

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: Vosoritid

Datum der Veröffentlichung: 3. Januar 2022

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abbildungsverzeichnis.....	4
Abkürzungsverzeichnis	5
Hintergrund	6
1 Fragestellung	7
2 Auswahl und Methodik der relevanten Studien.....	8
2.1 Studienbasis für die Nutzenbewertung.....	8
2.2 Beschreibung der eingeschlossenen Studie.....	9
2.3 Endpunkte.....	16
2.3.1 Mortalität.....	16
2.3.2 Morbidität	17
2.3.3 Lebensqualität.....	22
2.3.4 Sicherheit	29
2.3.5 Übersicht der Erhebungszeitpunkte.....	32
2.4 Statistische Methoden.....	32
2.5 Verzerrungspotential auf Studien- und Endpunktebene.....	34
3 Ergebnisse der eingeschlossenen Studie	35
3.1 Studiencharakteristika und Studienmedikation.....	35
3.2 Mortalität	39
3.3 Morbidität	39
3.4 Lebensqualität	41
3.5 Sicherheit.....	42
4 Diskussion der Methodik und Ergebnisse	46
4.1 Zulassungsstatus und Zulassungspopulation von Vosoritid	46
4.2 Design und Methodik der Studie.....	46
4.3 Mortalität	48
4.4 Morbidität	48
4.5 Lebensqualität	48
4.6 Sicherheit.....	50
5 Anforderung an eine qualitätsgesicherte Anwendung.....	51
6 Zusammenfassung der Nutzenbewertung	52
Referenzen.....	54
Anhang.....	56

Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis	8
Tabelle 2:	Charakterisierung der Studie 301	9
Tabelle 3:	Relevante Protokolländerungen in Studie 301	13
Tabelle 4:	Charakterisierung der Interventionen in Studie 301	14
Tabelle 5:	Zusammenfassung der Endpunktbewertung der Studie 301	16
Tabelle 6:	Domänen der jeweiligen Versionen der CBCL	21
Tabelle 7:	Erhebungszeitpunkte der berücksichtigten Endpunkte in Studie 301	32
Tabelle 8:	Verzerrungspotential der Studie 301	34
Tabelle 9:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studie 301	34
Tabelle 10:	Allgemeine Angaben Studie 301	35
Tabelle 11:	Charakterisierung der Studienpopulation in Studie 301	36
Tabelle 12:	Angaben zur Exposition mit der Studienmedikation in Studie 301	38
Tabelle 13:	Begleitmedikation > 25 % in einem Studienarm in Studie 301	38
Tabelle 14:	Protokollverletzungen in Studie 301	39
Tabelle 15:	Veränderung der Stehhöhe (z-Score) zu Studienwoche 52 in Studie 301, FAS	40
Tabelle 16:	Veränderung im PedsQL (selbstberichtet) zu Studienwoche 52 in Studie 301, FAS	41
Tabelle 17:	Veränderung im QoLISSY (selbstberichtet) zu Studienwoche 52 in Studie 301, FAS	42
Tabelle 18:	Ergebnisse der Nebenwirkungen – Zusammenfassung der UE bis Studienende in Studie 301, Sicherheitspopulation	43
Tabelle 19:	UE mit Inzidenz ≥ 10 % nach MedDRA-Systemorganklassen und Preferred Terms in einem Studienarm der Studie 301, Sicherheitspopulation	43
Tabelle 20:	UE von besonderem Interesse in Studie 301, Sicherheitspopulation	45
Tabelle 21:	Zusammenfassende Darstellung der Ergebnisse der Studie 301	52
Tabelle 22:	Veränderung im PedsQL (fremdberichtet) zu Studienwoche 52 in Studie 301, FAS	56
Tabelle 23:	Veränderung im QoLISSY (fremdberichtet) zu Studienwoche 52 in Studie 301, FAS	56

Abbildungsverzeichnis

Abbildung 1: Studienfluss in Studie BMN 111-901 [4]..... 15

Abkürzungsverzeichnis

ACH	Achondroplasie
AGV	Jährliche Wachstumsgeschwindigkeit
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse
CBCL	Child Behavior Checklist
CTCAE	Common Terminology Criteria for Adverse Events
EPAR	European Public Assessment Report
FAS	Full Analysis Set
G-BA	Gemeinsamer Bundesausschuss
GHD	Wachstumshormonmangel
GKV	Gesetzliche Krankenversicherung
HLT	High Level Term
ICC	Intraklassenkorrelationskoeffizient
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISS	Idiopathischer Kleinwuchs
KI	Konfidenzintervall
LS	Least Square
MANCOVA	multivariate Kovarianzanalyse
NEC	Not Elsewhere Classified
NIAID/FAAN	National Institute of Allergy and Infectious Disease and Food Allergy and Anaphylaxis Network
PedsQL	Pediatric Quality of Life Inventory
PT	Preferred Term
pU	pharmazeutischer Unternehmer
QoLISSY	Quality of Life in Short Stature Youth Questionnaire
RCT	Randomisierte kontrollierte Studie (randomized controlled trial)
SAP	Statistischer Analyseplan
SD	Standardabweichung
SGB	Sozialgesetzbuch
SMQ	Standard MedDRA Query
SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA
WeeFIM	Functional Independence Measure for Children

Hintergrund

Vosoritid 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 Vosoritid zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (pU) heran und bewertet die darin enthaltenen Daten nach Maßgabe der 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 Vosoritid in seiner Sitzung am 21. Dezember 2021 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 1. Oktober 2021 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung (GKV)
- Anzahl der GKV-Patientinnen und -Patienten in der Zielpopulation

Die Nutzenbewertung wird am 3. Januar 2022 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Vosoritid (Voxzogo®) 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 [7]:

Achondroplasie (ACH) bei Patientinnen und Patienten ab dem 2. Lebensjahr, bei denen die Epiphysen noch nicht geschlossen sind. Die Diagnose Achondroplasie sollte durch entsprechende Gentests bestätigt werden.

Gemäß Fachinformation beträgt die übliche Dosierung 15 µg/kg Körpergewicht. Aus praktischen Gründen und zur Berücksichtigung gewichtsbedingter Veränderungen der Pharmakokinetik beinhaltet die Fachinformation zusätzlich eine Übersicht von Volumina der Einzeldosis in Abhängigkeit vom Körpergewicht [7].

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Studien zum Wirkstoff				
BMN 111-301 ¹⁾	Ja	Ja	Ja	-
BMN 111-302 ¹⁾	Ja	Ja	Nein	Extensionsstudie zu BMN 111-301. Die mediane Behandlungsdauer betrug 481 Tage in der Placebo-Vosoritid-Gruppe und 841 in der Vosoritid-Vosoritid-Gruppe. Mit BMN 111-301 liegt Evidenz mit einer Kontroll-Gruppe über einen Zeitraum von 52 Wochen vor.
BMN 111-202 ¹⁾	Ja	Ja	Nein	Unkontrollierte Phase-II-Dosiseskalierungsstudie mit einer Studiendauer von 6 Monaten (mit einer optionalen Verlängerung um 18 Monate). Eingeschlossen wurden 35 Kinder und Jugendliche im Alter von 5 bis < 15 Jahren. 10 Personen erhielten zu Studienbeginn die lt. FI übliche Dosis von 15 µg/kg Vosoritid. In der Verlängerungsphase wurde bei zusätzlich 12 Personen die Dosis auf 15 µg/kg gesteigert. Die mediane Expositionszeit für 22 Personen mit 15 µg/kg Vosoritid liegt bei 504 Tagen. Mit BMN 111-301 liegt Evidenz mit einer Kontroll-Gruppe über einen Zeitraum von 52 Wochen vor.
BMN 111-205 ¹⁾	Ja	Ja	Nein	Extensionsstudie zu BMN 111-202. Die mediane Behandlungsdauer betrug max. 60 Monate. Ergebnisse liegen für Gruppen entsprechend der Startdosis mit 10 Personen für 15 µg/kg Vosoritid vor.
BMN 111-206 ¹⁾	Ja	Ja	Nein	RCT, in dem Säuglinge und Kinder bis 5 Jahre 15 µg/kg Vosoritid bzw. Placebo erhalten. Für die randomisierte Population liegt kein Datenschnitt mit getrennter Darstellung der Studienarme vor. Ausgewertet wurden 4 Sentinel-Patientinnen und -Patienten ²⁾ . Aufgrund der kleinen Stichprobengröße wird auf eine Darstellung verzichtet.
BMN 111-208 ¹⁾	Ja	Nein	Nein	Extensionsstudie zu BMN 111-206.

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
BMN 111-209	Nein	Nein	Nein	RCT, in dem Säuglinge (0 bis ≤ 12 Monate) 30 µg/kg Vosoritid + Standardtherapie bzw. Standardtherapie erhalten. Die Studie befindet sich in der Rekrutierungsphase. Es liegt kein Datenschnitt vor.
Weitere Studien				
BMN 111-901	Ja	Nein	Nein	Diese nicht-interventionelle Beobachtungsstudie diene als vorausgehende Beobachtungsphase für Studie BMN 111-301.

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

²⁾ Laut Studienprotokoll erhielten diese 4 Kinder Vosoritid und wurden vor Randomisierungsbeginn der übrigen Studienteilnehmenden hinsichtlich Kurzzeitsicherheit und Pharmakokinetik untersucht.

Abkürzungen: EPAR: European Public Assessment Report; FI: Fachinformation; pU: pharmazeutischer Unternehmer; RCT: Randomisierte kontrollierte Studie.

Zur Nutzenbewertung für Vosoritid herangezogene Studien und Daten

- Herstellerdossier zu Vosoritid
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [14]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie BMN 111-301 [5,6]
- Studienbericht inkl. Studienprotokoll und SAP der Studie BMN 111-901 [4]
- Publikation zur Studie BMN 111-301 [11]

2.2 Beschreibung der eingeschlossenen Studie

Die Unterlagen zur Nutzenbewertung für Vosoritid basieren auf der Zulassungsstudie BMN 111-301, im Folgenden als Studie 301 bezeichnet. Studie und Intervention werden in Tabelle 2–4 charakterisiert.

Tabelle 2: Charakterisierung der Studie 301

Charakteristikum	Beschreibung
Design und Studienablauf	<p>Bei Studie 301 handelt es sich um eine doppelblinde, kontrollierte, multizentrische Studie der Phase 3, in der Kinder und Jugendliche zwischen 5 bis 17 Jahren und mittels genetischer Testung nachgewiesener ACH randomisiert im Verhältnis 1:1 auf die Behandlungsarme 15 µg/kg Vosoritid und Placebo zugeteilt wurden. Die Randomisierung erfolgt stratifiziert für die Merkmale „Geschlecht“ (männlich oder weiblich) und „Pubertäre Entwicklung“ (Tanner-Stadium I oder Stadium > I)¹⁾. Auf die Stratifizierungsgruppen nach Geschlecht sollten jeweils annähernd 50 % ohne Überschreitung von 55 % in einer der Gruppen entfallen. Hinsichtlich der Tanner-Stadien war der Anteil mit Tanner-Stadium > I auf 20 % begrenzt²⁾.</p> <p>In die Studie wurden Kinder und Jugendliche eingeschlossen, die zuvor mindestens 6 Monate an der Beobachtungsstudie BMN 111-901 (siehe unten) teilnahmen.</p>

Charakteristikum	Beschreibung
	<p>Studie 301 teilt sich in die folgenden Phasen:</p> <ul style="list-style-type: none"> • Screening-Phase (Tag -30 bis Tag -1): Überprüfung der Ein- und Ausschlusskriterien. • Randomisierte Behandlungsphase (Tag 1 bis Woche 52): Verabreichung von Vosoritid 15 µg/kg oder Placebo als tägliche subkutane Injektion. • Nachbeobachtungsphase (Woche 53 bis Woche 56), sofern die Studienteilnehmenden nicht in die Extensionsstudie BMN 111-302 übergangen bzw. eine Teilnahme erst nach mehr als 4 Wochen erfolgte (nach der letzten Dosis des Studienmedikaments)³⁾: Nachbeobachtung der Sicherheit. <p>Die Studie ist abgeschlossen. Es liegt der finale Studienbericht vom 21.04.2020 vor.</p>
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Alter von 5 bis <18 Jahre. • ACH-Diagnose, klinisch belegt und durch Gentest bestätigt. • Vorliegen von Wachstumsdaten über mindestens 6 Monate und derzeit aktive Teilnahme an Studie BMN 111-901. • In der Lage zu gehen und ohne Hilfe zu stehen. • Betreuungspersonen sind bereit, täglich Injektionen zu verabreichen und das erforderliche Training zu absolvieren. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Hypochondroplasie oder eine andere Erkrankung der Kleinwüchsigkeit als ACH. • Eine der folgenden Erkrankungen: <ul style="list-style-type: none"> ○ Hypothyreose oder Hyperthyreose ○ Insulinpflichtiger Diabetes mellitus ○ Entzündliche Autoimmunerkrankung ○ Entzündliche Darmerkrankung ○ Autonome Neuropathie • Vorgeschichte mit einer der folgenden Erkrankungen: <ul style="list-style-type: none"> ○ Niereninsuffizienz, definiert als Serumkreatinin > 2 mg/dl ○ Chronische Anämie ○ Systolischer Blutdruck < 70 mmHg zu Baseline oder wiederkehrende symptomatische Hypotonie (definiert als Episoden mit niedrigem Blutdruck und Symptomen wie Schwindel, Ohnmacht) oder wiederholte symptomatische orthostatische Hypotonie ○ Herz- oder Gefäßerkrankungen, einschließlich der folgenden: <ul style="list-style-type: none"> ▪ Kardiale Funktionsstörung (abnormales Echokardiogramm einschließlich abnormal erhöhte linksventrikuläre Muskelmasse) bei Screening-Visite ▪ Hypertrophe Kardiomyopathie ▪ Pulmonale Hypertonie ▪ Kongenitale Herzerkrankung ▪ Zerebrovaskuläre Erkrankung ▪ Aorteninsuffizienz oder andere klinisch signifikante Herzklappenfehlfunktion ▪ Klinisch signifikante atriale oder ventrikuläre Arrhythmien • Klinisch signifikanter Befund oder eine Arrhythmie auf dem Screening-EKG, die auf abnorme kardiale Funktion, Reizleitung oder Fridericias-korrigiertes QT-Intervall > 450 ms hindeuteten. • Instabiler Zustand, der wahrscheinlich einen chirurgischen Eingriff während der Studie erfordert (einschließlich fortschreitender zervikaler medullärer Kompression oder schwerer unbehandelter Schlafapnoe). • Nachweis einer verminderten Wachstumsgeschwindigkeit (AGV < 1,5 cm/Jahr), beurteilt über einen Zeitraum von mindestens 6 Monaten oder des Schließens der Epiphysen (proximale Tibia, distaler Femur) durch bilaterale Röntgenaufnahmen der unteren Extremitäten, einschließlich anterior-posteriorer und lateraler Ansichten. • Dokumentierter Vitamin-D-Mangel.

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Aktuelle, dauerhafte Therapie mit blutdrucksenkenden Medikamenten, ACE-Hemmern, Angiotensin-II-Rezeptorenblockern, Diuretika, Betablockern, Kalziumkanalblockern, Herzglykoside, systemische Anticholinergika, GnRH-Agonisten, Medikamente, die eine kompensatorische Tachykardie beeinflussen können, oder andere Medikamente, die bekanntermaßen die Nieren- oder Tubulusfunktion verändern können. • Behandlung mit Wachstumshormon, insulinähnlichem Wachstumsfaktor 1 oder anabolen Steroiden in den letzten 6 Monaten oder Behandlung von mehr als 6 Monaten zu irgendeinem Zeitpunkt. • Mehr als einen Monat Behandlung mit oralen Kortikosteroiden (niedrig dosiertes, inhalatives Steroid für Asthma oder intranasale Steroide sind akzeptabel) in den vorangegangenen 12 Monaten. • Geplante oder erwartete gliedmaßenverlängernde Operationen während des Studienzeitraums. Personen mit vorangegangener gliedmaßenverlängernder Operation konnten in die Studie aufgenommen werden, wenn die Operation mindestens 18 Monate vor dem Screening stattgefunden hat und die Heilung vollständig und ohne Folgeerscheinungen war. • Geplante oder erwartete knochenchirurgische Eingriffe (d. h. Eingriffe mit Disruption der Knochenrinde, ausgenommen Zahnextraktion) während des Studienzeitraums. Personen mit vorherigen chirurgischen Eingriffen am Knochen konnten an der Studie teilnehmen, wenn der Eingriff mindestens 6 Monate vor dem Screening stattfand und die Heilung ohne Folgeerscheinungen abgeschlossen war. • Fraktur der Röhrenknochen oder Wirbelsäule innerhalb 6 Monate vor dem Screening. • AST oder ALT mindestens 3-fach ULN oder Gesamtbilirubin mindestens 2-fach ULN (außer bei bekannter Anamnese der Gilbert-Krankheit). • Vorgeschichte einer schweren, unbehandelten Schlafapnoe. • Neueinleitung einer Schlafapnoe-Behandlung in den letzten 2 Monaten vor dem Screening. • Vorgeschichte einer Hüftoperation oder Hüftdysplasie, die für ACH-Erkrankte untypisch ist. • Klinisch signifikante Hüftverletzung in den letzten 30 Tagen vor dem Screening. • Vorgeschichte einer verrutschten Oberschenkel epiphyse oder avaskulärer Nekrose des Femurkopfes. • Nicht in der Lage in Bauchlage flach zu liegen (für die Hüftuntersuchung erforderlich). • Abnorme Befunde zu Baseline bei der ärztlichen Untersuchung der Hüfte oder bei bildgebenden Untersuchungen, die als klinisch signifikant eingestuft wurden
Intervention und Zahl der Patientinnen und Patienten	<p>Eine Angabe zur Anzahl der gescreenten Patientinnen und Patienten wurde für Studie BMN 111-901 nicht identifiziert. Für Studie 301 wurden 124⁴⁾ Personen gescreent, von denen 121 auf die folgenden Studienarme randomisiert zugeteilt wurden:</p> <ul style="list-style-type: none"> • Placebo: n = 61 • Vosoritid 15 µg/kg: n = 60
Ort und Zeitraum der Durchführung	<p>Ort der Durchführung Die Studie wurde an 24 Studienzentren in 7 Ländern (USA, Australien, Deutschland, Spanien, Großbritannien, Japan, Türkei) durchgeführt.</p> <p>Zeitraum der Studie</p> <ul style="list-style-type: none"> • Erste Visite erste Testperson: 12.12.2016 • Erste Visite letzte Testperson: keine Angabe • Letzte Visite letzte Testperson: 30.10.2019

Charakteristikum	Beschreibung
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt Wachstumsgeschwindigkeit (AGV)</p> <p>Sekundärer Endpunkt</p> <ul style="list-style-type: none"> • Körpergröße (z-Score) • Verhältnis von oberem zu unterem Körpersegment • Verhältnis der Körperproportionen der Extremitäten • Körpergewicht • PedsQL • QoLISSY • WeeFIM <p>Sicherheit</p> <ul style="list-style-type: none"> • (Schwerwiegende) Unerwünschte Ereignisse • Klinisch signifikante Veränderungen der Vitalzeichen • Körperliche Untersuchungen • Laborparameter • Bildgebende Untersuchungen der Knochenmorphologie • CBCL • EKG • Klinische Beurteilung der Hüfte
Subgruppenanalysen	<p>Laut SAP waren folgende Subgruppenmerkmale für die Wirksamkeitsanalysen definiert:</p> <ul style="list-style-type: none"> • Geschlecht (männlich, weiblich) • Altersgruppe bei Studienbeginn (≥ 5 bis < 8, ≥ 8 bis < 11, ≥ 11 bis < 15, ≥ 15 bis < 18 Jahre) • Tanner-Stadium bei Studienbeginn (I, $> I$) • Randomisierungsstrata (männlich Tanner-Stadium I, weiblich Tanner-Stadium I, männlich Tanner-Stadium $> I$, weiblich Tanner-Stadium $> I$) • z-Score der Körpergröße zu Baseline (≤ -6, > -6 bis ≤ -5, > -5 bis ≤ -4, > -4) • AGV zu Baseline ($\leq 3,5$, $> 3,5$ bis $\leq 4,5$, $> 4,5$ cm/Jahr) <p>Zusätzlich stellt der pU gemäß Modulanforderungen die folgenden Post-hoc-Subgruppenanalysen dar:</p> <ul style="list-style-type: none"> • Ethnie (weiß, nicht-weiß) • Region (Nordamerika, Europa, Japan, Rest der Welt)

¹⁾ Im Originalprotokoll war eine Stratifizierung nach Geschlecht und Alter (< 11 oder ≥ 11 Jahre) festgelegt. Dies wurde mit Protokoll-Amendment 2 vom 27.04.2017 geändert.

²⁾ Da zum Zeitpunkt des Studienbeginns in Japan nur in begrenztem Umfang Erfahrungen mit Vosoritid bestanden, wurde die Aufnahme der Testpersonen zunächst gestaffelt, mit einem Zeitfenster von mindestens 2 Wochen zwischen den ersten 4 aufgenommenen Testpersonen. Um eine gleichmäßige Verteilung der japanischen Testpersonen zu gewährleisten, die Placebo und Vosoritid erhalten, wurde für die geringe Anzahl der geplanten Testpersonen (N = 7) eine separate stratifizierte Randomisierung durchgeführt. In diesem Randomisierungsschema wurde keine Obergrenze für das Tanner-Stadium > 1 angewendet.

³⁾ Bis einschließlich Protokoll-Amendment 3 vom 05.01.2018 umfasste die Sicherheitsbeobachtung die folgenden 2 Wochen nach der letzten Dosis. Dies wurde mit Protokoll-Amendment 4 vom 01.02.2019 geändert.

⁴⁾ Angabe aus Modul 4.

Abkürzungen: ACE: Angiotensin-Converting Enzym; ACH: Achondroplasie; AGV: Jährliche Wachstumsgeschwindigkeit; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CBCL: Child Behavior Checklist; GnRH: Gonadotropin-Releasing-Hormon; mmHg: Millimeter Quecksilbersäule; PedsQL: Pediatric Quality of Life Inventory; pU: pharmazeutischer Unternehmer; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; SAP: Statistischer Analyseplan; ULN: Obere Grenze des Normbereichs; WeeFIM: Functional Independence Measure for Children.

Protokolländerungen

Das erste Studienprotokoll ist auf den 01.09.2016 datiert. Nach Beginn der Studie am 24.11.2016 wurden insgesamt 4 Protokoll-Amendments erstellt. Wesentliche Änderungen in diesen Amendments sind in Tabelle 3 dargestellt.

Tabelle 3: Relevante Protokolländerungen in Studie 301

Amendment	Anzahl bis dahin eingeschlossener Personen	Wesentliche Änderungen	Kommentar
Amendment 1 (28.11.2016)	Keine Angabe.	Keine	-
Amendment 2 (27.04.2017)	Keine Angabe.	Die Obergrenze der Altersspanne in den Einschlusskriterien wurde von < 15 auf < 18 Jahre angehoben.	Die Altersgruppe der 15- bis 17-Jährigen ist von der FI als Zulassungspopulation umfasst.
		Das Stratifizierungsmerkmal „Alter“ wurde gegen das Merkmal „Tanner-Stadium“ getauscht.	Auswirkungen werden vor dem Hintergrund der Charakteristika der Studienpopulation in Kapitel 4.2 betrachtet.
		Es wurden Ereignisse von besonderem Interesse in der Sicherheitserhebung definiert.	Im SAP wurden später Ereignisse von besonderem Interesse in der Sicherheitserhebung definiert (siehe Kapitel 2.3.4).
Amendment 3 (05.01.2018)	Keine Angabe.	Das Ausschlusskriterium zum Verschluss der Wachstumsfuge wurde um eine reduzierte AGV in den letzten 6 Monaten erweitert.	Die Ergänzung einer reduzierten AGV sollte sicherstellen, dass für Personen, die sich dem Verschluss der Wachstumsfugen näherten, eine Studienteilnahme bereits nicht mehr möglich war. Ein Ausschluss von Personen mit reduzierter AGV findet sich ebenfalls in der FI [7].
		Die Erhebung der CBCL wurde eingeführt.	Aus dem Studienbericht ergibt sich, dass Baseline-Messungen nicht nachträglich erfassbar waren.
Amendment 4 (01.02.2019)	Keine Angabe.	Die finale Sicherheitsvisite wurde von 2 auf 4 Wochen nach der letzten Dosis verschoben.	Mit Einführung 2,5 Jahre nach Studienbeginn und Studienende am 30.10.2019 wird davon ausgegangen, dass für einen Großteil der Studienpopulation die Dauer der Nachbeobachtung 2 Wochen betrug.
		Eine zunächst nur in einem länderspezifischen Protokoll festgehaltene Rekrutierungsstrategie wurde in das globale Studienprotokoll aufgenommen: In Japan erfolgte die Rekrutierung gestaffelt, wobei zwischen den ersten 4 Personen ein Zeitfenster von mindestens 2 Wochen lag.	-

Abkürzungen: AGV: Jährliche Wachstumsgeschwindigkeit; CBCL: Child Behavior Checklist; FI: Fachinformation; SAP: Statistischer Analyseplan.

Charakterisierung der Intervention

Die in Studie 301 verabreichten Studieninterventionen sind in Tabelle 4 beschrieben.

Tabelle 4: Charakterisierung der Interventionen in Studie 301

Intervention	Kontrolle
<ul style="list-style-type: none"> • Vosoritid, geliefert in sterilen Glasfläschchen als Einzeldosis. • Als rekonstituierte Lösung farblos bis gelb mit 0,8 mg/ml oder 2 mg/ml Vosoritid. • Als einmalige Injektion von 15 µg/kg subkutan verabreicht und nach Möglichkeit jeden Tag etwa zur gleichen Zeit. • Dieselbe Injektionsstelle sollte nicht an 2 aufeinanderfolgenden Tagen verwendet werden. Es sollte zwischen den 4 Injektionsstellen (Oberschenkel, Oberarmrückseite, Bauch, Gesäß) rotiert werden. • Nach Verabreichung jeder Dosis erfolgte eine Beobachtung an den Tagen 1–3 mindestens 2 Stunden, an allen anderen Tagen 30 Minuten (bei klinischer Auffälligkeit auch länger). • Betreuungspersonen können Vosoritid zu Hause verabreichen, sobald dies vom Prüfarzt / von der Prüfarztin genehmigt und eine angemessene Schulung nachgewiesen wurde. 	<ul style="list-style-type: none"> • Placebo, geliefert in sterilen Glasfläschchen als Einzeldosis und mit dem Arzneimittel äußerlich vergleichbar. • Enthält alle Bestandteile des Arzneimittels mit Ausnahme des Wirkstoffs. • Als einmalige Injektion subkutan verabreicht und nach Möglichkeit jeden Tag etwa zur gleichen Zeit. • Dieselbe Injektionsstelle sollte nicht an 2 aufeinanderfolgenden Tagen verwendet werden. Es sollte zwischen den 4 Injektionsstellen (Oberschenkel, Oberarmrückseite, Bauch, Gesäß) rotiert werden. • Nach der Verabreichung jeder Dosis erfolgte eine Beobachtung an den Tagen 1–3 mindestens 2 Stunden, an allen anderen Tagen 30 Minuten (bei klinischer Auffälligkeit auch länger). • Betreuungspersonen können Placebo zu Hause verabreichen, sobald dies vom Prüfarzt / von der Prüfarztin genehmigt und eine angemessene Schulung nachgewiesen wurde.
<p>Nicht erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • Antihypertensive Medikamente • ACE-Hemmer • Angiotensin-II-Rezeptorblocker • Diuretika • Betablocker • Kalzium-Kanal-Blocker • Herzglykoside • Systemische Anticholinergika • GnRH-Agonisten • Medikamente, die eine kompensatorische Tachykardie beeinflussen können, oder andere Medikamente, die bekanntermaßen die Nieren- oder Tubulusfunktion verändern. • Alle Medikamente, die nach ärztlicher Einschätzung die Sicherheit oder Verträglichkeit der Teilnahme an der klinischen Studie beeinträchtigen können. <p>Erlaubte Begleitmedikation</p> <p>Es können alle Arzneimittel verschrieben werden, die nicht unter unerlaubter Begleitmedikation gelistet sind.</p>	

Abkürzungen: ACE: Angiotensin-Converting Enzym; GnRH: Gonadotropin-Releasing-Hormon.

Beobachtungsstudie BMN 111-901

Die in der Beobachtungsstudie BMN 111-901, im Folgenden als Studie 901 bezeichnet, gewonnenen Verlaufsdaten wurden zur Bestimmung der Wachstumsgeschwindigkeit als Baseline-Daten der Studienteilnehmenden verwendet. Die Studie startete am 20.04.2012. Im vorliegenden Studienbericht vom 12.12.2019 wird die Studie als laufend, die Studiendauer auf bis zu 7 Jahre gekennzeichnet.

Insgesamt nahmen 342 Testpersonen an der Studie teil. Teilnahmeberechtigt waren Säuglinge ab Geburt bis Jugendliche ≤ 17 Jahren mit genetisch bestätigter ACH-Diagnose. Die Ein- und Ausschlusskriterien decken sich (mit Ausnahme der Altersspanne) mit denen der Studie 301.

Von den 342 Testpersonen gingen 121 in Studie 301, 34 in Studie BMN 111-206 und 35 in Studie BMN 111-202 über. 152 Studienteilnehmende nahmen hingegen an keiner Arzneimittelstudie im weiteren Verlauf teil. Davon hatten 3 Personen (2,0 %) Studie 901 abgeschlossen, 71 (46,7 %) befanden sich noch in der Studie und 78 (51,3 %) hatten ihre Teilnahme abgebrochen. Die meisten Studienabbrüche waren auf den Widerruf der Einverständniserklärung zurückzuführen (N = 54, 69,2 % aller Studienabbrüche) (siehe Abbildung 1). Im Studienbericht wird angegeben, dass die meisten davon auf persönliche Gründe zurückzuführen waren (z. B. Zeitmangel, Entscheidung für Familie/Kinder, mangelndes Interesse, Angst vor dem Blutabnehmen).

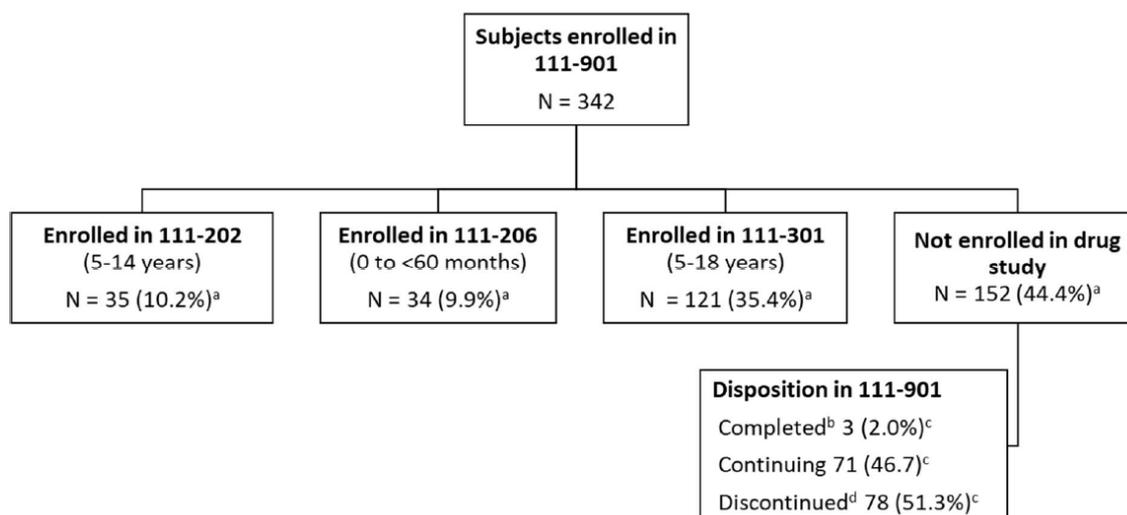


Abbildung 1: Studienfluss in Studie BMN 111-901 [4]

- ^{a)} Prozentsätze basierend auf Gesamtzahl aller in 111-901 eingeschriebenen Studienteilnehmenden im Nenner.
^{b)} Beendeten Studie BMN 111-901 ohne in eine Arzneimittelstudie aufgenommen zu werden.
^{c)} Prozentsätze basierend auf Anzahl der Studienteilnehmenden, die nicht in eine Arzneimittelstudie aufgenommen wurden, im Nenner.
^{d)} Aus Studie BMN 111-901 ausgeschieden und nicht in eine Arzneimittelstudie aufgenommen.

2.3 Endpunkte

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

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie 3

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Mortalität ¹⁾	Mortalität	Nein	Ja
Wachstumsgeschwindigkeit ²⁾ (AGV)	Morbidität	Ja	Nein ³⁾
Körpergröße		Ja	Ja
Verhältnis von oberem zu unterem Körpersegment		Ja	Nein ³⁾
Verhältnis der Körperproportionen		Ja	Nein ³⁾
Child Behavior Checklist ¹⁾		Nein	Ja ⁴⁾
PedsQL	Lebensqualität	Ja	Ja
QoLISSY		Ja	Ja
WeeFIM		Ja	Nein ⁵⁾
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

¹⁾ Im Rahmen der Sicherheit erhoben.

²⁾ Primärer Endpunkt.

³⁾ Aufgrund fehlender Patientenrelevanz.

⁴⁾ Ergebnisse aufgrund geringer Rücklaufquoten nicht dargestellt (siehe Kapitel 3.3).

⁵⁾ Die Anwendung des WeeFIM bei älteren Kindern und Jugendlichen wurde seitens des pU nicht plausibel begründet. Eine Auswertung für die Altersgruppe 5–7 Jahre liegt nicht vor.

Abkürzungen: AGV: Jährliche Wachstumsgeschwindigkeit; PedsQL: Pediatric Quality of Life Inventory; pU: pharmazeutischer Unternehmer; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; WeeFIM: Functional Independence Measure for Children.

2.3.1 Mortalität

Mortalität

Der Endpunkt „Mortalität“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Alle Todesfälle wurden innerhalb der Sicherheitserfassung erhoben. Sie wurden als Gesamtrate der in der Studie aufgetretenen Todesfälle dargestellt.

Erhebungszeitpunkte: Unerwünschte Ereignisse wurden während des Screenings, Tag 1, 2, 3 und 10, in Woche 6, 13, 26, 39 und 52 sowie Woche 54/56 oder in der Studienvsiste bei vorzeitigem Abbruch erhoben. Die finale Sicherheitsvisite wurde erst mit der Protokolländerung 3 vom 05.01.2018 von 2 auf 4 Wochen erhöht.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Die Erfassung der Mortalität ist patientenrelevant.

Validität

Der Endpunkt wird als valide bewertet.

2.3.2 Morbidität**Wachstumsgeschwindigkeit (AGV)**

Der Endpunkt „Wachstumsgeschwindigkeit“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt.

Operationalisierung*Beschreibung*

Für ein gegebenes Intervall [Datum1, Datum2] wurde die jährliche Wachstumsgeschwindigkeit (AGV) wie folgt bestimmt:

$$AGV \text{ (cm/Jahr)} = \frac{\text{Stehhöhe}_{\text{Datum 2}} - \text{Stehhöhe}_{\text{Datum 1}}}{\text{Datum 2} - \text{Datum 1}} \times 365,25$$

Die Erhebung der Stehhöhe ist im Endpunkt „Körpergröße“ (siehe unten) beschrieben.

Für Baseline wurde die AGV in der Run-in-Studie 901 (siehe oben) ermittelt, basierend auf den Messungen der Stehhöhe mindestens 6 Monate vor Eintritt in Studie 301 als Datum 1. Als Datum 2 wurde der Wert der Visite an Tag 1 der Studie 301 herangezogen. Für alle weiteren Erhebungszeitpunkte war die Stehhöhe zu Tag 1 der Studie 301 das Datum 1.

Die Veränderung in der AGV wurde für jede Person als Differenz zwischen AGV in Woche 52 und Baseline ermittelt, nach Behandlungsgruppe (Vosoritid oder Placebo) zusammengefasst und als Differenz der Mittelwerte zwischen den Behandlungsgruppen angegeben.

Erhebungszeitpunkte: Anthropometrische Messungen wurden während des Screenings, an Tag 1 und in Woche 13, 26, 39 und 52 durchgeführt, sowie im Falle eines Studienabbruchs, sofern dieser nach Woche 6 erfolgte oder die letzte Messung mehr als 2 Wochen zurück lag. Für Studie 901 wurden anthropometrische Messungen an Tag 1 durchgeführt. Die Messungen erfolgten dann regelmäßig alle 3 Monate (± 10 Tage) über bis zu 7 Jahre hinweg. Bei Studienabbruch zwischen 2 Messungen wurden, wenn möglich, auch noch Messungen zu einer Studienabschlussvisite erfasst.

Bewertung

Die Operationalisierung ist nachvollziehbar beschrieben.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Wachstumsgeschwindigkeit wird über die Änderung der Körpergröße pro Zeiteinheit erfasst. Die Rate des Wachstums ist jedoch kein Ausdruck der Symptomlinderung und bildet keine unmittelbar spürbare Symptomatik für die erkrankte Person ab. Die tatsächlich wahrnehmbare Änderung der Körpergröße ist in Studie 301 durch altersadjustierte Angaben der Stehhöhe erfasst. Wachstumsgeschwindigkeit bietet darüber hinaus keine relevanten Erkenntnisse, sodass diese in der Nutzenbewertung nicht dargestellt wird.

Validität

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

Körpergröße

Der Endpunkt „Körpergröße“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Körpergröße wurde als Stehhöhe erfasst. Zu jeder Studienvsiste wurde die Stehhöhe dreimal gemessen und der Mittelwert der Messungen für die Analysen verwendet. Im Falle des Vorliegens von nur 2 Werten wurde der Mittelwert aus 2 Messungen oder bei Einzelwerten auch dieser verwendet. Die Messungen wurden von geschultem Personal, vorzugsweise immer derselben Person durchgeführt. Die Messungen wurden möglichst zeitgleich zu jeder Visite (± 2 Stunden zu Erhebung des Baseline-Werts) durchgeführt und waren durch einen Leitfadens im Studienprotokoll standardisiert.

Für Stehhöhe wurden alters- und geschlechtsadjustierten z-Scores und Perzentilen berechnet. Die z-Scores spiegeln die Anzahl der Standardabweichungen (SD) eines jeden Werts von den nach Alter und Geschlecht standardisierten normalen Durchschnittswerten wider. Die Daten wurden als SD-Wert über oder unter der altersspezifischen Referenz (± 0) dargestellt.

Im SAP wurde festgelegt, dass als Vergleichsdaten der z-Scores eine Stichprobe gesunder Kinder in den USA mit den Erhebungszeiträumen 1963–1994 (Mädchen: N = 19.351, Jungen: N = 19.854) verwendet werden [10]. Um die Übertragbarkeit der Resultate auf den deutschen Versorgungskontext zu analysieren, wurden post hoc die z-Scores auch aus einer repräsentativen deutschen Referenzbevölkerung (N = 17.641) gebildet. Hierbei wurden Wachstumsdaten des Robert Koch-Instituts (Erhebungszeitraum Mai 2003 bis Mai 2006) [20] verwendet. Des Weiteren wurden Analysen des z-Scores referenziert auf eine europäische ACH-Kohorte (N = 466) durchgeführt [17].

Die Veränderung der z-Scores der Körpergröße wurde als Differenz der Mittelwerte zwischen den Behandlungsgruppen ausgewertet. Perzentile wurden nicht berechnet.

Erhebungszeitpunkte: Anthropometrische Messungen wurden während des Screenings, an Tag 1 und in Woche 13, 26, 39 und 52 durchgeführt, sowie im Falle eines Studienabbruchs, sofern dieser nach Woche 6 erfolgte oder die letzte Messung mehr als 2 Wochen zurück lag.

Bewertung

Die Operationalisierung ist nachvollziehbar beschrieben.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Anthropometrische Parameter können, insbesondere bei Kindern mit charakteristischen, krankheitsbedingten Wachstumsstörungen, als patientenrelevante Morbiditätsparameter eingeschätzt werden. Dabei werden Angaben, welche für Alter und Geschlecht adjustiert wurden, gegenüber absoluten Werten bevorzugt herangezogen. Der ebenfalls durchgeführte Vergleich auf Basis einer ACH-Kohorte bietet für die Nutzenbewertung hingegen keine weiteren Erkenntnisse, da die Krankheitslast bereits als Kontroll-Arm valide abgebildet ist.

Validität

Die a priori ermittelten z-Scores basieren auf einer Stichprobe gesunder Kinder aus den USA. Da zusätzlich Post-hoc-Analysen bezogen auf Wachstumsdaten einer repräsentativen deutschen Kohorte durchgeführt wurden, um auch die Übertragbarkeit auf die deutsche Versorgungssituation zu zeigen, werden diese Auswertungen ebenfalls herangezogen.

Vom pU wurde keine Evidenz zu einer Relevanzschwelle zu Gruppenunterschieden vorgelegt.

Verhältnis von oberem zu unterem Körpersegment

Der Endpunkt „Verhältnis von oberem zu unterem Körpersegment“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt.

Operationalisierung

Beschreibung

Das Verhältnis von oberem zu unterem Körpersegment wurde wie folgt berechnet:

$$\text{Verhältnis von oberem zu unterem Körpersegment} = \frac{\text{Sitzhöhe (cm)}}{\text{Stehhöhe (cm)} - \text{Sitzhöhe (cm)}}$$

Die Erhebung der anthropometrischen Messungen erfolgte analog der Beschreibungen im Endpunkt „Körpergröße“ (siehe oben). Der Endpunkt wurde als Gruppenunterschied zwischen Vosoritid und Placebo in den Veränderungen zu Baseline ausgewertet.

Erhebungszeitpunkte: Anthropometrische Messungen wurden während des Screenings, an Tag 1 und in Woche 13, 26, 39 und 52 durchgeführt, sowie im Falle eines Studienabbruchs, sofern dieser nach Woche 6 erfolgte oder die letzte Messung mehr als 2 Wochen zurück lag.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Aus dem Endpunkt „Verhältnis von oberem zu unterem Körpersegment“ kann keine unmittelbare Patientenrelevanz abgeleitet werden, da damit keine direkt für die erkrankte Person spürbare Symptomatik ausgedrückt wird. Zwar sind Einschränkungen in der Mobilität und Funktion durch ein disproportionalen Verhältnis der Körpersegmente denkbar, jedoch erlaubt der Endpunkt darauf keine direkten Rückschlüsse, da keine Informationen zum konkreten Zusammenhang vorgelegt wurden. Der Endpunkt wird daher für die Nutzenbewertung nicht berücksichtigt.

Validität

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

Verhältnis der Körperproportionen

Der Endpunkt „Verhältnis der Körperproportionen“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt.

Operationalisierung

Beschreibung

Der Endpunkt Verhältnis der Körperproportionen wurde in 4 Formen untersucht:

- Verhältnis von Oberarmlänge zu Unterarmlänge.
- Verhältnis von Oberschenkellänge zu Knie- bis Fersenzlänge.
- Verhältnis von Oberschenkellänge zu Schienbeinlänge.
- Verhältnis von Armspanne zu Stehhöhe.

Die Erhebung der anthropometrischen Messungen erfolgte analog der Beschreibungen im Endpunkt „Körpergröße“ (siehe oben). Der Endpunkt wurde als Gruppenunterschied zwischen Vosoritid und Placebo in den Veränderungen zu Baseline ausgewertet.

Erhebungszeitpunkte: Anthropometrische Messungen wurden während des Screenings, an Tag 1 und in Woche 13, 26, 39 und 52 durchgeführt, sowie im Falle eines Studienabbruchs, sofern dieser nach Woche 6 erfolgte oder die letzte Messung mehr als 2 Wochen zurück lag.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Analog zum Endpunkt „Verhältnis von oberem zu unterem Körpersegment“ kann aus den untersuchten Körperproportionenverhältnissen keine unmittelbare Patientenrelevanz als direkt für die erkrankte Person spürbare Symptomatik abgeleitet werden.

Validität

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

Child Behavior Checklist (CBCL)

Der Endpunkt „CBCL“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Bei den CBCL handelt es sich um elternberichtete Fragebögen zur Erfassung von Verhaltenssymptomen und emotionalen Symptomen bei Kindern und Jugendlichen. Es existieren 2 Versionen der CBCL, die abhängig vom Alter eingesetzt werden:

- „CBCL (1,5–5 Jahre)“ für Kinder im Alter von 18 Monaten bis ≤ 5 Jahren mit 100 Items; Bezugszeitraum sind die letzten 2 Monate.
- „CBCL (6–18 Jahre)“ für Kinder im Alter von 6 bis ≤ 17 Jahren mit 113 Items, Bezugszeitraum sind die letzten 6 Monate.

Die CBCL-Version für 1,5–5 Jahre wurde bei Kindern zwischen 5 bis < 6 Jahren eingesetzt, die Version für 6–18 Jahre bei allen übrigen Studienteilnehmenden. Erlebte ein Kind im Studienverlauf seinen 6. Geburtstag, wurde auf die Version 6–18 Jahre gewechselt. Wenn möglich, war eine gemeinsame Auswertung der Checklisten vorgesehen. Die Sprachskala der Checklisten wurde nicht erhoben, da diese nicht zur Bildung der vorgesehenen Verhaltensdomänen benötigt wird. Sofern länderspezifische Versionen des Fragebogens existierten, wurden diese verwendet.

Die Antwortoptionen befinden sich auf einer Likert-Skala von 0 (= nicht zutreffend), 1 (= etwas oder manchmal zutreffend) bis 2 (= sehr oder oft zutreffend). Höhere Werte bedeuten mehr berichtete Probleme. Ähnliche Items werden wie folgt gruppiert und zu Domänenwerten zusammengefasst:

Tabelle 6: Domänen der jeweiligen Versionen der CBCL

Version der Checkliste	Domänen
CBCL (1,5–5 Jahre)	Emotionale Reaktivität
	Ängstlich/Depressiv
	Körperliche Beschwerden
	Sozialer Rückzug
	Aufmerksamkeitsprobleme
	Aggressives Verhalten
	Gesamt
CBCL (6–18 Jahre)	Ängstlich/Depressiv
	Sozialer Rückzug / Depressiv
	Körperliche Beschwerden
	Regelverletzendes Verhalten
	Soziale Probleme
	Denk- und repetitive Probleme
	Aufmerksamkeitsprobleme
	Gesamt

Die Erhebung der CBCL wurde mit Protokoll-Amendment 3 (siehe Tabelle 3) eingeführt. Der Endpunkt wurde mittels deskriptiver Darstellung der Veränderungen zu Baseline ausgewertet. Statistische Analysen der Mittelwertdifferenzen (und Hedges' g bei signifikantem Unterschied) oder Responder-Analysen wurden weder geplant noch durchgeführt.

Erhebungszeitpunkte: Während des Screenings und in Woche 26 und 52.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Es wurde keine Rationale identifiziert, warum das Messinstrument nicht in der Studienvisite bei vorzeitigem Abbruch eingesetzt wurde.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Verhaltensprobleme und emotionale Funktionen werden als patientenrelevante Endpunkte bewertet. Es wurde keine Rationale des pU für die Einstufung als Sicherheitsendpunkt identifiziert. Der Endpunkt wird daher der Kategorie „Morbidity“ zugeordnet.

Validität

Die CBCL sind verschiedene, nach Altersgruppen aufgeteilte Instrumente, um internalisierendes Verhalten (wie Depression und Angst), externalisierendes Verhalten (wie Aggression), Stress, obsessiv-zwanghaftes Verhalten oder träges kognitives Tempo der Patientinnen und Patienten durch ihre betreuende Person zu bewerten. Es handelt sich um ein generisches Instrument zur Erfassung psychischer Störungen. Für alle Checklisten liegen umfangreiche Manuale vor, in denen Informationen zu Entwicklung, Auswertung und Validierung gegeben werden [2,3]. Angaben zu Reliabilität und Validität der Instrumente sind in diesen Manualen dokumentiert. Die Skalen wurden empirisch über Faktoranalysen generiert [1]. Die CBCL sind bereits vielfach in unterschiedlichen Indikationen eingesetzt worden, u. a. im Krankheitsbild der Epilepsie.

Ausreichende Informationen zu Änderungssensitivität der Instrumente konnten nicht identifiziert werden. Wie sensitiv Veränderungen aufgrund einer Intervention erfasst werden können, ist daher unklar.

2.3.3 Lebensqualität

Pediatric Quality of Life Inventory (PedsQL)

Der Endpunkt „PedsQL“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Das Instrument PedsQL erfasst die allgemeine gesundheitsbezogene Lebensqualität bei Kindern und Jugendlichen. Es kamen die altersgerechten, selbst- und fremdberichteten Versionen für die Altersgruppen 5–7 Jahre, 8–12 Jahre und 13–18 Jahre zum Einsatz. Die Version für die Altersgruppe 5–7 Jahre wurde ausschließlich fremdberichtet erhoben. Die Versionen zur Fremdeinschätzung wurden durch die Eltern/Betreuungspersonen verwendet. Der Bezugszeitraum beträgt einen Monat.

Die Fragebögen bestehen aus den folgenden 4 multidimensionalen Skalen:

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

Die Items wurden auf einer 5-Punkte-Likert-Skala von 0 ($\hat{=}$ nie) bis 4 ($\hat{=}$ fast immer) beantwortet. Die Werte werden anschließend auf eine Skala von 0 bis 100 linear transformiert, wobei höhere Werte

für eine bessere Lebensqualität stehen. Skalenwerte wurden nur berechnet, sofern mindestens 50 % der zugehörigen Items beantwortet wurden.

Die Auswertung findet in Form von drei zusammenfassenden Werten statt:

- Gesamtwert ($\hat{=}$ Summe aller Items geteilt durch die Anzahl der beantworteten Items);
- Mittelwert der psychosozialen Gesundheit ($\hat{=}$ Skalen „Emotionale Funktion“, „Soziale Funktion“ und „Schulische Funktion“);
- Mittelwert der physischen Gesundheit ($\hat{=}$ Skala „Physische Funktion“).

Der Endpunkt wurde mittels deskriptiver Darstellung der Veränderungen zu Baseline getrennt nach Eltern- und selbstberichteter Version ausgewertet. Statistische Analysen der Mittelwertdifferenzen (und Hedges' g bei signifikantem Unterschied) oder Responder-Analysen wurden weder geplant noch durchgeführt.

Erhebungszeitpunkte: Während des Screenings und in Woche 26 und 52.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Es wurde keine Rationale identifiziert, warum das Messinstrument nicht in der Studienvsited bei vorzeitigem Abbruch eingesetzt wurde.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Die Erhebung der gesundheitsbezogenen Lebensqualität mittels PedsQL ist patientenrelevant. Eine Selbstbewertung, sofern diese möglich ist, wird dabei gegenüber einer Fremdbewertung bevorzugt. Die Version zur Selbstbewertung wird daher für die vorliegende Altersgruppe ab 8 Jahren berücksichtigt. Eine ergänzende Berücksichtigung der fremdberichteten Einschätzung für die Altersgruppe 5–7 Jahre ist nicht möglich, da die Ergebnisse der fremdberichteten Versionen über alle Altersgruppen hinweg gemeinsam ausgewertet wurden. Diese Ergebnisse der Elternversionen werden daher lediglich im Anhang berichtet.

Validität

Der PedsQL ist ein etabliertes Instrument und wurde in den 1990er Jahren als generisches Instrument zur Erfassung der Lebensqualität bei pädiatrischen Populationen mit chronischen Erkrankungen entwickelt.

Der PedsQL ist ein modulares Konzept zur Erfassung der gesundheitsbezogenen Lebensqualität von gesunden Kindern und Jugendlichen und jenen mit akuten und chronischen Erkrankungen [23]. Das Instrument wurde an einer Population US-amerikanischer Kinder entwickelt und untersucht. Es bietet die Möglichkeit des Elternreports für 2- bis 18-Jährige und die des Selbstreports für Kinder und Jugendliche von 5 bis 18 Jahren. Die interne Konsistenz wurde an einer umfangreichen Stichprobe mit ca. 8.800 gesunden und ca. 850 chronisch kranken Kindern untersucht. Die Test-Retest-Reliabilität bei Kindern (7–12 Jahre) wird mit 0,86 angegeben. Die interne Validität wurde an einer Stichprobe mit ca. 5.500 Kindern (ca. 570 davon chronisch krank) untersucht. Es konnte gezeigt werden, dass Kinder mit chronischen Erkrankungen konsistent eine niedrigere Lebensqualität berichten als gesunde Kinder.

Zur Änderungssensitivität liegen Untersuchungen bei Personen mit orthopädischen [22], rheumatischen [24] sowie neurologischen Erkrankungen (chronischer Kopfschmerz) [13] vor. Es zeigten sich

bei Nachbeobachtungszeiten zwischen 70 Tagen und 7 Monaten bei Personen mit interventioneller Therapie statistisch signifikante Veränderungen des PedsQL-Gesamtwerts. In zwei Studien wurde die Effektstärke der Veränderung berechnet, welche mit 0,63–0,71 [13] bzw. 0,72–1,00 [22] als moderat bis hoch zu bewerten ist.

Zusammenfassend wird das Instrument als geeignet zur Erfassung der gesundheitsbezogenen Lebensqualität bewertet.

Quality of Life in Short Stature Youth Questionnaire (QoLISSY)

Der Endpunkt „QoLISSY“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Beim QoLISSY handelt es sich um ein Instrument zur Beurteilung der Lebensqualität von Kindern und Jugendlichen mit Kleinwuchs [12]. Es existiert eine Version zur direkten Befragung von Kindern und Jugendlichen in den Altersgruppen 8–12 Jahre und 13–18 Jahre sowie eine Version zur Befragung der Eltern von betroffenen Kindern und Jugendlichen in den Altersgruppen 4–7 Jahre, 8–12 Jahre sowie 13–18 Jahre. Die Versionen unterliegen folgendem Aufbau:

- 3 Hauptdomänen
 - Physische Domäne (6 Items)
 - Emotionale Domäne (8 Items)
 - Soziale Domäne (8 Items)
- 3 zusätzliche Domänen
 - Bewältigung („coping“) (10 Items)
 - Überzeugungen („beliefs“) (4 Items)
 - Behandlung („treatment“) (14 Items)

Die Domäne „Behandlung“ wurde nicht in Studie 301 erhoben. Die Elternversion enthält zusätzlich 2 Domänen, die die Zukunft des Kindes (5 Items) sowie den Einfluss der Erkrankung auf die Eltern (10 Items) thematisieren. Eine Frage zu Wachstumshormonen wurde im Vergleich zum Original-Fragebogen entfernt. Der Bezugszeitraum der Items umfasst die vergangenen 7 Tage.

Die Items werden jeweils auf einer 5-Punkte-Likert-Skala (von „nie / gar nicht“ (never / not at all) bis „immer / in höchstem Maße“ (always / extremely)) beantwortet. Die Rohwerte jeder Domäne werden als Summenwert gebildet und anschließend auf Skalen transformiert, die eine Spannweite von 0 bis 100 Punkten aufweisen. Aus den 3 Hauptdomänen wird ein Gesamtwert berechnet. Dieser entspricht dem Mittelwert der 22 Items, die für die 3 Hauptdomänen herangezogen werden. Höhere Werte bedeuten eine bessere Lebensqualität.

Der Endpunkt wurde mittels deskriptiver Darstellung der Veränderungen zu Baseline getrennt nach Eltern- und selbstberichteter Version ausgewertet. Statistische Analysen der Mittelwertdifferenzen (und Hedges' g bei signifikantem Unterschied) oder Responder-Analysen wurden weder geplant noch durchgeführt.

Erhebungszeitpunkte: Während des Screenings und in Woche 26 und 52.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Es wurde keine Rationale identifiziert, warum das Messinstrument nicht in der Studienvsiste bei vorzeitigem Abbruch eingesetzt wurde. Ebenso konnte keine Begründung für den Verzicht auf die Domäne „Behandlung“ identifiziert werden.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Die Erhebung der Lebensqualität mittels QoLISSY wird als patientenrelevant bewertet. Eine Selbstbewertung, sofern diese möglich ist, wird dabei gegenüber einer Fremdbewertung bevorzugt. Die Version zur Selbstbewertung wird daher für die vorliegende Altersgruppe ab 8 Jahren berücksichtigt. Eine ergänzende Berücksichtigung der fremdberichteten Einschätzung für die Altersgruppe 5–7 Jahre ist nicht möglich, da die Ergebnisse der fremdberichteten Versionen über alle Altersgruppen hinweg gemeinsam ausgewertet wurden. Diese Ergebnisse der Elternversionen werden daher lediglich im Anhang berichtet.

Validität

Entwicklung und Validierung des QoLISSY-Instruments umfassten mehrere Phasen. Nach Sichtung der Literatur wurden Fokusgruppen abgehalten, anschließend erfolgten Pilottests und Untersuchungen der psychometrischen Eigenschaften des Instruments. Basierend auf den Ergebnissen aus Fokusgruppen mit betroffenen Kindern und Jugendlichen ($n = 86$) sowie Eltern von Betroffenen ($n = 114$) wurde ein konzeptuelles Modell mit verschiedenen Domänen und Subdomänen gebildet. Ebenso wurde auf Basis dieser Ergebnisse die Formulierung von Items vorgenommen. Anschließend wurde der Fragebogen durch Anwendung bei Betroffenen ($n = 96$) und Eltern von Betroffenen ($n = 109$) pilotiert, indem psychometrische Eigenschaften des Fragebogens untersucht wurden. Zudem wurde ein „cognitive debriefing“ mit den Betroffenen durchgeführt. Die Ergebnisse dieser Untersuchungen wurden von den Entwicklerinnen und Entwicklern des Fragebogens herangezogen, um die Selektion der im Fragebogen verbleibenden Items vorzunehmen [12].

Zur Untersuchung der psychometrischen Eigenschaften des angepassten Fragebogens wurde eine multinationale Studie in Deutschland, Schweden, Spanien, Frankreich und dem Vereinigten Königreich durchgeführt. Es wurden 268 Kinder und Jugendliche mit idiopathischem Kleinwuchs (ISS) oder mit Wachstumshormonmangel (GHD) sowie 317 Eltern von Betroffenen eingeschlossen [12].

Zur Untersuchung der konvergenten Validität wurden Korrelationen zwischen den Domänenwerten des KIDSCREEN und den Domänenwerten sowie dem Gesamtwert des QoLISSY bestimmt. Es zeigten sich für alle Domänen des KIDSCREEN statistisch signifikante Korrelationen mit dem Gesamtwert der selbstaufgefüllten Version des QoLISSY ($r = 0,23–0,57$). Ebenso zeigten sich statistisch signifikante Korrelationen zwischen allen Domänen des KIDSCREEN mit den physischen ($r = 0,20–0,50$), emotionalen ($r = 0,25–0,58$) und sozialen Domänen ($r = 0,18–0,51$) sowie der Domäne „Überzeugungen“ ($r = 0,21–0,47$) des QoLISSY. Für die Domänen „Bewältigung“ und „Behandlung“ des QoLISSY zeigten sich statistisch signifikante Korrelationen mit einem Teil der Domänen des KIDSCREEN, die Korrelationen wiesen kein hohes Ausmaß auf ($r = 0,09–0,23$ bzw. $r = -0,01–0,24$). Für die elternberichtete Version des Fragebogens zeigten sich jeweils ähnliche Werte. Für die zusätzlichen Domänen der Elternversion zur Zukunft des Kindes und zum Einfluss der

Erkrankung auf die Eltern zeigten sich ebenfalls statistisch signifikante Korrelationen zu den meisten Domänen des KIDSSCREEN [12].

Zur Untersuchung der Known-Groups-Validität wurden Werte des QoLISSY für Personen mit verschiedenen Diagnosen (ISS GHD), unterschiedlichem Behandlungsstatus (behandelt; nicht behandelt) und unterschiedlicher Ausprägung des Kleinwuchses ($< 1,5$ SD; $1,5-2,5$ SD; $> 2,5$ SD verringerte Größe) zu einem Zeitpunkt miteinander verglichen. Die Ergebnisse der Untersuchungen deuten darauf hin, dass zwischen diesen Personengruppen mithilfe der physischen, sozialen und emotionalen Domänen sowie der Domäne „Überzeugung“ der selbstauszufüllenden Version des QoLISSY unterschieden werden kann. Auf Basis der vorliegenden Ergebnisse scheint durch die Domänen „Bewältigung“ und „Behandlung“ nicht zwischen den aufgeführten Personengruppen unterschieden werden zu können. Für die Version des Fragebogens, die von Eltern von Betroffenen ausgefüllt wird, zeigten sich jeweils ähnliche Werte. Die Ergebnisse deuten darauf hin, dass anhand der zusätzlichen Domänen der Elternversion zur Zukunft des Kindes und zum Einfluss der Erkrankung auf die Eltern ebenfalls zwischen verschiedenen Personengruppen unterschieden werden kann [12].

Zur Untersuchung der internen Konsistenz wurde Cronbachs α berechnet. Es zeigte sich für den Gesamtwert des selbstberichteten Fragebogens ein Cronbachs α von 0,95, für die einzelnen Domänen zeigten sich Werte zwischen 0,82 und 0,88. Zur Ermittlung der Test-Retest-Reliabilität wurde der Intraklassenkorrelationskoeffizient (ICC) berechnet, die zweite Beantwortung des Fragebogens erfolgte durchschnittlich innerhalb von 10 Tagen nach der ersten Beantwortung. Zur Einschätzung der Test-Retest-Reliabilität wurden Daten von 124 Kindern und Jugendlichen sowie 148 Eltern herangezogen. Für den Gesamtwert des QoLISSY wurde ein ICC von 0,88 ermittelt, die physische, soziale und emotionale Domäne sowie die Domäne „Überzeugung“ erreichten Werte zwischen 0,80 und 0,85. Für die Domäne „Behandlung“ zeigte sich ein ICC von 0,73, wohingegen für die Domäne „Bewältigung“ ein niedrigerer Wert von 0,56 ermittelt wurde. Für die elternberichtete Version des Fragebogens zeigten sich für die interne Konsistenz und die Test-Retest-Reliabilität jeweils ähnliche Werte. Für die zusätzlichen Domänen der Elternversion zur Zukunft des Kindes und zum Einfluss der Erkrankung auf die Eltern zeigten sich ebenfalls ausreichende bis hohe Werte für die interne Konsistenz und Test-Retest-Reliabilität [12].

In einer Studie, die in mehreren Zentren in Deutschland durchgeführt wurde, wurde u. a. die Änderungssensitivität des QoLISSY untersucht [9]. In die Studie wurden Kinder und Jugendliche zwischen 8 und 18 Jahren mit neu diagnostiziertem idiopathischem GHD (IGHD) oder mit niedriger Körperlänge bei Geburt bezogen auf das Reifealter („small for gestational age“) sowie Eltern von Betroffenen im Alter zwischen 4 und 18 Jahren eingeschlossen. Im Interventionsarm der Studie erhielten die Betroffenen ein humanes Wachstumshormon, im Kontroll-Arm erfolgte keine Behandlung. Der QoLISSY wurde zu Baseline sowie 12 Monate nach Baseline beantwortet. Es zeigte sich auf Basis einer multivariaten Kovarianzanalyse (MANCOVA) kein statistisch signifikanter Effekt für den Behandlungsstatus. Ebenso zeigte sich kein statistisch signifikanter Unterschied bei der mittels MANCOVA durchgeführten Untersuchung der Veränderungen der QoLISSY-Werte zwischen Baseline und Monat 12 für den Vergleich zwischen Personen mit Kleinwuchs zu Monat 12 ($n = 28$) und Personen, die zu Monat 12 eine normale Größe, d. h. eine Verringerung der Größe um weniger als 2 SD der Allgemeinbevölkerung, aufwiesen ($n = 20$). Für die elternberichtete Version zeigten sich ebenfalls keine statistisch signifikanten Unterschiede. Univariate Analysen zum selbstberichteten Fragebogen deuten allerdings darauf hin, dass sich die physischen (partial eta squared (η_p^2) = 0,15; $p < 0,01$), emotionalen ($\eta_p^2 = 0,20$; $p < 0,01$) und sozialen Aspekte ($\eta_p^2 = 0,24$; $p < 0,01$) der Lebensqualität von Behandelten ($n = 33$) zwischen Baseline und Monat 12 im Vergleich zu Unbehandelten ($n = 18$) verbesserten. In anderen Domänen und für die von Eltern auszufüllende Version zeigten sich keine

statistisch signifikanten Ergebnisse. Die Autorinnen und Autoren der Publikation diskutieren zwei Ursachen für die Ergebnisse ihrer Studie: Zum einen sei möglich, dass eine Veränderung der Lebensqualität innerhalb eines Jahres nicht von einer erhaltenen Behandlung oder dem Körperwachstum abhängt, zum anderen könnte der Fragebogen nicht sensitiv genug sein, um tatsächlich entstandene Veränderungen der Lebensqualität abzubilden, sodass weitere Studien zu diesem Aspekt notwendig erscheinen [9].

Studien zum QoLISSY bei Kindern mit ACH wurden nicht identifiziert. Die Ergebnisse der vorhandenen Studien bei Kindern mit anderen Ursachen für Kleinwuchs werden jedoch als übertragbar eingestuft. In der Gesamtschau scheint der QoLISSY damit ausreichend valide, um in Form des Gesamtwerts in der Nutzenbewertung berücksichtigt zu werden. In der Interpretation muss beachtet werden, dass auf Basis der vorliegenden Informationen keine abschließende Einschätzung der Änderungssensitivität getroffen werden kann.

Functional Independence Measure for Children (WeeFIM)

Der Endpunkt „WeeFIM“ wird in der Nutzenbewertung aufgrund nicht belegter Validität für ältere Kinder und Jugendliche nicht berücksichtigt. Eine Auswertung für die Altersgruppe 5–7 Jahre liegt nicht vor.

Operationalisierung

Beschreibung

Der WeeFIM wurde entwickelt, um funktionale Unabhängigkeit von Kindern im Alter von 6 Monaten bis 7 Jahren fremdberichtet zu messen. Das Instrument kann auch bei Kindern über 7 Jahren eingesetzt werden, solange ihre funktionalen Fähigkeiten (gemessen mit dem Instrument) unter denen liegen, die von 7-jährigen Kindern ohne Behinderungen erwartet werden. Es besteht aus einer 7-stufigen Ordinalskala, mit der anhand von 18 Items die Spanne von vollständiger Unabhängigkeit bis Abhängigkeit bewertet wird. Diese 18 Items repräsentieren die 6 Bereiche „Selbstversorgung“ (6 Items), „Kontrolle des Schließmuskels“ (2 Items), „Transfers“ (3 Items), „Fortbewegung“ (2 Items), „Kommunikation“ (2 Items) und „Soziale Kognition“ (3 Items). Diese werden in den 3 Domänen „Selbstpflege“, „Mobilität“ und „Kognition“ zusammengefasst.

Der Beobachtungszeitraum ist der gegenwärtige Zeitpunkt. Die verfügbaren Werte sind der Gesamtwert zwischen 18 und 126 und Domänenwerte. Ein höherer Wert steht für ein höheres Maß an Unabhängigkeit. In den Studienunterlagen wird ein Handbuch referenziert, in welchem Anwendung und Auswertung detailliert dargelegt sind [21].

Der Endpunkt wurde mittels deskriptiver Darstellung der Veränderungen zu Baseline ausgewertet. Statistische Analysen der Mittelwertdifferenzen (und Hedges' g bei signifikantem Unterschied) oder Responder-Analysen wurden weder geplant noch durchgeführt.

Erhebungszeitpunkte: Während des Screenings und in Woche 26 und 52.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Es wird davon ausgegangen, dass das Instrument in allen Altersgruppen eingesetzt wurde.

Es wurde keine Rationale identifiziert, warum das Messinstrument nicht in der Studienvisite bei vorzeitigem Abbruch eingesetzt wurde.

Patientenrelevanz

Entsprechend § 2 Satz 3 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“.

Die Erhebung der funktionalen Unabhängigkeit wird als patientenrelevant erachtet. Der pU liefert keine Rationale für die Zuordnung des WeeFIM zur Lebensqualität, der u. a. Gehfähigkeit und Kontrolle über die Darmtätigkeit und damit auch Aspekte der Morbidität erfasst. Das Instrument erhebt andererseits ebenfalls die Teilhabe und Partizipation im Alltag, die der Lebensqualität zugeordnet werden. Selbstversorgung ist in der „International Classification of Functioning, Disability and Health – Children & Youth Version“ (ICF-CY) der World Health Organization (WHO) [25] der Kategorie „Aktivitäten und Partizipation“ zugeordnet. Der Einordnung zur Lebensqualität wird in der Nutzenbewertung daher gefolgt.

Der pU diskutiert die Patientenrelevanz des WeeFIM, der für Kinder bis 7 Jahre entwickelt wurde, nicht vor dem Hintergrund einer festgelegten Studienpopulation im Alter von 5 bis 17 Jahren.

Validität

Der WeeFIM wurde 1990 auf Basis des „Adult Functional Independence Measure“ (FIM) entwickelt. Eine Beschreibung der zugrundeliegenden Entwicklungsschritte wurde nicht identifiziert. Die Relevanz und konzeptuelle Richtigkeit der Items für ihre Domänen wurde von 8 Fachleuten der pädiatrischen Versorgung bewertet. Während die Zuordnung der Items zu den Domänen und die Befragung zur Relevanz der Items positive Ergebnisse lieferte, wurde keine Domäne als konzeptuell richtig bewertet. Vor dem Hintergrund wird die Relevanz der jeweiligen Berufsgruppe der Bewertenden diskutiert, die sich jeweils nicht ausreichend repräsentiert sähe [16]. Eine Einbeziehung von Kindern oder deren Eltern ist nicht beschrieben.

Die psychometrischen Eigenschaften des WeeFIM wurden in verschiedenen Indikationen untersucht. Eine Übersichtsarbeit zu Instrumenten der Unabhängigkeit in der Selbstversorgung von Kindern mit Osteochondrodysplasien kam hinsichtlich des WeeFIM zu folgenden Ergebnissen: Für den WeeFIM wurde eine angemessene bis sehr gute Interrater- und Intrarater-Reliabilität ($ICC = 0,6-1,0$) für Kinder mit angeborenen oder genetischen Erkrankungen festgestellt. Es zeigte sich eine moderate bis hohe Konstruktvalidität, wobei diese Ergebnisse nicht spezifisch für Kinder mit Osteochondrodysplasien waren. Die Kinder in den berücksichtigten Studien waren zwischen 6 Monaten und 9,5 Jahren alt [15]. Eine weitere Übersichtsarbeit untersuchte den WeeFIM bei Kindern mit Cerebralparese. Die Arbeit attestiert dem WeeFIM bei exzellenter Studienqualität gute Ergebnisse in der internen Konsistenz und in der strukturellen Validität sowie bei ausreichender Studienqualität gute Ergebnisse in der Interrater-Reliabilität und in der Konstruktvalidität. Da insgesamt Informationen zu weniger als 5 Gütekriterien vorlagen, wurde dem Instrument dennoch ein limitiertes Evidenzniveau zugemessen [19].

Eine Studie untersuchte die Änderungssensitivität des WeeFIM anhand 5 statistischer Verfahren (Reliability Change Index, Proportional Change Index, Effektgröße, standardisierte Mittelwerte, gepaarte t-Tests). Dazu wurden Daten von 174 Kindern mit Entwicklungsstörungen zwischen 11 Monaten und 7 Jahren, die eine Frühförderung oder eine schulische Therapie erhielten, ausgewertet. Alle 5 Indizes deuteten darauf hin, dass der WeeFIM in der Lage ist, Veränderungen der funktionalen Fähigkeiten über einen Zeitraum zuverlässig zu dokumentieren [18].

Zusammenfassend lässt sich festhalten, dass die Entwicklung des WeeFIM unter Beteiligung von klinischen Fachleuten, nicht jedoch erkrankter Kinder und deren Eltern stattfand. Inwiefern den erfassten Aspekten daher eine besondere Bedeutung auch im Anwendungsgebiet der ACH zukommen, ist damit offen. Insgesamt zeigen sich für den WeeFIM positive Ergebnisse in den untersuchten psychometrischen Eigenschaften für die Population der Kinder.

Es ist zu beachten, dass das Instrument für Kinder bis 7 Jahre entwickelt wurde. Ein Einsatz bei älteren Kindern ist laut Manual möglich, sofern die Fähigkeiten hinter denen von 7-jährigen Kindern ohne Behinderungen zurückstehen. Gleichwohl sind in den Validierungsstudien bei Kindern mit Osteochondrodysplasien maximal Kinder in einem Alter von 9,5 Jahren eingeschlossen. Das mediane Alter der Studienpopulation liegt im Placebo-Arm bei 9,3 Jahren, im Vosoritid-Arm bei 7,8 Jahren (siehe Tabelle 11). Damit ist die Validität des Instruments für einen Großteil der Studienpopulation unklar.

Auch vor dem Hintergrund der inhaltlich unklaren Übertragbarkeit der im WeeFIM erfassten Aspekte auf ältere Kinder und Jugendliche und der fehlenden Diskussion seitens des pU zum Einsatz des WeeFIM in der gesamten Studienpopulation, wird auf eine Darstellung im Rahmen der Nutzenbewertung verzichtet. Eine Auswertung für die Altersgruppe 5–7 Jahre liegt nicht vor.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Sicherheit wird evaluiert durch das Auftreten unerwünschter Ereignisse (UE), schwerwiegender unerwünschter Ereignisse (SUE) oder klinisch signifikanter Änderungen der Vitalzeichen, bei körperlichen Untersuchungen, EKG-Befunden, Untersuchungen zur Knochenmorphologie, Beschaffenheit und Wachstum der Extremitäten und der Wirbelsäule, Laborparametern, bildgebenden Erhebungen, Hüftbefunden, Biomarkern, Immunogenität, Speichelkortisol-Werten, Serumprolaktin-Werten, FSH-/LH-Werten (Individuen > 8 Jahre oder Tanner-Stadium > I) und Daten der Messung physischer Parameter.

Als UE wurde jedes medizinische Auftreten definiert, welches mit dem Gebrauch des Medikaments oder anderen Interventionen des Studienprotokolls in Verbindung stand. Davon eingeschlossen sind:

- UE, die nicht bereits zuvor beobachtet wurden und im Studienverlauf auftraten.
- Verschlechterung (Schwere oder Häufigkeit oder Beschaffenheit) von bereits vor Studienbeginn bestehender Beschwerden.
- Komplikationen, die in Zusammenhang mit nicht-medikamentösen Interventionen des Studienprotokolls stehen.

Laut Studienprotokoll werden alle UE mit dem Medical Dictionary for Regulatory Activities (MedDRA), Version 22, kodiert und nach Systemorganklasse und Preferred Term (PT) dargestellt. ACH-bezogene UE wurden anhand der PT identifiziert und separat ausgewertet.

Die Schwere der UE und SUE wurden, soweit möglich, gemäß der Common Terminology Criteria for Adverse Events (CTCAE, v4.) des National Cancer Institute (NCI) erhoben:

- Mild (Grad 1): Asymptomatisch oder mild symptomatisch; klinische oder diagnostische Beobachtung ausreichend; keine Behandlungsnotwendigkeit.
- Mäßig (Grad 2): Minimale, lokale oder nicht-invasive Intervention erforderlich; zu altersentsprechenden Einschränkungen der täglichen Aktivität führend.

- Schwer (Grad 3): Schwer oder medizinisch signifikant aber nicht unmittelbar lebensbedrohlich; Krankenhausaufenthalt oder eine Verlängerung des Krankenhausaufenthaltes erforderlich; zu Behinderung führend; die tägliche selbstführsorgende Lebensaktivität einschränkend.
- Lebensbedrohlich (Grad 4): Studienteilnehmende/n einer unmittelbaren Todesgefahr aussetzend; dringend zwingend behandlungsbedürftig.
- Tod (Grad 5): zum Tode führend.

Als SUE wurden alle UE bewertet, die mindestens eines der folgenden Kriterien erfüllten:

- tödlich
- lebensbedrohlich
- einen Krankenhausaufenthalt oder eine Verlängerung des Krankenhausaufenthaltes erforderlich
- zu andauernder oder signifikanter Behinderung führend
- eine kongenitale Anomalie bzw. einen Geburtsfehler nach sich ziehend
- medizinisch signifikant oder eine Intervention erforderlich, um eines der oben genannten Ereignisse zu verhindern

In den Analysen sind UE und SUE mit Beginn oder Verschlimmerung nach Beginn der Studienmedikation eingeschlossen. Das Auftreten der UE wurde ab dem Zeitpunkt der ersten Verabreichung bis 4 Wochen nach der letzten Verabreichung des Studienmedikaments (oder 4 Wochen nach der Studienvisite bei vorzeitigem Abbruch, falls dies später war) bewertet. Nach Studieneinschluss, aber vor Gabe der ersten Studienmedikation, wurden SUE nur im Zusammenhang mit nicht-medikamentösen Interventionen des Studienprotokolls erhoben.

A priori im Studienprotokoll wurden mit Amendment 2 (27.04.2017) folgende UE von besonderem Interesse definiert:

- Knochenbrüche
- Verrutschte Oberschenkelepiphyse
- Avaskuläre Nekrose oder Knochennekrose

Im SAP werden folgende UE von besonderem Interesse definiert:

- Reaktion an der Einstichstelle (HLT)
- Hypotension (PT: Blutdruck ambulatorisch erniedrigt; Blutdruck erniedrigt; Blutdruck diastolisch erniedrigt; Blutdruck orthostatisch erniedrigt; Blutdruck systolisch erniedrigt; Blutdruck systolisch inspiratorisch erniedrigt; Hypotonie; Diastolische Hypotonie; Orthostasesyndrom)
- Änderung der Herzfrequenz (PT: Vorhoftachykardie; Posturales orthostatisches Tachykardiesyndrom; Rebound-Tachykardie; Sinustachykardie; Tachykardie supraventrikulär; Tachykardie; Paroxysmale Tachykardie; Tachykardie ventrikulär; Bradykardie)
- Überempfindlichkeit (SMQ (Standard MedDRA Query))
- Algorithmische Anaphylaxie (SMQ mit zeitlichen Restriktionen)
- Frakturen (HLT: Frakturen; NEC: Gliedmaßenfrakturen, Beckenfrakturen, Schädel- und Gesichtsfrakturen, Wirbelsäulenfrakturen, Thorakalkäfigfrakturen nicht spinal)
- Verrutschte Oberschenkelepiphyse (PT: Epiphysenerkrankung; Verletzung einer Epiphyse)
- Avaskuläre Nekrose oder Knochennekrose (PT: Osteonekrose; Osteonekrose des Kiefers; Osteonekrose des äußeren Gehörgangs; Nekrose ischämisch)

In Modul 4 wurde die Definition der UE von besonderem Interesse gegenüber der im SAP definierten aktualisiert, um alle Frakturen zu erfassen. Die ausgewählten PT umfassen den Begriff „Fraktur“ sowie PT von „Knochenfragmentierung“, „Knochenfissur“, „Skapulothorakale Dissoziation“ und „Thoraxwandflattern“.

Erhebungszeitpunkte: Während des Screenings, Tag 1, 2, 3 und 10, in Woche 6, 13, 26, 39 und 52, sowie Woche 54/56 oder in der Studiervisite bei vorzeitigem Abbruch. Die finale Sicherheitsvisite wurde erst mit der Protokolländerung 3 vom 05.01.2018 von 2 auf 4 Wochen erhöht.

Bewertung

Die Operationalisierung ist mit Einschränkungen nachvollziehbar. Die finale Sicherheitsvisite wurde erst mit Protokolländerung 3 vom 05.01.2018 von 2 auf 4 Wochen erhöht. Es ist davon auszugehen, dass für einen Großteil der Studienpopulation die Dauer der Nachbeobachtung 2 Wochen betrug. Mit Protokolländerung 2 vom 27.04.2017 wurden außerdem Ereignisse von besonderem Interesse in der Sicherheitserhebung definiert. Es ist unklar, ob die Ereignisse vor dem Protokoll-Amendment bereits erfasst wurden bzw. nachträglich erfasst werden konnten.

Patientenrelevanz

Das Auftreten unerwünschter Ereignisse ist patientenrelevant.

Entsprechend § 2 Satz 3 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“.

Validität

Die Erhebung der UE wird als valide bewertet. Bei der Interpretation der Ergebnisse ist zu beachten, dass ein Großteil der Studienpopulation 2 statt 4 Wochen hinsichtlich des Auftretens von Sicherheitsereignissen nach Behandlungsende nachbeobachtet wurde. Dieser Aspekt betrifft jedoch beide Studienarme gleichermaßen und schränkt die interne Validität daher nicht ein.

Die Benennung der UE von besonderem Interesse aus dem Studienprotokoll wurden durch den SAP geändert. Inwiefern die Listung des SAP als a priori betrachtet werden kann, ist aufgrund von bestehen Unklarheiten zum Zeitpunkt der Finalisierung des SAP (siehe Kapitel 2.4 und 2.5) mit den vorliegenden Informationen nicht abschließend bewertbar.

Positiv hervorzuheben ist die a priori geplante Trennung des pU von UE, die als Symptome von ACH betrachtet wurden.

2.3.5 Übersicht der Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten Endpunkte findet sich in Tabelle 7.

Tabelle 7: Erhebungszeitpunkte der berücksichtigten Endpunkte in der Studie 301

Endpunkt \ Studienvisite	Screening	Tag 1	Woche 13	Woche 26	Woche 39	Woche 52	VA
Mortalität ¹⁾	x	x	x	x	x	x	x
Körpergröße	x	x	x	x	x	x	x
CBCL	x	-	-	x	-	x	-
PedsQL	x	-	-	x	-	x	-
QoLISSY	x	-	-	x	-	x	-
Unerwünschte Ereignisse ¹⁾	x	x	x	x	x	x	x

¹⁾ Zu jeder Studienvisite (siehe Kapitel 2.3.1 und 2.3.4) erhoben.

Abkürzungen: CBCL: Child Behavior Checklist; PedsQL: Pediatric Quality of Life Inventory; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; VA: Vorzeitiger Abbruch.

2.4 Statistische Methoden

In Studie 301 wurde die Überlegenheit von Vosoritid gegenüber Placebo im Parallelgruppendesign getestet. Das Datum des Datenbankschlusses ist nicht berichtet. Die letzte SAP-Version wurde am 05.11.2019 finalisiert. Es ist damit unklar, ob der SAP vor Datenbankschluss finalisiert wurde.

Relevante Analysepopulationen

Full Analysis Set (FAS): Definiert nach dem Intention-to-Treat-Prinzip, umfasst alle randomisierten Studienteilnehmenden mit einer unterzeichneten Einverständniserklärung.

Sicherheitspopulation: Definiert als alle Personen des FAS, die darüber hinaus mindestens eine Dosis ihrer Studienintervention erhielten.

Datenschnitte

Es wurden keine Interim-Analysen durchgeführt. Vorliegend ist der finale Datenschnitt nach Studienende.

Präspezifizierung geplanter und durchgeführter Analysen

Im SAP sind zur Bestimmung von Gruppenunterschieden in der Körpergröße (z-Score) zu Woche 52 Kovarianzanalysen (ANCOVA) mit den folgenden fixen Effekten und Kovariaten beschrieben:

- Behandlungsgruppe
- Randomisierungsstratum (männlich Tanner-Stadium I, weiblich Tanner-Stadium I, männlich Tanner-Stadium > I, weiblich Tanner-Stadium > I)
- Alter zu Baseline
- AGV zu Baseline
- Körpergröße (z-Score) zu Baseline

Es wurde zweiseitig getestet.

Die Endpunkte „CBCL“, „PedsQL“ und „QoLISSY“ sowie die Endpunkte der Sicherheit waren zur deskriptiven Auswertung vorgesehen.

Fehlende Werte und Daten-Imputation

Für den primären Endpunkt „Wachstumsgeschwindigkeit“ wurde multiple Imputation vorgesehen, sofern Personen die Behandlung vorzeitig abbrachen. Sofern diese angewendet wurde, sollten fehlende Werte in der Auswertung der Stehhöhe ebenfalls durch multiple Imputation eingefügt werden. Geplant war die Anwendung von Pattern-Mixture-Modellen, bei denen davon ausgegangen wurde, dass die fehlenden Beobachtungen einem „Missing Not At Random“-Mechanismus folgen. Dafür waren für jede Imputation mindestens 5 Personen in derselben Behandlungsgruppe erforderlich, die die Behandlung zu einem ähnlichen Zeitpunkt vorzeitig abbrachen und für die nach dem Abbruch der Behandlung Messungen der Stehhöhe vorlagen. Für den Fall, dass nicht genügend fehlende Daten zur Verfügung standen, wurde im SAP ein alternatives Imputationsverfahren beschrieben, bei dem die fehlenden Werte für die Stehhöhe in Woche 52 (d. h. an Tag 365) durch Anwendung der Baseline-AGV (cm/Jahr) auf die letzte verfügbare Messung der Körpergröße imputiert wurden.

Hinsichtlich der eingesetzten Fragebögen wurde folgender Umgang mit fehlenden Werten festgelegt: Sofern im PedsQL mehr als 50 % der Items innerhalb einer Skala fehlten, wurden keine Skalenwerte berechnet; stattdessen wurden diese als fehlend betrachtet. Für den QoLISSY musste der Anteil vorhandener Werte 80 % zur Berechnung von Domänenwerten betragen. Für die CBCL sind keine Regeln zum Umgang mit fehlenden Werten beschrieben.

Sensitivitätsanalysen und ergänzende Analysen

Für die primäre Analyse wurden folgende Sensitivitätsanalysen und ergänzende Analysen geplant:

1. Washout-Modell, bei dem fehlende Daten für einen Endpunkt für beide Behandlungsgruppen durch beobachtete Daten aus der Placebo-Gruppe ersetzt wurden.
2. Studienabschluss-Modell, bei dem nur Personen mit beobachteten AGV-Werten nach 52 Wochen einbezogen wurden; fehlende Messungen der Stehhöhe nach 52 Wochen wurden nicht berücksichtigt.
3. Notfallmedikation-Modell, in dem Messungen ausgeschlossen wurden, die entweder nach Extremitätenverlängerung oder nach Einnahme von Wachstumshormonen/Gonadotropin-freisetzenden Hormonen erfolgten.

Im Studienbericht sind diese Analysen auch für den sekundären Endpunkt „Körpergröße“ (z-Score) enthalten.

Einschätzung der statistischen Auswertungen

Die durchgeführten Auswertungen folgten den präspezifizierten Beschreibungen. Es wurden keine Überprüfungen der ANCOVA-Modellannahmen vorgesehen und durchgeführt.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 8: Verzerrungspotential der Studie 301

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

¹⁾ Die Methode zur Erzeugung der Randomisierungssequenz ist in Modul 4 des Nutzendossiers benannt. Demnach sei die Sequenz von einem externen Drittanbieter und mittels SAS-Version 9.3 erstellt worden.

²⁾ Es zeigen sich unterschiedliche Verteilungen für das Merkmal „Alter“ zu Baseline zwischen den Studienarmen. Ursprünglich wurde eine Stratifizierung für Alter (< 11 oder ≥ 11 Jahre) vorgesehen, sodass Alter als potentieller Effektmofikator in Frage kommt. Für die Kategorien < 11 Jahre und ≥ 11 Jahre zeigen sich gleiche Verteilungen in den Studienarmen.

³⁾ Die Allokation wurde mittels Interactive Voice/Web Response System vorgenommen.

⁴⁾ Es bestehen Unklarheiten zur Finalisierung des SAP vor erster Kenntnis der Daten in den Studienarmen (siehe Kapitel 2.4).

Das Verzerrungspotential auf Studienebene wird als unklar eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 9 dargestellt. Auf eine eingehende Einschätzung des Verzerrungspotentials des Endpunkts „CBCL“ wird aufgrund der niedrigen Rücklaufquoten bereits zu Baseline (siehe Kapitel 3.3: CBCL) jedoch verzichtet.

Tabelle 9: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studie 301

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität	Ja	Ja	Nein	Nein	Gering
Körpergröße	Ja	Ja	Unklar ¹⁾	Nein	Unklar
PedsQL	Ja	Unklar ²⁾³⁾	Nein	Nein	Unklar
QoLISSY	Ja	Unklar ²⁾⁴⁾	Nein	Nein	Unklar
Unerwünschte Ereignisse	Ja	Ja	Nein	Unklar ⁵⁾	Gering

¹⁾ Es ist unklar, inwiefern das gewählte Analysemodell und der Modellaufbau vor erster Kenntnis der Daten in den Studienarmen festgelegt wurden, da Unklarheiten zur Finalisierung des SAP bestehen (siehe Kapitel 2.4).

²⁾ Die Analysepopulationen folgten dem ITT-Prinzip, umfassen aber nicht alle randomisierten Studienteilnehmenden. Angaben zu Gründen für die fehlenden Werte wurden nicht identifiziert. Fehlende Werte könnten mit dem Outcome-Status im Zusammenhang stehen, auch wenn keine Hinweise dafür identifiziert wurden (z. B. durch auffällig unterschiedliche Abbruchraten in den Studienarmen, auffällig unterschiedliche Rücklaufquoten in den Studienarmen).

³⁾ Die Anteile der Analysepopulation an der ITT-Population (FAS) liegen bei 89 % im Placebo-Arm und 86 % im Vosoritid-Arm.

⁴⁾ Die Anteile der Analysepopulation an der ITT-Population (FAS) liegen bei 95 % im Placebo-Arm und 90 % im Vosoritid-Arm.

⁵⁾ Es wurden für das im SAP definierte UE von besonderem Interesse „Algorithmische Anaphylaxie“ keine Ergebnisse identifiziert (siehe Kapitel 3.5).

Abkürzungen: FAS: Full Analysis Set; ITT: Intent to Treat; PedsQL: Pediatric Quality of Life Inventory; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; SAP: Statistischer Analyseplan; UE: Unerwünschtes Ereignis.

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

Die Verteilung der Testpersonen ist in Tabelle 10 dargestellt. Insgesamt wurden 121 Teilnehmende in die Studie aufgenommen. Alle Personen nahmen zuvor mindestens 6 Monate an Studie 901 teil. 61 Personen wurden in die Placebo-Gruppe, 60 Personen in die Vosoritid-Gruppe randomisiert. Diese Personen bilden sowohl das FAS als auch die Sicherheitspopulation.

Nach 52 Wochen Behandlung schlossen alle 61 Teilnehmenden der Placebo-Gruppe die Studie ab. In der Vosoritid-Gruppe beendeten 58 Teilnehmende die Studie, während 2 Personen aus der Studie ausschieden. Die mediane Behandlungsdauer beträgt dennoch in beiden Studienarmen 365 Tage. Es wurden keine Angaben zur medianen Beobachtungsdauer identifiziert. Da jedoch beide Personen mit Abbruch der Studienmedikation auch die Studie abbrachen, ist von keinem relevanten Unterschied zwischen medianer Behandlungsdauer und Beobachtungsdauer auszugehen.

Tabelle 10: Allgemeine Angaben Studie 301

Studie 301 Allgemeine Angaben	Vosoritid n (%)	Placebo n (%)
FAS ¹⁾	60 (100)	61 (100)
Sicherheitspopulation ¹⁾	60 (100)	61 (100)
Abbruch der Einnahme der Studienmedikation aufgrund von UE	2 (3,3)	0
Wunsch der Testperson	1 (1,7)	0
	1 (1,7)	0
Abbruch der Studie	2 (3,3)	0
Entzug der Einverständniserklärung	2 (3,3)	0
Übergang in Extensionsstudie BMN 111-302	58 (96,7)	61 (100)
Mediane Behandlungsdauer Tage (min; max)	365 (2; 384)	365 (355; 380)
Mediane Beobachtungsdauer Tage (min; max)	k. A.	k. A.

¹⁾ Definition, siehe Kapitel 2.4.

Abkürzungen: FAS: Full Analysis Set; k. A.: keine Angabe; UE: Unerwünschtes Ereignis.

Die Baseline-Charakteristika der Studienpopulation sind in Tabelle 11 dargestellt.

Im Placebo-Arm betrug das mittlere Alter 9,1 Jahre (SD: 2,5), im Vosoritid-Arm lag es mit 8,4 Jahren (SD: 2,4) etwas niedriger. Ebenso gab es im Vosoritid-Arm mehr Kinder im Alter von ≥ 5 bis < 8 Jahren ($n = 31$ (51,7 %)) als im Placebo-Arm ($n = 24$ (39,3 %)). Insgesamt waren 52,9 % der teilnehmenden Kinder und Jugendlichen männlich und 47,1 % weiblich. Die Verteilung in diesem Stratifizierungsmerkmal war in den beiden Behandlungsgruppen ähnlich.

Gleiches gilt für das Stratifizierungsmerkmal Tanner-Stadium: Jeweils 48 Kinder pro Studienarm waren vorpubertär (Tanner-Stadium I) und 13 bzw. 12 Personen im Placebo- bzw. Vosoritid-Arm waren im Tanner-Stadium $> I$. Jeweils 1 Kind pro Gruppe wurde in die Studie aufgenommen, bevor die Randomisierung nach Geschlecht und Tanner-Stadium stratifiziert wurde (eingeführt in Protokoll-Amendment 2, siehe Tabelle 3). Diese Kinder waren beide weiblich und ≤ 11 Jahre alt. In den Ergebnisauswertungen wurden sie entsprechend ihrer tatsächlichen Strata nach Tanner-Stadium und Geschlecht analysiert (d. h. weibliches Tanner-Stadium I).

Die mittlere Stehhöhe war im Vosoritid-Arm mit 100,2 cm (SD: 11,9) etwas niedriger als im Placebo-Arm mit 102,9 cm (SD: 11,0). Dies kann vermutlich auf das jüngere Alter der Studienteilnehmenden im Vosoritid-Arm zurückgeführt werden. In den z-Scores der Körpergröße zeigen sich mit einem Mittelwert von jeweils -5,1 (SD: 1,1) im Placebo- und im Vosoritid-Arm hingegen keine Unterschiede zwischen den Behandlungsgruppen. Auch in der AGV zeigen sich ähnliche Verteilungen in den Studienarmen, die im Mittelwert bei 4,1 cm/Jahr (SD: 1,2) im Placebo-Arm und 4,3 cm/Jahr (SD: 1,5) im Vosoritid-Arm lag. Auffällig ist die breite Streuung bei dieser Messung von 1,5 bis 6,7 cm/Jahr im Placebo-Arm und -0,1 bis 6,9 cm/Jahr im Vosoritid-Arm.

Tabelle 11: Charakterisierung der Studienpopulation in Studie 301

Studie 301 Charakterisierung der Studienpopulation	Vosoritid N = 60	Placebo N = 61
<i>Alter an Studientag 1 (Jahre)</i> Mittelwert (SD) Median (min; max)	8,4 (2,4) 7,8 (5,1; 13,1)	9,1 (2,5) 9,3 (5,1; 14,9)
<i>Altersgruppe (Jahre) an Studientag 1¹⁾, n (%)</i> ≥ 5 bis < 8 Jahre ≥ 8 bis < 11 Jahre ≥ 11 bis < 17 Jahre	31 (51,7) 17 (28,3) 12 (20,0)	24 (39,3) 24 (39,3) 13 (21,3)
<i>Geschlecht²⁾, n (%)</i> männlich weiblich	31 (51,7) 29 (48,3)	33 (54,1) 28 (45,9)
<i>Tanner-Stadium²⁾, n (%)</i> I > I	48 (80,0) 12 (20,0)	48 (78,7) 13 (21,3)
<i>Stratifizierungsstratum, n (%)</i> männlich Tanner-Stadium I weiblich Tanner-Stadium I männlich Tanner-Stadium > I weiblich Tanner-Stadium > I weiblich ≤ 11 Jahre	28 (46,7) 19 (31,7) 3 (5,0) 9 (15,0) 1 (1,7)	29 (47,5) 19 (31,1) 4 (6,6) 8 (13,1) 1 (1,6)
<i>Abstammung (genetisch), n (%)</i> kaukasisch/weiß asiatisch schwarz oder afro-amerikanisch mehrere	45 (75,0) 10 (16,7) 3 (5,0) 2 (3,3)	41 (67,2) 13 (21,3) 2 (3,3) 5 (8,2)
<i>BMI (z-Score)</i> Mittelwert (SD) Median (min; max)	1,9 (0,6) 1,9 (0,2; 3,2)	1,7 (0,6) 1,7 (0,0; 3,1)
<i>Körpergröße (z-Score)</i> Mittelwert (SD) Median (min; max)	-5,1 (1,1) -5,3 (-7,7; -1,1)	-5,1 (1,1) -5,2 (-7,9; -2,7)
<i>Stehhöhe (cm)</i> Mittelwert (SD) Median (min; max)	100,2 (11,9) 98,6 (80,1; 136,8)	102,9 (11,0) 104,6 (79,9; 129,3)
<i>AGV (cm/Jahr)</i> Mittelwert (SD) Median (min; max)	4,3 (1,5) 4,1 (-0,1; 6,9)	4,1 (1,2) 4,1 (1,5; 6,7)
Vorherige Einnahme von Medikamenten, n (%)	10 (16,7)	13 (21,3)

Studie 301 Charakterisierung der Studienpopulation	Vosoritid N = 60	Placebo N = 61
Vorerkrankungen nach MedDRA-Systemorganklasse / Preferred Term³⁾⁴⁾, n (%)		
Jegliche Vorerkrankungen	59 (98,3)	59 (96,7)
Chirurgische und medizinische Eingriffe	42 (70,0)	40 (65,6)
Einbringung eines Ohrtröhrchens	27 (45,0)	22 (36,1)
Adenoidektomie	20 (33,3)	19 (31,1)
Infektionen und parasitäre Erkrankungen	38 (63,3)	43 (70,5)
Otitis media	24 (40,0)	27 (44,3)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	38 (63,3)	37 (60,7)
Schlafapnoe-Syndrome	27 (45,0)	31 (50,8)
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen	27 (45,0)	31 (50,8)
Kongenitale, familiäre und genetische Erkrankungen	24 (40,0)	27 (44,3)
Kongenitale Verbiegung der langen Knochen	18 (30,0)	17 (27,9)
Erkrankungen des Nervensystems	18 (30,0)	24 (39,3)
Erkrankungen des Ohrs und des Labyrinths	20 (33,3)	19 (31,1)
Stoffwechsel- und Ernährungsstörungen	14 (23,3)	21 (34,4)
Vitamin-D-Mangel	12 (20,0)	20 (32,8)
Erkrankungen des Immunsystems	16 (26,7)	13 (21,3)

¹⁾ Ursprünglich war eine Stratifizierung für dieses Merkmal anstelle des Tanner-Stadiums vorgesehen. Dies wurde mit Protokoll-Amendment 2 geändert (siehe Tabelle 3).

²⁾ Stratifizierungsfaktor der Randomisierung.

³⁾ > 25 % in einem Studienarm.

⁴⁾ Bezugszeitraum unklar.

Abkürzungen: AGV: Jährliche Wachstumsgeschwindigkeit; MedDRA: Medical Dictionary for Regulatory Activities; SD: Standardabweichung.

Exposition mit der Studienmedikation

Angaben zur Exposition mit der Studienmedikation sind in Tabelle 12 zusammengefasst. Im Mittelwert erhielten die Studienteilnehmenden des Placebo-Arms 360 Dosen, die des Vosoritid-Arms 349 Dosen. Die geringere Anzahl im Vosoritid-Arm ist auf 2 Personen mit Abbruch der Studienmedikation zurückzuführen (vgl. Tabelle 10).

Jeweils 10 Personen pro Studienarm unterbrachen im Studienverlauf die Einnahme ihrer Studienmedikation. Dabei ist zu beachten, dass sich diese Angabe allein auf Unterbrechungen aufgrund von UE bezieht. Gleiches gilt für die Angabe der Dosisreduktionen mit jeweils 0 Ereignissen pro Studienarm. Inwiefern auch Dosismodifikationen aufgrund anderer Ursachen möglich waren und durchgeführt wurden, ist unklar. Keine Angaben wurden zur Anzahl von Dosiserhöhungen identifiziert. Die mittlere Vosoritid-Dosis entspricht mit 15 µg/kg/Tag jedoch der üblichen Dosis gemäß Fachinformation.

Tabelle 12: Angaben zur Exposition mit der Studienmedikation in Studie 301

Studie 301 Exposition mit der Studienmedikation	Vosoritid N = 60	Placebo N = 61
Anzahl verabreichter Dosen Mittelwert (SD) Median (min; max)	348,8 (65,1) 362 (2; 382)	359,6 (9,5) 362 (321; 373)
Dosisanpassungen ¹⁾ , n (%) Unterbrechung Erhöhung Reduktion	10 (16,7) k. A. 0	10 (16,4) k. A. 0
Durchschnittliche gewichtsadjustierte Dosis (µg/kg/Tag) Mittelwert (SD) Median (min; max)	14,9 (0,6) 15 (14; 19)	14,8 (0,6) 15 (13; 16)

¹⁾ Es werden lediglich Dosismodifikationen aufgrund von UE berichtet. Ob Dosismodifikationen aufgrund anderer Ursachen vorgenommen wurden, ist unklar.

Abkürzungen: k. A.: keine Angabe; SD: Standardabweichung; UE: Unerwünschtes Ereignis.

Begleitmedikation

Begleitmedikationen nahmen im Studienverlauf 57 Kinder und Jugendliche aus dem Placebo-Arm und 54 aus dem Vosoritid-Arm ein. Es zeigen sich in keiner Gruppe und für kein Arzneimittel, die von > 25 % in einem Studienarm eingenommen wurde, Unterschiede.

Tabelle 13: Begleitmedikation > 25 % in einem Studienarm in Studie 301

Studie 301 Begleitmedikation > 25 % in einem Studienarm	Vosoritid (N = 60)	Placebo N = 61)
ATC-Klasse Level 2 Bevorzugter Arzneimittelname	n (%)	n (%)
Jegliche Begleitmedikation	54 (90,0)	57 (93,4)
Vitamine Vitamin D und Analoga	36 (60,0) 17 (28,3)	39 (63,9) 22 (36,1)
Antiphlogistika und Antirheumatika Ibuprofen	27 (45,0) 27 (45,0)	35 (57,4) 31 (50,8)
Analgetika Paracetamol	30 (50,0) 29 (48,3)	25 (41,0) 24 (39,3)
Antibiotika zur systemischen Anwendung	22 (36,7)	20 (32,8)
Antipruriginosa, inkl. Antihistaminika, Anästhetika, etc.	19 (31,7)	17 (27,9)
Rhinologika	18 (30,0)	14 (23,0)
Impfstoffe	14 (23,3)	17 (27,9)
Husten- und Erkältungsmittel	8 (13,3)	16 (26,2)

Abkürzungen: ATC: Anatomisch-Therapeutisch-Chemisch.

Protokollverletzungen

Insgesamt gab es bei 46 Studienteilnehmenden (75,4 %) in der Placebo-Gruppe und bei 40 (66,7 %) in der Vosoritid-Gruppe mindestens eine wesentliche Protokollabweichung. Die meisten wesentlichen Protokollabweichungen standen im Zusammenhang mit einem „nicht durchgeführten Verfahren“ (33 (54,1 %) bzw. 28 (46,7 %) Personen; siehe Tabelle 14).

Die Mehrzahl der nicht gemäß Studienprotokoll durchgeführten Verfahren umfasste (in absteigender Reihenfolge) „Vitalparameter nach Verabreichung“, „EKG nach Verabreichung“, „Klinische Laboruntersuchungen“ und „Anthropometrische Messungen, die nicht gemäß den studienspezifischen Richtlinien durchgeführt wurden“ (8 Ereignisse bei 4 Personen der Placebo-Gruppe, 7 Ereignisse bei 5 Personen der Vosoritid-Gruppe).

Tabelle 14: Protokollverletzungen in Studie 301

Studie 301 Protokollverletzungen	Vosoritid N = 60	Placebo N = 61
Wesentliche Protokollverletzungen, n (%)	40 (66,7)	46 (75,4)
Verfahren nicht durchgeführt	28 (46,7)	33 (54,1)
Unregelmäßigkeit bei der Dosierung	12 (20,0)	11 (18,0)
Unstimmigkeit mit Studien-Kit	5 (8,3)	6 (9,8)
Außerhalb des Zeitfensters	2 (3,3)	3 (4,9)
Falsches Studien-Kit	2 (3,3)	1 (1,6)
Falsche Behandlung	1 (1,7)	1 (1,6)
Randomisierung in falsche Strata	0	1 (1,6)

3.2 Mortalität

Todesfälle wurden im Rahmen der Sicherheitserhebungen erfasst. Im Studienverlauf traten keine Todesfälle auf.

3.3 Morbidität

Körpergröße

Die Veränderung der Stehhöhe wurde während des 52-wöchigen Zeitraums als z-Score ermittelt. Fehlende Werte wurden wie a priori festgelegt imputiert. Der Anteil an fehlenden Werten ist nicht berichtet, jedoch brachen 2 Personen im Vosoritid-Arm die Studie ab (vgl. Tabelle 10). Da nicht genügend Daten für die multiple Imputation zur Verfügung standen (vgl. Kapitel 2.4), wurden fehlende Werte für die Stehhöhe in Woche 52 durch Anwendung der Baseline-AGV auf die letzte verfügbare Messung der Körpergröße berechnet.

Der pU legt eine Post-hoc-Auswertung der z-Scores mit einer deutschen Referenzpopulation vor (vgl. Kapitel 2.3.2: Körpergröße). Im Vergleich zu Baseline erreichten die Patientinnen und Patienten des Placebo-Arms in Studienwoche 52 eine Veränderung des z-Scores um -0,01 im LS Mean, die des Vosoritid-Arms um 0,28 (Tabelle 15). Der Unterschied der Veränderungen zwischen den Studienarmen beträgt 0,28 im LS Mean und ist statistisch signifikant. Die Auswertung auf Basis der US-amerikanischen Referenzbevölkerung zeigt ein vergleichbares Ergebnis.

Tabelle 15: Veränderung der Stehhöhe (z-Score) zu Studienwoche 52 in Studie 301, FAS

Studie 301 Veränderung der Stehhöhe	Vosoritid N = 60	Placebo N = 61
Deutsche Referenzpopulation¹⁾		
Z-Score zu Baseline		
Mittelwert (SD)	-5,69 (1,11)	-5,68 (1,09)
Median (min; max)	-5,84 (-7,7; -1,4)	-5,51 (-8,4; -3,3)
Z-Score zu Woche 52		
Mittelwert (SD)	-5,39 (1,11)	-5,66 (1,05)
Median (min; max)	-5,44 (-7,6; -1,3)	-5,59 (-8,1; -3,5)
Veränderung Woche 52 zu Baseline ²⁾		
LS Mean [95%-KI]	0,28 [0,20; 0,35]	-0,01 [-0,09; 0,07]
Gruppenunterschied der Veränderungen ²⁾		
LS Mean [95%-KI]		0,28 [0,19; 0,37]
p-Wert		< 0,0001
US-amerikanische Referenzpopulation		
Z-Score zu Baseline		
Mittelwert (SD)	-5,13 (1,11)	-5,14 (1,07)
Median (min; max)	-5,27 (-7,7; -1,1)	-5,15 (-7,9; -2,7)
Z-Score zu Woche 52		
Mittelwert (SD)	-4,89 (1,09)	-5,14 (1,09)
Median (min; max)	-4,86 (-7,5; -1,1)	-5,11 (-7,8; -2,8)
Veränderung Woche 52 zu Baseline ²⁾		
LS Mean [95%-KI]	0,27 [0,18; 0,36]	-0,01 [-0,10; 0,09]
Gruppenunterschied der Veränderungen ²⁾		
LS Mean [95%-KI]		0,28 [0,17; 0,39]
p-Wert		< 0,0001

¹⁾ Angaben aus Modul 4 – post hoc berechnet.

²⁾ ANCOVA mit Stratifizierungsstratum und Behandlungsgruppe als fixe Effekte sowie Alter zu Baseline, AGV zu Baseline und Stehhöhe (z-Score) zu Baseline als Kovariaten.

Abkürzungen: AGV: Jährliche Wachstumsgeschwindigkeit; ANCOVA: Kovarianzanalyse; FAS: Full Analysis Set; KI: Konfidenzintervall; LS: Least Square; SD: Standardabweichung.

Sensitivitätsanalysen und ergänzende Analysen

Die Ergebnisse zeigen sich in den Sensitivitätsanalysen und ergänzenden Analysen als robust:

1. Washout-Modell mit LS Mean Difference = 0,27 (95%-KI: [0,17; 0,38]), $p < 0,0001$.
2. Studienabschluss-Modell (N = 58 in Vosoritid-Arm) mit LS Mean Difference = 0,28 (95%-KI: [0,17; 0,38]), $p < 0,0001$.
3. Notfallmedikation-Modell (N = 60 im Placebo-Arm) mit LS Mean Difference = 0,27 (95%-KI: [0,16; 0,37]), $p < 0,0001$.

Subgruppenanalysen

Es wurden keine Interaktionstests vorgelegt. Der pU verweist vor dem Hintergrund in Modul 4 des Nutzendossiers auf ein Forest Plot der Subgruppenanalysen. Aufgrund fehlender Interaktionstests richtete sich die Bewertung daher danach, inwiefern sich in den Subgruppen ($n \geq 10$) gleichgerichtete Effekte zeigen und inwiefern die Konfidenzintervalle der Subgruppen (auch bei gleichgerichteten Effekten) überlappen.

Die Subgruppen zeigen auf Basis einer deskriptiven Beurteilung des Forest Plots einen positiven Effekt zugunsten von Vosoritid. Ausnahme bildet die Subgruppe „Tanner-Stadium > 1“, in der sich kein Unterschied zwischen den Studienarmen zeigt. Die Konfidenzintervalle in den Subgruppen „AGV ≤ 3,5 cm/Jahr“, „Alter ≥ 11 bis < 15 Jahre“ sowie Region „Rest der Welt“ umfassen zudem auch negative Werte und damit auch mögliche Effekte zugunsten von Placebo.

CBCL

Durch die Einführung des CBCL mit Protokoll-Amendment 3 liegen für die Version 1,5–5 Jahre Baseline-Werte nur von 3 Kindern aus dem Placebo-Arm und 6 aus dem Vosoritid-Arm vor. Für die Version 6–18 Jahre beläuft sich dies auf 6 Personen aus dem Placebo-Arm und 4 aus dem Vosoritid-Arm. Auch wenn die Anzahl der Kinder und Jugendlichen in den entsprechenden Altersgruppen zu Baseline nicht berichtet ist, ist von keinen ausreichenden Rücklaufquoten auszugehen. Die Ergebnisse können daher nicht berücksichtigt werden.

3.4 Lebensqualität

PedsQL

Der PedsQL wurde zur Erhebung der allgemeinen gesundheitsbezogenen Lebensqualität eingesetzt. Das Instrument wurde selbstberichtet für Kinder ab 8 Jahren (und fremdberichtet für Kinder von 5 bis 18 Jahren) erfasst. An Studientag 1 waren 37 Kinder aus dem Placebo-Arm und 29 aus dem Vosoritid-Arm ≥ 8 Jahre alt (vgl. Tabelle 11). Mit 35 bzw. 28 Kindern weist der PedsQL in der selbstberichteten Version daher eine sehr hohe Vollständigkeit zu Baseline und mit 33 bzw. 25 Kindern eine hohe Vollständigkeit in der Berechnung der Veränderung zu Woche 52 auf.

Im Placebo-Arm zeigten die Studienteilnehmenden mit 0,0 Punkten im Median von Baseline zu Woche 52 keine numerische Veränderung, während im Vosoritid-Arm eine numerische Verbesserung um 1,1 Punkte beobachtet wurde. Die Berechnung von Gruppenunterschieden der Veränderungen wurden weder geplant und durchgeführt.

Die Ergebnisse der Elternversion sind im Anhang berichtet.

Tabelle 16: Veränderung im PedsQL (selbstberichtet) zu Studienwoche 52 in Studie 301, FAS

Studie 301 Veränderung im PedsQL ¹⁾²⁾ (selbstberichtet)	Vosoritid N = 29 ³⁾	Placebo N = 37 ³⁾
<i>Gesamtwert zu Baseline</i>	n = 28	n = 35
Mittelwert (SD)	74,07 (11,87)	75,32 (14,98)
Median (min; max)	74,46 (50,0; 93,5)	73,91 (42,4; 96,7)
<i>Gesamtwert zu Woche 52</i>	n = 34	n = 43
Mittelwert (SD)	75,94 (12,26)	71,33 (13,15)
Median (min; max)	78,81 (45,7; 97,8)	70,65 (47,8; 96,7)
<i>Veränderung Woche 52 zu Baseline</i>	n = 25	n = 33
Mittelwert (SD)	0,85 (13,80)	-2,62 (15,06)
Median (min; max)	1,09 (-32,6; 34,8)	0,00 (-39,8; 22,8)

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

²⁾ Erhoben bei Kindern ab 8 Jahren.

³⁾ An Studientag 1 waren 37 Kinder aus dem Placebo-Arm und 29 aus dem Vosoritid-Arm ≥ 8 Jahre alt. Erlebte ein Kind im Studienverlauf seinen 8. Geburtstag, wurde der selbstberichtete Fragebogen in der nächsten Studienvisite erfasst.

Abkürzungen: FAS: Full Analysis Set; PedsQL: Pediatric Quality of Life Inventory; SD: Standardabweichung.

QoLISSY

Mit dem QoLISSY wurde die Lebensqualität von Kindern und Jugendlichen mit Kleinwuchs erfasst. Analog zum PedsQL gab es ab 8 Jahren eine Version zur selbstberichteten Erfassung (und für alle Studienteilnehmenden eine Version zur fremdberichteten Erfassung). Auffällig ist, dass zu Baseline Ergebnisse für 30 Personen im Vosoritid-Arm berichtet werden, während laut Baseline-Charakteristika (siehe Tabelle 11) 29 Studienteilnehmende mindestens 8 Jahre alt waren.

Es konnten für 35 von 37 Kindern aus dem Placebo-Arm und 26 der 29 Kinder aus dem Vosoritid-Arm, die zu Baseline ≥ 8 Jahre alt waren, eine Veränderung berechnet werden. Bis Woche 52 zeigten die Studienteilnehmenden aus dem Placebo-Arm im Median eine numerische Verbesserung um 1,4 Punkte und im Vosoritid-Arm ebenfalls eine numerische Verbesserung um 0,7 Punkte. Die Berechnung von Gruppenunterschieden der Veränderungen wurden weder für die selbst- noch die fremdberichtete Version geplant und durchgeführt.

Die Ergebnisse der Elternversion sind im Anhang berichtet.

Tabelle 17: Veränderung im QoLISSY (selbstberichtet) zu Studienwoche 52 in Studie 301, FAS

Studie 301 Veränderung im QoLISSY ¹⁾²⁾ (selbstberichtet)	Vosoritid N = 29 ³⁾	Placebo N = 37 ³⁾
<i>Gesamtwert zu Baseline</i>	n = 30 ⁴⁾	n = 36
Mittelwert (SD)	64,59 (17,57)	66,40 (16,05)
Median (min; max)	66,84 (20,5; 92,0)	66,50 (21,2; 90,6)
<i>Gesamtwert zu Woche 52</i>	n = 36	n = 44
Mittelwert (SD)	67,39 (16,41)	64,68 (19,14)
Median (min; max)	69,62 (26,0; 91,3)	67,71 (9,0; 91,3)
<i>Veränderung Woche 52 zu Baseline</i>	n = 26	n = 35
Mittelwert (SD)	4,34 (14,42)	-0,88 (19,02)
Median (min; max)	0,69 (-15,8; 41,0)	1,39 (-63,9; 39,2)

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

²⁾ Erhoben bei Kindern ab 8 Jahren.

³⁾ An Studientag 1 waren 37 Kinder aus dem Placebo-Arm und 29 aus dem Vosoritid-Arm ≥ 8 Jahre alt. Erlebte ein Kind im Studienverlauf seinen 8. Geburtstag, wurde der selbstberichtete Fragebogen in der nächsten Studienvisite erfasst.

⁴⁾ Es ist unklar, warum Ergebnisse für 30 Kinder berichtet werden, wohingegen laut Tabelle 11 29 im Studienarm ≥ 8 Jahre alt waren.

Abkürzungen: FAS: Full Analysis Set; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; SD: Standardabweichung.

3.5 Sicherheit

Im folgenden Kapitel sind Ergebnisse der Sicherheitsauswertungen für die kontrollierte Behandlungsperiode bis Studienende dargestellt.

UE wurden ab dem Zeitpunkt der ersten Verabreichung bis zur letzten Studienvisite 4 Wochen nach der letzten Verabreichung des Studienmedikaments / Visite bei vorzeitigem Abbruch ausgewertet (siehe Kapitel 2.3.4). Die mediane Behandlungsdauer beträgt jeweils 365 Tage in den Studienarmen (siehe Tabelle 10).

Die zusammenfassende Darstellung der UE ist in Tabelle 18 abgebildet. UE jeglichen Schweregrads traten bei 60 Personen (98 %) aus dem Placebo-Arm und 59 Personen (98 %) aus dem Vosoritid-Arm auf. Davon waren jeweils 3 Ereignisse mit einem CTCAE-Schweregrad ≥ 3 eingestuft und 4 Ereignisse (7 %) im Placebo-Arm bzw. 3 (5 %) im Vosoritid-Arm als SUE klassifiziert. Es führte ein UE im Vosoritid-Arm zum Abbruch der Studienmedikation. Effektschätzer zu Gruppenunterschieden wurden nicht vorgelegt.

Tabelle 18: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE bis Studienende in Studie 301, Sicherheitspopulation

Studie 301 Zusammenfassung der UE (Personen mit ≥ 1 Ereignis)	Vosoritid N = 60 n (%)	Placebo N = 61 n (%)	Effektschätzer [95%-KI] p-Wert
UE jeglichen Schweregrads	59 (98,3)	60 (98,4)	.. ¹⁾
UE CTCAE-Grad ≥ 3	3 (5,0)	3 (4,9)	k. A. ²⁾
SUE	3 (5,0)	4 (6,6)	k. A. ²⁾
UE, das zum Abbruch der Studienmedikation führte	1 (1,7)	0	k. A. ²⁾

¹⁾ Patientenrelevanz von Laborparametern unklar.

²⁾ Es wurden keine Effektschätzer vorgelegt.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; k. A.: keine Angabe; KI: Konfidenzintervall; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

Unerwünschte Ereignisse jeglichen Schweregrades

Nahezu alle Studienteilnehmenden aus dem Placebo-Arm (98 %) und aus dem Vosoritid-Arm (98 %) erlebten im Studienverlauf ein UE jeglichen Schweregrads. Die am häufigsten vertretenen Systemorganklassen waren „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (82 bzw. 90 %) und „Infektionen und parasitäre Erkrankungen“ (75 bzw. 63 %). Unterschiede ≥ 10 Personen zwischen den Studienarmen wurden in den PT „Reaktion am Verabreichungsort“ und „Schwellung an der Injektionsstelle“ mit jeweils höheren Inzidenzen im Vosoritid-Arm festgestellt (siehe Tabelle 19).

Es wurden vor Entblindung Preferred Terms identifiziert, die als Symptome der Grunderkrankung angesehen wurden (siehe Kapitel 2.3.4). 34 Personen (56 %) aus dem Placebo-Arm und 35 (58 %) aus dem Vosoritid-Arm erlebten ein UE, das in Zusammenhang mit ACH gesetzt wurde. Die betreffenden Preferred Terms zeigen jeweils eine vergleichbare Ereignisanzahl in den Studienarmen mit Ausnahme der PT „Arthralgie“ und „Ohrschmerzen“ mit numerisch leicht mehr Ereignissen in der Vosoritid-Gruppe sowie im PT „Kopfschmerzen“ mit numerisch leicht mehr Ereignissen in der Placebo-Gruppe.

Tabelle 19: UE mit Inzidenz ≥ 10 % nach MedDRA-Systemorganklassen und Preferred Terms in einem Studienarm der Studie 301, Sicherheitspopulation

Studie 301 UE mit Inzidenz ≥ 10 %	Vosoritid N = 60	Placebo N = 61
<i>MedDRA-Systemorganklasse Preferred Term</i>	<i>n (%)</i>	<i>n (%)</i>
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	54 (90,0)	50 (82,0)
Reaktion am Verabreichungsort	44 (73,3)	29 (47,5)
Erythem an der Injektionsstelle	41 (68,3)	40 (65,6)
Schwellung an der Injektionsstelle	23 (38,3)	6 (9,8)
Fieber	10 (16,7)	13 (21,3)
Urtikaria an der Injektionsstelle	8 (13,3)	2 (3,3)
Blauer Fleck an der Injektionsstelle	5 (8,3)	8 (13,1)
Blutung an der Injektionsstelle	2 (3,3)	7 (11,5)

Studie 301 UE mit Inzidenz \geq 10 %	Vosoritid N = 60	Placebo N = 61
MedDRA-Systemorganklasse Preferred Term	<i>n (%)</i>	<i>n (%)</i>
Infektionen und parasitäre Erkrankungen	38 (63,3)	46 (75,4)
Nasopharyngitis	16 (26,7)	18 (29,5)
Infektion der oberen Atemwege	8 (13,3)	10 (16,4)
Ohreninfektion ¹⁾	6 (10,0)	6 (9,8)
Grippe	6 (10,0)	3 (4,9)
Otitis media ¹⁾	6 (10,0)	6 (9,8)
Erkrankungen des Nervensystems	22 (36,7)	21 (34,4)
Kopfschmerzen ¹⁾	14 (23,3)	16 (26,2)
Erkrankungen des Gastrointestinaltrakts	20 (33,3)	24 (39,3)
Erbrechen	16 (26,7)	12 (19,7)
Diarrhö	6 (10,0)	2 (3,3)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	16 (26,7)	13 (21,3)
Arthralgie ¹⁾	9 (15,0)	4 (6,6)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	15 (25,0)	13 (21,3)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	14 (23,3)	16 (26,2)
Husten	7 (11,7)	8 (13,1)
Oropharyngeale Schmerzen	6 (10,0)	4 (6,6)
Erkrankungen des Ohrs und des Labyrinths	11 (18,3)	8 (13,1)
Ohrschmerzen ¹⁾	6 (10,0)	3 (4,9)
Erkrankungen der Haut und des Unterhautzellgewebes	9 (15,0)	8 (13,1)
Untersuchungen	8 (13,3)	3 (4,9)
Blutdruck erniedrigt	7 (11,7)	3 (4,9)
Stoffwechsel- und Ernährungsstörungen	3 (5,0)	8 (13,1)
Vitamin-D-Mangel	3 (5,0)	7 (11,5)

¹⁾ Preferred Term wurde a priori als ACH-bezogen identifiziert (siehe Kapitel 2.3.4). Alle Ereignisse dieses Preferred Terms werden im Studienbericht der Grunderkrankung ACH zugeordnet.

Abkürzungen: ACH: Achondroplasie; MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

Schwere unerwünschte Ereignisse

Jeweils 3 Personen aus dem Placebo- (5 %) und aus dem Vosoritid-Arm (5 %) erlebten im Studienverlauf ein UE mit dem CTCAE-Schweregrad 3. In keiner Systemorganklasse wurden Ereignisse bei mehr als 2 Personen (< 5 %) verzeichnet.

Es traten keine UE mit CTCAE-Schweregrad \geq 4 auf.

Schwerwiegende unerwünschte Ereignisse

Bei 4 Personen (7 %) aus dem Placebo-Arm und 3 (5 %) aus dem Vosoritid-Arm traten SUE auf. In keiner Systemorganklasse wurden Ereignisse bei mehr als 2 Personen (< 5 %) verzeichnet.

Unerwünschte Ereignisse, die zum Abbruch der Studienmedikation führten

Im Studienverlauf führte ein UE bei einer Person im Vosoritid-Arm zum Abbruch der Studienmedikation.

UE von besonderem Interesse

Im SAP wurden Ereignisse von besonderem Interesse definiert. Die Ergebnisse sind in Tabelle 20 präsentiert. Es traten bei 50 Personen im Placebo-Arm und 51 im Vosoritid-Arm Reaktionen an der Einstichstelle auf. 9 Personen mehr als im Placebo-Arm erlitten im Vosoritid-Arm eine Überempfindlichkeitsreaktion und 5 Personen mehr eine Hypotension. Es wurde eine Fraktur als UE von besonderem Interesse im Vosoritid-Arm aufgezeichnet. Keine Ereignisse wurden hinsichtlich „Änderung der Herzfrequenz“, „Avaskuläre Nekrose oder Knochennekrose“ oder „Verrutschte Oberschenkel-epiphyse“ festgestellt.

Im SAP wurde zusätzlich der SMQ „Algorithmische Anaphylaxie“ als UE von besonderem Ereignis festgelegt. Es wurden keine Ergebnisse für diesen SMQ identifiziert, jedoch berichtet der Studienbericht hinsichtlich der Überempfindlichkeit, dass kein Ereignis die NIAID/FAAN-Kriterien für Anaphylaxie erfüllte.

Tabelle 20: UE von besonderem Interesse in Studie 301, Sicherheitspopulation

Studie 301 UE von besonderem Interesse (Personen mit ≥ 1 Ereignis)	Vosoritid N = 60 n (%)	Placebo N = 61 n (%)
Reaktion an der Einstichstelle	51 (85,0)	50 (82,0)
Überempfindlichkeit	16 (26,7)	7 (11,5)
Hypotension	8 (13,3)	3 (4,9)
Frakturen	1 (1,7)	0
Änderung der Herzfrequenz	0	0
Avaskuläre Nekrose oder Knochennekrose	0	0
Verrutschte Oberschenkelepiphyse	0	0
Algorithmische Anaphylaxie ¹⁾	k. A.	k. A.

¹⁾ Laut Fließtext des Studienberichts erfüllte kein Ereignis der Überempfindlichkeit die NIAID/FAAN-Kriterien für Anaphylaxie. Es wurde keine Angabe für den prädefinierten SMQ identifiziert.

Abkürzungen: k. A.: keine Angabe; NIAID/FAAN: National Institute of Allergy and Infectious Disease and Food Allergy and Anaphylaxis Network; SMQ: Standard MedDRA Query; UE: Unerwünschtes Ereignis.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Vosoritid

Gemäß Fachinformation ist Vosoritid (Voxzogo®) zugelassen zur Behandlung der Achondroplasie (ACH) bei Patientinnen und Patienten ab dem 2. Lebensjahr, bei denen die Epiphysen noch nicht geschlossen sind. Die Diagnose Achondroplasie sollte durch entsprechende Gentests bestätigt werden. Die Therapie sieht die tägliche subkutane Injektion vor. Die übliche Dosis wird mit 15 µg/kg Körpergewicht angegeben. Laut Fachinformation soll die Behandlung mit Vosoritid beendet werden, sobald eine Wachstumsgeschwindigkeit von < 1,5 cm/Jahr und das Schließen der Epiphysen nachgewiesen ist [7].

Die vorliegende Studie 301 rekrutierte Kinder und Jugendliche im Alter von 5 bis < 18 Jahren. Eine Wachstumsgeschwindigkeit < 1,5 cm/Jahr oder ein Schließen der Epiphysen waren Ausschlusskriterien für eine Studienteilnahme, sodass die Anforderungen aus der Zulassungspopulation erfüllt sind. Die älteste teilnehmende Person war zu Baseline 14,9 Jahre alt. Inwiefern dies die obere Altersspanne der Zulassungspopulation darstellen könnte oder für das obere Altersspektrum der Zulassungspopulation keine Daten gewonnen werden konnten, wird seitens des pU nicht diskutiert und bleibt daher offen.

In der Nutzenbewertung konnten keine Informationen zur Zulassungspopulation der 2 bis < 5-Jährigen berücksichtigt werden. Die Wirksamkeit und Sicherheit von Vosoritid für diese Population wird im derzeit noch laufenden RCT BMN 111-206 untersucht, für welches zum Zeitpunkt dieser Nutzenbewertung keine Ergebnisse für die randomisierten Studienarme verfügbar waren. Die Fertigstellung dieser Studie wird auf Februar 2022 geschätzt [8].

An Studie 301 nahmen auch Studienzentren in Deutschland teil. Die Ergebnisse erscheinen daher auf den deutschen Versorgungskontext übertragbar.

4.2 Design und Methodik der Studie

Bei Studie 301 handelt es sich um eine randomisierte, doppelblinde, Placebo-kontrollierte, multizentrische Studie der Phase 3 im Parallelgruppendesign (1:1) mit zwei Behandlungsarmen. In der Studie wurde die Wirksamkeit und Sicherheit von Vosoritid gegenüber Placebo bei Kindern und Jugendlichen im Alter von 5 bis < 18 Jahren und genetisch nachgewiesener ACH untersucht. Primäres Zielkriterium war die Veränderung der jährlichen Wachstumsgeschwindigkeit von Baseline zu Woche 52.

An Studie 301 konnten Personen teilnehmen, die zuvor bereits mindestens 6 Monate in der Beobachtungsstudie BMN 111-901 erfasst wurden. Die Ein- und Ausschlusskriterien dieser Studie decken sich (mit Ausnahme der Altersspanne) mit denen der Studie 301. Die Beobachtungsstudie diente als Ausgangsstudie für mehrere Vosoritid-Studien. Dabei überlappen die Einschlusskriterien der Studie 301 (Altersgruppe 5 bis < 18 Jahre) mit denen der Studie BMN 111-202 (Altersgruppe 5 bis < 15 Jahre). Über den Zeitraum von ca. 1 Jahr liefen beide Studien parallel. Es wurden keine Informationen zur Entscheidung über eine Teilnahme an einer der beiden Studien identifiziert, sodass mögliche Selektionseffekte nicht ausgeschlossen werden können. Die Abbruchquote lag in der Beobachtungsstudie BMN 111-901 bei 23 %. Die Ausführungen zu den Gründen für den Studienabbruch lassen den Rückschluss zu, dass Personen mit einer höheren Compliance im Studienprogramm verblieben.

Für Studienteilnehmende, die nicht in die einarmige Extensionsstudie BMN 111-302 übergangen, war zudem eine Nachbeobachtung der Sicherheit über weitere 4 Wochen vorgesehen. Es gingen im Placebo-Arm alle Personen und im Vosoritid-Arm 58 von 60 Personen in die Extensionsstudie über.

Die übrigen 2 Personen im Vosoritid-Arm brachen die Studienteilnahme ab. Damit fand die 4-wöchige Nachbeobachtung der Sicherheit für keine Person innerhalb der Studie 301 statt.

Es wurden in den Placebo-Arm der Studie 61 Kinder und Jugendliche, in den Vosoritid-Arm 60 Kinder und Jugendliche randomisiert. Die Randomisierung in die Studienarme Vosoritid und Placebo erfolgte stratifiziert für „Geschlecht“ (männlich oder weiblich) und „Pubertäre Entwicklung“ (Tanner-Stadium I oder > I). Ursprünglich war eine Stratifizierung für „Alter“ (< 11 oder ≥ 11 Jahre) anstelle des Tanner-Stadiums vorgesehen. Dies wurde mit Protokoll-Amendment 2 vom 27.04.2017 ca. 4 Monate nach Studienstart am 12.12.2016 geändert. Zu diesem Zeitpunkt war bereits jeweils 1 Kind pro Studienarm in die Studie aufgenommen. Diese Kinder waren beide weiblich und ≤ 11 Jahre alt. In den Ergebnisauswertungen wurden sie entsprechend ihrer tatsächlichen Strata nach Tanner-Stadium und Geschlecht analysiert (d. h. weibliches Tanner-Stadium I). Auffällig ist eine unterschiedliche Verteilung im ehemaligen Stratifizierungsmerkmal Alter zu Baseline in den Studienarmen: Im Placebo-Arm betrug das mediane Alter 9,3 Jahre, im Vosoritid-Arm lag es mit 7,8 Jahren niedriger. Ebenso gab es im Vosoritid-Arm mehr Kinder im Alter von ≥ 5 bis < 8 Jahren (52 %) als im Placebo-Arm (39 %). Kein Unterschied zwischen den Studienarmen zeigt sich jedoch für die Gruppe der ≥ 11-Jährigen mit 21 % im Placebo-Arm und 20 % im Vosoritid-Arm. Somit sind auch für das ursprünglich geplante Stratifizierungsmerkmal „Alter“ (< 11 oder ≥ 11 Jahre) keine Unterschiede zwischen den Studienarmen zu beobachten. Das jüngere Ausgangsalter im Vosoritid-Arm und die damit vermutlich in Verbindung stehende ebenfalls niedrigere Stehhöhe zu Baseline (im Median 98,6 cm im Vergleich zu 104,6 cm im Placebo-Arm) erscheinen auch vor dem Hintergrund der gleichen Verteilungen im alterskorrigierten z-Score (im Median -5,2 im Placebo-Arm und -5,3 im Vosoritid-Arm) hinsichtlich der Ergebnisinterpretation der vorliegenden analytischen Auswertungen insgesamt vernachlässigbar. Hinsichtlich der deskriptiv ausgewerteten Endpunkte müssen diese Unterschiede zu Baseline jedoch berücksichtigt werden.

In den übrigen Baseline-Charakteristika zeigen sich gleiche Verteilungen in den Studienarmen. Gleiches gilt im Wesentlichen ebenfalls für die Begleitmedikationen und die Protokollverletzungen. Aufgrund des geringen Stichprobenumfangs erscheinen etwaige Unterschiede zwischen den Studienarmen nicht ungewöhnlich und können zufallsbedingt aufgetreten sein.

Die Therapie mit Vosoritid wurde in Studie 301 gemäß Fachinformation mit einer Dosis von 15,0 µg/kg täglich durchgeführt. Auch wenn Dosisunterbrechungen und -reduktionen nur im Zusammenhang mit UE berichtet werden und Angaben zu Dosiserhöhungen fehlen, spricht die mediane Dosis von 15,0 µg/kg/Tag für ein stabiles Behandlungsregime.

Die statistische Analyse wurde im SAP geplant. Das Datum des Datenbankschlusses ist nicht berichtet, sodass nicht bestimmt werden kann, ob der SAP vor Datenkenntnis erstellt wurde. Als Analysemodell wurde eine ANCOVA geplant. Angaben zu einer Überprüfung von Modellannahmen wurden nicht identifiziert. Ausführlich wurde der Umgang mit fehlenden Werten sowie darauf bezogene Sensitivitätsanalysen beschrieben. Auch wenn der Anteil fehlender Werte in der vorliegenden Analyse unklar ist, ermöglichen die Sensitivitätsanalysen eine abschließende Einschätzung zu möglichen Verzerrungen. Für die eingesetzten Messinstrumente wurden lediglich deskriptive Auswertungen vorgesehen. Auch mit Modul 4 des Nutzendossiers wurden keine Analysen nachgereicht.

Der pU legt mit Modul 4 Subgruppenanalysen, jedoch keine Interaktionstests vor. Da lediglich die statistische Analyse eines Endpunkts vorliegt, können keine übergreifenden Aussagen getroffen werden. Auffällig erscheint jedoch, dass sich für diesen Endpunkt kein Effekt in der Subgruppe „Tanner-Stadium > 1“ und in den negativen Bereich reichende Konfidenzintervalle in den

Subgruppen „AGV $\leq 3,5$ cm/Jahr“ sowie „Alter ≥ 11 bis < 15 Jahre“ zeigen. Dies könnte auf einen abnehmenden Effekt von Vosoritid bei älteren Kindern und Jugendlichen, die sich dem Wachstumsstopp nähern, hindeuten.

Das Verzerrungspotential auf Studienebene wird insgesamt mit unklar bewertet. Maßgeblich hierfür sind die beschriebenen Unklarheiten zum SAP.

4.3 Mortalität

Todesfälle wurden im Rahmen der Sicherheitserfassung dokumentiert. Im Verlauf der Studie 301 ist keine Person verstorben.

4.4 Morbidität

Zur Bewertung der Wirksamkeit in der Kategorie „Morbidität“ wurde der Endpunkt „Körpergröße (z-Score)“ herangezogen.

Die vom pU vorgelegten Endpunkte „Wachstumsgeschwindigkeit“, „Verhältnis von oberem zu unterem Körpersegment“ und „Verhältnis der Körperproportionen“ wurden hingegen nicht berücksichtigt, da eine Patientenrelevanz nicht nachvollzogen werden konnte. Ergebnisse aus dem Messinstrument „Child Behavior Checklist“ konnten aufgrund niedriger Rücklaufzahlen zu Baseline nicht herangezogen werden.

Körpergröße

Die Veränderung der Stehhöhe wurde als z-Score ermittelt. Vor dem Hintergrund der besseren Übertragbarkeit auf den deutschen Versorgungskontext wurde eine Post-hoc-Auswertung mit einer deutschen Referenzpopulation durchgeführt. Es zeigte sich für den Vosoritid-Arm eine statistisch signifikant stärkere Annäherung an den Normwert um 0,28 im Vergleich zum Placebo-Arm. Auch wenn der Anteil imputierter Daten in dieser Analyse nicht berichtet ist, wird dieses Ergebnis aufgrund der vorgelegten Sensitivitätsanalysen als robust bewertet. Zu Woche 52 erreichten damit die Studienteilnehmenden im Vosoritid-Arm im Median einen z-Score von -5,44 im Vergleich zu -5,59 im Placebo-Arm.

In Folge der beschriebenen Unklarheiten zum SAP wird das Verzerrungspotential für diesen Endpunkt mit unklar bewertet.

Zusammenfassende Einschätzung zur Morbidität

Es konnten lediglich Ergebnisse zum Endpunkt „Körpergröße“ berücksichtigt werden. Für diesen zeigte sich bei unklarem Verzerrungspotential ein statistisch signifikanter Vorteil zugunsten von Vosoritid gegenüber Placebo. Sofern Verzerrungen existieren, würden diese zugunsten von Vosoritid erwartet werden. Die Ergebnissicherheit ist aufgrund der geringen Studiengröße grundsätzlich limitiert.

4.5 Lebensqualität

Für die Bewertung der Lebensqualität konnten in der Nutzenbewertung die Ergebnisse des PedsQL und des QoLISSY in der jeweils selbstberichteten Version berücksichtigt werden. Diese kamen bei Kindern und Jugendlichen ab 8 Jahren zum Einsatz. Der WeeFIM wurde hingegen nicht herangezogen, da die Anwendung des Instruments bei älteren Kindern und Jugendlichen nicht begründet wurde. Eine Auswertung für die Gruppe der 5- bis 7-Jährigen liegt nicht vor.

PedsQL

Der PedsQL wurde zur Erhebung der allgemeinen gesundheitsbezogenen Lebensqualität eingesetzt. Eine Veränderung konnte für 89 % der Personen im Placebo-Arm bzw. 86 % im Vosoritid-Arm berechnet werden. Gründe für fehlende Werte wurden nicht identifiziert. Vor dem Hintergrund scheint problematisch, dass keine Erhebung des Instruments in der Studienvisite bei vorzeitigem Abbruch vorgesehen war. Es brachen 2 Personen aus dem Vosoritid-Arm im Studienverlauf ihre Teilnahme ab, sodass hierauf jedoch nicht der Großteil der fehlenden Werte zurückzuführen sein kann. In der Gesamtschau wird das Verzerrungspotential des Endpunkts mit unklar bewertet.

Im Placebo-Arm zeigten die Studienteilnehmenden von Baseline zu Woche 52 im Median mit 0 Punkten keine numerische Veränderung, während im Vosoritid-Arm eine numerische Verbesserung um 1 Punkt beobachtet wurde. Die beobachteten Veränderungen lassen sich aufgrund der deskriptiven Auswertung nicht hinsichtlich eines möglichen Gruppenunterschieds interpretieren. Auffällig sind die Abweichungen zwischen Mittelwert und Median, die für den Placebo-Arm im Mittelwert eine Verschlechterung um 3 Punkte zeigt.

Im Anhang wurden zusätzlich die Ergebnisse der Elternversionen des PedsQL dargestellt. Diese wurden für alle Kinder und Jugendlichen erfasst und über alle Altersgruppen hinweg gemeinsam ausgewertet. Im Gegensatz zur selbstberichteten Version sind somit auch die 5- bis 7-Jährigen erfasst. Auffällig ist hier, dass sich im Placebo-Arm im Median eine Verbesserung um 3 Punkte zeigt, während sich im Vosoritid-Arm eine Verschlechterung um 0,5 Punkte zeigt. Es ist unklar, inwiefern es sich hierbei um zufallsbedingte Schwankungen handelt, ob sich dies auf diskrepante Einschätzungen in der selbstberichteten und fremdberichteten Erfassung oder auf unterschiedliche Ergebnisse in der Gruppe der 5- bis 7-Jährigen zurückführen lässt. Eine separate Auswertung für diese jüngste Altersgruppe wurde nicht vorgelegt.

QoLISSY

Der QoLISSY diente der Erfassung der krankheitsspezifischen Lebensqualität. Der Bezugszeitraum umfasst die vergangenen 7 Tage. Mit zwei geplanten Erhebungen nach Baseline zu Woche 26 und 52 erlaubt der Fragebogen nur Aussagen für einen geringen Anteil der Studienzeit. Auch hier erfolgte keine Erhebung in der Studienvisite bei vorzeitigem Abbruch.

Es konnte für 95 % der Personen im Placebo-Arm und 90 % im Vosoritid-Arm eine Veränderung berechnet werden. Auch hier wurden keine Angaben zu den Gründen für fehlende Werte identifiziert, in dessen Konsequenz das Verzerrungspotential mit unklar bewertet wurde. Auffällig ist zudem, dass zu Baseline Ergebnisse für 30 Personen im Vosoritid-Arm berichtet werden, während nach Angaben der Baseline-Charakteristika 29 Studienteilnehmende mindestens 8 Jahre alt waren und damit die Analysepopulation bilden sollten.

Bis Woche 52 zeigten die Studienteilnehmenden aus dem Placebo- und dem Vosoritid-Arm im Median jeweils eine numerische Verbesserung um 1 Punkt. Bei Betrachtung des Mittelwerts zeigt sich hingegen eine Verschlechterung um 1 Punkt im Placebo-Arm und eine Verbesserung um 4 Punkte im Vosoritid-Arm. Analog zum PedsQL (siehe oben) wurde der Fragebogen ebenfalls in Form von Elternversionen eingesetzt, deren Ergebnisse im Anhang dargestellt sind. Hier zeigt sich für den Placebo-Arm ebenfalls im Median eine Verbesserung um 1 Punkt, während jedoch für den Vosoritid-Arm eine Verschlechterung um 2 Punkte beobachtet wurde. Da auch hier keine Analysen für die Altersgruppe 5–7 Jahre vorliegen, ist die Ursache dieser Diskrepanz unklar.

Zusammenfassende Einschätzung zur Lebensqualität

Für die Studienpopulation der 5- bis 7-Jährigen konnten keine Auswertungen für die Nutzenbewertung berücksichtigt werden. Die Ergebnisse des PedsQL und des QoLISSY in der Altersgruppe 8–17 Jahre liegen ausschließlich in deskriptiver Auswertung vor, sodass Gruppenunterschiede nicht bewertet werden können.

Für die Instrumente wurde ein unklares Verzerrungspotential festgestellt, dessen Auswirkung zugunsten oder zuungunsten von Vosoritid nicht bestimmbar ist. Die Ergebnissicherheit ist aufgrund der geringen Studiengröße grundsätzlich limitiert.

4.6 Sicherheit

Innerhalb der Behandlungsperiode bis Woche 52 trat nahezu bei allen Personen (jeweils 98 % pro Studienarm) mindestens ein UE auf. Jeweils 3 Ereignisse wurden mit einem Schweregrad 3 klassifiziert; es trat kein UE mit Schweregrad 4 auf. Mit 4 Ereignissen trat im Kontroll-Arm ein SUE mehr als im Interventionsarm mit 3 Ereignissen auf. Im Vosoritid-Arm führte ein UE zum Abbruch der Studienmedikation. Effektschätzer zu Gruppenunterschieden wurden nicht vorgelegt.

Hinsichtlich UE jedweden Schweregrads zeigten sich Unterschiede zwischen den Studienarmen von ≥ 10 Personen zwischen den Studienarmen in keiner Systemorganklasse, jedoch in den PT „Reaktion am Verabreichungsort“ und „Schwellung an der Injektionsstelle“ mit jeweils höheren Inzidenzen im Vosoritid-Arm.

Es wurden vor Entblindung Preferred Terms identifiziert, die als Symptome der Grunderkrankung angesehen wurden. Die Anteile der Personen, die ein UE in Zusammenhang mit ACH erlebten, war in den Studienarmen vergleichbar. Lediglich in den PT „Arthralgie“ und „Ohrschmerzen“ zeigten sich ungleiche Ereigniszahlen mit numerisch leicht mehr Ereignissen in der Vosoritid-Gruppe sowie im PT „Kopfschmerzen“ mit numerisch leicht mehr Ereignissen in der Placebo-Gruppe.

Im SAP wurden Ereignisse von besonderem Interesse definiert. Es zeigten sich keine wesentlichen Unterschiede zwischen Placebo- und Vosoritid-Arm hinsichtlich von Reaktionen an der Einstichstelle. Jedoch erlebten 9 Personen mehr im Vosoritid-Arm verglichen mit dem Placebo-Arm eine Überempfindlichkeitsreaktion und 5 Personen mehr eine Hypotension. Es wurde eine Fraktur als UE von besonderem Ereignis im Vosoritid-Arm aufgezeichnet. Keine Ereignisse wurden hinsichtlich „Änderung der Herzfrequenz“, „Avaskuläre Nekrose oder Knochennekrose“ oder „Verrutschte Oberschenkel-epiphyse“ festgestellt.

Im SAP wurde zusätzlich das SMQ „Algorithmische Anaphylaxie“ als UE von besonderem Interesse festgelegt; es wurden keine Ergebnisse für dieses SMQ identifiziert. Nach Angaben im Studienbericht erfüllte jedoch kein Ereignis die NIAID/FAAN-Kriterien für Anaphylaxie.

Zusammenfassende Einschätzung zur Sicherheit

Das Verzerrungspotential wurde in der Gesamtschau mit niedrig eingestuft. Insgesamt wurden nur geringe Unterschiede im Auftreten von UE zwischen den Studienarmen festgestellt, die sowohl numerisch zugunsten als auch zuungunsten von Vosoritid auftraten. Es wurden nur wenige schwere UE und SUE beobachtet.

Eine abschließende Bewertung des Sicherheitsprofils von Vosoritid im Vergleich zu Placebo ist aufgrund der geringen Stichprobe besonders in Hinblick auf schwere UE und SUE nicht möglich.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Vosoritid ist zugelassen zur Behandlung von Achondroplasie (ACH) bei Patientinnen und Patienten ab dem 2. Lebensjahr, bei denen die Epiphysen noch nicht geschlossen sind [7]. Die Nutzenbewertung von Vosoritid basiert auf der zulassungsbegründenden Studie 301, einer doppelblinden, kontrollierten, multizentrischen Studie der Phase 3, in der Kinder und Jugendliche zwischen 5 und 17 Jahren und mittels genetischer Testung nachgewiesener ACH randomisiert im Verhältnis 1:1 auf die Behandlungsarme 15 µg/kg Vosoritid und Placebo zugeteilt wurden.

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

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

Tabelle 21: Zusammenfassende Darstellung der Ergebnisse der Studie 301

Studie 301 Darstellung der Ergebnisse	Vosoritid (N = 60) vs. Placebo (N = 61) ¹⁾	Effekt
Mortalität		
Todesfälle	Keine Ereignisse aufgetreten.	n. b.
Morbidität		
Veränderung der Stehhöhe (z-Score) LS Mean Difference [95%-KI]; p-Wert ²⁾³⁾	0,28 [0,19; 0,37]; < 0,0001	↑
CBCL	- ⁴⁾	n. b.
Lebensqualität		
Veränderung im PedsQL-Gesamtwert zwischen Baseline und Studienwoche 52 ⁵⁾⁶⁾ Mittelwert (SD) Median (min; max)	n = 25 vs. n = 33 ⁷⁾ 0,85 (13,80) vs. -2,62 (15,06) 1,09 (-32,6; 34,8) vs. 0,00 (-39,8; 22,8)	n. b. ⁸⁾
Veränderung im QoLISSY-Gesamtwert zwischen Baseline und Studienwoche 52 ⁵⁾⁶⁾ Mittelwert (SD) Median (min; max)	n = 26 vs. n = 35 ⁷⁾ 4,34 (14,42) vs. -0,88 (19,02) 0,69 (-15,8; 41,0) vs. 1,39 (-63,9; 39,2)	n. b. ⁸⁾
Sicherheit		
UE CTCAE-Grad ≥ 3, n (%)	3 (5,0) vs. 3 (4,9)	n. b. ⁸⁾
SUE, n (%)	3 (5,0) vs. 4 (6,6)	n. b. ⁸⁾
UE, das zum Abbruch der Studie/Studienmedikation führte, n (%)	1 (1,7) vs. 0	n. b. ⁸⁾

Studie 301 Darstellung der Ergebnisse	Vosoritid (N = 60) vs. Placebo (N = 61)¹⁾	Effekt
<i>UE von besonderem Interesse, n (%)</i>		
Reaktion an der Einstichstelle	51 (85,0) vs. 50 (82,0)	n. b. ⁸⁾
Überempfindlichkeit	16 (26,7) vs. 7 (11,5)	n. b. ⁸⁾
Hypotension	8 (13,3) vs. 3 (4,9)	n. b. ⁸⁾
Frakturen	1 (1,7) vs. 0	n. b. ⁸⁾
Änderung der Herzfrequenz	Keine Ereignisse aufgetreten.	n. b. ⁸⁾
Avaskuläre Nekrose oder Knochennekrose	Keine Ereignisse aufgetreten.	n. b. ⁸⁾
Verrutschte Oberschenkelepiphyse	Keine Ereignisse aufgetreten.	n. b. ⁸⁾
Algorithmische Anaphylaxie	k. A.	n. b. ⁹⁾

¹⁾ Bildet sowohl das FAS (definiert nach dem Intention-to-Treat-Prinzip, umfasst alle randomisierten Studienteilnehmenden mit einer unterzeichneten Einverständniserklärung) als auch die Sicherheitspopulation (definiert als alle Personen des FAS, die darüber hinaus mindestens eine Dosis ihrer Studienintervention erhielten).

²⁾ ANCOVA mit Stratifizierungsstratum und Behandlungsgruppe als fixe Effekte sowie Alter zu Baseline, AGV zu Baseline und Stehhöhe (z-Score) zu Baseline als Kovariaten.

³⁾ Deutsche Referenzpopulation; post hoc berechnet.

⁴⁾ Aufgrund niedriger Rücklaufquoten konnte die Auswertung nicht berücksichtigt werden.

⁵⁾ Wertebereich zwischen 0 bis 100 Punkten, wobei höhere Werte eine bessere Lebensqualität repräsentieren.

⁶⁾ Selbstberichtet erhoben bei Kindern ab 8 Jahren.

⁷⁾ An Studientag 1 waren 37 Kinder aus dem Placebo- und 29 aus dem Vosoritid-Arm ≥ 8 Jahre alt.

⁸⁾ Es liegen ausschließlich deskriptive Auswertungen vor.

⁹⁾ Es wurden keine Angaben identifiziert.

Abkürzungen: AGV: Jährliche Wachstumsgeschwindigkeit; ANCOVA: Kovarianzanalyse; CBCL: Child Behavior Checklist; CTCAE: Common Terminology Criteria for Adverse Events; FAS: Full Analysis Set; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Square; n. b.: nicht bewertbar; PedsQL: Pediatric Quality of Life Inventory; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; SD: Standardabweichung; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. **Achenbach TM.** Child Behavior Checklist. In: Kreutzer JS, Deluca J, Caplan B (Eds). Encyclopedia of Clinical Neuropsychology. New York: Springer; 2011. S. 546-552.
2. **Achenbach TM, Rescorla LA.** Manual for ASEBA School-Age Forms and Profiles. Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2001.
3. **Achenbach TM, Rescorla LA.** Manual for the ASEBA Preschool Forms and Profiles: an integrated system of multi-informant assessment; child behavior checklist for ages 1 1/2 - 5; language development survey; caregiver-teacher report form. Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2000.
4. **BioMarin.** A multicenter, multinational clinical assessment study for pediatric patients with achondroplasia: BMN 111-901; clinical study report [unveröffentlicht]. 2019.
5. **BioMarin.** A phase 3 randomized, double-blind, placebo-controlled, multicenter study to evaluate the efficacy and safety of BMN 111 in children with achondroplasia: BMN 111-301; Appendices 16, 16.1.1 Protocol and Protocol Amendments [unveröffentlicht]. 2019.
6. **BioMarin.** A phase 3 randomized, double-blind, placebo-controlled, multicenter study to evaluate the efficacy and safety of BMN 111 in children with achondroplasia: BMN 111-301; clinical study report [unveröffentlicht]. 2020.
7. **Biomarin.** VOXZOGO 0,4 mg/- 0,56 mg/- 1,2 mg Pulver und Lösungsmittel zur Herstellung einer Injektionslösung [online]. 08.2021. [Zugriff: 06.10.2021]. URL: <https://www.fachinfo.de>.
8. **BioMarin Pharmaceutical.** A clinical trial to evaluate the safety and efficacy of BMN 111 in infants and young children with achondroplasia [online]. NCT03583697. In: ClinicalTrials.gov. 29.06.2021. [Zugriff: 26.11.2021]. URL: <https://ClinicalTrials.gov/show/NCT03583697>.
9. **Bloemeke J, Silva N, Bullinger M, Witt S, Dörr HG, Quitmann J.** Psychometric properties of the quality of life in short statured youth (QoLISSY) questionnaire within the course of growth hormone treatment. Health Qual Life Outcomes 2019;17(1):49.
10. **Centers for Disease Control and Prevention, National Center for Health Statistics.** Data table of stature-for-age charts [online]. Hyattsville (USA): US Department of Health & Human Services; 2001. [Zugriff: 05.11.2021]. URL: https://www.cdc.gov/growthcharts/html_charts/statage.htm.
11. **Chan ML, Qi Y, Larimore K, Cherukuri A, Seid L, Jayaram K, et al.** Pharmacokinetics and exposure-response of vosoritide in children with achondroplasia. Clin Pharmacokinet 2021; Online ahead of print.
12. **Chaplin J, European QoLISSY Group.** Quality of life in short stature youth: the QoLISSY questionnaire user's manual. Berlin: Pabst Science Publishers; 2013.
13. **Connelly M, Rapoff MA.** Assessing health-related quality of life in children with recurrent headache: reliability and validity of the PedsQLTM 4.0 in a pediatric headache sample. J Pediatr Psychol 2006;31(7):698-702.

14. **European Medicines Agency (EMA).** Voxzogo (vovoritide): European public assessment report EMEA/H/C/005475 [online]. 09.09.2021. Amsterdam (NED): EMA. [Zugriff: 06.10.2021]. URL: https://www.ema.europa.eu/en/documents/assessment-report/voxzogo-epar-public-assessment-report_en.pdf.
15. **Ireland P, Johnston LM.** Measures of self-care independence for children with osteochondrodysplasia: a clinimetric review. *Phys Occup Ther Pediatr* 2012;32(1):80-96.
16. **McCabe MA, Granger CV.** Content validity of a pediatric functional independence measure. *Appl Nurs Res* 1990;3(3):120-122.
17. **Merker A, Neumeyer L, Hertel NT, Grigelioniene G, Makitie O, Mohnike K, et al.** Growth in achondroplasia: development of height, weight, head circumference, and body mass index in a European cohort. *Am J Med Genet A* 2018;176(8):1723-1734.
18. **Ottenbacher KJ, Msall ME, Lyon N, Duffy LC, Ziviani J, Granger CV, et al.** The WeeFIM instrument: its utility in detecting change in children with developmental disabilities. *Arch Phys Med Rehabil* 2000;81(10):1317-1326.
19. **Pavao SL, Silva FP, Dusing SC, Rocha NA.** Clinical tools designed to assess motor abilities in children with cerebral palsy. *Dev Neurorehabil* 2017;20(3):149-159.
20. **Robert Koch-Institut (RKI).** Beiträge zur Gesundheitsberichterstattung des Bundes: Referenzperzentile für anthropometrische Maßzahlen und Blutdruck aus der Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland (KiGGS); 2. erweiterte Auflage [online]. Berlin (GER): RKI; 2013. [Zugriff: 05.11.2021]. (Gesundheitsberichterstattung des Bundes). URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/KiGGS_Referenzperzentile.pdf?__blob=publicationFile.
21. **Uniform Data System for Medical Rehabilitation.** The WeeFIM II® Clinical Guide; Version 6.4. 2016.
22. **Varni JW, Seid M, Knight TS, Uzark K, Szer IS.** The PedsQL 4.0 Generic Core Scales: sensitivity, responsiveness, and impact on clinical decision-making. *J Behav Med* 2002;25(2):175-193.
23. **Varni JW, Seid M, Rode CA.** The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care* 1999;37(2):126-139.
24. **Varni JW, Seid M, Smith Knight T, Burwinkle T, Brown J, Szer IS.** The PedsQL in pediatric rheumatology: reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. *Arthritis Rheum* 2002;46(3):714-725.
25. **World Health Organization (WHO).** International classification of functioning, disability and health: children & youth version. Genf: WHO; 2007.

Anhang

PedsQL

Tabelle 22: Veränderung im PedsQL (fremdberichtet) zu Studienwoche 52 in Studie 301, FAS

Studie 301 Veränderung im PedsQL ¹⁾²⁾ (fremdberichtet)	Vosoritid N = 60	Placebo N = 61
<i>Gesamtwert zu Baseline</i>	n = 59	n = 59
Mittelwert (SD)	71,45 (16,19)	72,03 (16,79)
Median (min; max)	71,74 (30,4; 100,0)	73,86 (30,4; 100,0)
<i>Gesamtwert zu Woche 52</i>	n = 56	n = 59
Mittelwert (SD)	71,44 (15,96)	73,29 (16,37)
Median (min; max)	75,55 (25,0; 96,7)	72,83 (35,9; 98,9)
<i>Veränderung Woche 52 zu Baseline</i>	n = 56	n = 57
Mittelwert (SD)	-0,07 (17,29)	1,10 (12,74)
Median (min; max)	-0,54 (-75,0; 38,1)	2,96 (-38,0; 21,7)

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

²⁾ Erhoben durch Eltern/Betreuungspersonen für Kinder und Jugendliche von 5 bis 17 Jahre.

Abkürzungen: FAS: Full Analysis Set; PedsQL: Pediatric Quality of Life Inventory; SD: Standardabweichung.

QoLISSY

Tabelle 23: Veränderung im QoLISSY (fremdberichtet) zu Studienwoche 52 in Studie 301, FAS

Studie 301 Veränderung im QoLISSY ¹⁾²⁾ (fremdberichtet)	Vosoritid N = 60	Placebo N = 61
<i>Gesamtwert zu Baseline</i>	n = 60	n = 61
Mittelwert (SD)	57,00 (17,39)	55,23 (21,56)
Median (min; max)	56,25 (14,6; 97,2)	58,33 (13,5; 92,4)
<i>Gesamtwert zu Woche 52</i>	n = 57	n = 60
Mittelwert (SD)	57,07 (17,30)	58,34 (20,32)
Median (min; max)	59,38 (1,2; 90,6)	61,81 (10,1; 89,2)
<i>Veränderung Woche 52 zu Baseline</i>	n = 57	n = 60
Mittelwert (SD)	-0,27 (14,84)	2,85 (15,00)
Median (min; max)	-1,73 (-62,0; 34,4)	1,22 (-29,5; 50,7)

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

²⁾ Erhoben durch Eltern/Betreuungspersonen für Kinder und Jugendliche von 5 bis 17 Jahre.

Abkürzungen: FAS: Full Analysis Set; QoLISSY: Quality of Life in Short Stature Youth Questionnaire; SD: Standardabweichung.