgenommen, die Einstufung des Schweregrads gemäß CTCAE (Common Terminology Criteria for Adverse Events) des National Cancer Institute (NCI), Version 3.0.</p> <p>Verschiedene UE wurden u. a. wie folgt zusammengefasst:</p> <ul style="list-style-type: none"> • Patienten mit UE • Patienten mit therapiebedingten (treatment emergent) UE • Patienten mit UE operationalisiert nach dem Begriffscluster der Systemorganklassen • Patienten mit SUE

Endpunktkategorie Sicherheit

- Patienten mit Therapieabbruch wegen UE

Ein therapiebedingtes (treatment emergent) UE war definiert als ein UE, dass nach Erhalt der initialen Studienmedikation bis zu sieben Tage nach der letzten Einnahme der Studienmedikation auftrat, oder ein vorher existierendes UE, welches sich im Studienzeitraum verschlechterte. Vor der Studie aufgetretene UE wurden bei der Visite zu Studienbeginn vor Verabreichung der ersten Studienmedikation erhoben und separat von den therapiebedingten UE erfasst.

Therapiebedingte UE beinhalteten folgendes:

- Jegliche Diagnosen, Zeichen, Symptome oder abnormale Laborwerte, die zu Studienbeginn noch nicht vorhanden, entdeckt oder beanstandet wurden und/oder zu Studienbeginn noch nicht im electronic Clinical Report Form dokumentiert wurden.
- Jegliche Diagnose, Zeichen, Symptom oder abnormale Laborwerte, die zu Studienbeginn notiert wurden und sich im Schweregrad oder der Intensität verschlechterten oder in der Häufigkeit erhöhten.

Ein UE wurde als schwerwiegend eingestuft, sofern es fatal oder lebensbedrohlich war, eine Verlängerung der Hospitalisierung nach sich zog, anhaltend war oder zu einer Behinderung oder Handlungsunfähigkeit führte, eine kongenitale Anomalie oder einen Geburtsfehler erzeugte oder aus anderen Gründen medizinisch bedeutsam war (medizinische oder chirurgische Eingriffe erforderlich, um eines der vorgenannten Ereignisse zu verhindern).

Für UE, die nicht mittels CTCAE erfasst wurden, wurde hinsichtlich des Schweregrades die folgende Einteilung vorgenommen:

- mild: keine Beeinträchtigung von Routineaktivitäten
- moderat: Beeinträchtigung von Routineaktivitäten
- schwer: Routineaktivitäten können nicht mehr ausgeführt werden
- lebensbedrohlich: lebensbedrohliches oder zur Behinderung führendes UE
- fatal: UE mit dem Endpunkt Tod

Jedes UE wurde vom Studienleiter hinsichtlich des Zusammenhangs mit der Studienmedikation in eine von drei Kategorien eingeordnet: nicht im Zusammenhang stehend, möglicherweise im Zusammenhang stehend und wahrscheinlich im Zusammenhang stehend.

Hyperammonämie bzw. hyperammonämische Krisen (siehe Endpunkt HAC, Tabelle 8) sollte als ein UE betrachtet werden, sofern der Prüfarzt dieses abnormale Ergebnis nicht einem ungenauen Testergebnis zuordnete. Sollte eine HAC auftreten, wurde der Prüfarzt gefragt, den Auslöser zu bewerten, sofern bekannt. Waren bei einer HAC die Kriterien für ein SUE erfüllt, so wurde das Ereignis entsprechend auch als SUE gewertet.

In der Studie HPN-100-006 sollte ein unabhängiges DSMB die Sicherheitsdaten begutachten, wenn ca. die Hälfte der geplanten Stichprobe eine Behandlung erhalten haben. Außerdem wurde das DSMB benachrichtigt, sobald die Kriterien für einen Studienabbruch erfüllt waren oder wenn HAC auftraten (siehe „stopping rules“ in Tabelle 2). Bei den supportiven Studien HPN-100-007 (Sicherheitserweiterung der Studie HPN-100-006) und HPN-100-011 war kein DSMB vorgesehen und somit auch keine Überprüfung der Sicherheitsdaten nach Einschluss einer gewissen Anzahl von Patienten. Bei den übrigen supportiven Studien gab es ein unabhängiges DSMB, das die Sicherheitsdaten begutachtet hat.

Insgesamt ist die Operationalisierung der Erhebung der Sicherheitsendpunkte auch in den supportiven Studien mit geringfügigen Abweichungen vergleichbar.

Endpunktkategorie Sicherheit	
	<p>Erhebungszeitpunkte <u>Studie HPN-100-006</u> Die Patienten wurden an jeder Studienvisite (Tag 1, 7, 14, 15, 21, 28 und 29) hinsichtlich UE befragt und diese bis zu 30 Tage nach Einnahme der letzten Studienmedikation erfasst. Alle SUE, die zu jeder Zeit mit der Erklärung der Einwilligung und bis zu 30 Tage nach der letzten Studienmedikation auftraten, mussten dem Sponsor oder Bevollmächtigtem innerhalb von 24 Stunden berichtet werden.</p> <p><u>Supportive Studien</u> Studie HPN-100-005: Die Patienten wurden bei jeder Studienvisite (Switch-over-Phase Tag 1, 7, 15, Sicherheitserweiterung Monat 0, 1, 2, 3, 6, 9, 12) hinsichtlich UE befragt und diese bis zu 30 Tage nach Einnahme der letzten Studienmedikation erfasst. Studie HPN-100-007: Die Patienten wurden bei jeder Studienvisite (Woche 1, Monat 2, 3, 6, 9, 12) hinsichtlich UE befragt und diese bis zu 30 Tage nach Einnahme der letzten Studienmedikation erfasst. Studie HPN-100-009: Die Patienten wurden bei jeder Studienvisite (Tag 1, 2, 3, 4, 7, Monat 1, 2, 3, 4, 5, 6, 9, 12, 15, 18, 21) hinsichtlich UE befragt und diese bis zu 30 Tage nach Einnahme der letzten Studienmedikation erfasst. Studie HPN-100-011: Anschließend an die vorhergehenden Studien wurden die Patienten bei jeder Studienvisite (alle Teilnehmer der Studien HPN-100-005, HPN-100-007, HPN-100-012, die die Studienvisite nach 12 Monaten absolviert haben: jeweils nach 6 Monaten, nach Ablauf von 12 Monaten jährlich) hinsichtlich UE befragt und diese bis zu 30 Tage nach Einnahme der letzten Studienmedikation erfasst. Studie HPN-100-012: Die Patienten wurden bei jeder Studienvisite (Switch-over-Phase Tag 1, 2, 10, Sicherheitserweiterung Woche 1, Monat 1, 3, 6, 9, 12) hinsichtlich UE befragt und diese bis zu 30 Tage nach Einnahme der letzten Studienmedikation erfasst.</p> <p>Alle supportiven Studien: Alle SUE, die zu jeder Zeit mit der Erklärung der Einwilligung und bis zu 30 Tage (Studie HPN-100-011: 7 Tage) nach der letzten Studienmedikation auftraten, mussten dem Sponsor oder Bevollmächtigtem innerhalb von 24 Stunden berichtet werden.</p> <p>Die Analyse der UE erfolgte auf Basis der Safety-Population.</p>
Bewertung	<p>Operationalisierung Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patienten-relevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“</p>

2.5.4 Statistische Methoden

Die folgenden Ausführungen beziehen sich auf die pivotale Studie HPN-100-006.

Fallzahlplanung

Eine Stichprobengröße von 44 auswertbaren Probanden sollte mit einer Power von 90 % zeigen, dass das Verhältnis der Mittelwerte der 24-Stunden-AUC für Ammoniak im Blut zwischen Glycerolphenylbutyrat und NaPBA 1,25 nicht übersteigt. Dies setzte ein einseitiges Alpha von 0,025, eine Standardabweichung der Intrasubjekt-Differenzen (natürliche Log-Skala) von 0,225 und ein erwartetes Verhältnis der Gruppennittel von 1 voraus.

Interimsanalysen

Es wurde keine formale Interimsanalyse geplant. Allerdings wurden ausgewählte Daten zur Sicherheit dem Data and Safety Monitoring Board (DSMB) zur Begutachtung vorgelegt. Die Studie sollte gestoppt werden, wenn zwei Patienten das gleiche Ereignis des CTCAE-Grades 4 erlitten (außer Hyperammonämie) oder bei 25 % von je 10 eingeschlossenen Patienten oder bei 5 Patienten (je nachdem, was zuerst zutrifft) eine Hyperammonämie eintrat. Sollte aus Sicherheitsgründen für einzelne Studienteilnehmer eine Entblindung erforderlich sein, dann sollten die Auswertungen durch einen Statistiker erfolgen, der nicht in die Studie involviert war.

Analysepopulationen

Für die Studie wurden vier Analysepopulationen definiert:

1. Safety-Population: umfasst alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Die Safety-Population wird für die Analysen der Baselinewerte, Therapie-treue, Demographie und Sicherheit herangezogen.
2. Intention-to-Treat (ITT): umfasst alle Patienten, die mindestens eine Dosis einer Studienmedikation (NaPBA oder Glycerolphenylbutyrat) erhalten haben. Die ITT-Population wird für Auswertungen zur Wirksamkeit und Pharmakokinetik herangezogen.
3. Modified Intention-to-Treat (MITT): umfasst alle Patienten der ITT-Population mit auswertbaren AUC-Werten für Ammoniak in beiden Behandlungsperioden und mit mindestens 4 Ammoniakproben, davon eine nach 8 oder 12 Stunden. Ausgewählte Analysen der Wirksamkeit und Pharmakokinetik werden mit dieser Population wiederholt.
4. Per-Protokoll (PP): umfasst alle Patienten der ITT-Population, die beide Medikamente erhalten haben und folgende Kriterien erfüllen:
 - o Kalkulierbare AUC für Ammoniak für beide Behandlungsperioden;
 - o mindestens 4 Ammoniakproben, davon eine nach 8 oder 12 Stunden;
 - o Ammoniakprobe zum Zeitpunkt Null, die nicht mehr als 60 Minuten nach der Dosierung und dem Frühstück entnommen wurde sowie 24-Stunden-Ammoniakprobe, die nicht mehr als 60 Minuten nach der Dosierung und dem Frühstück entnommen wurde;
 - o Compliance der Einnahme der Studienmedikation $\geq 80\%$ an den Tagen 14 und 28;
 - o keine Einnahme von Natriumbenzoat an den Tagen 14 und 28.

Statistische Modelle

Für kontinuierliche Variablen wurden Anzahl der Patienten, Mittelwerte, Standardabweichung, Median, Minimum und Maximum, für dichotome Variablen wurden Häufigkeiten und Prozentanteile ausgegeben.

Für den primären Endpunkt (24-Stunden-AUC-Ammoniakkonzentration im Blut) wurde ein alpha-Fehler von 0,025 (einseitig) festgelegt, für die Einschätzung der Nichtunterlegenheit wurde ein zweiseitiges 95%-Konfidenzintervall (95%-KI) angegeben. Für sekundäre Endpunkte wurden die p-Werte mittels Hochberg-Prozedur adjustiert.

Der primäre Endpunkt wurde sowohl für die Einheit $\mu\text{mol/l}$ wie für $\mu\text{g/dl}$ für die ITT-Population berechnet. Der Konversionsfaktor von $\mu\text{g/dl}$ auf $\mu\text{mol/l}$ war 0,5872. Bevor die Ammoniakwerte für den primären Endpunkt ausgewertet wurden, wurden die Werte anhand der Referenzbereiche aus einem Standardlabor normalisiert. Das war notwendig, da die Werte (um zeitnahe Ergebnisse zur Verfügung zu haben) in den lokalen Labors der Studienzentren gemessen wurden, dort aber etwas unterschiedliche Referenzwerte verwendet wurden. Die Kalkulation der AUC erfolgte mithilfe der Trapezregel.

Für den Gruppenvergleich wurde ein ANOVA-Modell mit den Faktoren Behandlung, Sequenz, Studienteilnehmer innerhalb einer Behandlungssequenz als Zufallsfaktor sowie Periode berechnet. Hierzu wurden Ammoniakwerte in den natürlichen Logarithmus transformiert. Das 95%-KI wurde mittels der Methode der kleinsten Quadrate der Ammoniakmittelwerte berechnet. Anschließend wurden die Werte rücktransformiert. Nichtunterlegenheit wurde festgestellt, wenn die obere Grenze des Konfidenzintervalls kleiner oder gleich 1,25 war; lag die obere Grenze unter 1 wurde Überlegenheit gefolgert. Außerdem wurden ein t-Test und ein Wilcoxon-Rangsummentest für den Gruppenvergleich durchgeführt (an originalskalierten Werten).

Umgang mit fehlenden Werten

Für den primären Endpunkt wurden Imputationsregeln festgelegt. Eine AUC wurde demnach als nicht kalkulierbar klassifiziert, wenn entweder der (1) Prä-Dosiswert und der 24-Stunden-Wert fehlten oder (2) das Sampling-Intervall weniger als 12 Stunden betrug. In diesen Fällen wurden keine Werte imputiert, sondern als Missing eingestuft.

Bei einer kalkulierbaren AUC, aber fehlenden Einzelmesswerten für Ammoniak, wurde folgendes Vorgehen festgelegt: Fehlte entweder der Prä-Dosiswert oder der 24-Stunden-Wert, dann wurde der jeweils fehlende Wert imputiert. Dafür wurde der nominale Wert der Prä-Dosis- oder der 24-Stunden-Messung verwendet, da von einem „steady state“ der Medikation ausgegangen wurde. Dazwischen fehlende Werte wurden nicht imputiert.

Bei als nicht-kalkulierbar eingestuften AUC-Werten wurde ein Resampling durchgeführt, d. h. es wurden andere Studienteilnehmer mit vergleichbarer Dosierung der Studienmedikation identifiziert (mit 20 % Abweichungstoleranz) und von diesen zufällig ausgewählte Werte für die Imputation verwendet. Für den Fall, dass bei einem Studienteilnehmer die AUC-Werte für beide Behandlungsperioden nicht kalkulierbar waren (NaPBA und Glycerolphénybutyrat), so wurden die fehlenden Werte nicht imputiert, sondern der Studienteilnehmer wurde ausgeschlossen.

Sensitivitätsanalysen

Sensitivitätsanalysen wurden für den primären Endpunkt vorgenommen, um den Effekt des Ausschlusses von Patienten mit mehr als einem fehlenden Wert auf den Endpunkt einzuschätzen. Dafür wurden zwei Sensitivitätsanalysen geplant: Eine, bei der alle Studienteilnehmer ausgeschlossen werden mit mehr als einem fehlenden Wert beim primären Endpunkt (insgesamt 8 Messwerte) und eine weitere, bei der alle Studienteilnehmer ausgeschlossen werden, die nicht-kalkulierbare AUC-Werte haben.

Subgruppen

A priori wurden im SAP keine Subgruppenanalysen festgelegt. Bei den Wirksamkeitsanalysen wurden für den primären Endpunkt Subgruppen auf Basis des Zeitpunkts der UCD-Diagnose post hoc festgelegt. Die Stichproben wurden in zwei Gruppen geteilt: Studienteilnehmer mit einer UCD-Diagnose vor dem 2. Lebensjahr und Studienteilnehmer mit einer Diagnose nach dem 2. Lebensjahr.

Amendment des SAP

Der SAP wurde einmal geändert. Die Änderung bezieht sich auf die Definition der MITT-Population. Ursprünglich umfasste die MITT-Population alle Patienten der ITT-Population mit auswertbaren AUC-Werten für Ammoniak in beiden Behandlungsperioden für 12 Stunden. Dies wurde geändert auf mit mindestens 4 Ammoniakproben, davon eine nach 8 oder 12 Stunden.

2.6 Ergebnisse zum Zusatznutzen

Für die Nutzenbewertung der Wirksamkeitsendpunkte wurde die pivotale Studie HPN-100-006 berücksichtigt. Diese Studie ist eine randomisierte, aktiv kontrollierte, doppelt-verblindete Cross-over-Studie, bei der erwachsene Studienteilnehmer mit Harnstoffzyklusstörungen mit NaPBA oder Glycerolphenylbutyrat behandelt wurden.

Für die Sicherheitsaspekte wurden zusätzlich zur pivotalen Studie auch die supportiven einarmigen Beobachtungsstudien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE berücksichtigt. Die Ergebnisse zu neuropsychologischen Tests und zur Lebensqualität aus den supportiven Studien werden im Anhang dargestellt.

2.6.1 Charakterisierung der Studienpopulation

2.6.1.1 Demographischen Eigenschaften und Baseline-Charakteristika der pivotalen Studie HPN-100-006

In der pivotalen Studie HPN-100-006 wurden insgesamt 46 Studienteilnehmer eingeschlossen. Ein Studienteilnehmer in Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation – wegen Verletzung eines Einschlusskriteriums – ausgeschlossen (Blutammoniak-Wert > 100 µmol/l). Ein Studienteilnehmer in Behandlungsarm A (NaPBA → Glycerolphenylbutyrat) erlitt unerwünschte Ereignisse aufgrund von Non-Compliance der Diät und wurde nach dem ersten Studientag ausgeschlossen. Zusätzlich erhielt ein Patient, der in Behandlungsarm A (NaPBA → Glycerolphenylbutyrat) randomisiert wurde, aufgrund eines Randomisierungsfehlers die falsche Reihenfolge der beiden Medikamente

(tatsächliche Reihenfolge wie in Behandlungsarm B: Glycerolphenylbutyrat → NaPBA). Eine Übersicht der Stichprobe der Studie ist in Tabelle 10 dargestellt.

Tabelle 10: Allgemeine Angaben zu der Studienpopulation in der Studie HPN-100-006

HPN-100-006	Tatsächliche Studien-Sequenz	
	NaPBA → Glycerolphenylbutyrat (N=22) n (%)	Glycerolphenylbutyrat → NaPBA (N=24) ¹⁾ n (%)
Übersicht Studienteilnehmer		
Randomisierte Patienten	22 (100) ¹⁾	24 (100) ¹⁾
Patienten ohne Behandlung	0	1 (4,2) ²⁾
Patienten mit Behandlung	22 (100)	23 (95,8)
Studie abgeschlossen	21 (95,5)	23 (95,8)
Studienabbruch, n (%)	1 (4,5)	1 (4,2)
unerwünschte Ereignisse	1 (4,5) ³⁾	0
Protokollverletzung	0	0
konnte nicht nachverfolgt werden	0	0
Einwilligung zurückgezogen	0	0
andere	0	1 (4,2) ²⁾
Safety-Population	NaPBA: 45 (97,8); Glycerolphenylbutyrat: 44 (95,7)	
ITT-Population	45 (97,8)	
MITT-Population	43 (93,5)	
Per-Protokoll-Population	43 (93,5) ¹⁾	

¹⁾ Ein Studienteilnehmer in Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) erhielt die Studienmedikation in der falschen Reihenfolge und wurde aus der Per-Protokoll-Analyse ausgeschlossen.

²⁾ Ein Studienteilnehmer in Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation – wegen Verletzung eines Einschlusskriteriums – ausgeschlossen (Blutammoniak-Wert > 100 µmol/l).

³⁾ Ein Studienteilnehmer in Behandlungsarm A (NaPBA → Glycerolphenylbutyrat) erlitt unerwünschte Ereignisse aufgrund von Non-Compliance der Diät und wurde nach dem ersten Studientag ausgeschlossen.

Abkürzungen: ITT: Intention-to-Treat; MITT: Modified Intention-to-Treat; NaPBA: Natriumphenylbutyrat.

In Tabelle 11 und Tabelle 12 finden sich die demographischen Eigenschaften, wie Geschlecht, Abstammung und Alter, sowie die Baseline-Charakteristika, wie Art des Enzymdefektes der Harnstoffzyklusstörung, Beginn der Harnstoffzyklusstörung, Methode zur Bestimmung der Harnstoffzyklusstörung, Dauer der vorangehenden NaPBA-Behandlung, Darreichungsform der vorangehenden NaPBA-Behandlung, tägliche NaPBA-Dosis, Anzahl von Patienten mit einer hyperammonämischen Krise (HAC) in den vergangenen 12 Monaten. In den beiden Tabellen sind die Studienteilnehmer nach Behandlungssequenz eingeteilt dargestellt. Da es sich um einen Cross-over-RCT handelt, haben alle Studienteilnehmer eine zweiwöchige Periode mit den beiden Studienmedikationen (Glycerolphenylbutyrat und NaPBA) erhalten. Es ist nicht davon auszugehen, dass ein Unterschied in den demographischen Eigenschaften oder den Baseline-Charakteristika zwischen den Gruppen einen Einfluss auf die Reihenfolge der Studienmedikationen hat. Zudem waren die Populationen der beiden Studienarme (Behandlungsarm A: NaPBA → Glycerolphenylbutyrat und Behandlungsarm B Glycerolphenylbutyrat → NaPBA) überwiegend vergleichbar.

Tabelle 11: Demographische Eigenschaften der ITT-Population in der Studie HPN-100-006

HPN-100-006	NaPBA → Glycerolphenylbutyrat (N=23) ¹⁾	Glycerolphenylbutyrat → NaPBA (N=22) ²⁾
<i>Geschlecht, n (%)</i>		
männlich	7 (30,4)	7 (31,8)
weiblich	16 (69,6)	15 (68,2)
<i>Alter, in Jahren</i>		
MW (SD)	36,9 (14,8)	28,4 (10,7)
Median (min;max)	32 (18;75)	23,5 (18;55)
<i>Abstammung, n (%)</i>		
weiß	18 (78,3)	17 (77,3)
schwarz oder afroamerikanisch	0	3 (13,6)
asiatisch	0	1 (4,5)
amerikanisch indianisch / Alaska-Eingeborener Eingeborener aus Hawaii oder anderen pazifischen Inseln	1 (4,3)	1 (4,5)
hispano- oder lateinamerikanisch	0	0
andere	3 (13)	0
	1 (4,3)	0

¹⁾ Ein Studienteilnehmer, der in den Behandlungsarm A (NaPBA → Glycerolphenylbutyrat) randomisiert wurde, erhielt die Studienmedikation tatsächlich in der falschen Reihenfolge (Glycerolphenylbutyrat → NaPBA).

²⁾ Ein Studienteilnehmer im Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation – wegen Verletzung eines Einschlusskriteriums – ausgeschlossen. Dieser wurde für die ITT-Population nicht berücksichtigt.

Abkürzungen: MW: Mittelwert; NaPBA: Natriumphenylbutyrat; SD: Standardabweichung.

Tabelle 12: Charakterisierung der Safety-Population in der Studie HPN-100-006

HPN-100-006	NaPBA → Glycerolphenylbutyrat (N=22) ¹⁾	Glycerolphenylbutyrat → NaPBA (N=23) ¹⁾
<i>Art des Enzymdefektes, n (%)</i>		
Mangel an:		
OTC	20 (90,9)	20 (87)
ASS	0	3 (13)
CPS	2 (9,1)	0
AL	0	0
ARG	0	0
Ornithin-Translokase	0	0
<i>Beginn der Harnstoffzyklusstörung, n (%)</i>		
neonatal (0 bis ≤ 30 Tage)	1 (4,5)	4 (17,4)
infantil (> 30 Tage bis ≤ 2 Jahre)	3 (13,6)	2 (8,7)
Kindheit (> 2 Jahre)	9 (40,9)	11 (47,8)
Erwachsenenalter (≥ 18 Jahre)	9 (40,9)	6 (26,1)
<i>Methode der UCD-Diagnostik, n (%)</i>		
Aminosäurenanalyse	4 (18,2)	5 (21,7)
DNA-Mutationsanalyse	12 (54,5)	12 (52,2)
Enzymanalyse der Leber	2 (9,1)	0
andere	4 (18,2)	6 (26,1)
<i>Dauer der NaPBA-Behandlung, in Monaten¹⁾</i>		
MW (SD)	121,6 (94,8)	135,5 (101,7) ²⁾
Median (min;max)	126 (0,2;300)	114,0 (0,2;300) ²⁾

HPN-100-006	NaPBA → Glycerolphenylbutyrat (N=22) ¹⁾	Glycerolphenylbutyrat → NaPBA (N=23) ¹⁾
Tägliche NaPBA-Dosis (vor Studienbeginn), in g MW (SD) Median (min;max)	14,3 (6,1) 15 (1,5;24)	14,7 (7,6) 15 (3;36)
Patienten mit mindestens einer HAC in den 12 Monaten vor Studieneintritt, n (%) ³⁾	4 (18,2)	5 (21,7)

¹⁾ Ein Studienteilnehmer, der in Behandlungsarm A (NaPBA → Glycerolphenylbutyrat) randomisiert wurde, erhielt die Studienmedikation tatsächlich in der falschen Reihenfolge (Glycerolphenylbutyrat → NaPBA) und wurde daher bei der Safety-Population Behandlungsarm B zugeordnet. Die Angaben aus der Tabelle waren aus den Unterlagen des pU nur für die Safety-Population zu entnehmen.

²⁾ In der Gruppe Glycerolphenylbutyrat → NaPBA wurden bei der Erfassung der Dauer der NaPBA-Behandlung vor Studieneintritt nur Daten von 22 Studienteilnehmern berichtet.

³⁾ Zu den auslösenden Faktoren zählten u. a. Non-Compliance mit der UCD-Medikation (NaPBA), Non-Compliance mit der Diät, Infektionen, interkurrente Erkrankungen oder andere Faktoren.

Abkürzungen: AL: Argininosuccinat-Lyase; ARG: Arginase 1; ASS: Argininosuccinat-Synthetase; CPS: Carbamoylphosphat-Synthetase 1; HAC: hyperammonämische Krise/n; MW: Mittelwert; NaPBA: Natriumphénylbutyrat; OTC: Ornithin-Transcarbamylase; SD: Standardabweichung; UCD: Harnstoffzyklusstörungen.

2.6.1.2 Medikation, Therapietreue und Proteinzufuhr der pivotalen Studie HPN-100-006

Weiterhin nahmen fast alle Studienteilnehmer während des Studienverlaufs eine neben der Studienbehandlung verabreichte Begleitmedikation ein (während NaPBA-Medikation: n=42 (93,3 %); während Glycerolphenylbutyrat-Medikation: n=43 (97,7 %)). Die Begleitmedikation wurde eingeteilt nach WHO-DD B2 (Version September 2009). Zu den häufigsten Begleitmedikationen (≥ 20 %) gehörten andere Psychostimulanzien und Nootropika (v. a. Citrullin; während NaPBA-Medikation: n=29 (64,4 %); während Glycerolphenylbutyrat-Medikation: n=29 (65,9 %)), Kalzium (während NaPBA-Medikation: n=11 (24,4 %); während Glycerolphenylbutyrat-Medikation: n=11 (25 %)) sowie Multivitamine, plain (während NaPBA-Medikation: n=9 (20 %); während Glycerolphenylbutyrat-Medikation: n=9 (20,5 %)). Ungleichheiten bei der Begleitbehandlung zwischen den Interventionen schienen während der jeweiligen Behandlung nicht vorzuliegen.

Die Studienmedikation war während der beiden Interventionen (NaPBA und Glycerolphenylbutyrat) vergleichbar. Ebenso war die Therapietreue („Compliance“) während der beiden Interventionen vergleichbar. Sowohl die Menge der eingenommenen Studienmedikation als auch die Therapietreue („Compliance“) wurden an den beiden Therapietagen 14 und 28 während des stationären Aufenthaltes erfasst (Kontrolle der Medikation). Demnach nahmen alle Studienteilnehmer mehr als 80 % der geplanten Medikation ein und alle bis auf ein Teilnehmer (während NaPBA Medikation) nahmen über 90 % der geplanten Medikation ein.

An den beiden Studientagen 14 und 28 wurde zudem die Proteinzufuhr erfasst und somit überprüft, ob die Studienteilnehmer sich an die Diät gehalten haben. Die Proteinzufuhr war während der beiden Interventionen (NaPBA und Glycerolphenylbutyrat) vergleichbar. Die Information zur Medikation, Therapietreue und Proteinzufuhr sind Tabelle 13 zu entnehmen.

Tabelle 13: Allgemeine Angaben zur Medikation, Therapietreue und Proteinzufuhr und in der Studie HPN-100-006

HPN-100-006	NaPBA (N=45 ¹⁾)	Glycerolphenylbutyrat (N=44 ¹⁾)
Begleitmedikation²⁾		
Patienten mit mind. einer Begleitmedikation, n (%)	42 (93,3)	43 (97,7)
Anzahl Patienten, n (%)		
andere Psychostimulanzien und Nootropika	29 (64,4)	29 (65,9)
Kalzium	11 (24,4)	11 (25)
Multivitamine, plain	9 (20)	9 (20,5)
Exposition der Studienmedikation³⁾		
<i>Tatsächliche gesamte Tagesdosis⁴⁾, in g/d</i>		
MW (SD)	14 (6,3) ⁴⁾	13,5 (6)
Median (min;max)	15 (1,5;36 ³⁾) ⁴⁾	14,4 (1,7;34,3)
<i>Tägliche Gesamtdosis⁴⁾, in g/m²/d</i>		
MW (SD)	7,9 (3,4) ⁴⁾	7,6 (3,2)
Median (min;max)	8,0 (0,7;16,1) ⁴⁾	7,9 (0,7;15,4)
<i>Tatsächliche gesamte Tagesdosis⁴⁾, in mg/kg/d</i>		
MW (SD)	203,6 (95,9) ⁴⁾	196,1 (90,7)
Median (min;max)	203,8 (13,9;457,5) ⁴⁾	195,7 (15,2;431,4)
<i>Anteil von Phenylbutyrat⁴⁾, in g/d</i>		
MW (SD)	12,3 (5,6) ⁴⁾	12,5 (5,5)
Median (min;max)	13,2 (1,3;31,7) ⁴⁾	13,3 (1,5;31,8)
Therapietreue („Compliance“)		
<i>Tatsächliche Tagesdosis / vorgeschriebene Tagesdosis × 100⁴⁾</i>		
MW (SD)	99,7 (1,9) ⁴⁾	100,3 (1)
Median (min;max)	100 (87,5;100,5) ⁴⁾	100 (99,2;103,8)
<i>Prozentklassen⁴⁾</i>		
< 80 %	0 ⁴⁾	0
80 % bis < 90 %	1 (2,2) ⁴⁾	0
90 % bis < 100 %	1 (2,2) ⁴⁾	7 (15,9)
100 %	41 (91,1) ⁴⁾	29 (65,9)
> 100 %	1 (2,2) ⁴⁾	8 (18,2)
fehlende Angaben	1 (2,2) ⁴⁾	0
Diätetische Proteinzufuhr		
<i>Baseline, in g/d</i>		
MW (SD)	46,8 (13,4)	46,3 (13,1)
Median (min;max)	45,0 (17,5;75,7)	45,0 (17,5;75,7)
<i>Während des Aufenthalts⁴⁾, in g/d</i>		
MW (SD)	45,9 (21,4) ⁴⁾	47,7 (22,8)
Median (min;max)	38,0 (12,4;112,6) ⁴⁾	41,9 (19,6;128,9)

¹⁾ Angaben beziehen sich auf die Safety-Population.

²⁾ Für die Kodierung wurde das WHO Drug Dictionary, Anatomical Therapeutic Chemical Codes (WHO-DD B2 Version September 2009) verwendet. Angegeben sind nur die am häufigsten verwendeten Medikamente (≥ 20 %).

³⁾ Die maximale Tagesdosis betrug 20 g NaPBA (entsprechen 17,4 ml Glycerolphenylbutyrat). Bei vier Patienten wurde eine höhere Dosis als 20 g genehmigt, um den Ammoniakspiegel zu kontrollieren.

⁴⁾ Die Werte beziehen sich auf die beiden stationären Studientage 14 und 28. Da ein Patient an Studientag 1 die Studie abgebrochen hat, liegen für diesen Patienten keine Werte vor. Daher beziehen sich die Werte für NaPBA nur auf N=44.

Abkürzungen: MW: Mittelwert; NaPBA: Natriumphenylbutyrat; SD: Standardabweichung; WHO: World Health Organization.

2.6.1.3 Charakteristika der supportiven Studien

Neben der pivotalen Studie HPN-100-006 wurden folgende supportiven einarmigen Beobachtungsstudien berücksichtigt: HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE.

Relevante Informationen zu den Studiencharakteristika dieser supportiven Studien finden sich in Tabelle 14. Die Populationen dieser Studien unterschieden sich vor allem hinsichtlich des Alters. Wie bereits in Abschnitt 2.2 dargestellt, wurden in der Studie HPN-100-009 Kinder zwischen 0 und 2 Jahren untersucht. Dabei wurden zum einen Neugeborene bis zum Alter von 2 Monaten untersucht (n=16) und zum anderen Babys ab 2 Monaten bis Kleinkinder bis 2 Jahre (n=10). Da das Medikament Glycerolphenylbutyrat ab einem Alter von 2 Monaten in Deutschland zugelassen ist, werden hier nur die Ergebnisse für Studienteilnehmer ab 2 Monaten dargestellt. Bei den Studien HPN-100-012 und HPN-100-012SE wurden Kinder im Alter zwischen 29 Tagen und 6 Jahren eingeschlossen. Tatsächlich war zum Zeitpunkt des Screenings kein Studienteilnehmer jünger als 2 Monate, sodass die Daten aller Studienteilnehmer bei der Darstellung berücksichtigt wurden. Bei den Studien HPN-100-005 und Studien HPN-100-005SE wurden Kinder und Jugendliche in einer Altersspanne von 6 bis 17 Jahren eingeschlossen.

Die einarmige Beobachtungsstudie HPN-100-007 stellt die Sicherheitserweiterung der pivotalen Studie HPN-100-006 dar. Studienteilnehmer aus Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) wechselten nach Abschluss der Studie HPN-100-006 für eine Teilnahme an der Sicherheitserweiterung (HPN-100-007) wieder zurück auf das Studienmedikament Glycerolphenylbutyrat. Zusätzlich wurden für die Sicherheitserweiterung neue Studienteilnehmer eingeschlossen, die vorher nicht an der Studie HPN-100-006 teilgenommen haben. Dabei wurden nicht nur erwachsene Studienteilnehmer (wie in HPN-100-006), sondern auch Studienteilnehmer mit einer Altersspanne von 6 bis 17 Jahren berücksichtigt.

Die Studie HPN-100-011 stellt eine weitere Sicherheitserweiterung der Studien HPN-100-012SE (Kinder im Alter zwischen 29 Tagen und 6 Jahren), HPN-100-005SE (Kinder und Jugendliche im Alter zwischen 6 und 17 Jahren) und HPN-100-007 (Kinder und Jugendliche im Alter zwischen 6 und 17 Jahren sowie Erwachsene) dar. Eine Übersicht über die Verteilung der Studienteilnehmer aus der Studie HPN-100-011 ist in Abbildung 2 dargestellt.

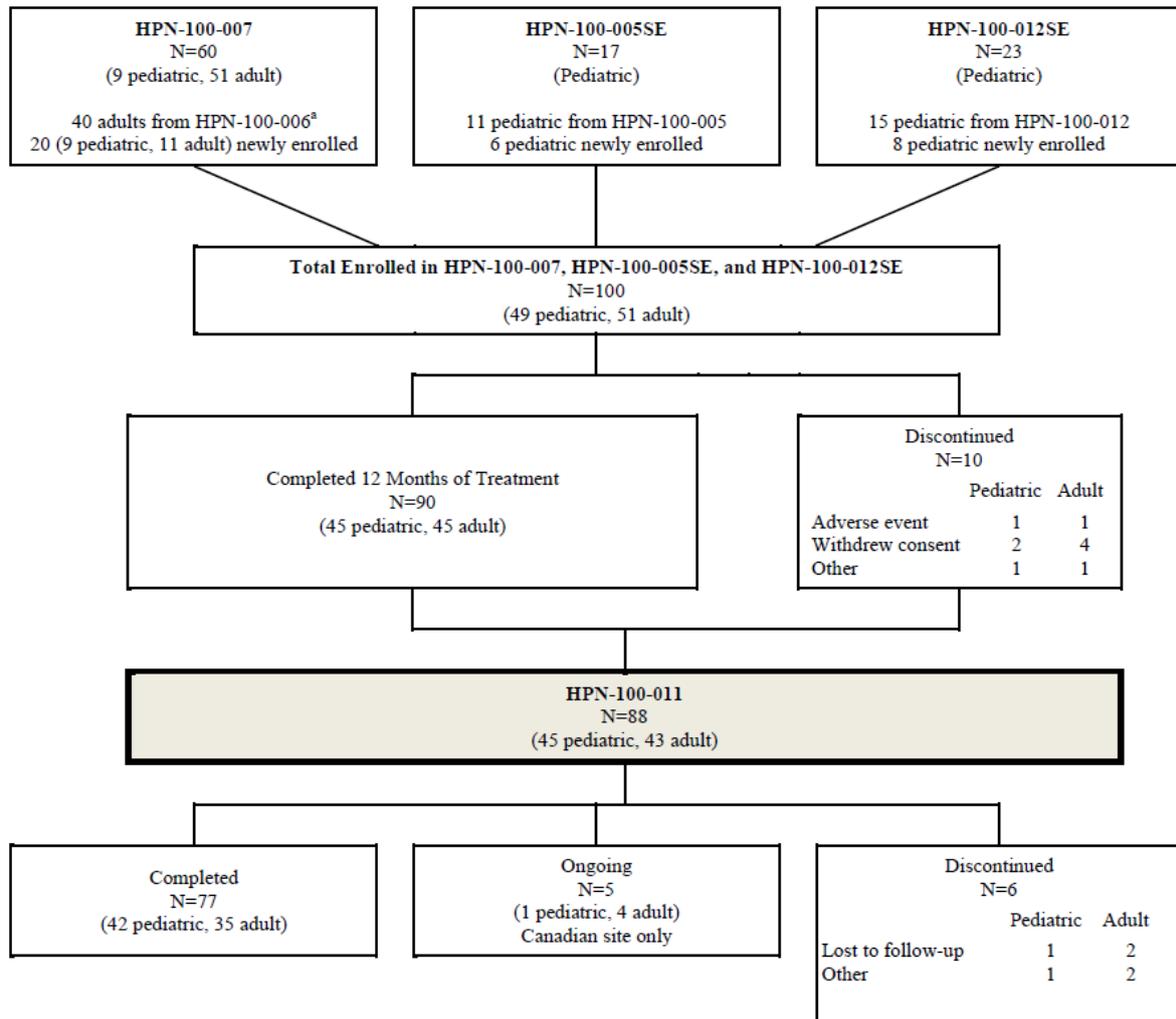


Abbildung 2: Patientenfluss der Studie HPN-100-011 [27]

Tabelle 14: Charakterisierung der Studienpopulation der supportiven einarmigen Beobachtungsstudien

Merkmal	HPN-100-005	HPN-100-005SE	HPN-100-007	HPN-100-009 ¹⁾	HPN-100-011	HPN-100-012	HPN-100-012SE
	≥ 6 bis < 18 J.	≥ 6 bis < 18 J.	≥ 6 J.	≥ 2 M. bis < 2 J.	alle Altersgr.	bis < 6 J.	bis < 6 J.
<i>Übersicht Studienteilnehmer, n (%)</i>							
eingeschlossene Patienten	11 (100)	17 (100)	60 (100)	10 (100)	88 (100)	15 (100)	23 (100)
Safety-Population	11 (100)	17 (100)	60 (100)	10 (100)	88 (100)	15 (100)	23 (100)
Studie abgeschlossen	11 (100)	16 (94,1)	53 (88,3)	6 (60)	77 (87,5)	15 (100)	21 (91,3)
Studie noch nicht beendet	-	-	-	-	5 (5,7)	-	-
Studie abgebrochen	0	1 (5,9)	7 (11,7)	4 (40)	6 (6,8)	0	2 (8,7)
<i>Studienabbruch, n (%)</i>							
unerwünschte Ereignisse		0	1 (1,7)	1 (10)	0		1 (4,3)
Protokollverletzung		0	0	0	0		0
konnte nicht nachverfolgt werden		0	0	1 (10)	3 (3,4)		0
Einwilligung zurückgezogen		1 (5,9)	5 (8,3)	0	0		0
andere		0	1 (1,7)	2 (20) ²⁾	3 (3,4)		1 (4,3)
<i>Altersklasse, n (%)</i>							
≥ 2 Monate bis < 2 Jahre	0	0	0	10 (100)	-	4 (26,7) ⁸⁾	7 (30,4) ⁸⁾
≥ 2 bis < 6 Jahre	0	0	0	0	16 (18,2) ⁵⁾	11 (73,3)	16 (69,6)
≥ 6 bis < 12 Jahre	7 (63,6)	11 (64,7)	6 (10)	0	22 (25)	0	0
≥ 12 bis < 18 Jahre	4 (36,4)	6 (35,3)	3 (5)	0	7 (8)	0	0
≥ 18 Jahre	0	0	51 (85)	0	43 (48,9)	0	0
<i>Alter (Jahre)</i>							
MW (SD)	10,2 (3,95)	10,0 (3,48)	28,8 (13,93)	0,82 (0,46) ³⁾	20 (15,76)	2,87 (1,89)	2,7 (1,85)
Median (min;max)	10 (6;17)	9 (6;17)	25 (6;60)	0,75 (0,36;1,73) ³⁾	17 (1;61)	3 (0;5)	3 (0;5)
<i>Geschlecht, n (%)</i>							
männlich	1 (9,1)	3 (17,6)	19 (31,7)	5 (50)	30 (34,1)	8 (53,3)	11 (47,8)
weiblich	10 (90,9)	14 (82,4)	41 (68,3)	5 (50)	58 (65,9)	7 (46,7)	12 (52,2)
<i>Abstammung, n (%)</i>							
weiß	9 (81,8)	14 (82,4)	48 (80)	6 (60)	74 (84,1)	12 (80)	19 (82,6)

Merkmal	HPN-100-005	HPN-100-005SE	HPN-100-007	HPN-100-009 ¹⁾	HPN-100-011	HPN-100-012	HPN-100-012SE
	≥ 6 bis < 18 J.	≥ 6 bis < 18 J.	≥ 6 J.	≥ 2 M. bis < 2 J.	alle Altersgr.	bis < 6 J.	bis < 6 J.
schwarz oder afroamerikanisch	1 (9,1)	1 (5,9)	4 (6,7)	2 (20)	5 (5,7)	2 (13,3)	2 (8,7)
asiatisch	1 (9,1)	1 (5,9)	4 (6,7)	0	4 (4,5)	0	1 (4,3)
amerikanisch indianisch / Alaska- Eingeborener	0	1 (5,9)	1 (1,7)	0	2 (2,3)	0	0
Eingeborener aus Hawaii oder anderen pazifischen Inseln	0	0	0	0	-	0	0
hispano- oder lateinamerikanisch	-	0	3 (5)	-	-	0	0
andere	0	0	0	2 (20)	3 (3,4)	1 (6,7)	1 (4,3)
<i>Art des Enzymdefektes, n (%)</i>							
Mangel an:							
OTC	9 (81,8)	14 (82,4)	49 (81,7)	2 (20)	61 (69,3)	3 (20)	6 (26,1)
ASS	1 (9,1)	2 (11,8)	4 (6,7)	2 (20)	9 (10,2)	3 (20)	6 (26,1)
CPS	0	0	1 (1,7)	2 (20)	0	0	0
AL	1 (9,1)	1 (5,9)	2 (3,3)	3 (30)	13 (14,8)	8 (53,3)	10 (43,5)
ARG	0	0	1 (1,7)	1 (10)	2 (2,3)	1 (6,7)	1 (4,3)
Ornithin-Translokase	0	0	3 (5)	0	3 (3,4)	0	0
<i>Beginn Harnstoffzyklusstörung, n (%)</i>							
neonatal (0 bis ≤ 30 Tage)	3 (27,3)	4 (23,5)	8 (13,3)	8 (80)	24 (27,3)	13 (86,7)	18 (78,3)
infantil (> 30 Tage bis ≤ 2 Jahre)	3 (27,3)	5 (29,4)	9 (15)	2 (20)	18 (20,5)	2 (13,3)	5 (21,7)
Kindheit (> 2 Jahre)	5 (45,5)	8 (47,1)	28 (46,7)	-	46 (52,3) ⁶⁾	-	-
Erwachsenenalter (> 18 Jahre)	-	-	15 (25)	-	-	-	-
<i>Dauer der NaPBA-Behandlung, in Monaten</i>							
MW (SD)	74,7 (48,2)	62,5 (45,7)	126,79 (96,95)	- ⁴⁾	87,38 (90,74)	19,29 (17,15)	20,84 (18,08)
Median (min;max)	76 (0,5;162)	68 (0,5;162)	120 (0,2;300)	- ⁴⁾	48 (0,2;300)	19 (0,2;62)	18,5 (0,2;62)
<i>Tägliche NaPBA-Dosis (vor Studien- beginn), in g/d</i>							
MW (SD)	12,4 (4,4)	11,3 (4,6)	14,08 (6,51)	4,89 (2,26) ⁴⁾	- ⁷⁾	5,27 (2,45)	- ⁹⁾
Median (min;max)	10,5 (8;20)	10,5 (2,3;20)	13,5 (1,5;36)	4,8 (1,8;8,0) ⁴⁾	- ⁷⁾	5,1 (0,9;9)	- ⁹⁾

Merkmal	HPN-100-005	HPN-100-005SE	HPN-100-007	HPN-100-009 ¹⁾	HPN-100-011	HPN-100-012	HPN-100-012SE
	≥ 6 bis < 18 J.	≥ 6 bis < 18 J.	≥ 6 J.	≥ 2 M. bis < 2 J.	alle Altersgr.	bis < 6 J.	bis < 6 J.
<i>Tatsächliche Glycerolphenylbutyrat-Dosis, in g/d</i>							
MW (SD)	11,9 (4,2)	10,8 (4,4)	13,4 (5,8)	4,5 (1,1)	11,1 (5,6)	5,2 (2,3)	5,1 (2,3)
Median (min;max)	9,9 (7,6;19,1)	10 (2,1;19,1)	13,9 (1,7;34,3)	4,3 (2,9;6,3)	9,9 (1,7;34,3)	5,0 (1,0;8,6)	5,3 (0,8;8,6)
<i>Anteil von Phenylbutyrat, in g/d</i>							
MW (SD)	11,0 (3,9)	10 (4,1)	12,4 (5,4)	4,2 (1)	10,3 (5,2)	4,8 (2,1)	4,7 (2,1)
Median (min;max)	9,2 (7,0;17,8)	9,3 (1,9;17,7)	12,9 (1,5;31,8)	4 (2,7;5,8)	9,2 (1,5;31,8)	4,6 (0,9;8,0)	4,9 (0,7;8,0)
Patienten mit mindestens einer Begleitmedikation, <i>n</i> (%)	11 (100)	16 (94,1)	60 (100)	10 (100)	87 (98,9)	15 (100)	23 (100)

¹⁾ In der Studie HPN-100-009 wurden Kinder zwischen 0 und 2 Jahren untersucht. Es gab zwei Studienberichte, getrennt nach Altersgruppe: Zum einen für Neugeborene bis zum Alter von 2 Monaten und zum anderen für Babys ab 2 Monaten bis Kleinkinder bis 2 Jahre. Da das Medikament Glycerolphenylbutyrat erst ab einem Alter von 2 Monaten in Deutschland zugelassen ist, werden hier nur die Ergebnisse für Studienteilnehmer ab 2 Monaten dargestellt.

²⁾ Zwei Studienteilnehmer haben die Studie aufgrund einer Lebertransplantation abgebrochen.

³⁾ Im Studienbericht ist das Alter in Monaten angegeben und wurde hier in Jahren umgerechnet.

⁴⁾ In der Studie HPN-100-009 war eine vorausgehende NaPBA-Medikation kein Einschlusskriterium. Werte zur Dauer waren den Unterlagen des pU nicht zu entnehmen. Die Werte zur Dosierung beziehen sich auf 7 Studienteilnehmer, die vorher entweder mit NaPBA oder Natriumbenzoat mediziert wurden.

⁵⁾ In der Studie HPN-100-011 wird im Studienbericht nicht zwischen den Altersgruppen 0 bis 2 Jahre und 2 bis 6 Jahre differenziert, der Wert bezieht sich also auf Kinder zwischen 0 und < 6 Jahren.

⁶⁾ In der Studie HPN-100-011 wird im Studienbericht beim Beginn der Harnstoffzyklusstörung nicht zwischen den Altersgruppen Kindheit und Erwachsenenalter differenziert.

⁷⁾ Die Studie HPN-100-011 stellt eine weitere Sicherheitserweiterung der supportiven Studien HPN-100-005SE, HPN-100-007 und HPN-100-012SE dar. Dementsprechend wurden die Studienteilnehmer nicht mit NaPBA vormediziert, sondern mit Glycerolphenylbutyrat.

⁸⁾ In den Studien HPN-100-012 und HPN-100-012SE wurde ab einem Alter von 29 Tagen rekrutiert. Zum Zeitpunkt des Screenings waren die jüngsten Studienteilnehmer mindestens 2 Monate alt.

⁹⁾ Aus den Unterlagen des pU waren zu der Studie HPN-100-012SE keine Werte zur NaPBA-Dosierung vor Studienbeginn zu entnehmen.

Abkürzungen: AL: Argininosuccinat-Lyase; Altersgr.: Altersgruppen; ARG: Arginase 1; ASS: Argininosuccinat-Synthetase; CPS: Carbamoylphosphat-Synthetase 1; MW: Mittelwert; NaPBA: Natriumphenylbutyrat; OTC: Ornithin-Transcarbamylase; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

2.6.2 Mortalität

Todesfälle

Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst (Abschnitt 2.6.5). In der vierwöchigen Cross-over-Studie HPN-100-006 traten keine Todesfälle auf.

2.6.3 Morbidität

24-Stunden-AUC-Ammoniakkonzentration im Blut

Der Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut ist als primärer Endpunkt der Studie HPN-100-006 definiert. Wie in Abschnitt 2.5.3 dargelegt, liegen keine Daten vor, die die Validität des Endpunktes 24-Stunden-AUC-Ammoniakkonzentration im Blut als Surrogatparameter für patientenrelevante Endpunkte belegen. Der Endpunkt wird als primärer Endpunkt der Studie ergänzend dargestellt (siehe Tabelle 8).

In dem Cross-over-RCT wurde der Ammoniakspiegel im Blut über 24 Stunden an Tag 14 (Ende der ersten Studienphase) und an Tag 28 (Ende der zweiten Studienphase) ermittelt, somit waren die beiden Erhebungstage am Ende der zweiwöchigen Behandlungsphase. Es ist davon auszugehen, dass ein „steady state“ zu diesen Zeitpunkten erreicht wurde. Die Studie war als Nichtunterlegenheitsstudie ausgelegt (definierte Nichtunterlegenheits-Grenze: 1,25).

In die Analyse der ITT-Population gingen 44 Studienteilnehmer der insgesamt 46 eingeschlossenen Patienten ein. Ein Studienteilnehmer wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation ausgeschlossen. Ein weiterer wurde nach der ersten Dosis NaPBA aufgrund von Nebenwirkungen ausgeschlossen (kein Wert zur 24-Stunden-AUC-Ammoniakkonzentration im Blut schätzbar).

Im Durchschnitt wiesen die Studienteilnehmer unter der NaPBA-Medikation 976,6 $\mu\text{mol h/l}$ (SD: 865,4 $\mu\text{mol h/l}$) 24-Stunden-AUC-Ammoniakkonzentration im Blut gegenüber 865,9 $\mu\text{mol h/l}$ (SD: 660,5 $\mu\text{mol h/l}$) 24-Stunden-AUC-Ammoniakkonzentration im Blut unter Glycerolphenylbutyrat-Medikation auf. Das Verhältnis der geometrischen Mittel betrug 0,91. Für die Beurteilung der Nichtunterlegenheit ist die obere Grenze des 95%-KI der Verhältnisse der geometrischen Mittel relevant und diese lag mit 1,034 unterhalb der a priori festgelegten Nichtunterlegenheitsschwelle von 1,25. Die Ergebnisse sind in Tabelle 15 dargestellt.

Bei der MITT-Analyse wurden 43 Studienteilnehmer berücksichtigt, da zwei Studienteilnehmer nicht ausreichend auswertbare AUC-Werte für Ammoniak aufwiesen (Voraussetzungen für die MITT-Population, siehe 2.5.4). Auch bei dieser Auswertung lag die obere Grenze des 95%-KI des Verhältnisses der geometrischen Mittel (1,024) unterhalb der Nichtunterlegenheitsschwelle.

In die Auswertung der PP-Analyse gingen ebenso 43 Studienteilnehmer ein. Zwei Studienteilnehmer erfüllten die Voraussetzungen für die Analyse nicht (siehe 2.5.4). Die obere Grenze des 95%-KI des Verhältnisses der geometrischen Mittel (1,030) dieser Analyse lag ebenso unterhalb der Nichtunterlegenheitsschwelle.

Tabelle 15: Ergebnisse zum Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut in der Studie HPN-100-006

24-Stunden-AUC-Ammoniakkonzentration im Blut ¹⁾ , in µmol h/l	Glycerolphenylbutyrat (N=44) ¹⁾			NaPBA (N=44) ¹⁾			Glycerolphenylbutyrat vs. NaPBA		
	N	MW (SD)	Median (min; max)	N	MW (SD)	Median (min; max)	MW (SD)	Median (min; max)	Verhältnis der geometrischen Mittel [95%-KI]; p-Wert ²⁾
ITT-Population	44 ³⁾	865,9 (660,5)	672,6 (206; 3351)	44 ³⁾	976,6 (865,4)	652,5 (302; 4666)	-111 (579)	-47 (-2953; 1007)	0,91 [0,799;1,034]; 0,211
MITT-Population	43 ⁴⁾	867,7 (668,2)	655,5 (206; 3351)	43 ⁴⁾	985,6 (873,5)	674,2 (302; 4666)	-118 (583,8)	-74 (-2953; 1007)	0,90 [0,789;1,024]; 0,192
PP-Population	43 ⁵⁾	868,3 (668,1)	655,5 (206; 3351)	43 ⁵⁾	985,5 (873,6)	674,2 (302; 4666)	-117,2 (584,2)	-74 (-2953; 1007)	0,90 [0,792;1,030]; 0,196

¹⁾ Die dargestellten Ergebnisse enthalten Imputationen für fehlende Werte; die dargestellten Werte sind rücktransformierte Werte auf die Originalskala (Varianzanalyse wurde mit logarithmierten Werten berechnet).

²⁾ Der p-Wert stammt aus einem gepaarten t-Test (Test auf Überlegenheit).

³⁾ In der ITT-Population wurden 44 der 46 Studienteilnehmer berücksichtigt: Ein Studienteilnehmer wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation ausgeschlossen. Ein weiterer wurde nach dem ersten Studientag aufgrund von Nebenwirkungen ausgeschlossen.

⁴⁾ In der MITT-Population wurden 43 der 46 Studienteilnehmer berücksichtigt: Ein Studienteilnehmer wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation ausgeschlossen. Zwei weitere Studienteilnehmer erfüllten die Voraussetzungen der MITT-Population nicht (siehe 2.5.4).

⁵⁾ In der PP-Population wurden 43 der 46 Studienteilnehmer berücksichtigt: Ein Studienteilnehmer wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation ausgeschlossen. Zwei weitere Studienteilnehmer erfüllten die Voraussetzungen der PP-Population nicht (siehe 2.5.4).

Abkürzungen: AUC: area under the curve; ITT: Intention-to-Treat; KI: Konfidenzintervall; MITT: Modified Intention-to-Treat; MW: Mittelwert; NaPBA: Natriumphénylbutyrat; PP: Per-Protokoll; SD: Standardabweichung.

In der dargestellten Analyse der ITT-Population wurden fehlende Werte imputiert (siehe Abschnitt 2.5.4). Sieben der 44 Studienteilnehmer hatten mindestens einen fehlenden Wert während der Glycerolphenylbutyrat-Medikation, während der NaPBA-Medikation waren es 5 Studienteilnehmer. Insgesamt fehlten 9 von 352 geplanten Messwerten unter Glycerolphenylbutyrat-Medikation und 7 von 352 geplanten Messwerten unter NaPBA-Medikation. Die fehlenden Messwerte konnten für die 24-Stunden-AUC-Ammoniakkonzentration im Blut mit der Methode für kalkulierbare AUC-Werte imputiert werden. Bei einem Studienteilnehmer fehlten die Messwerte an Tag 28, sodass für diesen Studienteilnehmer die Methode für nicht-kalkulierbare AUC-Werte angewendet wurde. Durch diese Imputationen konnte der Endpunkt für die 44 Studienteilnehmer der ITT-Population berechnet werden.

Die Autoren haben zwei Sensitivitätsanalysen berechnet. Bei der ersten wurden alle Studienteilnehmer ausgeschlossen, von denen mehr als ein Messwert fehlte; dies traf auf zwei Studienteilnehmer zu. Die Ergebnisse der Sensitivitätsanalyse mit 42 Studienteilnehmern lieferte vergleichbare Ergebnisse (Verhältnis der geometrischen Mittel [95%-KI]: 0,92 [0,8;1,047]) wie bei der oben dargestellten ITT-Analyse. Bei der zweiten Analyse wurden alle Studienteilnehmer ausgeschlossen, die nicht-kalkulierbare AUC-Werte hatten. Gemäß Studienbericht entsprach die Population der zweiten Sensitivitätsanalyse der Population der ITT-Analyse (siehe oben).

Anzahl und Schweregrad von HAC

Hyperammonämische Krisen (HAC) wurden im Rahmen der Sicherheitsbewertung erfasst (Abschnitt 2.6.5). Eine HAC ist eine potentiell lebensbedrohliche Krise und wird als patienten-relevant erachtet. In der Studie HPN-100-006 trat eine HAC auf.

Ein Patient aus Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) hat an Studientag 28 während der NaPBA-Medikation eine HAC erlitten. Der Prüfarzt bewertete diese mit Schweregrad 3 und als im Zusammenhang mit einer Non-Compliance mit der Studienmedikation stehend. Die Ergebnisse sind in Tabelle 16 zusammengefasst.

Tabelle 16: Ergebnisse zum Endpunkt Anzahl und Schweregrad hyperammonämischer Krisen in der Studie HPN-100-006 – Safety-Population

HPN-100-006 Safety-Population	Glycerolphenyl- butyrat ¹⁾		NaPBA ¹⁾		Glycerolphenylbutyrat vs. NaPBA
	N ²⁾	Patienten mit Ereignis, n(%)	N ²⁾	Patienten mit Ereignis, n(%)	RR [95%-KI] p-Wert ⁴⁾
Patienten mit HAC	44	0	45	1 (2,2) ³⁾	nicht berechnet

¹⁾ Alle Ereignisse für Studienteilnehmer, die mindestens eine Dosis der Studienmedikation erhalten haben.

²⁾ Anzahl Patienten in der Auswertung.

³⁾ Ein Patient aus Behandlungsarm B (Glycerolphenylbutyrat → NaPBA) erlitt an Studientag 28 eine HAC mit Schweregrad 3, die laut Prüfarzt im Zusammenhang mit fehlender Compliance zur Studienmedikation aufgetreten ist.

Abkürzungen: HAC: hyperammonämische Krise/n; KI: Konfidenzintervall; NaPBA: Natriumphenylbutyrat; RR: Relatives Risiko.

Neuropsychologische Tests

Aufgrund fehlender Informationen zu verschiedenen Testgütekriterien, wie z. B. der Änderungsensitivität, bleibt die Aussagekraft der Ergebnisse zu den Endpunkten der neuropsychologischen Tests unklar. Die Ergebnisse werden nicht dargestellt. Die Charakterisierung der Endpunkte wurden ergänzend im Anhang dargestellt.

2.6.4 Lebensqualität

In der pivotalen Studie HPN-100-006 wurde der SF-36 an Studientag 1 erhoben (Baseline-Erhebung). Zu den weiteren Beobachtungszeitpunkten wurde der Fragebogen nicht nochmals erhoben (z. B. an den Studientagen 14 und 28, nach der zweiwöchigen Behandlungsdauer). Somit sind keine validen Vergleiche zwischen den beiden Behandlungsperioden (NaPBA und Glycerolphenylbutyrat) aus der pivotalen Studie HPN-100-006 ableitbar.

Die Verlaufsdaten zur Lebensqualität aus dem SF-36 (Baseline-Erhebung der Lebensqualität aus der pivotalen Studie HPN-100-006 und Verlaufsdaten aus der einarmigen Beobachtungsstudie HPN-100-007) werden im Anhang dargestellt.

Aufgrund fehlender Informationen zu verschiedenen Testgütekriterien, wie z. B. der Änderungsensitivität, bleibt die Aussagekraft der Ergebnisse zum SF-15 unklar. Die Ergebnisse werden nicht dargestellt. Die Charakterisierung des Endpunkts SF-15 wurde ergänzend im Anhang dargestellt.

2.6.5 Sicherheit

Die nachfolgenden Darstellungen zur Sicherheit von Glycerolphenylbutyrat beziehen sich auf die Safety-Population der Studie HPN-100-006 (Tabellen 17 und 18) sowie ergänzend auf die supportiven Studien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE (Tabellen 19 bis 21).

2.6.5.1 Studie HPN-100-006

Die Safety-Population umfasst alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben.

Ein unabhängiges DSMB hat die Sicherheit der Studie HPN-100-006 überwacht. Nachdem ungefähr 50 % der Studienteilnehmer die Medikation erhalten haben, hat das DSMB die Daten begutachtet und keine Sicherheitsbedenken geäußert. Auch nach Abschluss der Studie wurden die Daten bewertet und keine Sicherheitsbedenken geäußert.

Es traten keine Todesfälle im Studienzeitraum auf.

Tabelle 17: Zusammenfassung der Ergebnisse für unerwünschte Ereignisse in der Studie HPN-100-006 – Safety-Population

Studie HPN-100-006	Patienten mit Ereignis ¹⁾ , n(%)		Glycerolphenylbutyrat vs. NaPBA
	Glycerolphenylbutyrat (N=44)	NaPBA (N=45)	RR [95%-KI]; p-Wert
UE	27 (61,4)	23 (51,1)	-
SUE	1 (2,3)	1 (2,2)	1,023 [0,066;15,847]; 0,9872
UE, das zum Abbruch der Studienmedikation führte	0	1 (2,2) ²⁾	nicht berechnet
UE CTCAE-Grad \geq 3	0	1 (2,2) ³⁾	nicht berechnet

¹⁾ Alle Ereignisse für Studienteilnehmer, die mindestens eine Dosis der Studienmedikation erhalten haben (Safety-Population).

²⁾ Ein Patient brach die Studie an Tag 1 aufgrund erhöhter Ammoniakwerte ab.

³⁾ Bei einem Studienteilnehmer kam es zu einer hyperammonämischen Krise (Noncompliance mit der Medikamenteneinnahme).

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; NaPBA: Natriumphenylbutyrat; RR: Relatives Risiko; (S)UE: (schwerwiegende/s) unerwünschte/s Ereignis/se.

Jeweils 51 bzw. 61 % der Patienten in den jeweiligen Behandlungsarmen berichteten mindestens ein UE. Bei zwei Patienten traten schwerwiegende unerwünschte Ereignisse (SUE) auf. Ein Patient im NaPBA-Arm hatte eine Hyperammonämie und wurde hospitalisiert; dieses Ereignis wurde als UE mit CTACE-Grad 3 eingestuft. Ein Patient im Glycerolphenylbutyrat-Arm hatte eine akute Gastroenteritis und wurde ebenfalls hospitalisiert.

Tabelle 18: Zusammenfassung der Ergebnisse für unerwünschte Ereignisse gemäß MedDRA-Systemorganklasse und Preferred Term mit einer Inzidenz $\geq 5\%$ in der Studie HPN-100-006 in einem der Studienarme der Safety-Population

Studie HPN-100-006	Patienten mit Ereignis¹⁾, n (%)	
	Glycerolphenylbutyrat (N=44)	NaPBA (N=45)
MedDRA-Systemorganklasse Preferred Term		
Erkrankungen des Gastrointestinaltrakts	16 (36,4)	13 (28,9)
Bauchbeschwerden	0	3 (6,7)
Bauchschmerzen	3 (6,8)	2 (4,4)
Diarrhö	7 (15,9)	3 (6,7)
Dyspepsie	2 (4,5)	3 (6,7)
Flatulenz	6 (13,6)	1 (2,2)
Übelkeit	1 (2,3)	3 (6,7)
Erbrechen	3 (6,8)	2 (4,4)
Erkrankungen des Nervensystems	7 (15,9)	7 (15,6)
Schwindel	0	4 (8,9)
Kopfschmerzen	6 (13,6)	4 (8,9)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	5 (11,4)	2 (4,4)
Fatigue	3 (6,8)	1 (2,2)
Stoffwechsel-und Ernährungsstörungen	3 (6,8)	4 (8,9)
Verminderter Appetit	3 (6,8)	2 (4,4)
Untersuchungen	3 (6,8)	2 (4,4)

¹⁾ Alle Ereignisse für Studienteilnehmer, die mindestens eine Dosis der Studienmedikation erhalten haben (Safety-Population).
Abkürzungen: NaPBA: Natriumphenylbutyrat; MedDRA: Medical Dictionary for Regulatory Activities.

Unterschiede in der Häufigkeit von jeglichen UE zwischen den Behandlungsarmen zeigten sich für einige gastrointestinale Symptome (Diarrhö, Flatulenz), die jeweils häufiger während der Behandlung mit Glycerolphenylbutyrat auftraten. Patienten unter NaPBA hatten häufiger Schwindel, Patienten unter Glycerolphenylbutyrat hatten häufiger Kopfschmerzen.

2.6.5.2 Supportive Studien

In den folgenden Tabellen werden die Sicherheitsdaten der supportiven Studien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE in einer Kurzübersicht ergänzend aufgelistet.

Die Operationalisierung der Erhebung der Sicherheitsendpunkte war in allen Studien mit geringfügigen Abweichungen vergleichbar. Der wesentliche Unterschied bestand darin, dass in der Studie HPN-100-007 (Sicherheitserweiterung der Studie HPN-100-006) sowie in der Studie HPN-100-011 kein DSMB vorgesehen war und damit auch keine Überprüfung der Sicherheitsdaten nach Einschluss einer gewissen Anzahl von Patienten.

In den folgenden Tabellen 20 und 21 findet sich außerdem eine Auswahl von UE, die entsprechend dem Risk Management Plan des EPAR als spezifisch für Glycerolphenylbutyrat anzusehen sind.

Tabelle 19: Kurzübersicht der Ergebnisse für unerwünschte Ereignisse und hyperammonämische Krisen in den supportiven Studien auf Patientenebene

Merkmal	HPN-100-005 ¹⁾	HPN-100-005SE	HPN-100-007	HPN-100-009	HPN-100-011	HPN-100-012 ⁹⁾	HPN-100-012SE
	≥ 6 bis < 18 J. (N=11)	≥ 6 bis < 18 J. (N=17)	≥ 6 J. (N=60)	≥ 2 M. bis < 2 J. (N=10)	alle Altersgr. (N=88)	bis < 6 J. (N=15)	bis < 6 J. (N=23)
<i>Altersklasse, n (%)</i>							
≥ 2 Monate bis < 2 Jahre	0	0	0	10 (100)	-	4 (26,7) ¹⁰⁾	7 (30,4) ¹⁰⁾
≥ 2 bis < 6 Jahre	0	0	0	0	16 (18,2) ⁷⁾	11 (73,3)	16 (69,6)
≥ 6 bis < 12 Jahre	7 (63,6)	11 (64,7)	6 (10)	0	22 (25)	0	0
≥ 12 bis < 18 Jahre	4 (36,4)	6 (35,3)	3 (5)	0	7 (8)	0	0
≥ 18 Jahre	0	0	51 (85)	0	43 (48,9)	0	0
<i>Dauer der Exposition, in Tagen</i>							
Mittelwert (SD)	- ²⁾	363,5 (13,46)	333,6 (83,05)	9,13 (6,84) ⁵⁾	613,6 (367,35)	- ¹¹⁾	345 (63,93)
Median (min; max)	- ²⁾	365 (329;388)	360 (22;387)	9,33 (0,2;20,3) ⁵⁾	674 (1;1696)	- ¹¹⁾	360 (118; 396)
Jegliches UE, n (%)	4 (36,4)	16 (94)	59 (98)	10 (100)	74 (84,1)	6 (40)	23 (100)
Jegliches SUE, n (%)	0	3 (18)	12 (20)	6 (60)	30 (34,1)	0	11 (48)
UE, das zum permanenten Abbruch der Studienmedikation führte, n (%)	0	0	2 (3,5) ⁴⁾	1 (10) ⁶⁾	2 (2,3) ⁸⁾	0	1 (4) ¹²⁾
UE CTCAE-Grad ≥ 3, n (%)	0	1 (6)	11 (18)	4 (40) ⁶⁾	24 (27,2)	0	8 (35)
<i>Hyperammonämische Krisen,</i>							
Anzahl, n (%)	0	3 (17,6)	9 (15)	3 (30)	20 (22,7)	0	7 (30,0)
Mittelwert (SD)		0,18 (0,393) ³⁾	0,2 (0,514) ³⁾	0,8 (1,39)	0,49 (1,259) ³⁾		0,52 (1,123) ³⁾
min;max		0;1	0;2	0;4	0;8		0;5

¹⁾ Zweiwöchige Switch-Over-Studie (Woche 1: NaPBA, Woche 2: Glycerolphenylbutyrat). Die extrahierten Daten beziehen sich auf die Behandlungsphase mit Glycerolphenylbutyrat.

²⁾ In der Studie HPN-100-005 schlossen alle 11 Studienteilnehmer die zweiwöchige Behandlung ab. Die Behandlung mit Glycerolphenylbutyrat dauerte eine Woche.

³⁾ Bezogen auf Anzahl der Krisen pro Patient (6).

⁴⁾ Bei je einem Patienten Widerruf der Einwilligungserklärung sowie Abbruch nach Dosisreduktion aufgrund diverser fortbestehender UE.

⁵⁾ In der Studie HPN-100-009 wurden die Angaben zur Dauer der Exposition in Monaten angegeben.

⁶⁾ In der Studie HPN-100-009 ist ein Patient an Komplikationen eines Eingriffs verstorben, nicht kausal durch Studienmedikation bedingt.

⁷⁾ In der Studie HPN-100-011 wird im Studienbericht nicht zwischen den Altersgruppen 0 bis 2 Jahre und 2 bis 6 Jahre differenziert, der Wert bezieht sich also auf Kinder zwischen 0 und < 6 Jahren.

⁸⁾ Bei je einem Patienten Absetzen der Studienmedikation nach Hospitalisierung aufgrund von Hypokapnie bzw. aufgrund von Panzytopenie.

⁹⁾ Zehntägige Switch-Over-Studie (Tag 1: NaPBA, Tag 2 bis 10: Glycerolphenylbutyrat). Die extrahierten Daten beziehen sich auf die Behandlungsphase mit Glycerolphenylbutyrat.

¹⁰⁾ In den Studien HPN-100-012 und HPN-100-012SE wurde ab einem Alter von 29 Tagen rekrutiert. Zum Zeitpunkt des Screenings waren die jüngsten Studienteilnehmer mindestens 2 Monate alt.

¹¹⁾ In der Studie HPN-100-012 schlossen alle 15 Studienteilnehmer die zehntägige Behandlung ab. Die Behandlung mit Glycerolphenylbutyrat dauerte neun Tage.

¹²⁾ Widerruf der Einwilligungserklärung durch die Eltern nach einer hyperammonämischen Krise.

Abkürzungen: Altersgr.: Altersgruppen; CTCAE: Common Terminology Criteria for Adverse Events; MW: Mittelwert; NaPBA: Natriumphenylbutyrat; SD: Standardabweichung; (S)UE: (schwerwiegende/s) unerwünschte/s Ereignis/se.

Tabelle 20: Kurzübersicht über eine Auswahl von Ergebnissen für unerwünschte Ereignisse, die entsprechend dem Risk Management Plan des EPAR als spezifisch für Glycerolphenylbutyrat eingeschätzt sind, gemäß MedDRA-Systemorganklasse und Preferred Term mit einer Inzidenz ≥ 10 % in den supportiven Studien

MedDRA-Systemorganklasse Preferred Term, n (%)	HPN-100-005 ¹⁾	HPN-100-005SE	HPN-100-007	HPN-100-009	HPN-100-011	HPN-100-012 ²⁾	HPN-100-012SE
	≥ 6 bis < 18 J. (N=11)	≥ 6 bis < 18 J. (N=17)	≥ 6 J. (N=60)	≥ 2 M. bis < 2 J. (N=10)	alle Altersgr. (N=88)	bis < 6 J. (N=15)	bis < 6 J. (N=23)
Erkrankungen des Gastrointestinaltrakts	3 (27,3)	7 (41)	35 (58)	7 (70)	27 (30,7)	3 (20)	15 (65)
Bauchschmerzen	2 (18,2)	3 (18)	6 (10)				
Diarrhö			11 (18)	1 (10)	10 (11,4)		6 (26)
Übelkeit			13 (22)				
Erbrechen		4 (24)	19 (32)	4 (40)	11 (12,5)	2 (13,3)	15 (52)
Verstopfung				2 (20)			
Pneumatisis intestinalis				1 (10)			
Teething				1 (10)			
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort		4 (24)	16 (27)	3 (30)			9 (39)
Fatigue			7 (12)				
Fieber		2 (12)		3 (30)			9 (39)

MedDRA-Systemorganklasse Preferred Term, n (%)	HPN-100-005 ¹⁾	HPN-100-005SE	HPN-100-007	HPN-100-009	HPN-100-011	HPN-100-012 ²⁾	HPN-100-012SE
	≥ 6 bis < 18 J. (N=11)	≥ 6 bis < 18 J. (N=17)	≥ 6 J. (N=60)	≥ 2 M. bis < 2 J. (N=10)	alle Altersgr. (N=88)	bis < 6 J. (N=15)	bis < 6 J. (N=23)
Stoffwechsel- und Ernährungsstörungen		5 (29)	20 (33)	5 (50)	31 (35,2)		13 (57)
Verminderter Appetit		2 (12)	8 (13)				3 (13)
Hyperammonämie		3 (18)	9 (15)	3 (30)	21 (23,9)		7 (30)
Hypophagie				2 (20)			
Metabolische Azidose				2 (20)			
Erkrankungen des Nervensystems		3 (18)	27 (45)	1 (10)	23 (26,1)		7 (30)
Schwindel			9 (15)				
Kopfschmerzen		2 (12)	10 (17)		9 (10,2)		
Motorische Entwicklungsstörung				1 (10)			
Untersuchungen		8 (47)		3 (30)	27 (30,7)		6 (26)
Aminosäurewert reduziert				1 (10)			
Ammoniakwert erhöht				1 (10)	10 (11,4)		
ALT erhöht		2 (12)					
Anionenlücke erhöht		2 (12)					
AST erhöht		2 (12)					
Kohlensäure reduziert				1 (10)			
Alkalische Phosphatase erhöht							

¹⁾ Zweiwöchige Switch-Over-Studie (Woche 1: NaPBA, Woche 2: Glycerolphenylbutyrat). Die extrahierten Daten beziehen sich auf die Behandlungsphase mit Glycerolphenylbutyrat.

²⁾ Zehntägige Switch-Over-Studie (Tag 1: NaPBA, Tag 2 bis 10: Glycerolphenylbutyrat). Die extrahierten Daten beziehen sich auf die Behandlungsphase mit Glycerolphenylbutyrat.

Abkürzungen: ALT: Alaninaminotransferase; Altersgr.: Altersgruppen; AST: Aspartataminotransferase; MedDRA: Medical Dictionary for Regulatory Activities; MW: Mittelwert; NaPBA: Natriumphenylbutyrat; SD: Standardabweichung.

Tabelle 21: Kurzübersicht der Ergebnisse für unerwünschte Ereignisse CTCAE-Grad ≥ 3 gemäß MedDRA-Systemorganklasse und Preferred Term in den supportiven Studien

MedDRA-Systemorganklasse Preferred Term, n (%)	HPN-100-005 ¹⁾	HPN-100-005SE	HPN-100-007	HPN-100-009	HPN-100-011	HPN-100-012 ²⁾	HPN-100-012SE
	≥ 6 bis < 18 J. (N=11)	≥ 6 bis < 18 J. (N=17)	≥ 6 J. (N=60)	≥ 2 M. bis < 2 J. (N=10)	alle Altersgr. (N=88)	bis < 6 J. (N=15)	bis < 6 J. (N=23)
Jegliches UE CTCAE-Grad 3	0	1 (6)	11 (18)	2 (20)	23 (26,1)	0	8 (34,8)
UE CTCAE-Grad 4 oder 5	0	0	0	2 (20) ³⁾	1 (1,1)	0	0
Erkrankungen des Gastrointestinaltrakts			2 (3)	1 (10)	4 (4,5)		2 (8,7)
Pneumatis intestinalis				1 (10)			
Bauchschmerzen			1 (2)		3 (3,4)		
Übelkeit			1 (2)				
Erbrechen					1 (1,1)		2 (8,7)
Hepatobiliäre Erkrankungen			1 (2)		1 (1,1)		
Schmerzen der Gallenblase			1 (2)				
Biliäre Dyskinesie					1 (1,1)		
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort		1 (6)		1 (10)	2 (2,3)		
Fieber		1 (6)		1 (10)			
Ödem					1 (1,1)		
Schmerzen					1 (1,1)		

MedDRA-Systemorganklasse Preferred Term, n (%)	HPN-100-005 ¹⁾	HPN-100-005SE	HPN-100-007	HPN-100-009	HPN-100-011	HPN-100-012 ²⁾	HPN-100-012SE
	≥ 6 bis < 18 J. (N=11)	≥ 6 bis < 18 J. (N=17)	≥ 6 J. (N=60)	≥ 2 M. bis < 2 J. (N=10)	alle Altersgr. (N=88)	bis < 6 J. (N=15)	bis < 6 J. (N=23)
Stoffwechsel- und Ernährungsstörungen			8 (13)	3 (30)	13 (14,8)		7 (30,4)
Verminderter Appetit			1 (2)				
Hyperammonämie			7 (12)	2 (20)	13 (14,8)		6 (26,1)
Hypophagie				1 (10)			1 (4,3)
Dehydratation					1 (1,1)		
Erkrankungen des Nervensystems			1 (2)		1 (1,1)		1 (4,3)
Periphere Neuropathie			1 (2)				
Konvulsion					1 (1,1)		1 (4,3)
Untersuchungen				1 (10)	2 (2,3)		
Ammoniakwert erhöht				1 (10)			
Aminosäurewert erhöht					1 (1,1)		
Gewicht reduziert					1 (1,1)		
Psychiatrische Erkrankungen			1 (2)		1 (1,1)		
Psychose			1 (2)				
Psychologische Faktoren mit Einfluss auf die Erkrankung					1 (1,1)		

¹⁾ Zweiwöchige Switch-Over-Studie (Woche 1: NaPBA, Woche 2: Glycerolphenylbutyrat). Die extrahierten Daten beziehen sich auf die Behandlungsphase mit Glycerolphenylbutyrat.

²⁾ Zehntägige Switch-Over-Studie (Tag 1: NaPBA, Tag 2 bis 10: Glycerolphenylbutyrat). Die extrahierten Daten beziehen sich auf die Behandlungsphase mit Glycerolphenylbutyrat.

³⁾ Ein Patient an Komplikationen eines Eingriffs verstorben, nicht kausal durch Studienmedikation bedingt

Abkürzungen: Altersgr.: Altersgruppen; MedDRA: Medical Dictionary for Regulatory Activities; MW: Mittelwert; NaPBA: Natriumphénylbutyrat; SD: Standardabweichung; (S)UE: (schwerwiegende/s) unerwünschte/s Ereignis/se.

Die Sicherheitsdaten aus den supportiven Studien spiegeln UE wider, die unter der Behandlung mit Glycerolphenylbutyrat aufgetreten sind. Der Anteil der Patienten mit jeglichem UE lag zwischen 84 und 100 %, der Anteil mit SUE zwischen 18 und 60 % (siehe Tabelle 19; in den Switch-over-Studien HPN-100-005 und HPN-100-012 waren die Anteile wegen der kurzen Laufzeit entsprechend niedriger). Die von jeglichem UE am häufigsten betroffenen Systemorganklassen waren Erkrankungen des Gastrointestinaltrakts, Allgemeine Erkrankungen und Beschwerden am Verabreichungsort, Stoffwechsel- und Ernährungsstörungen, Erkrankungen des Nervensystems und Untersuchungen. Das Spektrum der UE in den supportiven Studien war dabei je nach Altersgruppe unterschiedlich; bspw. traten bei Kindern Infektionserkrankungen (u. a. Erbrechen, Diarrhö, Fieber) häufig auf (siehe Tabelle 20).

SUE waren mit 48 bzw. 60 % häufiger in den Studien mit Kindern bis 6 Jahren (und HPN-100-012SE und Studien HPN-100-009) im Vergleich zu den Studien mit Kindern über 6 Jahren und Erwachsenen (HPN-100-005SE und Studien HPN-100-007) mit 18 bzw. 20 % (siehe Tabelle 19). Das häufigste UE CTCAE-Grad ≥ 3 war Hyperammonämie (zwischen 12 und 30 % der Studienteilnehmer betroffen). Eine Übersicht findet sich in Tabelle 21).

Ein Patient in der Studie HPN-100-009 verstarb im Studienzeitraum an den Folgen eines Eingriffs. Bei 5 Patienten wurde die Therapie mit Glycerolphenylbutyrat abgebrochen, die Gründe dafür waren unterschiedlich: bei zwei Patienten wurde die Einwilligung widerrufen, bei den übrigen Patienten wurde die Behandlung aufgrund unterschiedlicher fortbestehender UE abgebrochen.

2.6.6 Subgruppenanalysen

Post hoc festgelegte Subgruppenanalysen wurden vom pU lediglich für den primären Endpunkt präsentiert. Dieser wird nur ergänzend dargestellt und ist nicht bewertungsrelevant. Aus diesen Gründen wird auf die Darstellung der Subgruppenanalysen verzichtet.

3 Methodische Anmerkungen und Bewertung der Unterlagen

3.1 Zulassungsstatus und Zulassungspopulation von Glycerolphenylbutyrat

Gemäß Zulassung wird Glycerolphenylbutyrat (Ravicti®) angewendet als Zusatzbehandlung bei erwachsenen und pädiatrischen Patienten im Alter von ≥ 2 Monaten mit Harnstoffzyklusstörungen einschließlich Mangel an Carbamoylphosphat-Synthetase 1 (CPS), Ornithin-Transcarbamylase (OTC), Argininosuccinat-Synthetase (ASS), Argininosuccinat-Lyase (AL), Arginase 1 (ARG) und Ornithin-Translokase (Hyperammonämie-Hyperornithinämie-Homocitrullinurie-Syndrom), die durch diätetische Eiweißrestriktion und/oder Aminosäuresubstitution allein nicht behandelt werden können. Glycerolphenylbutyrat unterhält eine Zulassung als Arzneimittel zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlamentes und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. Die Zulassung durch die EMA (European Medicines Agency) besteht seit dem 27. November 2015.

Die vorliegende Nutzenbewertung zu Glycerolphenylbutyrat basiert auf den Zulassungsstudien HPN-100-006, HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-012 und HPN-100-012SE sowie den weiteren Studien HPN-100-009 und HPN-100-011. Von einer Übertragbarkeit der Zulassungspopulation auf den deutschen Versorgungskontext ist auszugehen. Bei der pivotalen Studie HPN-100-006 handelt es sich um ein Cross-over-RCT, während die anderen genannten supportiven Studien ausschließlich einarmige Beobachtungsstudien sind. Die pivotalen Studie und die supportiven Studien wurden zur Erfassung der Wirksamkeit und Sicherheit herangezogen.

In der vierwöchigen pivotalen Studie HPN-100-006 wurde die Wirksamkeit und Sicherheit von Glycerolphenylbutyrat im Vergleich zu NaPBA unter Berücksichtigung der individuellen proteinarmen Diät sowie der Aminosäuresubstitution untersucht. Dafür wurden erwachsene Patienten mit Harnstoffzyklusstörungen mit den Subtypen Mangel an Carbamoylphosphat-Synthetase 1 (CPS), Ornithin-Transcarbamylase (OTC) und Argininosuccinat-Synthetase (ASS) eingeschlossen. Patienten mit Harnstoffzyklusstörungen der übrigen drei Subtypen (Mangel an Argininosuccinat-Lyase (AL), Arginase 1 (ARG) und N-Acetyl-Glutamat-Synthetase (NAGS) bzw. Ornithin-Translokase) sowie Kinder und Jugendliche mit Harnstoffzyklusstörungen wurden in der pivotalen Studie nicht eingeschlossen, sind aber Teil der Zulassungspopulation.

In den supportiven Studien wurden Kinder (ab 2 Monaten bis < 18 Jahren) und Erwachsene hinsichtlich der Sicherheit (primärer Endpunkt in allen supportiven Studien) und Wirksamkeit von Glycerolphenylbutyrat kurzfristig (Switch-over- bzw. Transition-Phase) und langfristig (in den meisten Studien 12 Monate) unter Berücksichtigung der individuellen proteinarmen Diät sowie der Aminosäuresubstitution untersucht. Dabei wurden UCD-Patienten aus verschiedenen Altersklassen eingeschlossen (das Alter variierte zwischen den Studien). Außerdem wurden alle UCD-Subtypen eingeschlossen, mit Ausnahme des Subtyps NAGS-Mangel in den Studien HPN-100-009 und HPN-100-012 bzw. HPN-100-012SE. Die Population der supportiven Studien ist somit annähernd vergleichbar mit der Zulassungspopulation. Alle supportiven Studien sind einarmige Beobachtungsstudien, dementsprechend wurden die Verlaufsdaten für die Nutzenbewertung berücksichtigt.

3.2 Design und Methodik der berücksichtigten Studien

Pivotal Studie HPN-100-006

Bei der pivotalen Studie HPN-100-006 handelt sich um eine randomisierte, aktiv kontrollierte, doppelt-verblindete Cross-over-Studie der Phase III im zweiarmigen Parallelgruppendesign zur Ermittlung der Wirksamkeit und Sicherheit von Glycerolphenylbutyrat im Vergleich zu NaPBA zur Behandlung von Erwachsenen mit Harnstoffzyklusstörungen unter Berücksichtigung der individuellen proteinarmen Diät sowie der Aminosäuresubstitution. Die Studie wurde als Nichtunterlegenheitsstudie hinsichtlich des primären Endpunktes 24-Stunden-AUC-Ammoniakkonzentration im Blut geplant. Die Studienteilnehmer wurden zufällig und unter Einsatz eines computergestützten zentralen Schemas randomisiert in die beiden Studienarmen NaPBA + Placebo → Glycerolphenylbutyrat + Placebo oder Glycerolphenylbutyrat + Placebo → NaPBA + Placebo zugeordnet. Placebo-Präparate waren für beide Studienmedikamente je nach Behandlungsphase nötig, da sich die beiden Studienmedikamente in der Darreichungsform unterschieden (Glycerolphenylbutyrat: farblose bis gelbliche geruchs- und geschmacklose Flüssigkeit; NaPBA: Tabletten à 500 mg Wirkstoff, aversiver Geruch und Geschmack, auch in Pulverform dargereicht). Die Studie wurde an 19 Studienzentren in den USA und Kanada durchgeführt. Am 12. Oktober 2009 wurde die Studie begonnen und am 9. September 2010 abgeschlossen.

Da es sich um eine Cross-over-Studie handelt, wird überprüft, ob die Voraussetzungen für eine Cross-over-Studie gegeben sind. Dazu gehört u. a. der Ausschluss von Carry-over-Effekten von der ersten in die zweite Behandlungssequenz, d. h. zu Beginn der zweiten Behandlungssequenz sollten die Baselinewerte mit denen der ersten Behandlungssequenz in beiden Gruppen vergleichbar sein. In der Regel wird dies durch eine Wash-out-Phase zwischen den beiden Behandlungssequenzen sichergestellt. In der Studie HPN-100-006 wurde auf eine Wash-out-Phase verzichtet, da ein Aussetzen der Medikation zu lebensbedrohlichen Situationen wie z. B. hyperammonämischen Krisen führen kann. Daher muss der Einfluss der Restmedikation zu Beginn der Behandlungsphasen abgeschätzt werden. Laut Fachinformation [50] von NaPBA wird das Medikament innerhalb von 24 Stunden zu 80–100 % abgebaut. Im Gegensatz zu NaPBA wird Glycerolphenylbutyrat langsamer verstoffwechselt [25]. Dementsprechend ist auch die Halbwertszeit bei Glycerolphenylbutyrat höher als bei NaPBA. Dennoch kann hier nicht von einem bedeutsamen Carry-over-Effekt ausgegangen werden, da die Halbwertszeiten beider Medikamente insgesamt relativ kurz sind [9] und der primäre Endpunkt jeweils am Ende der beiden Behandlungsperioden (Tag 14 und 28) erhoben wurde. Für die übrigen Endpunkte (Mortalität, HAC und Sicherheit) kann nicht ausgeschlossen werden, dass bei einem Ereignis zu Beginn der Behandlungsperioden noch ein Einfluss des Medikaments aus der vorangehenden Behandlungsperiode vorlag, da es keine Wash-out-Phase gab.

Eine weitere Voraussetzung für Cross-Over-Studien ist, dass es sich um eine chronische Erkrankung handelt, die durch die Prüfintervention nicht geheilt wird, die aber auch nicht während der Behandlungsphase progredient oder regredient verläuft. Bei den Harnstoffzyklusstörungen handelt es sich um einen angeborenen Enzymdefekt, der nur durch eine Lebertransplantation geheilt werden kann. Da die Behandlung durch die Studienmedikamente Glycerolphenylbutyrat und NaPBA symptomatisch ist, ist nicht davon auszugehen, dass die Erkrankung während des kurzen Behandlungszeitraums (vier Wochen) progredient oder regredient verläuft.

Die Studie HPN-100-006 wurde doppelt verblindet durchgeführt. Hierbei ist zu beachten, dass das Medikament NaPBA einen aversiven Geruch und Geschmack hat [9,44]. Es ist unklar, inwiefern

die gustatorischen und olfaktorischen Eigenschaften der Placebo-Medikamente dem Originalpräparat (NaPBA) glichen. Sollten die Eigenschaften nicht vergleichbar gewesen sein, könnten sich die Studienteilnehmer ihrer Gruppenzugehörigkeit entsprechend klarwerden. Da die Patienten im Durchschnitt bereits mehr als 10 Jahre vor Studienbeginn mit NaPBA mediziert wurden (siehe Abschnitt 2.6.1), ist davon auszugehen, dass die Patienten den Geschmack und Geruch von NaPBA sehr genau kannten. Gleichzeitig wird den Patienten jedoch auch bewusst gewesen sein, dass eine Non-Compliance der Medikation zu lebensbedrohlichen Situationen wie z. B. hyperammonämischen Krisen führen kann. Daher ist trotz einer potentiellen Entblindung von einer hohen Motivation zur regelmäßigen Einnahme der Studienmedikation auszugehen, um mögliche hyperammonämische Krisen zu verhindern. Von einem Einfluss auf das Verzerrungspotential auf Studien- und Endpunktebene (Mortalität und Morbidität) ist jedoch nicht auszugehen. Es ist allerdings nicht auszuschließen, dass eine mögliche Entblindung einen Einfluss auf die unerwünschten Ereignisse haben könnte. Daher wird das Verzerrungspotential bei diesem Endpunkt als unklar klassifiziert.

Die Randomisierungssequenz, die Verdeckung der Gruppenzuteilung und die Verblindung der Behandler erscheint bei der Studie HPN-100-006 adäquat, es ergeben sich zudem keine weiteren Hinweise auf ein Verzerrungspotential. Insgesamt ist daher auf Studienniveau von einem niedrigen Verzerrungspotential auszugehen.

Es wurden in der Studie HPN-100-006 46 Studienteilnehmer auf die beiden Behandlungsarme randomisiert. Ein Studienteilnehmer wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation – wegen Verletzung eines Einschlusskriteriums – ausgeschlossen. Dieser Studienteilnehmer wurde bei den Analysen der Studie nicht berücksichtigt und von der ITT-Population ausgeschlossen. Dieses Vorgehen wird als kritisch erachtet, da formal die ITT-Population alle randomisierten Patienten umfasst. Ein weiterer Studienteilnehmer erlitt unerwünschte Ereignisse aufgrund von Non-Compliance der Diät und wurde nach dem ersten Studientag ausgeschlossen.

Generell ist anzumerken, dass die Patientencharakteristika nach Behandlungssequenz (NaPBA → Glycerolphenylbutyrat gegenüber NaPBA → Glycerolphenylbutyrat) dargestellt sind, jedoch alle Studienteilnehmer eine zweiwöchige Periode mit den beiden Studienmedikationen (Glycerolphenylbutyrat und NaPBA) erhalten haben. Es ist nicht davon auszugehen, dass ein Unterschied in den demographischen Eigenschaften oder den Baseline-Charakteristika zwischen den Gruppen einen Einfluss auf die Reihenfolge der Studienmedikationen hat.

Insgesamt haben 44 Studienteilnehmer die vierwöchige Studie abgeschlossen. Die Population bestand zu 68,9 % (n=31) aus weiblichen und zu 31,1 % (n=14) aus männlichen Studienteilnehmern, die im Mittel 32,7 Jahre alt und überwiegend kaukasischer Abstammung (n=31, 77,8 %) waren. Hinsichtlich der krankheitsspezifischen Merkmale wiesen die Studienteilnehmer fast ausschließlich den Subtypen OTC-Mangel auf (n=40, 88,9 %). Die Erkrankung wurde überwiegend erst nach dem 2. Lebensjahr (n=20, 44,4 %) bzw. nach dem 18. Lebensjahr (n=15, 33,3 %) diagnostiziert. Im Mittel wurden die Studienteilnehmer bereits über 10 Jahre mit NaPBA mediziert (MW (in Monaten) = 128,6).

Der Enzymdefekt OTC-Mangel ist der häufigste Mangel bei Harnstoffzyklusstörungen [24], was auch in der vorliegenden Population der Fall ist. Ungewöhnlich ist jedoch das Geschlechterverhältnis der Studienteilnehmer. OTC-Mangel ist mit dem X-Chromosom assoziiert und tritt daher häufiger bei Männern als bei Frauen auf [24]. Dies konnte auch in einer deutschen epidemiologischen Untersuchung gezeigt werden [19]. In einer Studie aus den USA [16] betrug das Männer-

Frauen-Verhältnis bei 142 Studienteilnehmern mit OTC-Mangel 86 Männer zu 76 Frauen (entspricht 1,13), während in der vorliegenden Studie HPN-100-006 das Verhältnis aller Studienteilnehmer 14 Männer zu 31 Frauen (entspricht 0,45) betrug. Allerdings unterschieden sich die beiden Studien hinsichtlich des Alters: Während in der Enns-Studie [16] die Teilnehmer ohne Altersbeschränkung rekrutiert wurden, wurden in der Studie HPN-100-006 nur erwachsene Studienteilnehmer eingeschlossen. Da die Überlebensrate bei Männern mit OTC-Mangel niedriger ist als bei Frauen mit OTC-Mangel [16,48], könnte es sein, dass im Erwachsenenalter weniger Männer für Studien rekrutiert werden können. Dies sind mögliche Erklärungsansätze für die ungleiche Geschlechterverteilung. Es ist jedoch nicht auszuschließen, dass es noch andere (studien-spezifische) Gründe für dieses unausgeglichene Geschlechterverhältnis gibt.

In der Studie HPN-100-006 betrug der Anteil von Phenylbutyrat bei den beiden Studienmedikationen im Mittel bei Glycerolphenylbutyrat 12,5 g pro Tag und bei NaPBA 12,3 g pro Tag. Die Therapietreue war insgesamt hoch. An den beiden Studientagen 14 und 28 nahmen alle Studienteilnehmer mehr als 80 % der geplanten Medikation ein und alle bis auf einen Teilnehmer (während NaPBA-Medikation) nahmen über 90 % der geplanten Medikation ein. Neben der Studienmedikation wurden die Studienteilnehmer angehalten eine individuelle proteinarme Diät sowie ggf. eine Aminosäuresubstitution einzuhalten. Aus diesem Grund war die Begleitmedikation während der Studienphase bei nahezu allen Studienteilnehmern gegeben (während NaPBA-Medikation: n=42 (93,3 %), während Glycerolphenylbutyrat-Medikation: n=43 (97,7 %)).

Die Dauer der Studie HPN-100-006 betrug vier Wochen. Für eine valide Einschätzung der Wirksamkeit und der Sicherheit wäre eine längere Studiendauer erforderlich. Vor allem hinsichtlich des Morbiditätsendpunktes HAC wäre dieser Aspekt relevant, da hyperammonämische Krisen bei stabiler Medikation und Einhaltung der Diät relativ selten auftreten.

Prinzipiell ist es bei der vorliegenden Fragestellung möglich, RCT mit einer längeren Studiendauer sowie bei Patienten mit Harnstoffzyklusstörungen im Alter zwischen 2 Monaten und 18 Jahren durchzuführen. Es wird nicht nachvollziehbar dargelegt, warum sich der pU bei der Durchführung des RCT auf eine sehr kurze Beobachtungszeit beschränkte und nur im begrenzten Umfang Morbiditäts- und Lebensqualitätsdaten erhoben hat. Insbesondere die Erhebung von Morbiditäts- und Lebensqualitätsendpunkten ist über einen längeren Zeitraum möglich, wie in den einarmigen supportiven Beobachtungsstudien gezeigt wurde.

Da der pU als einzige verfügbare Evidenz für Langzeiteffekte sowie für Patienten mit Harnstoffzyklusstörungen unter 18 Jahren einarmige Beobachtungsstudien vorgelegt hat, werden diese in der Nutzenbewertung als supportive Studien hinsichtlich des Verlaufs berücksichtigt.

Supportive Studien

Mit dem Nutzendossier legte der pU mehrere einarmigen Beobachtungsstudien mit der Medikation Glycerolphenylbutyrat mit dem primären Ziel der Erfassung der Sicherheit vor. Einarmige Beobachtungsstudien bilden in diesem Zusammenhang den Krankheitsverlauf während der Glycerolphenylbutyrat-Intervention ab und werden als supportive Studien berücksichtigt, um ergänzend Sicherheitsaspekte sowie die Endpunkte neuropsychologische Tests und Lebensqualität im Langzeitverlauf darzustellen. Da es sich um Studien ohne Kontrollgruppe im Parallelgruppendesign handelt, wird bei diesen Studien von einem hohen Verzerrungspotential auf Studien- und Endpunktebene ausgegangen.

Einschlusskriterium bei den supportiven Studien war mit einer Ausnahme (HPN-100-009) eine stabile Medikation mit NaPBA. Die Studien unterschieden sich vor allem hinsichtlich der Altersklassen. Kinder unter 6 Jahren wurden in 4 Studien untersucht (HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE), Kinder und Jugendliche zwischen 6 und 17 Jahren wurden in 4 Studien untersucht (HPN-100-005, HPN-100-005SE, HPN-100-007 und HPN-100-011) und Erwachsene wurden in 2 Studien untersucht (HPN-100-007 und HPN-100-011). Dabei ist anzumerken, dass teils dieselben Studienteilnehmer an mehreren Studien teilnahmen, denn die Studien HPN-100-005SE, HPN-100-007, HPN-100-012SE waren jeweils Sicherheitserweiterungen der Studien HPN-100-005, HPN-100-006, HPN-100-012, wobei bei den Sicherheitserweiterungen zusätzlich noch neue Studienteilnehmer rekrutiert wurden. Die Studie HPN-100-011 wiederum stellt eine gemeinsame weitere Sicherheitserweiterung der Studien HPN-100-005SE, HPN-100-007, HPN-100-012SE dar.

Die Populationen variierten hinsichtlich der Stichprobengrößen (10 bis 88 Studienteilnehmer), wobei die Populationen überwiegend weiblich waren und eine kaukasische Abstammung aufwiesen. Die Art des Enzymdefekts war überwiegend OTC-Mangel. Aufgrund des unterschiedlichen Alters der Studienteilnehmer in den Studien variierten der diagnostizierte Beginn der Harnstoffzyklusstörung (z. B. kann bei Studien mit Kindern und Jugendlichen der Beginn der Harnstoffzyklusstörung nicht erst im Erwachsenenalter diagnostiziert werden), die Dauer der NaPBA-Vorbehandlung und die tägliche Studienmedikation (aufgrund von Unterschieden in Körpergewicht und -größe).

Die supportiven Studien hatten alle als primären Endpunkt das Auftreten unerwünschter Ereignisse festgelegt und waren somit primär zur Untersuchung der Sicherheit ausgelegt. Dementsprechend war im Gegensatz zur pivotalen Studie (vier Wochen) die Studiendauer bei den supportiven Studien länger. Abgesehen von den Kurzzeitstudien (HPN-100-005 und HPN-100-012) betrug der Beobachtungszeitraum in den anderen Studien überwiegend 12 Monate (HPN-100-005: 12 Monate, HPN-100-007: 12 Monate, HPN-100-009: 6 bis 24 Monate, HPN-100-011: unbegrenzt, HPN-100-012: 12 Monate). Die Aussagekraft dieser Daten ist trotz der längeren Laufzeit eingeschränkt, da es keine vergleichende Kontrollgruppe im Parallelgruppendesign gab.

3.3 Wirksamkeit

Für den Nachweis der Wirksamkeit von Glycerolphenylbutyrat bei Patienten mit Harnstoffzyklusstörungen reichte der pU Daten zu Mortalität und Morbidität (24-Stunden-AUC-Ammoniakkonzentration im Blut, tägliche durchschnittliche Ammoniakkonzentration im Blut, neuropsychologische Tests, Anzahl und Schweregrad hyperammonämischer Krisen) ein. Für die Nutzenbewertung werden Mortalität sowie Anzahl und Schweregrad hyperammonämischer Krisen berücksichtigt, da diese als patientenrelevant eingestuft werden. Der Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut wird nicht als bewertungsrelevant eingestuft. Da es sich bei diesem Endpunkt jedoch um den primären Endpunkt der pivotalen Studie HPN-100-006 handelt, wird der Endpunkt ergänzend dargestellt. Die neuropsychologischen Tests wurden ausschließlich in den supportiven einarmigen Beobachtungsstudien erfasst. Da diese Studien keine valide Kontrollgruppe im Parallelgruppendesign enthalten, werden die Endpunkte zu den neuropsychologischen Tests im Anhang ergänzend dargestellt. Aufgrund fehlender Informationen zu verschiedenen Testgütekriterien, wie z. B. der Änderungssensitivität, bleibt die Aussagekraft der Ergebnisse unklar und sie werden daher nicht dargestellt.

Mortalität

In der vierwöchigen Cross-over-Studie HPN-100-006 traten während der beiden Behandlungsperioden (NaPBA und Glycerolphenylbutyrat) keine Todesfälle auf. Todesfälle sind außerdem im Rahmen der Sicherheitsbewertung erfasst.

24-Stunden-AUC-Ammoniakkonzentration im Blut

Der primäre Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut der Studie HPN-100-006 ist ein Laborparameter und nicht bewertungsrelevant. Wie in Abschnitt 2.5.3 dargelegt, liegen keine Daten vor, die die Validität des Endpunktes 24-Stunden-AUC-Ammoniakkonzentration im Blut als Surrogatparameter für patientenrelevante Endpunkte belegen.

Die Studie war als Nichtunterlegenheitsstudie ausgelegt, dabei wurde die Nichtunterlegenheitsgrenze auf den Wert 1,25 festgelegt. Eine Begründung der Schwelle auf einer Evidenzbasis war den Unterlagen des pU nicht zu entnehmen. Die EMA betrachtet diese festgelegte Grenze als zu liberal, da davon auszugehen ist, dass die erwartete obere Grenze des 95%-KI deutlich unterhalb der festgelegten Schwelle liegt [17].

Im Durchschnitt wiesen die Studienteilnehmer unter NaPBA-Medikation im Blut eine 24-Stunden-AUC-Ammoniakkonzentration von 976,6 $\mu\text{mol h/l}$ auf gegenüber 865,9 $\mu\text{mol h/l}$ unter Glycerolphenylbutyrat-Medikation. Das Verhältnis der geometrischen Mittel betrug 0,91, wobei die obere Grenze des 95%-KI des Verhältnisses der geometrischen Mittel 1,034 betrug und somit unterhalb der Nichtunterlegenheitsschwelle lag. Das Verzerrungspotential für den Endpunkt wird als gering bewertet.

Anzahl und Schweregrad von HAC

In der vierwöchigen Cross-over-Studie HPN-100-006 trat insgesamt eine HAC während der Behandlungsperiode mit NaPBA auf. Todesfälle sind außerdem im Rahmen der Sicherheitsbewertung erfasst. Der Prüfarzt bewertete die HAC mit Schweregrad 3 und als im Zusammenhang mit einer Non-Compliance mit der Studienmedikation stehend. Aufgrund der geringen Anzahl der Ereignisse konnte kein Relatives Risiko berechnet werden. Das Verzerrungspotential für den Endpunkt wird als gering bewertet.

Zusammenfassende Einschätzung zur Wirksamkeit

Die vorliegenden Ergebnisse zu Todesfällen sowie Anzahl und Schweregrad von HAC sind aufgrund der kurzen Laufzeit der Studie und der geringen Inzidenz von Ereignissen nicht belastbar. Aufgrund der fehlenden validen Kontrollgruppe im Parallelgruppendesign sowie der fehlenden Informationen zu verschiedenen Testgütekriterien wurden die Ergebnisse zum Endpunkt neuropsychologische Tests nicht dargestellt. Es lassen sich daher keine belastbaren Aussagen zu den o. g. patientenrelevanten Endpunkten von Glycerolphenylbutyrat oder NaPBA ableiten.

3.4 Lebensqualität

In der pivotalen Studie HPN-100-006 wurde die Lebensqualität mittels des Fragebogens SF-36 erfasst. Da der Fragebogen in der Cross-over-Studie ausschließlich an Studientag 1 (Baseline-Erhebung) erhoben wurde, lassen sich keine validen Vergleiche zwischen den beiden Behandlungsperioden (NaPBA und Glycerolphenylbutyrat) ableiten. Aussagen zur Lebensqualität lassen sich daher aus der Studie HPN-100-006 nicht ableiten.

In den supportiven einarmigen Beobachtungsstudien wurden der SF-15 und der SF-36 verwendet. Die Ergebnisse zum Endpunkt Lebensqualität aus den einarmigen Beobachtungsstudien sind ohne eine prospektive parallele Vergleichsgruppe nicht aussagekräftig. Die Verlaufsdaten aus dem SF-36 werden im Anhang ergänzend dargestellt. Die Ergebnisse des SF-15 wurden nicht dargestellt, da die Aussagekraft der Ergebnisse aufgrund fehlender Informationen zu Testgütekriterien eingeschränkt ist.

3.5 Sicherheit

Zur Bewertung der Sicherheit wurden neben der vierwöchigen pivotalen Studie HPN-100-006 auch die supportiven einarmigen Beobachtungsstudien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE mit teils deutlich längerer Beobachtungsdauer berücksichtigt. Die Sicherheitsaspekte werden dementsprechend getrennt dargestellt.

Pivotaler Studie HPN-100-006

UE traten während der beiden Behandlungsperioden etwa gleich häufig auf. Jeweils ein SUE trat in den beiden Behandlungsperioden auf, Todesfälle traten nicht auf. Im NaPBA-Arm brach ein Patient die Studie an Tag 1 aufgrund erhöhter Ammoniakwerte ab.

Unter Glycerolphenylbutyrat traten insgesamt mehr Erkrankungen des Gastrointestinaltrakts auf (v. a. Diarrhö und Flatulenz). Hinsichtlich der Erkrankungen des Nervensystems klagten mehr Patienten während der Glycerolphenylbutyrat-Medikation unter Kopfschmerzen, Patienten unter NaPBA hatten häufiger Schwindel.

Insgesamt wurde Glycerolphenylbutyrat gut toleriert. Es trat ein UE mit CTCAE-Grad 3 auf (eine HAC während NaPBA-Medikation), ansonsten hatten alle anderen UE den CTCAE-Grad 1 oder 2. Allerdings empfiehlt die EMA, bei Patienten mit eingeschränkter Leberfunktion die Therapie mit Glycerolphenylbutyrat mit der niedrigsten empfohlenen Dosis zu beginnen [17].

Supportive Studien

Erwartungsgemäß lag der Anteil der Patienten mit UE in den über einen längeren Zeitraum durchgeführten supportiven Studien höher als in der pivotalen Studie. Der Anteil der Patienten mit jeglichem UE lag in den langfristigen Studien zwischen 84 und 100 %, der Anteil mit SUE zwischen 18 und 60 % (in den reinen Switch-over-Studien HPN-100-005 und HPN-100-012 waren die Anteile wegen der kurzen Laufzeit entsprechend niedriger). Das Spektrum der UE in den supportiven Studien variierte in den Altersgruppen; bspw. traten bei Kindern Infektionserkrankungen (u. a. Erbrechen, Diarrhö, Fieber) häufiger auf.

Das am häufigsten registrierte UE mit CTCAE-Grad ≥ 3 war Hyperammonämie (zwischen 12 und 30 % der Studienteilnehmer), was den Krankheitsverlauf in dieser Indikation reflektiert. Hyperammonämien waren am häufigsten in den beiden Studien HPN-100-009 (Kinder zwischen 2 Monaten und 2 Jahren) und HPN-100-012SE (Kinder bis 6 Jahre) zu beobachten. Gerade in diesem jungen Alter, müssen Ernährung und Medikation engmaschig kontrolliert werden, da durch rasches Wachstum oft Anpassungen nötig sind. Außerdem könnte in dieser Altersgruppe auch der Anteil von Studienteilnehmern mit schwerwiegenderem Krankheitsbild höher liegen. Bei Studien mit Erwachsenen ist davon auszugehen, dass UCD-Patienten mit einem sehr hohen Schweregrad nicht eingeschlossen werden, da diese bereits eine Lebertransplantation erhalten haben oder aufgrund der Folgen der Erkrankung nicht an einer Studie teilnehmen können (z. B. ausgeprägte Gedeihstörungen oder frühzeitiges Versterben).

Ein Patient war im Studienzeitraum an den Folgen eines Eingriffs im Alter von 4 Monaten verstorben. Bei fünf Patienten wurde die Therapie mit Glycerolphenylbutyrat aus unterschiedlichen Gründen abgebrochen (Widerruf der Einwilligung sowie unterschiedliche fortbestehende UE), es gab aber kein dominierendes UE, das zum Abbruch der Therapie geführt hat.

Zusammenfassende Einschätzung zur Sicherheit

Die pivotale Studie hat gezeigt, dass Glycerolphenylbutyrat und NaPBA vergleichbar gut toleriert wurde. Einige gastrointestinale Symptome traten in der Interventionsgruppe numerisch häufiger auf, diese führten aber nicht zum Therapieabbruch. Die Aussagekraft der pivotalen Studie ist aufgrund der kurzen Dauer (vier Wochen) und des unklaren Verzerrungspotentials eingeschränkt.

Langzeitdaten wurden vom pU durch die supportiven einarmigen Beobachtungsstudien vorgelegt. Diese Beobachtungsstudien mit hohem Verzerrungspotential wiesen keine Kontrollgruppe im Parallelgruppendesign auf, daher ist die Aussagekraft der Studien deutlich eingeschränkt. SUE und UE mit CTCAE-Grad ≥ 3 traten häufiger bei Kindern unter 6 Jahren auf im Vergleich zu älteren Kindern oder Erwachsenen.

Vor dem Hintergrund eines größeren Anteils von Patienten mit SUE in den supportiven Studien im Vergleich zur pivotalen Studie weist die EMA darauf hin, dass keine definitiven Schlussfolgerungen zur Inzidenz von SUE bzw. zum Therapieabbruch aufgrund SUE gezogen werden können, weil die Anzahl der Patienten relativ klein sei und die Langzeitstudien keine Kontrollgruppe aufwiesen. Daher wurde die Einrichtung eines Registers (Laufzeit 10 Jahre) als Voraussetzung für die Marktzulassung vereinbart, insbesondere um pädiatrische Patienten über einen langen Zeitraum beobachten zu können [17].

Im Risk Management Plan der EMA werden die gastrointestinalen Ereignisse als „important identified risks“ eingestuft [17].

4 Anforderung an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen. Die Behandlung mit Glycerolphenylbutyrat sollte von einem Arzt verschrieben werden, der über Erfahrung mit der Behandlung von Harnstoffzyklusstörungen verfügt.

5 Zusammenfassung der Bewertung

Glycerolphenylbutyrat (Ravicti®) ist zugelassen als Zusatztherapie bei erwachsenen und pädiatrischen Patienten im Alter von ≥ 2 Monaten mit Harnstoffzyklusstörungen einschließlich Mangel an Carbamoylphosphat-Synthetase 1 (CPS), Ornithin-Transcarbamylase (OTC), Argininosuccinat-Synthetase (ASS), Argininosuccinat-Lyase (AL), Arginase 1 (ARG) und Ornithin-Translokase (Hyperammonämie-Hyperornithinämie-Homocitrullinurie-Syndrom), die durch diätetische Eiweißrestriktion und/oder Aminosäuresubstitution allein nicht behandelt werden können [25]. Die Nutzenbewertung von Glycerolphenylbutyrat basiert auf der pivotalen Studie HPN-100-006 (randomisierte, aktiv kontrollierte, doppelt-verblindete Cross-over-Studie der Phase III im zweiarmigen Parallelgruppendesign) sowie den supportiven einarmigen Beobachtungsstudien HPN-100-005, HPN-100-005SE, HPN-100-007, HPN-100-009, HPN-100-011, HPN-100-012 und HPN-100-012SE.

In die pivotalen Studie HPN-100-006 wurden erwachsene Patienten mit drei UCD-Subtypen (CPS-, OTC- und ASS-Mangel) eingeschlossen. In den einarmigen Beobachtungsstudien wurden darüber hinaus auch Kinder (ab 2 Monaten bis zu 18 Jahren) und erwachsene UCD-Patienten mit den übrigen Subtypen berücksichtigt, wobei bei zwei supportiven Studien ein Subtyp nicht berücksichtigt wurde. Insgesamt decken damit die Populationen über alle Studien hinweg annähernd das Spektrum der Zulassungspopulation ab.

Die Ergebnisse der Studien werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Glycerolphenylbutyrat 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.

Tabelle 22: Zusammenfassende Darstellung der Studienergebnisse.

Endpunkt	Glycerolphenylbutyrat vs. NaPBA ¹⁾ Effektschätzer [95%-KI]; p-Wert	Effekt
Mortalität		
Todesfälle	HPN-100-006: keine Ereignisse aufgetreten	-
Morbidität		
24-Stunden-AUC-Ammoniakkonzentration im Blut ²⁾	HPN-100-006: k.A. ²⁾	-
Anzahl und Schweregrad von HAC	HPN-100-006: RR: n.b. ³⁾	-
Lebensqualität		
SF-36	Verlaufsdaten: siehe Anhang	-
Sicherheit: pivotalen Studie HPN-100-006		
UE, n (%)	Glycerolphenylbutyrat: 27 (61,4) NaPBA: 23 (51,1)	-
UE CTCAE-Grad ≥ 3	RR ⁴⁾ : n.b.	-
SUE	RR: 1,023 [0,066;15,847]; 0,9872	\leftrightarrow
Studienabbruch aufgrund UE	RR ⁵⁾ : n.b.	-

Endpunkt	Glycerolphenylbutyrat vs. NaPBA ¹⁾ Effektschätzer [95%-KI]; p-Wert	Effekt
Sicherheit: supportive Studien		
UE, n (%)	Studie HPN-100-005: 4 (36,4) Studie HPN-100-005SE: 16 (94,0) Studie HPN-100-007: 59 (98,0) Studie HPN-100-009: 10 (100) Studie HPN-100-011: 74 (84,1) Studie HPN-100-012: 6 (40,0) Studie HPN-100-012SE: 23 (100)	.. ⁶⁾
UE CTCAE-Grad \geq 3, n (%)	Studie HPN-100-005: 0 Studie HPN-100-005SE: 1 (6,0) Studie HPN-100-007: 11 (18,0) Studie HPN-100-009: 4 (40,0) Studie HPN-100-011: 24 (27,2) Studie HPN-100-012: 0 Studie HPN-100-012SE: 8 (35,0)	.. ⁶⁾
SUE, n (%)	Studie HPN-100-005: 0 Studie HPN-100-005SE: 3 (18,0) Studie HPN-100-007: 12 (20,0) Studie HPN-100-009: 6 (60,0) Studie HPN-100-011: 30 (34,1) Studie HPN-100-012: 0 Studie HPN-100-012SE: 11 (48,0)	.. ⁶⁾
Studienabbruch aufgrund UE, n (%)	Studie HPN-100-005: 0 Studie HPN-100-005SE: 0 Studie HPN-100-007: 2 (3,5) Studie HPN-100-009: 1 (10,0) Studie HPN-100-011: 2 (2,3) Studie HPN-100-012: 0 Studie HPN-100-012SE: 1 (4,0)	.. ⁶⁾

¹⁾ Die dargestellten Ergebnisse der Kategorien Mortalität und Morbidität beziehen sich auf die Studie HPN-100-006. Für die Endpunkte Mortalität sowie Anzahl und Schweregrad von HAC wurden die Daten der Safety-Population (NaPBA: N=45, Glycerolphenylbutyrat: N=44) herangezogen und für die Daten des Endpunktes 24-Stunden-AUC-Ammoniakkonzentration im Blut die ITT-Population (NaPBA: N=44, Glycerolphenylbutyrat: N=44). Ein Studienteilnehmer wurde nach der Randomisierung, aber vor der Einnahme von Studienmedikation – wegen Verletzung eines Einschlusskriteriums – ausgeschlossen. Ein weiterer Studienteilnehmer erlitt UE aufgrund von Non-Compliance der Diät und wurde nach dem ersten Studientag ausgeschlossen.

²⁾ Der Endpunkt 24-Stunden-AUC-Ammoniakkonzentration im Blut ist der primäre Endpunkt der Studie HPN-100-006. Dieser Endpunkt wird als nicht bewertungsrelevant eingestuft und daher für die Nutzenbewertung nicht herangezogen. Der Endpunkt wurde im Abschnitt 2.6.3 ergänzend dargestellt.

³⁾ In der Studie HPN-100-006 trat eine HAC auf. Ein Patient erlitt während NaPBA-Medikation an Studientag 28 eine HAC mit Schweregrad 3, die laut Prüfarzt im Zusammenhang mit fehlender Compliance zur Studienmedikation aufgetreten ist. Ein Relatives Risiko war aufgrund der niedrigen Inzidenz nicht berechenbar.

⁴⁾ Bei einem Studienteilnehmer kam es zu einer HAC aufgrund von Non-Compliance mit der Medikamenteneinnahme (während NaPBA-Medikation). Ein Relatives Risiko war aufgrund der niedrigen Inzidenz nicht berechenbar.

⁵⁾ Ein Studienteilnehmer aus Behandlungsarm A (NaPBA → Glycerolphenylbutyrat) erlitt aufgrund von Non-Compliance der Diät UE, die zum Therapieabbruch führten, und wurde nach dem ersten Studientag ausgeschlossen. Ein Relatives Risiko war aufgrund der niedrigen Inzidenz nicht berechenbar.

⁶⁾ Die Ergebnisse zu UE für die supportiven Studien werden nur deskriptiv für die Einzelstudien während der Medikation mit Glycerolphenylbutyrat dargestellt, da eine Kontrollgruppe im Parallelgruppendesign nicht vorliegt. Ein Effekt kann daher nicht abgeleitet werden.

Abkürzungen: AUC: area under the curve; CTCAE: Common Terminology Criteria for Adverse Events; HAC: hyperammonämische Krise/n; ITT: Intention-to-Treat; k.A.: keine Angabe/n; KI: Konfidenzintervall; NaPBA: Natriumphenylbutyrat n.b.: nicht berechnet; RR: Relatives Risiko; SF-36: 36-Item Short Form Health Survey; (S)UE: (schwerwiegende/s) unerwünschte/s Ereignis/se.

Referenzen

1. **Achenbach T, Rescorla L.** The Child Behavior Checklist (CBCL) [online]. In: ProQolid. 1997. [Zugriff: 28.05.2018]. URL: <https://eprovide.mapi-trust.org/instruments/child-behavior-checklist>.
2. **Achenbach T, Rescorla L.** Manual for the ASEBA Preschool Forms and Profiles. Burlington; University of Vermont Department of Psychiatry; 2000.
3. **Achenbach T, Rescorla L.** Manual for the ASEBA School-Age Forms and Profiles. Burlington; University of Vermont, Research Center for Children, Youth, and Families; 2001.
4. **Anderson VA, Anderson P, Northam E, Jacobs R, Mikiewicz O.** Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychol* 2002;8(4):231-240.
5. **Batshaw ML, Brusilow S, Waber L, Blom W, Brubakk AM, Burton BK, et al.** Treatment of inborn errors of urea synthesis: activation of alternative pathways of waste nitrogen synthesis and excretion. *N Engl J Med* 1982;306(23):1387-1392.
6. **Batshaw ML, Tuchman M, Summar M, Seminara J.** A longitudinal study of urea cycle disorders. *Mol Genet Metab* 2014;113(1-2):127-130.
7. **Bayley N.** Bayley Scales of Infant and Toddler Development, Third Edition (Bayley III) [online]. In: Proqolid. 2018. [Zugriff: 24.05.2018]. URL: https://eprovide.mapi-trust.org/instruments/bayley-scales-of-infant-and-toddler-development-third-edition#content_validity_documentation.
8. **Bayley N.** Technical Report 1: Using the Bayley Scales of Infant and Toddler Development, Third Edition, to Assess Individuals with Severe Delays. San Antonio; Harcourt Assessment 2006.
9. **Canadian Agency for Drugs and Technologies in Health (CADTH).** Glycerol Phenylbutyrate (Ravicti) [online]. Ottawa (CAN): CADTH; 2017. [Zugriff: 23.03.2018]. (Common Drug Reviews). URL: <https://www.cadth.ca/glycerol-phenylbutyrate>.
10. **Canivez GL, Konold TR, Collins JM, Wilson G.** Construct Validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and Structural Validity. *School Psychology Quarterly* 2009;24(4):252-265.
11. **Chen X, Origasa H, Ichida F, Kamibeppu K, Varni JW.** Reliability and validity of the Pediatric Quality of Life Inventory (PedsQL) Short Form 15 Generic Core Scales in Japan. *Qual Life Res* 2007;16(7):1239-1249.
12. **Chlebowski C.** Wechsler Memory Scale All Versions In: *Encyclopedia of Clinical Neuropsychology*. New York; Springer; 2011. 2688-2690.
13. **Daseking M, Petermann F.** Verhaltensinventar zur Beurteilung exekutiver Funktionen für das Kindergartenalter. Deutschsprachige Adaptation des Behavior Rating Inventory of Executive Function® - Preschool Version (BRIEF®-P) von Gerard A. Gioia, Kimberly Andrews Espy und Peter K. Isquith. Göttingen; Huber, Hogrefe; 2013. URL: <https://www.testzentrale.de/shop/testkit-inkl-50-nutzungen-und-manual-76031.html>.

14. **Delis C, Kramer J, E K.** California Verbal Learning Test - Second Edition, Adult Version. San Antonio; Pearson; 2000.
15. **Drechsler R, Steinhausen H-C.** Verhaltensinventar zur Beurteilung exekutiver Funktionen. Deutschsprachige Adaptation des Behavior Rating Inventory of Executive Function (BRIEF) von G. A. Gioia, P. K. Isquith, S. C. Guy und L. Kenworthy und der Self-Report Version (BRIEF-SR) von S. C. Guy, P. K. Isquith und G. A. Gioia. Göttingen; Huber, Hogrefe; 2013. URL: <https://www.testzentrale.de/shop/verhaltensinventar-zur-beurteilung-exekutiver-funktionen.html>.
16. **Enns GM, Berry SA, Berry GT, Rhead WJ, Brusilow SW, Hamosh A.** Survival after treatment with phenylacetate and benzoate for urea-cycle disorders. N Engl J Med 2007;356(22):2282-2292.
17. **European Medicines Agency (EMA).** Ravicti (International non-proprietary name: glycerol phenylbutyrate): European public assessment report EMEA/H/C/003822/0000 [online]. London (GBR): EMA; 24.09.2015. [Zugriff: 23.03.2018]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/003822/WC500199159.pdf.
18. **European Medicines Agency (EMA).** Refusal of the marketing authorisation for Heparesc (human heterologous liver cells). Outcome of re-examination; EMA/684158/2015 [online]. 30.10.2015. London (GBR): EMA; 2015. [Zugriff: 23.03.2018]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/003750/WC500188784.pdf.
19. **Forschungsstelle für pädiatrische Epidemiologie bei der Deutschen Gesellschaft für Kinder- und Jugendmedizin.** Erhebungseinheit für seltene pädiatrische Erkrankungen in Deutschland: ESPED - Jahresbericht 2015 [online]. Düsseldorf (GER): Heinrich-Heine-Universität Düsseldorf; 2015. [Zugriff: 23.03.2018]. URL: <http://www.esped.uni-duesseldorf.de/esped/resources/files/Jahresbericht%202015-2.pdf>.
20. **Gioia GA, Isquith PK, Guy SC, Kenworthy L.** BRIEF: Behavior Rating Inventory of Executive Function [Presentation] [online]. 2009. [Zugriff: 28.05.2018]. URL: <https://www.parinc.com/WebUploads/samplerpts/BRIEF.ppt>.
21. **Gioia GA, Isquith PK, Retzlaff PD, Espy KA.** Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. Child Neuropsychol 2002;8(4):249-257.
22. **Gordon B.** Test Review: Wechsler, D. (2002). The Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III). San Antonio, TX: The Psychological Corporation. Canadian Journal of School Psychology 2017;19(1-2):205-220.
23. **Häberle J, Boddaert N, Burlina A, Chakrapani A, Dixon M, Huemer M, et al.** Diagnostik und Therapie von Harnstoffzyklusstörungen; S3-Leitlinie [online]. AWMF-Register Nr. 027-006. 04.01.2016. Berlin (GER): Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF); 2012. [Zugriff: 23.03.2018]. URL: <http://www.awmf.org/leitlinien/detail/II/027-006.html>.

24. **Häberle J, Boddaert N, Burlina A, Chakrapani A, Dixon M, Huemer M, et al.** Suggested guidelines for the diagnosis and management of urea cycle disorders. *Orphanet J. Rare Dis.* 2012;7:32.
25. **Horizon Pharma Ireland.** Ravicti 1,1g/ml [online]. 15.06.2017. [Zugriff: 23.03.2018]. URL: <http://www.fachinfo.de>.
26. **Horizon Therapeutics.** An open label study of the safety, efficacy and pharmacokinetics of glycerol phenylbutyrate (GPB; RAVICTI®) in pediatric subjects under two years of age with urea cycle disorders (UCDs): HPN-100-009; clinical study report [unveröffentlicht]. 2017.
27. **Hyperion Therapeutics.** Long term use of HPN-100 in urea cycle disorders: HPN-100-011; clinical study report. Interims analysis 07 February 2014 [unveröffentlicht]. 2014.
28. **Hyperion Therapeutics.** A Phase 2, fixed-sequence, open-label, switch-over study of the safety and tolerability of HPN-100 compared to sodium phenylbutyrate in children 6-17 years of age with urea cycle disorders, with a long-term safety extension HPN-100-005; clinical study report [unveröffentlicht]. 2011.
29. **Hyperion Therapeutics.** A phase 2, fixed-sequence, open-label, switch-over study of the safety and tolerability of HPN-100 compared to sodium phenylbutyrate in children 6-17 years of age with urea cycle disorders, with a long-term safety extension: HPN-100-005 (safety extension); clinical study report [unveröffentlicht]. 2011.
30. **Hyperion Therapeutics.** A phase 2, open-label, switch-over, dose-escalation study of the safety and tolerability of Hpn-100 (Glyceryl Tri [4 Phenylbutyrate]) compared to buphenyl® (Sodium Phenylbutyrate) in patients with urea cycle disorders: UP-1204-003; Clinical Study Report, Final version 2.0 (Amendment 1) [unveröffentlicht]. 2011.
31. **Hyperion Therapeutics.** A phase 3, open-label study of the safety of HPN-100 for the long-term treatment of urea cycle disorders: HPN-100-007; clinical study report [unveröffentlicht]. 2011.
32. **Hyperion Therapeutics.** A phase 3, randomized, double-blind, cross-over, active-controlled study of the efficacy and safety of HPN-100, Glyceryl tri-(4-phenylbutyrate), for the treatment of adults with urea cycle disorders: HPN-100-006; clinical study report [unveröffentlicht]. 2011.
33. **Hyperion Therapeutics.** A switch-over, open-label study of the safety, pharmacokinetics, and efficacy of HPN-100, in pediatric subjects under 6 years of age with urea cycle disorders (UCDs): HPN-100-012 (Safety Extension); clinical study report [unveröffentlicht]. 2013.
34. **Hyperion Therapeutics.** A switch-over, open-label study of the safety, pharmacokinetics, and efficacy of HPN-100, in pediatric subjects under 6 years of age with urea cycle disorders (UCDs): HPN-100-012; clinical study report [unveröffentlicht]. 2012.
35. **Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).** Allgemeine Methoden; Vers. 5.0 [online]. 10.07.2017. Köln (GER): IQWiG; 2017. [Zugriff: 23.03.2018]. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.
36. **Isquith PK, Crawford JS, Espy KA, Gioia GA.** Assessment of executive function in preschool-aged children. *Ment Retard Dev Disabil Res Rev* 2005;11(3):209-215.

37. **Jacobs ML, Donders J.** Criterion validity of the California Verbal Learning Test-(CVLT-II) after traumatic brain injury. *Archives of clinical neuropsychology* 2007;22(2):143-149.
38. **Lee B, Diaz GA, Rhead W, Lichter-Konecki U, Feigenbaum A, Berry SA, et al.** Blood ammonia and glutamine as predictors of hyperammonemic crises in patients with urea cycle disorder. *Genetics in Medicine* 2015;17(7):561-568.
39. **Macha T, Petermann F.** Bayley Scales of Infant and Toddler Development, Third Edition – Deutsche Fassung. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie* 2015;63(2):139-143.
40. **Mathias CW, Greve KW, Bianchini KJ, Houston RJ, Crouch JA.** Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment* 2002;9(3):301-308.
41. **Merker B, Podell K.** Grooved pegboard test. In *Encyclopedia of clinical neuropsychology*. New York; Springer; 2011. 1176-1178.
42. **Msall M, Batshaw ML, Suss R, Brusilow SW, Mellits ED.** Neurologic outcome in children with inborn errors of urea synthesis. Outcome of urea-cycle enzymopathies. *N Engl J Med* 1984;310(23):1500-1505.
43. **Msall M, Monahan PS, Chapanis N, Batshaw ML.** Cognitive development in children with inborn errors of urea synthesis. *Acta Paediatr Jpn* 1988;30(4):435-441.
44. **Pena-Quintana L, Llarena M, Reyes-Suarez D, Aldamiz-Echevarria L.** Profile of sodium phenylbutyrate granules for the treatment of urea-cycle disorders: patient perspectives. *Patient Preference Adherence* 2017;11:1489-1496.
45. **Petermann F, Ricken G, Fritz A, Schuck K, Preuß U, Lipsius M (Eds.)** Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III); Deutsche Version. 3., überarbeitete und erweiterte Auflage. Frankfurt/M.: Pearson; 2014.
46. **Ryan JJ, Carruthers CA, Miller LJ, Souheaver GT, Gontkovsky ST, Zehr MD.** Exploratory factor analysis of the Wechsler Abbreviated Scale of Intelligence (WASI) in adult standardization and clinical samples. *Applied Neuropsychology* 2003;10(4):252-256.
47. **Stano JF.** Wechsler abbreviated scale of intelligence. *Rehabilitation Counseling Bulletin* 2004;48(1): 56-57.
48. **Summar ML, Dobbelaere D, Brusilow S, Lee B.** Diagnosis, symptoms, frequency and mortality of 260 patients with urea cycle disorders from a 21-year, multicentre study of acute hyperammonaemic episodes. *Acta Paediatr* 2008;97(10):1420-1425.
49. **Swedish Orphan Biovitrum.** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Glycerolphénylbutyrat (RAVICTI) [unveröffentlicht]. 17.07.2017.
50. **Swedish Orphan Biovitrum International.** Ammonaps 500 mg Tabletten [online]. 11.2016. [Zugriff: 23.03.2018]. URL: <http://www.fachinfo.de>.
51. **Trites R.** Grooved pegboard test. User Instruction; Modell 32025. Lafayette; Lafayette Instruments; 2011.
52. **Trites R.** Neuropsychological Test Manual. Ottawa; Royal Ottawa Hospital; 1977.

53. **U.S. Food and Drug Administration (FDA).** Ravicti: 203284Orig1s000: medical review(s) [online]. 27.11.2012. Silver Spring (USA): FDA. [Zugriff: 23.03.2018]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/203284Orig1s000MedR.pdf.
54. **U.S. Food and Drug Administration (FDA).** Ravicti: 203284Orig1s000; statistical review(s) [online]. 23.01.2013. Silver Spring (USA): FDA. [Zugriff: 23.03.2018]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/203284Orig1s000StatR.pdf.
55. **Van Wyhe K.** Wechsler Abbreviated Scale of Intelligence: Preliminary normative data for 12-15-year-old English-and Afrikaans-speaking Coloured learners in the Western Cape (Doctoral dissertation, University of Cape Town) [online]. 2012. [Zugriff. URL: https://open.uct.ac.za/bitstream/item/12279/thesis_hum_2012_van_wyhe_k.pdf?sequence=1].
56. **Varni JW, Limbers CA, Burwinkle TM.** How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. *Health Qual Life Outcomes* 2007;5:1.
57. **Wambach D, Lamar M, Swenson R, Penney DL, Kaplan E, Libon DJ.** Digit span. In: *Encyclopedia of Clinical Neuropsychology*. New York; Springer; 2011. 844-849.
58. **Warschausky S.** Wechsler Preschool and Primary Scale of Intelligence. In: *Encyclopedia of Clinical Neuropsychology*. New York: Springer; 2011. S. 2690-2693.
59. **Wechsler D.** Wechsler Abbreviated Scale of Intelligence. New York; Harcourt Brace & Company; 1999.
60. **Wechsler D.** WMS-III Administration and Scoring Manual. San Antonio; The Psychological Corporation: Harcourt Brace; 1997.
61. **Yi A.** California Verbal Learning Test (California Verbal Learning Test-II). In: *Encyclopedia of clinical neuropsychology*. New York; Springer; 2011.

Anhang

Charakterisierung der ergänzend berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz

Laut SGB V § 35a Absatz 1 Satz 11 1. Halbs. gilt der medizinische Zusatznutzen durch die Zulassung als belegt. Nachweise gemäß § 5 Absatz 1 bis 6 sind nicht verpflichtend vorzulegen. Zur Beurteilung der Patientenrelevanz der vorgelegten Endpunkte wurden das Dossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Das Ergebnis der Betrachtung der Endpunkte zu den Neuropsychologischen Tests sowie der Lebensqualität finden sich in den folgenden Tabellen 23 und 24.

Tabelle 23: Ergänzende Beschreibung der Endpunkte der Kategorie Morbidität

Endpunktkategorie Morbidität	
Endpunkt	Neuropsychologische Tests: Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III)
Operationalisierung	<p>Im Rahmen der neuropsychologischen Tests wurde in der Studie HPN-100-009 die Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) eingesetzt. Dem Studienbericht der Studie HPN-100-009 sind keine Details zum Testverfahren zu entnehmen.</p> <p>Folgend dargestellt sind Details zum Testverfahren, die anderen Quellen als den Unterlagen des pU entnommen wurden [7,8]. Die BSID-III dienen der Untersuchung des Entwicklungsniveaus von Kindern im Alter von 1 bis 42 Monaten und messen anhand folgender Skalen die kognitiven, sprachlichen und motorischen Funktionen:</p> <ul style="list-style-type: none"> • Sozial-emotionale Entwicklung (Fragebogen): Erfassen von sozial-emotionalen Begebenheiten, z. B. altersentsprechende emotionale Regulation, Interesse an Dingen und Geschehnissen, Kommunikationsbedürfnis usw. • Sprachskala (97 Items): besteht aus den Unterskalen rezepptive Kommunikation und expressive Kommunikation. • Kognitive Skala (91 Items): u. a. Sensomotorik, Explorationsverhalten und Handhabung von Gegenständen usw. • Motorische Skala (zusammen 138 Items): besteht aus den Unterskalen Feinmotorik und Grobmotorik <ul style="list-style-type: none"> ○ Feinmotorik (66 Items): u. a. Greifentwicklung, sensomotorische Integration (z. B. Augenbewegungen, Umblättern einer Buchseite) usw. ○ Grobmotorik: (72 Items): u. a. Motorik der Gliedmaßen und des Rumpfes, wie statische Haltungskontrolle (Kopfkontrolle, Sitzen, Stehen) usw. • Alltagsfertigkeiten (Adaptive Behavior): gliedert sich in unterschiedliche Subskalen (u. a. Kommunikation, Teilnahme an der Gesellschaft, Gesundheit und Sicherheit usw.). <p>Höhere Skalenwerte weisen auf eine bessere Funktion hin.</p> <p>In der Studie HPN-100-009 haben die Testleiter ein entsprechendes Training in der Abnahme des Bayley III absolviert, außerdem sind die Testleiter Psychologen oder andere ausgebildete Fachkräfte (z. B. Psychometriker), die von einem Psychologen supervidiert wurden.</p> <p>Erhebungszeitpunkte</p> <p>Der BSID-III wurde in der Studie HPN-100-009 zu folgenden Zeitpunkten erhoben:</p> <ul style="list-style-type: none"> • Patienten mit einem Alter < 12 Monate bei Studieneinschluss: im Alter von 12 und 24 Monaten • Patienten zwischen 12 und 18 Monaten bei Studieneinschluss: zum Zeitpunkt des Studieneinschlusses und im Alter von 24 Monaten • Patienten zwischen 18 und 24 Monaten bei Studieneinschluss: zum Zeitpunkt des Studieneinschlusses und 6 Monate später

Endpunktkategorie Morbidität	
Bewertung	<p>Operationalisierung Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Es ist unklar, ob alle Studienzentren sich bei der Testdurchführung an ein Testmanual gehalten haben.</p> <p>Validität und Patientenrelevanz BSID-III ist ein standardisierter Test, der die fein- und grobmotorischen und kognitiven Fähigkeiten von Kindern bis zum Alter von 42 Monaten bewertet [7,8]. Die Reliabilitätskoeffizienten (Cronbachs alpha) der Untertests der BSID-III-Skalen liegen zwischen $r=0,86$ (Feinmotorik) und $r=0,91$ (Grobmotorik) und belegen eine gute Reliabilität. Die inhaltliche Validität und Konstruktvalidität konnten nachgewiesen werden. Zur <i>Known Groups Validity</i> liegen Nachweise vor [7,8]. Eine Validierung der deutschen Version ist in Ansätzen erfolgt [39]. Keine Informationen liegen zur Veränderungssensitivität vor. Kognitive, sprachliche und motorische Funktionen werden als patientenrelevant eingeschätzt. Aufgrund fehlender Informationen zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität bleibt die Relevanz von Veränderungen im BSID-III unklar.</p> <p>Es lagen BSID-III-Daten nur von maximal 3 der 10 eingeschlossenen Patienten (Studienkohorte 2 Monate bis < 2 Jahre) vor. Aufgrund dieses geringen Anteils sowie der fehlenden Angaben zur Änderungssensitivität werden die Daten zum BSID-III als nicht valide betrachtet. Daher werden die Ergebnisse nicht dargestellt.</p>
Endpunkt	Neuropsychologische Tests: Wechsler Preschool and Primary Scales of Intelligence – Third Edition (WPPSI-III)
Operationalisierung	<p>Der Intelligenztest Wechsler Preschool and Primary Scales of Intelligence – Third Edition (WPPSI-III) wurde im Rahmen einer neuropsychologischen Testbatterie in den Studien HPN-100-012 und HPN-100-012SE bei Kindern im Alter zwischen 3 und 6 Jahren erhoben. Dem Studienbericht der Studie HPN-100-012SE sind keine Details zum Testverfahren zu entnehmen (auch keine Literaturangabe).</p> <p>In der Studie HPN-100-012SE wurden Kinder im Alter zwischen 29 Tagen und 6 Jahren eingeschlossen. Bei der Durchführung des Testverfahrens wurde allerdings die Stichprobe auf Kinder ab 3 Jahren begrenzt. Dem Studienbericht ist zu entnehmen, dass der WPPSI-III erst ab einem Alter von 2,5 Jahren normiert ist. Weiterhin ist angegeben, dass die Messungen des WPPSI-III bei Kindern unter 3 Jahren nicht reliabel seien und Messungen in der Altersgruppe vor allem in Beziehung zu Entwicklungsfaktoren wie Sprach- oder Motorikentwicklung stehen würden. Daher wurden Kinder unter 3 Jahren nicht mit dem WPPSI-III untersucht.</p> <p>Folgend dargestellt sind Details zum Testverfahren, die anderen Quellen als den Unterlagen des pU entnommen wurden [22,45,58]. Der WPPSI-II dient der Überprüfung der kognitiven Leistungsfähigkeit von Kindern im Alter von 2 Jahren und 6 Monaten bis zu einem Alter von 7 Jahren und 3 Monaten. Der Test ist nochmal unterteilt in 2 Altersgruppen: eine Version für jüngere Kinder (2 Jahre und 6 Monate bis 3 Jahre und 11 Monate) und eine Version für ältere Kinder (4 Jahre bis 7 Jahre und 3 Monate). Die beiden Versionen beinhalten unterschiedliche Tests:</p> <ul style="list-style-type: none"> • Tests für jüngere Kinder: Vier Kernuntertests und zwar zwei aus dem Verbalteil (Passiver Wortschatz und Allgemeines Wissen) und zwei aus dem Handlungsteil (Mosaik-Test und Figuren legen). Aus diesen vier Kernuntertests kann anschließend der generelle Intelligenzquotient abgeleitet werden. Ein optionaler Untertest (Aktiver Wortschatz) kann zusätzlich durchgeführt werden. Dauer: ca. 30–45 Minuten. • Tests für ältere Kinder: Sieben Kernuntertests und zwar drei aus dem Verbalteil (Allgemeines Wissen, Wortschatz-Test und Begriffe erkennen), drei aus dem Handlungsteil (Mosaik-Test, Matrizen-Test und Bildkonzepte) sowie der Test Symbole kodieren. Aus diesen sieben Kernuntertests kann anschließend der generelle Intelligenzquotient abgeleitet werden. Ergänzend können zudem fünf optionale Untertests (Symbol-Suche, Allgemeines Verständnis, Gemeinsamkeiten finden,

Endpunktkategorie Morbidität	
	<p>Figuren legen und Bilder ergänzen) und zwei zusätzliche Untertests (Passiver Wortschatz und Aktiver Wortschatz) erhoben werden. Dauer: ca. 40–60 Minuten.</p> <p>Der WPPSI-II wurde an einer Stichprobe von 1700 Kindern (stratifiziert nach Abstammung, Ausbildung der Eltern und geographische Region) normiert und in neun Subgruppen eingeteilt. Im Folgenden werden kurz die Aufgaben der Kernuntertests aufgelistet:</p> <p>Kernuntertests für jüngere Kinder:</p> <ul style="list-style-type: none"> • Passiver Wortschatz: Das Kind sieht sich vier Bilder an und soll auf das Bild zeigen, das der Testleiter benennt. • Allgemeines Wissen: Das Kind beantwortet Fragen über allgemein bekannte Ereignisse, Sachverhalte, Orte und Persönlichkeiten. • Mosaik-Test: Mithilfe von ein- oder zweifarbigen Würfeln soll das Kind unterschiedlich komplexe Mustervorlagen (Modell oder Bildvorlage) innerhalb einer vorgegebenen Zeitspanne nachbauen. • Figuren legen: Dem Kind werden auf standardisierte Weise Teile eines Puzzles vorgelegt. Das Kind setzt sie innerhalb einer festgelegten Zeit zu einem bedeutungshaltigen Ganzen zusammen. <p>Kernuntertests für ältere Kinder:</p> <ul style="list-style-type: none"> • Allgemeines Wissen: Das Kind beantwortet Fragen über allgemein bekannte Ereignisse, Sachverhalte, Orte und Persönlichkeiten. • Wortschatz-Test: Das Kind gibt Definitionen für die vom Testleiter vorgegebenen Worte. • Begriffe erkennen: Das Kind entschlüsselt einen allgemeinen Begriff, der vom Testleiter umschrieben wird. • Mosaik-Test: Mithilfe von ein- oder zweifarbigen Würfeln soll das Kind unterschiedlich komplexe Mustervorlagen (Modell oder Bildvorlage) innerhalb einer vorgegebenen Zeitspanne nachbauen. • Matrizen-Test: Das Kind betrachtet eine unvollständige Matrix und wählt das fehlende Teil aus vier bzw. fünf Antwortmöglichkeiten. • Bildkonzepte: Das Kind soll aus zwei Bildreihen (mit je zwei oder drei Bildern) jeweils ein Bild auswählen, um daraus eine Gruppe mit einer gemeinsamen Eigenschaft zu bilden. • Symbole kodieren: Einer Serie einfacher geometrischer Figuren soll das Kind abstrakte Symbole zuordnen. Es zeichnet in einer begrenzten Zeit die Symbole in die dazugehörigen Felder, indem es einen vorgegebenen Schlüssel verwendet. <p>Die jeweiligen Kernuntertests sind nach Verbal- und Handlungsteil unterteilt. Die Rohwerte aus den Kernuntertests werden in normierte Wertpunkte transformiert, sodass ein individuelles Testprofil für jedes Kind erstellt werden kann. Die Wertpunkte können dann in Intelligenzquotienten für den Verbal- und Handlungsteil sowie einen generellen Intelligenzquotienten überführt werden, die jeweils standardisiert einen Mittelwert von 100 und eine Standardabweichung von 15 aufweisen.</p> <p>Die Ergebnisse liefern eine Messung für die kognitiven Fähigkeiten und bieten zudem u. a. Informationen bei Entwicklungsauffälligkeiten.</p> <p>Erhebungszeitpunkte</p> <p>Der WPPSI-III wurde in den Studien HPN-100-012 und HPN-100-012SE bei Kindern im Alter zwischen 3 und 6 Jahren erhoben im Rahmen der neuropsychologischen Testbatterie zu den Zeitpunkten Tag 1 (Baseline-Erhebung der Studie HPN-100-012) und Monat 12 (letzte Erhebung der Studie HPN-100-012SE).</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Bei Intelligenztests kann die Durchführungsobjektivität nur eingehalten werden, wenn die Testerhebung anhand eines Manuals durchgeführt wird. Es ist unklar, ob alle Studienzentren sich bei der Testdurchführung an ein Testmanual gehalten haben (in</p>

Endpunktkategorie Morbidität	
	<p>den Studienberichten der Studien HPN-100-012 und HPN-100-012SE wurde auf kein Testmanual verwiesen).</p> <p>Validität und Patientenrelevanz</p> <p>Der WPPSI-III ist ein individuell durchgeführtes Testverfahren, um die Intelligenz bei Kindern im Alter von 2 Jahren und 6 Monaten bis zu einem Alter von 7 Jahren und 3 Monaten zu erfassen. Damit repräsentiert das Testverfahren ein Maß für die kognitive Leistungsfähigkeit [22,45,58].</p> <p>Die psychometrischen Eigenschaften des Testverfahrens sind als sehr gut einzuschätzen. Die Reliabilitäts-Koeffizienten zur internen Konsistenz für die Subtests rangieren zwischen 0,83 und 0,95, für die zusammengefassten Werte (Intelligenzquotienten für den Verbal- und Handlungsteil sowie einen generellen Intelligenzquotienten) bewegen sich die Reliabilitäts-Koeffizienten zur internen Konsistenz zwischen 0,89 und 0,96. Die Test-Retest-Reliabilität (für ein mittleres Intervall von 26 Tagen) für die Subtests war im Bereich 0,74 bis 0,90. Für die zusammengefassten Werte (Intelligenzquotienten) war die Test-Retest-Reliabilität im Bereich zwischen 0,86 und 0,91. Die Interrater-Reliabilität war insgesamt auch sehr hoch (Reliabilitäts-Koeffizienten zw. 0,97 und 0,99).</p> <p>Hinsichtlich der Validität wurde die interne Struktur des Tests sowie der Zusammenhang mit anderen Intelligenzmaßen untersucht. Die interne Struktur scheint gegeben zu sein, da die einzelnen Untertests signifikant miteinander zusammenhängen. Außerdem wurde die Faktorenstruktur untersucht, die eine Zwei-Faktoren-Lösung nahelegen. Der Intelligenzquotient des WPPSI-III hängt u. a. mit dem BSID-II (Bayley Scales of Infant Development – II) oder dem DAS (Differential Abilities Scales) zusammen, sodass die Konstruktvalidität gegeben ist.</p> <p>Insgesamt scheinen die psychometrischen Eigenschaften des WPPSI-III hinsichtlich Reliabilität und Validität gegeben zu sein. Angaben zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität konnten den Unterlagen des pU nicht entnommen werden.</p> <p>Der WPPSI-III kann Informationen bei Entwicklungsauffälligkeiten liefern und könnte daher in der vorliegenden Fragestellung als ein Index für Entwicklungs- und Gedeihstörungen sowie als Hinweis auf kognitive Beeinträchtigungen durch HAC herangezogen werden. Somit ist der Endpunkt als patientenrelevant einzuschätzen.</p> <p>In der Studie HPN-100-012SE wurden 23 Studienteilnehmer eingeschlossen, davon waren 16 zum Zeitpunkt des Studieneinschlusses zwischen 2 und 6 Jahren alt (siehe Abschnitt 2.6.1). Der WPPSI-III wurde zum Zeitpunkt der Baseline von 10 Studienteilnehmern erhoben. Eine Rücklaufquote basierend auf diesen Daten liegt unter 70 %. Eine genaue Rücklaufquote ist nicht bestimmbar, da den Unterlagen des pU keine Stichprobengröße für Studienteilnehmer zwischen 3 und 6 Jahren zu entnehmen ist. Aufgrund dieser Einschränkung und der möglichen Einschränkungen bezüglich der Durchführungsobjektivität sowie der fehlenden Informationen zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität werden die Ergebnisse nicht dargestellt.</p>
Endpunkt	Neuropsychologische Tests: Wechsler Abbreviated Scale of Intelligence (WASI)
Operationalisierung	<p>Der Intelligenztest Wechsler Abbreviated Scale of Intelligence (WASI) wurde im Rahmen einer neuropsychologischen Testbatterie in den Studien HPN-100-005SE, HPN-100-007 und HPN-100-011 bei Kindern, Jugendlichen (zwischen 6 und 17 Jahren) und erwachsenen Studienteilnehmern erhoben. In dem Studienbericht der Studie HPN-100-011 sind Details zum Testverfahren angegeben [59]. Es handelt sich um ein verkürztes Testverfahren zur Erfassung der generellen intellektuellen Fähigkeiten und basiert auf zwei Subtests zur verbalen Fähigkeit (Wortschatz-Test) und zum logischen Denken (Matrizen-Test).</p> <p>Der Wortschatz-Test wurde entworfen, um Sprachwissen und Sprachverständnis zu messen. Der Test beinhaltet 4 Bilder, die die Studienteilnehmer benennen sollten, und 38 verbale visuell und oral präsentierte Items, die die Studienteilnehmer definieren sollten.</p>

Endpunktkategorie Morbidität	
	<p>Der Matrizen-Test wurde entworfen, um beispielsweise fluide Intelligenz, räumliches Vorstellungsvermögen, Fähigkeit Teil-Ganzes Beziehungen zu erkennen und parallele Verarbeitung zu messen. In diesem Test betrachten die Studienteilnehmer 35 unvollständige Matrizen und werden aufgefordert, die Matrize mit einer aus fünf verschiedenen Antwortmöglichkeiten zu ergänzen.</p> <p>Für beide Tests wurde die Anzahl der richtigen Antworten gezählt und mittels des WASI-Manuals in T-Werte transformiert. T-Werte sind standardisierte Werte mit einem Mittelwert von 50 und einer Standardabweichung von 10.</p> <p>Die Rohwerte beider Tests wurden aufsummiert und anhand des WASI-Manuals in einen Standardwert-Intelligenzquotienten mit dem Mittelwert 100 und der Standardabweichung 15 transformiert.</p> <p>Der WASI ist für die Altersspanne 6 bis 89 Jahre normiert. Der Test wurde ab einem Alter von 6 Jahren bei den Studienteilnehmern eingesetzt.</p> <p>Folgend dargestellt sind Details zum Testverfahren, die anderen Quellen als den Unterlagen des pU entnommen wurden [47]. Der WASI ist ein Screening-Instrument und enthält Subtests, die leicht abgewandelt in den Testverfahren Wechsler Adult Intelligence Scale – Third Edition (WAIS-III) und Wechsler Intelligence Scale for Children – Third Edition (WISC-III) enthalten sind. Der WASI enthält in der Regel vier Subtests, kann aber auch in einer gekürzten Version mit zwei Tests angewendet werden. Die gekürzte Version dieses Screeninginstruments kommt in den vorgelegten Studien zum Einsatz.</p> <p>Der WASI wurde an 2.245 repräsentativen Kindern und Erwachsenen (stratifiziert nach Geschlecht, Abstammung und geographischer Region) normiert.</p> <p>Erhebungszeitpunkte</p> <p>Der WASI wurde in den Studien HPN-100-005SE und HPN-100-007 im Rahmen der neuropsychologischen Testbatterie zu den Zeitpunkten 0 (Baseline) und 12 Monate erhoben. Bei der Studie HPN-100-011, die eine Sicherheitsweiterung u. a. der Studien HPN-100-005SE und HPN-100-007 darstellt, wurde der Test im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und dann alle 12 Monate (bis zum Studienende) erhoben.</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Bei Intelligenztests kann die Durchführungsobjektivität nur eingehalten werden, wenn die Testerhebung anhand eines Manuals durchgeführt wird. Es ist unklar, ob alle Studienzentren sich bei der Testdurchführung an ein Testmanual gehalten haben (auf das Testmanual wurde nur für die Berechnung der Werte Bezug genommen). Die Informationen zu dem Test sind dem Studienbericht der Studie HPN-100-011 entnommen. In den Studienberichten der Studien HPN-100-005SE und HPN-100-007 sind keine ausführlicheren Informationen zu dem Test zu entnehmen.</p> <p>Validität und Patientenrelevanz</p> <p>Der WASI ist ein standardisiertes Screeninginstrument, das in der vorliegenden Version zwei Subtests enthält und damit eine verkürzte Version des WASI darstellt. Der Wortschatz-Test erfasst Sprachverständnis, während der Matrizen-Test logisches bzw. schlussfolgerndes Denken erfasst. Die Werte beider Tests, aufaddiert und transformiert, bilden zudem einen Intelligenzquotienten [47].</p> <p>Hinsichtlich der psychometrischen Eigenschaften wurden Reliabilität und Validität untersucht (vor allem bei der Version mit vier Untertests). Bei der Test-Halbierungs-Reliabilität fanden sich zwischen den Testhälften keine Unterschiede, außerdem rangierten die Reliabilitäts-Koeffizienten zwischen 0,81 und 0,97 für Kinder und zwischen 0,84 und 0,98 für Erwachsene [47]. Die Test-Retest-Reliabilitäten waren außerdem stets größer als 0,87 [46,55].</p> <p>Hinsichtlich der Validität ist anzumerken, dass der WASI mit anderen Intelligenzmaßen wie z. B. dem Wechsler Individual Achievement Test oder Wide Range Intelligence Test zusammenhängt [10]. Der WASI wird aufgrund seiner Kürze häufig in verschiedenen</p>

Endpunktkategorie Morbidität	
	<p>klinischen Populationen verwendet z. B. bei geistiger Behinderung oder bei Gehirnverletzungen. Insgesamt ist die Version mit vier Subtests eher geeignet als die Version mit zwei Subtests, um einen validen Messwert zu erhalten.</p> <p>Insgesamt scheinen die psychometrischen Eigenschaften des WASI hinsichtlich Reliabilität und Validität gegeben zu sein. Es ist nicht auszuschließen, dass die Reliabilität und Validität bei der Version mit zwei Subtests (wie hier in den Studien angewendet) eingeschränkt ist. Angaben zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität konnten den Unterlagen des pU nicht entnommen werden.</p> <p>Der WASI kann Informationen bei Entwicklungsauffälligkeiten liefern und wird auch bei Populationen mit Gehirnverletzungen eingesetzt. In der vorliegenden Fragestellung könnte der WASI somit als ein Hinweis auf kognitive Beeinträchtigungen durch hyperammonämische Krisen herangezogen werden. Somit ist der Endpunkt als patientenrelevant einzuschätzen.</p> <p>Aufgrund der möglichen Einschränkungen bezüglich der Durchführungsobjektivität sowie den fehlenden Informationen zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität werden die Ergebnisse nicht dargestellt.</p>
Endpunkt	Neuropsychologische Tests: The Child Behavior Checklist (CBCL)
Operationalisierung	<p>Emotionale Funktionen und Verhalten wurden in den Studien HPN-100-005SE, HPN-100-007, HPN-100-011, HPN-100-012 und HPN-100-012SE mit der Child Behavior Checklist (CBCL) bewertet.</p> <p>Die CBCL ist ein Instrument zur Beurteilung pathologischen Verhaltens und sozialer Kompetenz bei Kindern und Jugendlichen. Für die verschiedenen Altersklassen existieren zwei Versionen dieses Instruments:</p> <ul style="list-style-type: none"> • Version für Kinder von 1,5 bis 5 Jahren [2] • Version für Kinder von 6 bis 18 Jahren [3] <p>Folgend dargestellt sind Details zum Testverfahren, die anderen Quellen als den Unterlagen des pU entnommen wurden [1,2,3].</p> <p><i>Version für Kinder von 1,5 bis 5 Jahren</i></p> <p>Die Fragebogenversion für die Altersklasse der 1,5- bis 5-Jährigen beinhaltet 100 Items zu Verhaltensproblemen plus Beschreibungen von Beeinträchtigungen und Dingen, die die Eltern am meisten über das Kind beunruhigen sowie Beschreibungen zu positiven Aspekten zu dem Kind. Der Fragenbogen umfasst 12 Domänen:</p> <p>Problem-Skalen</p> <ul style="list-style-type: none"> • Emotionale Reaktivität • Ängstlich/depressiv • Körperliche Beschwerden • Sozialer Rückzug • Schlafprobleme • Aufmerksamkeitsprobleme • Aggressives Verhalten <p>DSM-orientierte Skalen</p> <ul style="list-style-type: none"> • Depressive/affektive Symptome • Angstsymptome • Allgegenwärtige Entwicklungsprobleme/Autismus-Spektrum-Probleme • Unaufmerksamkeits-Hyperaktivitätssymptome • Oppositionelle Verhaltenssymptome <p>Der Bezugszeitraum sind die letzten 2 Monate. Zusätzlich wird bei Kindern zwischen 18 und 35 Monaten anhand von 9 Items die Sprachentwicklung beurteilt.</p>

Endpunktkategorie Morbidität	
	<p><i>Version für Kinder von 6 bis 18 Jahren</i> 120 Items zu Verhaltensproblemen und 20 Items zu sozialer Kompetenz der Fragebogenversion für die 6- bis 18-Jährigen lassen sich folgenden Domänen zuordnen:</p> <p>Kompetenz and adaptive Skalen</p> <ul style="list-style-type: none"> • Aktivitäten • Soziales • Schule <p>Problemskalen</p> <ul style="list-style-type: none"> • Ängstlich/depressiv • Rückzöglich/depressiv • Körperliche Beschwerden • Soziale Probleme • Denkprobleme • Aufmerksamkeitsstörungen • Regelverletzendes Verhalten • Aggressives Verhalten <p>DSM-orientierte Skalen:</p> <ul style="list-style-type: none"> • Depressive/affektive Symptome • Angstsymptome • Körperliche Probleme • Unaufmerksamkeits-Hyperaktivitätssymptome • Oppositionelle Verhaltenssymptome • Dissoziale Symptome <p>Der Bezugszeitraum sind die letzten 6 Monate.</p> <p>Die Fragebögen werden von den Eltern entweder eigenständig oder Interview-basiert ausgefüllt. Für die Altersgruppe der 6- bis 18-Jährigen steht zudem ein Lehrerfragebogen und ein Fragebogen zum Selbstberichten für Jugendliche zur Verfügung. Je nach Frage werden die Antworten auf einer 3- oder 4-Punkt-Likert-Skala gegeben oder die Fragen mit ja/nein beantwortet. Darüber hinaus gibt es offene Fragen. Für jede Domäne wird ein Domänen-Score berechnet. Zudem können für beide Versionen übergeordnete Skalen berechnet werden (Gesamtauffälligkeit, Internalisierende Auffälligkeiten, Externalisierende Auffälligkeiten). Höhere Werte bedeuten mehr berichtete Probleme. Normwerte stehen im Manual zur Verfügung.</p> <p>Erhebungszeitpunkte</p> <ul style="list-style-type: none"> • HPN-100-005SE: Baseline und nach 12 Monaten • HPN-100-007 (nur Kohorte der 6- bis 17-Jährigen): Baseline und nach 12 Monaten • HPN-100-011 (nur bei Kindern ≥ 5 Jahre): Baseline, nachfolgend alle 12 Monate und zum Zeitpunkt der Studienbeendigung • HPN-100-012 bzw. HPN-100-012SE: (nur bei 3- bis 5-Jährigen): Baseline und nach 12 Monaten oder bei frühzeitiger Studienbeendigung <p>Ergebnisdarstellung</p> <p>Die Ergebnisse wurden als T-Werte berichtet. Für die Studie HPN-100-005SE und Studie HPN-100-007 wurden die Ergebnisse für Intern. Probleme, Extern. Probleme, Affektive Symptome, Körperliche Symptome, Unaufmerksamkeits-Hyperaktivitätssymptome, Oppositionelle Verhaltenssymptome, Dissoziale Symptome und Gesamtauffälligkeit dargestellt.</p> <p>Für die Studien HPN-100-012 bzw. HPN-100-012SE und HPN-100-011 wurden die T-Werte aller Domänen berichtet.</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Den Unterlagen des pU ist nicht zu entnehmen, wie die Durchführungsobjektivität eingehalten</p>

Endpunktkategorie Morbidität	
	<p>wurde. Es wurde angegeben, dass der Fragebogen teils Interview-basiert durchgeführt wurde. Es ist unklar, ob alle Studienzentren sich bei der dieser Art der Testdurchführung an ein Testmanual gehalten haben.</p> <p>Validität und Patientenrelevanz Emotionale Funktionen und Verhalten werden als patientenrelevante Endpunkte bewertet. Untersuchungen des Instrumentes weisen auf eine ausreichend hohe Test-Retest-Reliabilität der einzelnen Domänen der CBCL-Versionen hin (Version für 1,5- bis 5-Jährige: N=68, Pearson Korrelationskoeffizienten 0,7–0,9; Version für 6- bis 18-Jährige: N=73; Pearson Korrelationskoeffizienten 0,8–0,9). Die Inhaltsvalidität und Kriteriumsvalidität wurde ausreichend beschrieben. Daten zur diskriminanten Validität und Konstruktvalidität liegen vor [1,2,3]. Informationen zur Änderungssensitivität konnten ebenso wie Angaben zur MID nicht identifiziert werden. Damit kann die Relevanz von Veränderungen in den CBCL-Domänen nicht beurteilt werden. Aufgrund der möglichen Einschränkungen bezüglich der Durchführungsobjektivität sowie den fehlenden Informationen zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität werden die Ergebnisse nicht dargestellt.</p>
Endpunkt	Neuropsychologische Tests: Behavior Rating Inventory of Executive Function® (BRIEF)
Operationalisierung	<p>Es handelt sich um ein Instrument zur Einschätzung von so genannten exekutiven Funktionen bei Kindern und Jugendlichen im Alter von 5 bis 18 Jahren. Exekutive Funktionen beziehen sich auf die Kontrolle von Kognition, Emotionen und Verhalten. Der BRIEF liegt in drei Versionen vor: für Eltern, Lehrer und zur Selbstbeurteilung (ab 11 Jahren). Er umfasst insgesamt 86 Fragen verteilt auf acht Skalen (Hemmen, Umstellen, emotionale Kontrolle, Initiative, Arbeitsgedächtnis, Planen/Strukturieren, Ordnen/Organisieren, Überprüfen) und drei Indizes. Der Verhaltensregulations-Index wird aus den Skalen Hemmen, Umstellen und emotionale Kontrolle gebildet, der Kognitive Regulations-Index aus den Skalen Initiative, Arbeitsgedächtnis, Planen/Strukturieren, Ordnen/Organisieren und Überprüfen. Aus den beiden Indizes wird der Exekutive Gesamtwert gebildet. Die Ausfüllzeit beträgt 10–15 Minuten. Um tendenziöse Beurteilungen zu identifizieren, wurden noch jeweils ein Negativitäts- und ein Inkonsistenzscore entwickelt (Validitätsskalen). Diese wurden im Rahmen der Studie jedoch nicht erhoben. Der BRIEF wurden entsprechend dem Manual des Instruments angewendet.</p> <p>Der BRIEF wird eingesetzt, um Kinder und Jugendliche mit einem breiten Spektrum von angeborenen oder erworbenen neurologischen Konditionen in den Settings Zuhause und in der Schule zu untersuchen. Das Instrument wurde in den Jahren 2000 bis 2005 von den Autoren Gerard A. Gioia, Kimberly A. Espy und Peter K. Isquith entwickelt. Es steht ein Manual sowie eine Software zur Auswertung zur Verfügung.</p> <p>Für die Analyse wird der T-Score durch Transformation der Rohwerte für jede Skala kalkuliert, um einen Vergleich mit einer Referenzgruppe zu ermöglichen. Ein höherer Score indiziert schlechtere exekutive Funktionen. Ein T-Score von ≥ 65 wird als Grenzwert für Auffälligkeiten gewertet. Als mittlerer Normwert wird ein T-Score von 50 (SD 10) angegeben. Eine Änderung im T-Score von 5 Punkten (0,5 SD) gilt als klinisch signifikanter Unterschied.</p> <p>Erhebungszeitpunkte</p> <ul style="list-style-type: none"> • HPN-100-005SE: Baseline und nach 12 Monaten • HPN-100-007 (nur Kohorte der 6- bis 17-Jährigen): Baseline und nach 12 Monaten • HPN-100-011 (nur bei Kindern ≥ 5 Jahre): Baseline, nachfolgend alle 12 Monate und zum Zeitpunkt der Studienbeendigung
Bewertung	<p>Operationalisierung Die Operationalisierung des Endpunkts ist weitgehend nachvollziehbar. Es bleibt unklar, warum die Validitätsskalen zum Negativitäts- und ein Inkonsistenzscore nicht miterhoben wurden. Tendenziöse Beurteilungen konnten so nicht identifiziert werden.</p>

Endpunktkategorie Morbidität	
	<p>Validität und Patientenrelevanz</p> <p>Das Instrument ist in mehreren Studien, u. a. bei Kindern mit ADHS, Autismus, Epilepsie und bei Frühgeborenen untersucht worden. Nach Angaben des Hogrefe-Verlags [15] liegt die interne Konsistenz der Eltern- bzw. Lehrerversion bei 0,79 und 0,89. Die Retest-Reliabilität wird mit über 0,8 angegeben. Die Interrater-Reliabilität wird mit 0,3 (im Vergleich Lehrer-Eltern) angegeben [20]. Die klinische Validität wurde in Faktorenanalysen anhand verschiedener Störungsbilder an 374 Kindern untersucht [21]. Die Normwerte wurden aus Beurteilungen von rund 2.100 repräsentativ ausgewählten Lehrern und Eltern abgeleitet.</p> <p>Hinsichtlich der Konstruktvalidität konnte ein Zusammenhang des BRIEF mit bereits anderen Instrumenten wie z. B. der Tower-of-London-Aufgabe gezeigt werden [4]. Der BRIEF ist ein Verhaltensinventar, das ein Profil zu relevanten Verhaltensweisen liefert und einen Vergleich im Zeitverlauf ermöglicht. Das Instrument kann damit patientenrelevante Informationen liefern. Die psychometrischen Eigenschaften des BRIEF hinsichtlich Reliabilität und Validität erscheinen ausreichend. Angaben zur Änderungssensitivität konnten den Unterlagen des pU nicht entnommen werden. Einschränkung muss angemerkt werden, dass vom pU keine Informationen zu psychometrischen Eigenschaften des BRIEF in der Zielpopulation vorgelegt wurden. Aufgrund dieser Einschränkungen sowie der fehlenden Erhebung der Validitätskalen (Negativitäts- und ein Inkonsistenzscore) werden die Ergebnisse nicht dargestellt.</p>
Endpunkt	Neuropsychologische Tests: Behavior Rating Inventory of Executive Function®-Preschool Edition (BRIEF-P)
Operationalisierung	<p>Es handelt sich um eine Version des BRIEF bei Vorschulkindern im Alter von 2–5 Jahren und 11 Monaten. Es handelt sich um einen Elternfragebogen mit einer Ausfüllzeit 10–15 Minuten. Im Unterschied zum BRIEF werden 63 Items in fünf Skalen erfasst (Hemmung, Umstellfähigkeit, emotionale Kontrolle, Arbeitsgedächtnis und Planen/Organisieren), aus denen die Indizes Inhibitorische Selbstkontrolle, Flexibilität und Metakognitive Entwicklung gebildet werden. Aus den fünf Skalen wird der Exekutive Gesamtwert gebildet. Die Auswertung erfolgt analog zum BRIEF. T-Werte ≥ 65 werden als klinisch signifikant erachtet. Dem Studienbericht der Studie HPN-100-012SE sind keine weiteren Details zum Testverfahren zu entnehmen.</p> <p>Erhebungszeitpunkte</p> <p>HPN-100-012 bzw. HPN-100-012SE: (nur bei 3-bis 5-Jährigen): Baseline und nach 12 Monaten oder bei frühzeitiger Studienbeendigung</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist im Wesentlichen nachvollziehbar.</p> <p>Validität und Patientenrelevanz</p> <p>Nach Angaben des Hogrefe-Verlags [13] liegt die interne Konsistenz der Eltern- bzw. Erzieherversion bei 0,75 und 0,94. Die Interrater-Reliabilität wird mit 0,56 für den Gesamtwert angegeben. Alters- und geschlechtsgetrennte T-Werte und Prozentränge wurden für Kinder im Vorschulalter (N = 460) ermittelt.</p> <p>Untersuchungen zur Validität liegen analog zum BRIEF durch Zusammenhänge mit anderen Instrumenten (z. B. CBCL) vor [36].</p> <p>Die psychometrischen Eigenschaften des BRIEF-P hinsichtlich Reliabilität und Validität erscheinen ausreichend. Angaben zur Änderungssensitivität konnten den Unterlagen des pU nicht entnommen werden. Einschränkung muss angemerkt werden, dass vom pU keine Informationen zu psychometrischen Eigenschaften des BRIEF-P in der Zielpopulation vorgelegt wurden. Aufgrund dieser Einschränkungen werden die Ergebnisse nicht dargestellt.</p>

Endpunktkategorie Morbidität	
Endpunkt	Neuropsychologische Tests: Grooved Pegboard
Operationalisierung	<p>Der Test Grooved Pegboard wurde im Rahmen einer neuropsychologischen Testbatterie bei erwachsenen Studienteilnehmern in den Studien HPN-100-007 und HPN-100-011 erhoben. Im Studienbericht der Studie HPN-100-011 sind Details zum Testverfahren angegeben [52]. Es handelt sich um ein Testverfahren zur Händigkeit, das motorische und visuelle Fähigkeiten bei Erwachsenen, Jugendlichen und Kindern erfasst. Bei dem Test wird den Studienteilnehmern ein Testapparat vorgelegt, der 25 Vertiefungen enthält. Jede Vertiefung enthält eine zufällig angeordnete Schräge, sodass ein passender angeschrägter Stift in seiner Orientierung nach der Einkerbung ausgerichtet werden muss. Um den Test erfolgreich zu bewältigen, werden komplexe visuomotorische Fähigkeiten verlangt.</p> <p>Die Aufgabe des Studienteilnehmers besteht darin, die 25 Vertiefungen so schnell wie möglich mit den Stiften auszufüllen. Der Test wird zunächst mit der dominanten Hand durchgeführt. Im zweiten Durchgang wird die nicht-dominante Hand genutzt. Es wird die Zeit erfasst, die benötigt wird, um die Aufgabe abzuschließen. Anschließend werden altersspezifische Normwerte (für die mittlere Bearbeitungszeit sowie die Standardabweichung der Bearbeitungszeit) aus dem Manual entnommen, die für die Berechnung von standardisierten Z-Werten verwendet werden. Die Z-Werte haben dann einen Mittelwert von 0 und eine Standardabweichung von 1. Höhere Z-Werte sind ein Indikator für bessere Leistung.</p> <p>Aus dem Nutzerhandbuch des Testherstellers des Grooved Pegboard [51] wird außerdem klar, dass Normwerte für Kinder ab 5 Jahren vorliegen. Der Test wurde vor allem für Lateralisierungseinschränkungen bei Morbus Parkinson eingesetzt. Für einen Lateralisierungsindex werden die Bearbeitungszeiten der dominanten und nicht-dominanten Hand in Beziehung gesetzt [41].</p> <p>Erhebungszeitpunkte</p> <p>Der Test Grooved Pegboard wurde in Studie HPN-100-007 im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und 12 Monate erhoben. Bei der Studie HPN-100-011, die eine Sicherheitserweiterung u. a. der Studie HPN-100-007 darstellt, wurde der Test im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und dann alle 12 Monate (bis zum Studienende) erhoben.</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Es ist unklar, ob alle Studienzentren sich bei der Testdurchführung an ein Testmanual gehalten haben. Die Informationen zu dem Test sind dem Studienbericht der Studie HPN-100-011 entnommen. Im Studienbericht der Studie HPN-100-007 sind keine ausführlicheren Informationen zu dem Test zu entnehmen.</p> <p>Validität und Patientenrelevanz</p> <p>Der Test Grooved Pegboard ist ein standardisierter Test, der visuomotorische Fähigkeiten (insbesondere Auge-Hand-Koordination und motorische Geschwindigkeit) erfasst, die eine wichtige Hirnfunktion repräsentieren.</p> <p>Hinsichtlich der psychometrischen Eigenschaften wurde u. a. Reliabilität untersucht. Die Test-Retest-Reliabilitäten waren 0,69 für die dominante und 0,76 für die nicht-dominante Hand. Dies könnte an Lerneffekten liegen, die vor allem durch Verbesserungen zwischen dem ersten und zweiten Mal der Testdurchführung zu beobachten waren.</p> <p>Der Test Grooved Pegboard wird auch eingesetzt, um kognitive Verlangsamung durch Krankheitsprogression vor allem bei neurologischen Erkrankungen zu erfassen [41].</p> <p>Der pU hat keine Daten zur Reliabilität, zur Validität, zur Änderungssensitivität und zum klinisch relevanten Unterschied (MID) dargelegt. Aufgrund fehlender Informationen zu diesen Gütekriterien bleibt die Aussagekraft des Grooved Pegboard Tests unklar. Hinsichtlich der Patientenrelevanz bleibt unklar, welche Rolle die Händigkeit sowie die visuomotorischen Fähigkeiten im Krankheitsverlauf spielen. Somit sind die</p>

Endpunktkategorie Morbidität	
	Patientenrelevanz und die Validität nicht abschließend beurteilbar. Es werden daher keine Ergebnisse dargestellt.
Endpunkt	Neuropsychologische Tests: The California Verbal Learning Test (CVLT-II)
Operationalisierung	<p>Der California Verbal Learning Test in der zweiten Version (CVLT-II) wurde im Rahmen einer neuropsychologischen Testbatterie bei erwachsenen Studienteilnehmern in den Studien HPN-100-007 und HPN-100-011 erhoben. In dem Studienbericht der Studie HPN-100-011 sind Details zum Testverfahren angegeben [14]. Der Test erfasst verbale Lern- und Gedächtnisfähigkeiten mit verschiedenen Testwerten.</p> <p>In der Lernphase wurden den Studienteilnehmern eine Liste A mit 16 Wörtern vorgelesen, die zu vier verschiedenen Kategorien gehören (z. B. Früchte, Spielzeuge). Die Liste A wurde insgesamt fünf Mal vorgelesen und nach jedem Vorlesen wurden die Studienteilnehmer aufgefordert, die Liste wiederzugeben (sofortiger Abruf). Anschließend wurde eine Interferenz-Liste B (mit 16 Wörtern) einmalig vorgelesen. Dann wurden die Studienteilnehmer aufgefordert, Liste A erneut ohne Hinweise wiederzugeben (kurzfristig verzögerter Abruf) und schließlich mit Hinweisen (unterstützter kurzfristig verzögerter Abruf). Es folgte eine Reihe von anderen non-verbale Tests für 20 Minuten (z. B. Grooved Pegboard, Digit Span). Weiterhin wurde erneut nach Liste A gefragt (langfristig verzögerter Abruf, langfristig unterstützter verzögerter Abruf sowie ja/nein-Wiedererkennen-Diskriminierbarkeit). Weitere 10 Minuten später folgte die letzte Aufgabe: eine Aufgabe mit erzwungener Entscheidung (nach Liste-A-Wörtern) von verschiedenen Items.</p> <p>Zur Baseline wurde die Standard-Version des CVLT-II genutzt und in den Verlaufsmessungen die alternative Version (in der alternativen Version werden andere Wörter genutzt, um einen Langzeiterinnerungseffekt bei der Testwiederholung zu vermeiden). Der Gesamtscore für den sofortigen Abruf der ersten fünf Durchgänge wurde in T-Werte transformiert. Die Werte für die Lernkurve, kurzfristig und langfristig verzögerten bzw. unterstützten verzögerten Abruf und „total recognition discrimination“ wurden in Z-Werte mittels Computersoftware transformiert.</p> <p>Folgend dargestellt sind Details zum Testverfahren, die anderen Quellen als den Unterlagen des pU entnommen wurden [61]. Der CVLT-II wurde an 1.087 Personen im Alter zwischen 16 und 89 Jahren normiert.</p> <p>Erhebungszeitpunkte</p> <p>Der CVLT-II wurde in Studie HPN-100-007 im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und 12 Monate erhoben. Bei der Studie HPN-100-011, die eine Sicherheitserweiterung u. a. der Studie HPN-100-007 darstellt, wurde der Test im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und dann alle 12 Monate (bis zum Studienende) erhoben.</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Es ist unklar, ob alle Studienzentren sich bei der Testdurchführung an ein Testmanual gehalten haben. Die Informationen zu dem Test sind dem Studienbericht der Studie HPN-100-011 entnommen. Dem Studienbericht der Studie HPN-100-007 sind keine ausführlicheren Informationen zu dem Test zu entnehmen.</p> <p>Validität und Patientenrelevanz</p> <p>Der CVLT-II ist ein standardisierter Test, der episodische verbale Lern- und Gedächtnisfähigkeiten erfasst, die eine wichtige Hirnfunktion repräsentieren. Hinsichtlich der psychometrischen Eigenschaften wurden Reliabilität und Validität untersucht. Die Reliabilitätskoeffizienten der Test-Halbierungs-Reliabilität rangieren zwischen 0,68 und 0,94 (je nach betrachtetem Testwert) [61]. Der CVLT-II hing eng mit der Wechsler Abbreviated Scale of Intelligence (WASI) zusammen. Der CVLT-II wurde an diversen klinischen Populationen untersucht u. a. Populationen mit Gehirnerkrankungen [37].</p>

Endpunktkategorie Morbidität	
	<p>Insgesamt scheinen die psychometrischen Eigenschaften des CVLT-II hinsichtlich Reliabilität und Validität gegeben zu sein. Angaben zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität konnten den Unterlagen des pU nicht entnommen werden.</p> <p>Aufgrund der möglichen Einschränkungen bezüglich der Durchführungsobjektivität sowie den fehlenden Informationen zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität werden die Ergebnisse nicht dargestellt.</p>
Endpunkt	Neuropsychologische Tests: Digit Span
Operationalisierung	<p>Der Test Digit Span wurde im Rahmen einer neuropsychologischen Testbatterie bei erwachsenen Studienteilnehmern in den Studien HPN-100-007 und HPN-100-011 erhoben. In dem Studienbericht der Studie HPN-100-011 sind Details zum Testverfahren angegeben [60]. Der Digit Span ist ein Untertest der Wechsler Memory Scale – Third Addition (WMS-III) und wurde dem Test entnommen. Der Digit Span erfasst Aufmerksamkeit und Kurzzeitgedächtnisfähigkeit. Der Digit Span erfasst mit der Kurzzeitgedächtnisfähigkeit ein eng verwandtes Konstrukt zum Testwert „sofortiger Abruf“ des CVLT-II (siehe oben).</p> <p>Bei der Testdurchführung liest der Studienleiter den Testpersonen eine Reihe von Zahlen vor, die die Testpersonen anschließend in einer spezifischen Reihenfolge wiedergeben sollen (vorwärts oder rückwärts). Der Test beginnt mit zwei Zahlen. Alle zwei Durchgänge erhöht sich die Anzahl der Zahlen bis maximal neun Zahlen in der Vorwärts-Version und acht Zahlen in der Rückwärts-Version. Die Anzahl der korrekten Antworten wird zusammengezählt (maximal 30). Die Rohwerte werden dann anhand der Normen des Testmanuals in standardisierte Werte mit einem Mittelwert von 10 und einer Standardabweichung von 3 transformiert.</p> <p>Folgend dargestellt sind Details zum Testverfahren, die anderen Quellen als den Unterlagen des pU entnommen wurden [57]. Der Digit Span erfasst nicht nur die Kurzzeitgedächtnisfähigkeit (Digit span vorwärts), sondern auch die Arbeitsgedächtnisfähigkeit (Digit span rückwärts), da beim Rückwärts-Wiedergeben eine mentale Operation (Umsortieren der Zahlen) aus den Zahlen aus dem Kurzzeitgedächtnispeicher (phonologische Schleife) durchgeführt werden muss. Die aktuellste Version des Wechsler Intelligenztests (wurde in den Studien der vorliegenden Nutzenbewertung nicht verwendet, sondern mit dem WASI eine Kurz-Version, die den Digit Span nicht enthielt) trennt die beiden Konzepte sogar und hat für Digit span vorwärts und rückwärts jeweils separate Skalen.</p> <p>Erhebungszeitpunkte</p> <p>Der Digit Span wurde in Studie HPN-100-007 im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und 12 Monate erhoben. Bei der Studie HPN-100-011, die eine Sicherheitserweiterung u. a. der Studie HPN-100-007 darstellt, wurde der Test im Rahmen der neuropsychologischen Testbatterie für Erwachsene zu den Zeitpunkten 0 (Baseline) und dann alle 12 Monate (bis zum Studienende) erhoben.</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Bei Intelligenztests (Digit Span ist ein Subtest eines Intelligenztests) kann die Durchführungsobjektivität nur eingehalten werden, wenn die Testerhebung anhand eines Manuals durchgeführt wird. Es ist unklar, ob alle Studienzentren sich bei der Testdurchführung an ein Testmanual gehalten haben (auf das Testmanual wurde nur für die Berechnung der Werte Bezug genommen). Die Informationen zu dem Test sind dem Studienbericht der Studie HPN-100-011 entnommen. Dem Studienbericht der Studie HPN-100-007 sind keine ausführlicheren Informationen zu dem Test zu entnehmen.</p> <p>Validität und Patientenrelevanz</p> <p>Der Digit Span ist ein standardisierter Test, der die Kurzzeitgedächtnisfähigkeit und Arbeitsgedächtnisfähigkeit erfasst.</p>

Endpunktkategorie Morbidität	
	<p>Der Digit Span ist ein Untertest der Wechsler Memory Scale – Third Addition (WMS-III). Die Wechsler Memory Scale gilt als ein reliables und valides Testverfahren [12]. Der Digit Span wurde als Testverfahren in verschiedenen klinischen Populationen mit Gehirnverletzungen angewendet und kann einen Hinweis auf kognitive Beeinträchtigungen liefern [40].</p> <p>Insgesamt können die psychometrischen Eigenschaften des Digit Span hinsichtlich Reliabilität und Validität nicht abschließend beurteilt werden, da keine separaten Angaben zum Digit Span vorliegen, sondern nur zur gesamten Wechsler Memory Scale. Angaben zum klinisch relevanten Unterschied (MID) und zur Änderungssensitivität konnten den Unterlagen des pU nicht entnommen werden.</p> <p>Der Digit Span wurde bei Populationen mit Gehirnverletzungen eingesetzt. In der vorliegenden Fragestellung könnte der Digit Span somit als ein Hinweis auf kognitive Beeinträchtigungen durch hyperammonämische Krisen herangezogen werden. Somit ist der Endpunkt als patientenrelevant einzuschätzen.</p> <p>Aufgrund der möglichen Einschränkungen bezüglich der Durchführungsobjektivität sowie den fehlenden Informationen zu Testgütekriterien wie der Änderungssensitivität werden die Ergebnisse nicht dargestellt.</p>

Tabelle 24: Ergänzende Beschreibung der Endpunkte der Kategorie Lebensqualität

Endpunktkategorie Lebensqualität	
Endpunkt	SF-36
Operationalisierung	<p>Zur Operationalisierung der Lebensqualität wurde die Version 2 des SF-36 (SF-36v2) eingesetzt. Der SF-36-Fragebogen beinhaltet acht Dimensionen: körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperlicher Schmerz, allgemeine Gesundheitswahrnehmung, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden sowie zwei Summenskalen (körperliche Summenskala (PCS) und mentale Summenskala (MCS)). Die PCS schließt die Domänen-Skalen für körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen und allgemeine Gesundheitswahrnehmung ein. In die MCS gehen die Domänen-Skalen für Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden ein. Es können Werte von 0 bis 100 angenommen werden, wobei ein höherer Wert einen besseren Gesundheitszustand widerspiegelt.</p> <p>Für die Kalkulation der Scores wurde die Quality Metric Health Outcomes Scoring™ Software 4.0 eingesetzt. Mit fehlenden Daten wurde wie folgt verfahren: ein Score wurde nur kalkuliert, wenn mindestens 50 % der Items beantwortet wurden. In diesem Fall wurden fehlende Items aus den Mittelwerten der vorhandenen Items kalkuliert. Andernfalls (also > 50 % fehlende Items) wurde kein Score berechnet und auch kein Summenscore gebildet.</p> <p>Erhebungszeitpunkte</p> <p>Studie HPN-100-006 (RCT) und HPN-100-007 (Sicherheitserweiterung der Studie HPN-100-006): Tag 1 in Studie HPN-100-006 (= Baseline für HPN-100-007), Monat 3, 6, 9, 12 (bei Kindern/Jugendlichen wurde der SF-15 angewendet); 51 Erwachsene, 6 Kinder 6–11 Jahre, 3 Kinder 12–17 Jahre.</p>
Bewertung	<p>Operationalisierung</p> <p>Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz</p> <p><i>Validität:</i></p> <p>Der SF-36v2 ist ein häufig genutzter, valider, generischer Fragebogen zur gesundheitsbezogenen Lebensqualität. Für den PCS und MCS beschreibt das Manual eine MID von 3 Punkten.</p>

Endpunktkategorie Lebensqualität	
	<i>Patientenrelevanz:</i> Die Erfassung der Lebensqualität ist entsprechend § 2 Satz 3 AM-NutzenV als patientenrelevanter Endpunkt einzuschätzen.
Endpunkt	SF-15
Operationalisierung	<p>Der SF-15 wurde in den Studien HPN-100-005 und HPN-100-005SE erhoben und liegt in 5 Versionen vor, jeweils als selbstberichteter bzw. Elternfragebogen für die Altersgruppen 5–7, 8–12 und 13–18 Jahre. Aus dem Studienbericht der Studie HPN-100-005SE geht hervor, dass die Items auf einer 5-Punkte-Likertskala bewertet (0=nie, 4=fast immer) bzw. auf einer 3-Punkte-Skala für Kinderselbstberichte bis 7 Jahre bewertet (0=überhaupt nicht, 3=oft), invertiert und linear auf eine 100-Punkte-Skala transformiert wurden (0=100, 1=75, 2=50, 3=25, 4=0). Dimensionenscores wurden aus den Mittelwerten der Antworten jeweils für psychosoziale Scores sowie physikalischer Score kalkuliert. Zusätzlich wurde ein Gesamtscore kalkuliert. Bei mehr als 50 % Missings wurden keine Scores berechnet. Der SF-15 wurde separat für Kinder (selbstberichtet) und Eltern ausgewertet. Dabei wurden die Altersgruppen 6–11 Jahre (N=11) und 12–17 Jahre (N=6) zusammengefasst.</p> <p>Der SF-15 wurde von ausgebildetem medizinischen Fachpersonal gemäß Testmanual abgenommen.</p> <p>Erhebungszeitpunkte Studien HPN-100-005 und HPN-100-005SE (Kinder 6–17 Jahre): Tag 1 (= Baseline für SE), Monat 3, 6, 9, 12.</p>
Bewertung	<p>Operationalisierung Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz Der SF-15 ist eine generische Kurzform des PedsQL-Instruments [56]. Es enthält für die Domäne physikalische Funktion 5 Items, emotionale Funktion 4 Items, soziale Funktion 3 Items und Schulfunktion 3 Items. Höhere Scores bedeuten bessere Lebensqualität. Während der PedsQL in seiner generischen Form (mit 23 Items) in einer Reihe von Studien validiert wurde, liegt für die Kurzversion mit 15 Items keine entsprechende Studie für den englischen Sprachraum vor. In einer japanischen Studie zeigte sich eine ausreichende Reliabilität und Validität [11]. Informationen zur MID und zur Änderungssensitivität des SF-15 liegen nicht vor.</p> <p>Aufgrund der fehlenden Informationen zur MID und zur Änderungssensitivität werden die Ergebnisse nicht dargestellt.</p>

Ergebnisse zum Zusatznutzen der ergänzend dargestellten Endpunkte

Lebensqualität: SF-36

Tabelle 25: Ergebnisse zum Endpunkt SF-36 in der Studie HPN-100-007 für erwachsene Studienteilnehmer – Safety-Population

HPN-100-007 ¹⁾ Safety-Population SF-36	Baseline		Monat 12		Veränderung zu Baseline	
	N ²⁾	MW (SD)	N ²⁾	MW (SD)	N ²⁾	MW (SD)
Körperliche Summenskala (PCS)	49	48,8 (10,2)	45	49,3 (10,6)	45	-0,0 (7,5)
Mentale Summenskala (MCS)	49	50,4 (11,1)	45	47,5 (12,3)	45	-2,4 (9,7)

¹⁾ Altersgruppe Erwachsene

²⁾ Anzahl Patienten in der Auswertung. Rücklaufquote ≥ 88 %.

Abkürzungen: MW: Mittelwert; SD: Standardabweichung; SF-36: 36-Item Short Form Health Survey.