

Nutzenbewertung



Gemeinsamer
Bundesausschuss

von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V

**Bewertung von Arzneimitteln für seltene Leiden nach § 35a
Absatz 1 Satz 11 i. V. m. Kapitel 5 § 12 Nr. 1 Satz 2 Verfo**

Wirkstoff: Somatrogen

Datum der Veröffentlichung: 1. Juli 2022

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abkürzungsverzeichnis	4
Hintergrund	5
1 Fragestellung	6
2 Auswahl und Methodik der relevanten Studien	7
2.1 Studienbasis für die Nutzenbewertung	7
2.2 Beschreibung der eingeschlossenen Studie	8
2.3 Endpunkte	14
2.3.1 Mortalität	15
2.3.2 Morbidität	15
2.3.3 Lebensqualität	17
2.3.4 Sicherheit	20
2.3.5 Übersicht der Erhebungszeitpunkte	23
2.4 Statistische Methoden	23
2.5 Verzerrungspotential auf Studien- und Endpunktebene	24
3 Ergebnisse der eingeschlossenen Studie	26
3.1 Studiencharakteristika und Studienmedikation	26
3.2 Mortalität	28
3.3 Morbidität	28
3.4 Lebensqualität	29
3.5 Sicherheit	30
4 Diskussion der Methodik und Ergebnisse	34
4.1 Zulassungsstatus und Zulassungspopulation von Somatrogon	34
4.2 Design und Methodik der Studie	34
4.3 Mortalität	35
4.4 Morbidität	36
4.5 Lebensqualität	36
4.6 Sicherheit	37
5 Anforderung an eine qualitätsgesicherte Anwendung	39
6 Zusammenfassung der Nutzenbewertung	40
Referenzen	42
Anhang	44

Tabellenverzeichnis

Tabelle 1: Übersicht über die Studienbasis	7
Tabelle 2: Charakterisierung der Studie CP-4-006	8
Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie CP-4-006	12
Tabelle 4: Charakterisierung der Intervention der Studie CP-4-006	13
Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie CP-4-006	14
Tabelle 6: Erhebungszeitpunkte der berücksichtigten Endpunkte in der Studie CP-4-006	23
Tabelle 7: Verzerrungspotential der Studie CP-4-006	24
Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studie CP-4-006	25
Tabelle 9: Allgemeine Angaben zur Studie CP-4-006	26
Tabelle 10: Charakterisierung der Studienpopulation der Studie CP-4-006; Sicherheitspopulation	27
Tabelle 11: Begleitmedikation mit einer Häufigkeit von $\geq 20\%$ in beiden Behandlungsarmen in der Studie CP-4-006; Sicherheitspopulation	28
Tabelle 12: Veränderung der Körpergröße (z-Score) zu Monat 12 in der Studie CP-4-006; FAS	29
Tabelle 13: Ergebnisse zu Responderanalysen für eine Verbesserung im QoLISSY (Gesamtskala) von ≥ 15 Punkten in der Studie CP-4-006; PRO Analysis Set	30
Tabelle 14: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE in der Studie CP-4-006; Sicherheitspopulation	31
Tabelle 15: UE mit Inzidenz $\geq 10\%$ oder bei ≥ 10 Patientinnen und Patienten und mit einer Inzidenz von $\geq 1\%$ in einem Studienarm nach MedDRA-Systemorganklassen und Preferred Terms in der Studie CP-4-006; Sicherheitspopulation	31
Tabelle 16: UE von besonderem Interesse (jeglicher Schweregrad) in der Studie CP-4-006; Sicherheitspopulation	33
Tabelle 17: Zusammenfassende Darstellung der Ergebnisse der Studie CP-4-006	40
Tabelle 18: Veränderung der Jährlichen Wachstumsgeschwindigkeit zu Monat 12 in der Studie CP-4-006; FAS	44

Abkürzungsverzeichnis

AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse
CDC	Center for Disease Control and Prevention
eCRF	elektronisches Fallberichtsformular
EPAR	European Public Assessment Report
FAS	Full Analysis Set
G-BA	Gemeinsamer Bundesausschuss
GH	Wachstumshormon
GHD	Wachstumshormonmangel
GKV	Gesetzliche Krankenversicherung
HLT	High-Level Term
HV	Wachstumsgeschwindigkeit
ICC	Intraklassenkorrelationskoeffizient
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interaktives Web Response Technologiesystem
ISS	Idiopathischer Kleinwuchs
KI	Konfidenzintervall
LLT	Lowest Level Term
LS	Least Squares
MANCOVA	multivariate Kovarianzanalyse
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities
MNAR	Missing Not At Random
PASS	Post-Authorisation Safety Study
PRO	Patient-Reported Outcome
PT	Preferred Term
pU	pharmazeutischer Unternehmer
QoLISSY	Quality of Life in Short Stature Youth Questionnaire
r-hGH	rekombinantes humanes Wachstumshormon
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SD	Standardabweichung
SGB	Sozialgesetzbuch
SMQ	Standardised MedDRA Query
SOC	Systemorganklasse
SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA

Hintergrund

Somatrogon 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 Kapitel 5 § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Somatrogon zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (pU) heran und bewertet die darin enthaltenen Daten nach Maßgabe der in Kapitel 5 § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Somatrogon in seiner Sitzung am 28. Juni 2022 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 31. März 2022 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 1. Juli 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

Somatrogon (Ngenla®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 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 Patientinnen und Patienten in folgender Indikation eingeschlossen [12]:

Kinder und Jugendliche ab einem Alter von 3 Jahren mit Wachstumsstörung durch unzureichende Ausschüttung von Wachstumshormon.

Gemäß Fachinformation beträgt die zugelassene Dosierung 0,66 mg/kg Körpergewicht, verabreicht einmal wöchentlich als subkutane Injektion [12].

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				
CP-4-006 (C0311009) ¹⁾	Ja	Ja	Ja	-
CP-4-001	Nein	Nein	Nein	Die Studienpopulation weicht vom AWG ab (gesunde Freiwillige).
CP-4-003	Nein	Nein	Nein	Die Studienpopulation weicht vom AWG ab (Erwachsene mit GHD).
CP-4-004	Nein	Nein	Nein	Die Fallzahl ist im Vergleich zum berücksichtigten RCT (CP-4-006) gering und zudem ist die Behandlung nicht durchgehend FI-konform. Offene, randomisierte, multizentrische Phase-II-Studie (Sicherheits- und Dosisfindungsstudie) zur wöchentlichen Gabe von Somatrogen (0,25 mg/kg (N = 13); 0,48 mg/kg (N = 15) und 0,66 mg/kg (N = 14)) im Vergleich zu einer täglichen Genotropin-Gabe (0,034 mg/kg, N = 11) bei vorpubertären Kindern mit GHD mit einer Behandlungsdauer von 12 Monaten und einer anschließenden offenen einarmigen Extensionsstudie (noch laufend).
CP-4-005	Nein	Nein	Nein	Die Studienpopulation weicht vom AWG ab (Erwachsene mit GHD).
CP-4-007	Nein	Nein	Nein	Die Studienpopulation weicht vom AWG ab (gesunde erwachsene Männer).
CP-4-009	Nein	Nein	Nein	Die Fallzahl ist im Vergleich zum berücksichtigten RCT (CP-4-006) gering und zudem ist die Behandlung nicht durchgehend FI-konform. Offene, randomisierte, multizentrische Phase-III-Studie zur Untersuchung der Wirksamkeit und Sicherheit der wöchentlichen Gabe von Somatrogen (n = 22) im Vergleich zu einer täglichen Genotropin-Gabe (n = 22) bei japanischen, vorpubertären Kindern mit GHD. Drei Dosiserhöhungen alle 2 Wochen (0,25, 0,48 und 0,66 mg/kg); anschließend 46 Wochen Dosis von 0,66 mg/kg/Woche.
CP-4-011	Nein	Nein	Nein	Die Studienpopulation weicht vom AWG ab (gesunde erwachsene Männer).

Studiename (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
C0311002	Nein	Nein	Nein	Kurze Behandlungsdauer von 12 Wochen. Randomisierte, multizentrische, offene Phase-III-Studie mit Crossover zur Untersuchung der von den Studienteilnehmenden empfundenen Behandlungsbelastung bei der wöchentlichen Injektion von Somatrogen im Vergleich zur täglichen Injektion von Genotropin bei Kindern mit GHD.

¹⁾ Zulassungsrelevante Studie gemäß EPAR [6].

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

Zur Nutzenbewertung für Somatrogen herangezogene Studien und Daten

- Herstellerdossier zu Somatrogen [16].
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [6].
- Studienbericht inkl. Studienprotokoll, statistischem Analyseplan (SAP) und Zusatzanalysen der Studie CP-4-006 (C0311009) [9,10,11,13,14,15].
- Publikation zur Studie CP-4-006 [4].

2.2 Beschreibung der eingeschlossenen Studie

Die Unterlagen zur Nutzenbewertung für Somatrogen basieren auf der Zulassungsstudie CP-4-006 (C0311009). Die Studie und die Intervention werden in den Tabellen 2–4 charakterisiert.

Tabelle 2: Charakterisierung der Studie CP-4-006

Charakteristikum	Beschreibung
Design und Studienablauf	<p>Die Studie CP-4-006 (C0311009) ist eine offene, randomisierte, aktiv kontrollierte, multizentrische Phase-III-Studie im Parallelgruppendesign zur Untersuchung der Wirksamkeit und Sicherheit von wöchentlich verabreichtem Somatrogen im Vergleich zu täglich verabreichtem Genotropin® bei pädiatrischen Patientinnen und Patienten mit Wachstumsstörungen aufgrund einer unzureichenden Wachstumshormon-Sekretion. Die Patientinnen und Patienten wurden im Verhältnis 1:1 auf die Behandlungsarme Somatrogen oder Genotropin randomisiert, stratifiziert für die drei Faktoren:</p> <ul style="list-style-type: none"> • Region (Westeuropa, Israel, Australien, Neuseeland, Kanada, USA vs. Mittel- und Osteuropa, Griechenland, Türkei, Lateinamerika, Asien (mit Ausnahme von Indien und Vietnam) vs. Indien und Vietnam). • Höchste gemessene GH-Konzentration beim Screening (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml)¹⁾. • Chronologisches Alter (≥ 3 bis ≤ 7 vs. > 7 Jahre).

Charakteristikum	Beschreibung
	<p>Die Studie gliedert sich in folgende Phasen:</p> <ul style="list-style-type: none"> • 2-monatige Screeningphase²⁾, in der die Geeignetheit der Patientinnen und Patienten zum Einschluss in die Studie geprüft wurde. • 12-monatige Behandlungsphase, in der entweder wöchentlich Somatrogen oder täglich Genotropin subkutan verabreicht wurden. • Langfristige, einarmige, offene Extensionsphase, in der alle Patientinnen und Patienten mit Somatrogen behandelt wurden. Studienteilnehmende, die während der 12-monatigen Behandlungsphase Somatrogen erhielten, wurden in der Extensionsstudie mit derselben Dosis (mg/kg/Woche) Somatrogen weiterbehandelt. Studienteilnehmende, die während der 12-monatigen Behandlungsphase Genotropin erhielten, wurden auf Somatrogen umgestellt und begannen die Behandlung mit einer Dosis von 0,66 mg/kg/Woche nicht später als einen Tag nach Beendigung der Genotropin-Behandlung. Laut Studienprotokoll wird die Extensionsstudie bis zur Marktzulassung fortgeführt.
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Präpubertäre Kinder im Alter von ≥ 3 Jahren und noch nicht 11 Jahre (10 Jahre und 364 Tage) für Mädchen oder noch nicht 12 Jahre (11 Jahre und 364 Tage) für Jungen (zum Zeitpunkt der Unterschrift der Einwilligungserklärung) mit entweder einem isolierten GHD oder einem GHD als Teil eines multiplen Hypophysenhormon-Mangels. • Bestätigte Diagnose des GHD durch zwei verschiedene GH-Stimulationstests mit einer maximalen GH-Konzentration von ≤ 10 ng/ml, gemessen durch ein lokales oder zentrales Labor. Dabei wurden folgende etablierte Untersuchungsverfahren verwendet: Insulin-Toleranz-Test mit einer Serum-Cortisol-Antwort auf Hypoglykämie / Arginin-Test / Clonidin-Test / Glucagon-Test / L-Dopa-Test. Der globale medizinische Monitor kann frühere lokale Laborergebnisse akzeptieren; vorbehaltlich einer Vorabgenehmigung und wenn die Tests wie empfohlen durchgeführt wurden (die vorherige spontane nächtliche GH-Sekretion sollte ≤ 10 ng/ml sein). • Das Knochenalter ist nicht älter als das chronologische Alter und sollte < 10 Jahre für Mädchen und < 11 Jahre für Jungen sein. • Keine vorherige Behandlung mit r-hGH (therapienaive Personen). • Verminderte Körpergröße und Wachstumsgeschwindigkeit bestimmt durch: <ul style="list-style-type: none"> ○ Jährliche Wachstumsgeschwindigkeit unter der 25. Perzentile für das chronologische Alter (HV $< -0,7$ des z-Score) und Geschlecht gemäß OPKO HV (Tanner, Prader und Hermanussen)-Rechner. ○ Das Intervall zwischen zwei Messungen der Körpergröße sollte mindestens 6 Monate betragen, aber nicht länger als 18 Monate vor Studieneinschluss zurückliegen. • IGF-1-Spiegel von mindestens 1 SD unter dem durchschnittlichen IGF-1-Spiegel zu Baseline, standardisiert nach Alter und Geschlecht (IGF-1 z-Score ≤ -1)³⁾, gemäß den Referenzwerten des zentralen Labors. Ein einziger Wiederholungstest ist erlaubt (nach Absprache mit dem globalen medizinischen Monitor), falls alle anderen Kriterien erfüllt werden. • Normierte Berechnung der glomerulären Filtrationsrate basierend auf der aktualisierten Schwartz-Formel für pädiatrische Patientinnen und Patienten. • Bei Kindern mit mehreren Hormonmängeln müssen die verwendeten Ersatztherapien für die betreffenden Organe der Hypothalamus-Hypophysen-Achse stabil eingestellt sein (keine Änderung der Dosis), d. h. für mindestens 3 Monate vor der Unterschrift der Einwilligungserklärung. • Normaler 46,XX-Karyotyp bei Mädchen.

Charakteristikum	Beschreibung
	<p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Kinder mit einer Vorgeschichte von Leukämie, Lymphom, Sarkom oder einer anderen Form von Krebs. • Vorgeschichte von Strahlen- oder Chemotherapie. • Untergewichtige Kinder, definiert als BMI < -2 (z-Score) für Alter und Geschlecht. • Kinder mit psychosozialem Kleinwuchs. • Kinder, die zu klein und/oder zu leicht, bezogen auf das Gestationsalter geboren wurden (Geburtsgewicht und/oder Geburtslänge < -2 (z-Score) für Gestationsalter). • Vorliegen von Anti-hGH-Antikörpern zum Zeitpunkt des Screenings. • Jede klinisch bedeutsame Anomalie, die das Wachstum oder die Fähigkeit, das Wachstum zu beurteilen, beeinträchtigen könnte, wie z. B. chronische Krankheiten wie Niereninsuffizienz, Rückenmarksbestrahlung, usw. • Patientinnen und Patienten mit Typ 1- und Typ 2-Diabetes, die nach Einschätzung der Prüffärztin / des Prüffarztes keine Therapie entsprechend dem Behandlungsstandard erhalten oder „nicht-compliant“ gegenüber ihrer verschriebenen Therapie sind oder deren Stoffwechsel nur unzureichend eingestellt ist. • Chromosomale Anomalien einschließlich Turner-Syndrom, Laron-Syndrom, Noonan-Syndrom, Prader-Willi-Syndrom, Silver-Russell-Syndrom, SHOX-Mutationen/-Deletionen und Skelettdysplasien. • Gleichzeitige Verabreichung anderer Therapien, die das Wachstum beeinflussen können, wie z. B. anabole Steroide oder Sexualsteroiden, mit der Ausnahme von Medikamenten bei (ADHS) oder Hormonersatztherapien (Thyroxin, Hydrokortison, Desmopressin [DDAVP®]). • Kinder, die eine Glukokortikoid-Therapie (z. B. gegen Asthma) benötigen und die dauerhaft eine Dosis von über 400 µg/Tag inhaliertem Budesonid oder einem Äquivalent einnehmen. • Schwerwiegende Erkrankungen und/oder Vorliegen einer Kontraindikation gegenüber einer r-hGH-Therapie. • Mehr als eine geschlossene Epiphysenfuge. • Andere Gründe für Kleinwuchs, wie Zöliakie, nicht eingestellte primäre Hypothyreose und Rachitis.
Intervention und Zahl der Patientinnen und Patienten	<p>Screening: N = 536 Randomisierung: N = 228 Intervention: Somatrogen N = 109 Kontrolle: Genotropin N = 115</p>
Ort und Zeitraum der Durchführung	<p>Ort der Durchführung Die Studie wurde an 84 Zentren in 21 Ländern durchgeführt: Argentinien, Australien, Bulgarien, Belarus, Georgien, Griechenland, Großbritannien, Indien, Israel, Kanada, Kolumbien, Mexico, Neuseeland, Polen, Russland, Spanien, Südkorea, Taiwan, Türkei, Ukraine, USA.</p> <p>Zeitraum der Durchführung</p> <ul style="list-style-type: none"> • Einschluss erste/r Patient/in, erste Visite: 19.04.2017 • Einschluss letzte/r Patient/in: k. A. • Letzte/r Visite/Kontakt letzte/r Patient/in: 23.08.2019 <p>Datenschnitte und Auswertungen Datenschnitt: 23.08.2019 (Datenschnitt nachdem alle Personen Monat 12 abgeschlossen hatten). Es wurde keine Interimsanalyse durchgeführt.</p>

Charakteristikum	Beschreibung
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt Jährliche Wachstumsgeschwindigkeit in cm/Jahr nach 12 Monaten Behandlung.</p> <p>Sekundäre Endpunkte</p> <p><u>Auxologie-Wirksamkeitsendpunkte</u></p> <ul style="list-style-type: none"> • Jährliche Wachstumsgeschwindigkeit in cm/Jahr nach 6 Monaten Behandlung. • Veränderung der standardisierten Körpergröße (z-Score) nach 6 und 12 Monaten im Vergleich zu Baseline. • Veränderung der Knochenreifung (berechnet als Knochenalter / chronologisches Alter) nach 12 Monaten im Vergleich zum Knochenalter zu Baseline. <p><u>Biochemische Marker Endpunkte</u></p> <ul style="list-style-type: none"> • IGF-1-Spiegel und IGF-1 (z-Score) an Tag 4 (-1) nach der Somatrogen-Dosis über die Studienbesuche hinweg. • IGFBP-3-Spiegel und IGFBP-3 (z-Score) an Tag 4 (-1) nach der Somatrogen-Dosis über die Studienbesuche hinweg. <p>Sicherheitsendpunkte</p> <ul style="list-style-type: none"> • UE und SUE. • Inzidenz der Bildung von Anti-Somatrogen-Antikörper (einschließlich Charakterisierung der Antikörper und neutralisierenden Eigenschaften). • Bewertung der lokalen Reaktion an der Injektionsstelle. • Parameter des Glukosestoffwechsels: Nüchternblutzucker, Nüchterninsulinspiegel, HbA1c. • Schilddrüsenstatus. • Lipidprofil. • Alle anderen sicherheitsrelevanten hämatologischen und biochemischen Parameter sowie Urinanalyse. • Körperliche Untersuchung. • Ergebnisse der Fundoskopie – falls durchgeführt (normal/abnormal). • Vitalparameter • EKG. <p>Zusätzliche Endpunkte</p> <ul style="list-style-type: none"> • Gesundheitsbezogene Lebensqualität (gemessen durch den QoLISSY-Kernfragebogen zu Studienbeginn und Monat 12 in bestimmten Ländern⁴). • Nur in USA erhoben: Anteil der erfolgreichen Einzelinjektionen an der Gesamtzahl der Einzelinjektionen mit dem Somatrogen-Fertigpen bei Patientinnen und Patienten in den Wochen 1, 2, 3, 4, 5 und 6, basierend auf dem PAT. • Nur in USA erhoben: Anteil der erfolgreichen Einzelinjektionen an der Gesamtzahl der Einzelinjektionen mit dem Somatrogen-Fertigpen bei Patientinnen und Patienten in Woche 1, basierend auf dem OAT. • Kommentare zum PAT in Bezug auf erfolgreiche oder erfolglose Injektionsversuche. • Kommentare zum OAT in Bezug auf erfolgreiche oder erfolglose Injektionsversuche. • Informationen aus der Inspektion der zurückgegebenen Geräte.
Subgruppenanalysen	<p>Gemäß SAP-Version 4.0 vom 14.06.2019 wurden für die Wachstumsgeschwindigkeit (cm/Jahr) nach 12 Monaten, die Wachstumsgeschwindigkeit (cm/Jahr) nach 6 Monaten und die Veränderung der Körpergröße (z-Score) nach 6 und 12 Monaten für folgende Subgruppen Analysen durchgeführt:</p> <ul style="list-style-type: none"> • Alter (< 7 vs. ≥ 7 Jahre) • Geschlecht (männlich vs. weiblich)

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Höchste gemessene GH-Konzentration (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml) • Region⁵⁾ (Region 1: Westeuropa, Israel, Griechenland, Australien, Neuseeland, Kanada, USA vs. Region 2: Zentral- und Osteuropa, Türkei, Lateinamerika, Asien (außer Indien und Vietnam) vs. Region 3: Indien, Vietnam) Post hoc wurden für die gesundheitsgezogene Lebensqualität (QoLISSY) und die Sicherheitsendpunkte Analysen für die genannten Subgruppen durchgeführt.

¹⁾ Der Anteil der Patientinnen und Patienten mit GH-Konzentrationen von > 7 bis ≤ 10 ng/ml wurde bei der Randomisierung auf 35–40 % der Gesamtstichprobe begrenzt.

²⁾ Verzögert sich die Untersuchung der Patientin / des Patienten aufgrund einer harmlosen Krankheit oder eines unvorhergesehenen harmlosen Zustands (z. B. Pharyngitis, virale gastrointestinale Probleme, kleiner Unfall oder Trauma usw.) oder ein technisches Problem im Zusammenhang mit dem Screening (z. B. Verzögerungen bei den Laborergebnissen), wird die Dauer der Untersuchung verlängert; jedoch nicht um mehr als 4 Wochen (insgesamt 12 Wochen Untersuchungszeitraum).

³⁾ Gemäß der Rundungsregel können IGF-1-Werte von $\leq -0,95$ akzeptiert werden.

⁴⁾ Der QoLISSY-Fragebogen wurde nur in den folgenden Ländern ausgefüllt, für die ein validiertes übersetztes Instrument existiert: USA, Australien, Neuseeland, Belarus, Russland, Ukraine, Großbritannien und Spanien.

⁵⁾ Für den primären Endpunkt „Wachstumsgeschwindigkeit in cm/Jahr nach 12 Monaten“ war die Subgruppenanalyse nach Region nicht im SAP (Version 4.0 vom 14.06.2019), aber im Studienprotokoll (Version 5 vom 27.04.2018) a priori geplant.

Abkürzungen: ADHS: Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung; GH: Wachstumshormon; GHD: Wachstumshormonmangel; hGH: humanes Wachstumshormon; HV: Wachstumsgeschwindigkeit; IGF-1: Insulinähnlicher Wachstumsfaktor 1; IGFBP-3: Insulinähnliches Wachstumsfaktor-bindendes Protein 3; k. A.: keine Angabe; OAT: Observer Assessment Tool; PAT: Participant Assessment Tool; QoLISSY: Quality of Life in Short Stature Youth; r-hGH: rekombinantes humanes Wachstumshormon; SAP: Statistischer Analyseplan; SD: Standardabweichung; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

Protokolländerungen

Das Originalprotokoll (USA-Version 1.0) ist vom 06.10.2016. Der Einschluss der ersten Patientin / des ersten Patienten in die Studie erfolgte am 19.04.2017. Nach diesem Datum erschienen Protokollamendment 1.0 (Version 2.0 vom 17.10.2017, Version 3.0 vom 01.12.2017 und Version 4.0 vom 16.01.2018) und Protokollamendment 2.0 (Version 5.0 vom 27.04.2018, die die finale Protokollversion darstellt). Die relevanten Protokolländerungen sind in Tabelle 3 dargestellt. Es wurden einige länderspezifische Protokolländerungen auf Wunsch einer bestimmten Gesundheitsbehörde erstellt. Diese enthalten keine für die Nutzenbewertung relevanten Änderungen.

Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie CP-4-006

Amendment	Relevante Änderungen
Amendment 1 Version 2.0 (USA) vom 17.10.2017 Einschluss ¹⁾ : k. A.	Dieses Protokoll, das einen Abschnitt zur Dosisescalation enthielt, wurde der FDA vorgelegt, anschließend zurückgezogen und nicht umgesetzt. Das Protokoll wurde überarbeitet, um den Abschnitt über die Dosisescalation zu entfernen.
Amendment 1 Version 3.0 (USA) vom 01.12.2017 Einschluss ¹⁾ : k. A.	Diese Version wurde nur bei der FDA eingereicht und nicht umgesetzt.
Amendment 1 Version 4.0 (USA) vom 16.01.2018 Einschluss ¹⁾ : k. A.	Keine relevanten Änderungen vorgenommen.

Amendment	Relevante Änderungen
Amendment 2 Version 5.0 (USA) vom 27.04.2018 Einschluss ¹⁾ : k. A.	Einführung der einarmigen, offenen Extensionsphase für Patientinnen und Patienten, die die 12 Monate Behandlung erfolgreich abgeschlossen haben, um zusätzliche Daten zur Langzeitsicherheit und Wirksamkeit zu sammeln.

¹⁾ Anzahl der bis zu diesem Zeitpunkt eingeschlossenen Personen.

Abkürzungen: FDA: U. S. Food and Drug Administration; k. A.: keine Angabe.

Charakterisierung der Intervention

Tabelle 4: Charakterisierung der Intervention der Studie CP-4-006

Intervention	Kontrolle
Somatrogon <ul style="list-style-type: none"> • Injektionslösung mit 20 oder 50 mg/ml Somatrogon in einem vorgefüllten Multidosis-Pen. • 0,66 mg/kg/Woche. • Einmal wöchentlich subkutane Injektion in den Bereich der Oberarme, des Gesäßes, der Oberschenkel oder des Bauches (8 Stellen) (vorzugsweise morgens). • Dieselbe Injektionsstelle sollte nicht an 2 aufeinanderfolgenden Tagen verwendet werden. Es sollte zwischen den 8 Injektionsstellen rotiert werden. 	Genotropin® (Wirkstoff: Somatropin) <ul style="list-style-type: none"> • 0,034 mg/kg/Tag (oder 0,24 mg/kg/Woche, gleichmäßig verteilt auf 7 Injektionen über eine Woche). • Tägliche subkutane Injektion in den Bereich der Oberarme, des Gesäßes, der Oberschenkel oder des Bauches (8 Stellen) mittels Multidosis-Pen (vorzugsweise abends / im Bett). • Dieselbe Injektionsstelle sollte nicht an 2 aufeinanderfolgenden Tagen verwendet werden. Es sollte zwischen den 8 Injektionsstellen rotiert werden.
Dosismodifikation Die Dosis von Somatrogon und Genotropin wird alle 3 Monate anhand des Körpergewichts der Studienteilnehmenden ermittelt. Die Dosis wird durch das IRT bestimmt und beinhaltet eine automatische Auf- oder Abrundung auf die nächstliegende Pen-Stufe (0,2-Stufen bei 20mg/ml-Pens und 0,5-Schritte bei Pens mit 50 mg/ml).	
Dosisreduktion Patientinnen und Patienten, die einen IGF-1-Spiegel von > +2 (z-Score) aufwiesen (in der Somatrogon-Gruppe bei 2 unabhängigen Messungen an Tag 4 (-1) nach der Injektion; in der Genotropin-Gruppe bei 2 unabhängigen Messungen) wurden gebeten, innerhalb von 4 bis 6 Wochen nach Messung des erhöhten IGF-1-Spiegels zu einer außerplanmäßigen Studiervisite zu erscheinen. Wenn der IGF-1-Spiegel weiterhin > +2 (z-Score) betrug, wurde die Dosis um 15 % reduziert (d. h. auf 0,56 mg/kg/Woche für Somatrogon und 29 µg/kg/Tag für Genotropin). Die Patientin / Der Patient wurde mindestens 4 Wochen mit der neuen Dosis behandelt und danach wurde eine erneute Messung des IGF-1-Spiegels vorgenommen. Sofern der IGF-1-Spiegel zum Zeitpunkt der erneuten Messung wieder > +2 (z-Score) betrug, wurde die Dosis erneut um 15 % reduziert (d. h. auf 0,48 mg/kg/Woche für Somatrogon und auf 24,7 µg/kg/Tag für Genotropin). Lag der IGF-1-Spiegel nach 2 Dosisreduktionen weiterhin > + 2 (z-Score), dann wurde die weitere Behandlung vom Medizinischen Monitor (ggf. in Absprache mit dem DSMB) individuell festgelegt. Bei UE, die schwerwiegend und arzneimittelbedingt sind, wird eine Dosisreduzierung nach Diskussion mit dem Medizinischen Monitor und dem DSMB eingeleitet. Die Dosis sollte in ähnlicher Weise in einem 2-stufigen Ansatz reduziert werden wie oben beschrieben.	

Intervention	Kontrolle
<p>Nicht erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • Jede hormonell wirksame Medikation außer der Ersatztherapie bei Hypophyseninsuffizienz. Die Glukokortikoid-Ersatzdosis sollte bei Kindern, die Somatrogen und Genotropin und Glukokortikoid-Behandlungen erhalten, vorsichtig angepasst werden, um sowohl einen Hypoadrenalismus als auch eine wachstumshemmende Wirkung zu vermeiden. • Jede neue Langzeit-Glukokortikoid-Behandlung; eine kurzfristige Glukokortikoid-Behandlung ist zulässig. • Anabole Steroide (einschließlich Testosteronersatztherapie). • Medikamente zur Gewichtsreduktion. • Psychopharmaka, die typischerweise mit Gewichtsveränderungen und/oder Diabetes in Verbindung gebracht werden mit Ausnahme von Medikamenten zur Behandlung von ADHS. <p>Erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • Medikamente, die zu Beginn der Studie zur Kontrolle bestehender medizinischer Zustände und/oder die während der Studie zur Behandlung von UE eingenommen werden. • Eine sorgfältige Überwachung ist ratsam, wenn Somatrogen in Kombination mit anderen Medikamenten verabreicht wird, von denen bekannt ist, dass sie durch CYP450-Leberenzyme metabolisiert werden. Bei Patientinnen und Patienten mit Diabetes mellitus, die eine medikamentöse Therapie benötigen, kann es erforderlich sein die Dosis von Insulin und/oder oralen/injizierbaren Medikamenten anzupassen, wenn die Behandlung mit Somatrogen begonnen wird. 	

Abkürzungen: ADHS: Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung; DSMB: Data Safety Monitoring Board; IGF-1: Insulinähnlicher Wachstumsfaktor 1; IRT: Interaktives Web Response Technologiesystem; UE: Unerwünschtes Ereignis.

2.3 Endpunkte

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

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie CP-4-006

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Todesfälle ¹⁾	Mortalität	Nein	Ja
Jährliche Wachstumsgeschwindigkeit ²⁾	Morbidität	Ja	Nein ³⁾
Körpergröße (z-Score)		Ja	Ja
QoLISSY	Lebensqualität	Ja	Ja
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

¹⁾ Im Rahmen der Sicherheit erhoben.

²⁾ Primärer Endpunkt.

³⁾ Der Endpunkt wird als nicht patientenrelevant eingeschätzt und daher nicht für die Nutzenbewertung herangezogen. Die Ergebnisse des Endpunkts werden in Analogie zum Nutzenbewertungsverfahren zu Vosoritid [7] im Anhang dargestellt.

Abkürzungen: pU: pharmazeutischer Unternehmer; QoLISSY: Quality of Life in Short Stature Youth.

2.3.1 Mortalität

Todesfälle

Der Endpunkt „Todesfälle“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Anzahl an Todesfällen wurde im Rahmen der Sicherheitserfassung kontinuierlich über die gesamte Studiendauer erhoben.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Todesfälle stellen einen patientenrelevanten Endpunkt entsprechend § 2 Satz 3 Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) dar.

Validität

Die Erhebung des Endpunkts wird als valide erachtet.

2.3.2 Morbidität

Jährliche Wachstumsgeschwindigkeit

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

Operationalisierung

Beschreibung

Die Wachstumsgeschwindigkeit (HV) in cm/Jahr nach 12 Monaten wurde wie folgt bestimmt:

$$HV \text{ (cm/Jahr)} = \frac{\text{Körpergröße zu Monat 12} - \text{Körpergröße zu Baseline}}{\text{Datum Monat 12} - \text{Datum Baseline}} \times 365,25$$

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

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“.

Die 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 CP-4-006 durch altersadjustierte Angaben der Körpergröße erfasst. Die Wachstumsgeschwindigkeit bietet darüber hinaus keine relevanten Erkenntnisse, sodass diese in der Nutzenbewertung nicht dargestellt wird. In Analogie zum Nutzenbewertungsverfahren zu Vosoritid [7] werden die Ergebnisse jedoch im Anhang dargestellt.

Validität

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

Körpergröße (z-Score)

Der Endpunkt „Körpergröße (z-Score)“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Körpergröße wurde als Stehhöhe erfasst. Zu jeder Studienvisite wurde die Körpergröße dreimal mithilfe eines kalibrierten, an der Wand befestigten Stadiometers gemessen und der Mittelwert der Messungen für die Analysen verwendet. Die Messungen wurden von geschultem Personal, vorzugsweise immer derselben Person, durchgeführt und sollten idealerweise bei jeder Studienvisite zur gleichen Tageszeit erfolgen (vorzugsweise am Morgen). Die Messungen waren durch einen Leitfaden im Studienprotokoll standardisiert und der Zeitpunkt der Messung und die Initialen des Untersuchers wurden im elektronischen Fallberichtsformular (eCRF) vermerkt.

Für die Körpergröße wurden alters- und geschlechtsadjustierte z-Scores und Perzentilen mithilfe der 2000 CDC Growth Chart berechnet [8]. Die z-Scores spiegeln die Anzahl der Standardabweichungen (SD) eines jeden Werts von den nach Alter und Geschlecht standardisierten normalen Durchschnittswerten wider. Ein z-Score von 0 bedeutet, dass der gemessene Wert dem Mittelwert der Referenzpopulation entspricht. Bei +1 bzw. -1 liegt der Wert eine SD über bzw. unter dem Mittelwert. In der Studie CP-4-006 wurden die Daten als SD über oder unter der altersspezifischen Referenz ($\cong 0$) dargestellt.

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“.

Die Körpergröße kann, insbesondere bei Kindern mit charakteristischen, krankheitsbedingten Wachstumsstörungen, als patientenrelevanter Morbiditätsparameter eingeschätzt werden. Dabei werden Angaben, welche für Alter und Geschlecht adjustiert wurden, gegenüber absoluten Werten bevorzugt herangezogen.

Validität

Die ermittelten z-Scores basieren auf einer US-Referenzpopulation des CDC [8]. Es wurden die Daten der National Health Examination Surveys und der National Health and Nutrition Examination Surveys zur Generierung von alters- und geschlechtsspezifischen Referenzperzentilen für anthropometrische Maßzahlen verwendet. Die Surveys basieren hauptsächlich auf Untersuchungen von mehreren Stichproben der US-amerikanischen Allgemeinbevölkerung in den Jahren zwischen 1963 und 1994. Länderspezifische Referenzpopulationen wurden nicht berücksichtigt. Die in der Studie CP-4-006 herangezogenen z-Scores werden als geeignet zur Standardisierung anthropometrischer Maße der betreffenden Studienpopulation erachtet.

2.3.3 Lebensqualität

Quality of Life in Short Stature Youth (QoLISSY)

Der Endpunkt „QoLISSY“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

In der Studie CP-4-006 wurde bei Kindern und Jugendlichen ≥ 7 Jahren die selbstberichtete Kinderversion des QoLISSY eingesetzt, während für jüngere Kinder im Alter von 3 bis 7 Jahren die Eltern den Eltern-Proxy-Fragebogen ausfüllten.

Der QoLISSY-Fragebogen ist ein Instrument zur Beurteilung der Lebensqualität von Kindern und Jugendlichen mit Kleinwuchs [2,3]. Es existiert eine Version zur direkten Befragung von Kindern und Jugendlichen in den Altersgruppen 8–12 und 13–18 Jahre sowie eine Version zur Befragung der Eltern von betroffenen Kindern und Jugendlichen in den Altersgruppen 4–7, 8–12 sowie 13–18 Jahre.

Der QoLISSY besteht aus:

- 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 Elternversion enthält zusätzlich 2 Domänen, die die Zukunft des Kindes (5 Items) sowie den Einfluss der Erkrankung auf die Eltern (11 Items) thematisieren.

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.

In der Studie CP-4-006 wurden a priori geplant nur die 3 Hauptdomänen des QoLISSY-Fragebogens erfasst und ein Gesamtwert sowie die 3 Einzeldomänen dargestellt. Zudem wurde der Fragebogen im Rahmen der Studie nur in den folgenden Ländern, für die eine validierte Übersetzung vorliegt, eingesetzt: USA, Australien, Neuseeland, Belarus, Russland, Ukraine, Großbritannien und Spanien.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Es wurde keine Begründung identifiziert, warum nur die 3 Hauptdomänen (physische, emotionale und soziale), nicht aber die zusätzlichen Domänen „Bewältigung“, „Überzeugungen“, „Behandlung“ sowie die 2 elternspezifischen Domänen (Zukunft des Kindes und Einfluss der Erkrankung auf die Eltern) in der Studie eingesetzt wurden. Zudem haben bereits 7-jährige Kinder den Fragebogen selbst ausgefüllt, obwohl eine direkte Befragung laut Entwicklung erst ab 8 Jahren vorgesehen ist [3]. Der pU begründet diese Abweichung damit, dass im Rahmen der Studie Kindern in diesem Alter zugetraut wurde, eine valide Selbstbewertung abzugeben. Laut SAP konnten die Eltern beim Ausfüllen des Fragebogens helfen, sodass dieses Vorgehen in der vorliegenden

Konstellation als akzeptabel eingeschätzt wird. Da keine Informationen vorliegen, wie viele Personen in welchen Ländern in die Studie eingeschlossen wurden, ist dementsprechend unklar, für wie viele Personen ein validierter Fragebogen vorlag.

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 Kinder und Jugendlichen ≥ 7 Jahre berücksichtigt. Zudem wird die Version der Eltern für Kinder im Alter von 3 bis 7 Jahren berücksichtigt.

Validität

Entwicklung und Validierung des QoLISSY-Instruments umfassten mehrere Phasen. Nach einem Literaturreview wurden Fokusgruppen abgehalten, anschließend erfolgten Pilottests und Untersuchungen der psychometrischen Eigenschaften des Instruments. Basierend auf den Ergebnissen aus Fokusgruppen mit Kindern und Jugendlichen ($n = 86$) mit idiopathischem Kleinwuchs (ISS) oder mit Wachstumshormonmangel (GHD) 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 [3].

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 ISS oder GHD sowie 317 Eltern von Betroffenen eingeschlossen [3].

Zur Untersuchung der konvergenten Validität wurden Korrelationen zwischen den Domänenwerten des KIDSCREEN-Fragebogens und den Domänenwerten sowie dem Gesamtwert des QoLISSY bestimmt. Es zeigten sich für alle Domänen des KIDSCREEN statistisch signifikante Korrelationen mit dem Gesamtwert der selbstausgefüllten 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 ($r = 0,18-0,51$) Domänen 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 KIDSCREEN [3].

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 [3].

Zur Untersuchung der internen Konsistenz wurde Cronbach's α berechnet. Es zeigte sich für den Gesamtwert des selbstberichteten Fragebogens ein Cronbach's 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 [2,3].

In einer Studie, die in mehreren Zentren in Deutschland durchgeführt wurde, wurde u. a. die Änderungssensitivität des QoLISSY untersucht [1]. 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 Kontrollarm 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 ($\eta_p^2 = 0,24$; $p < 0,01$) Aspekte 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 [1].

Der QoLISSY erscheint ausreichend valide, um in Form des Gesamtwerts der 3 Einzeldomänen 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.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt „Unerwünschte Ereignisse“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Ein unerwünschtes Ereignis (UE) war in der Studie CP-4-006 definiert als jedes neu aufgetretene medizinische Ereignis unabhängig davon, ob ein kausaler Zusammenhang mit der Studienmedikation besteht oder nicht. Ein UE kann jedes subjektive Anzeichen, jedes Symptom oder jede Diagnose, jede klinisch signifikante Abweichung von Ausgangslaborwerten oder Vitalzeichen oder jede Verschlechterung einer bestehenden Krankheit sein. UE wurden gemäß Medical Dictionary for Regulatory Activities (MedDRA), Version 22.0, kodiert. Der Schweregrad der UE wurde anhand von 3 Kategorien unterschieden. Hierbei handelt es sich um eine studienindividuelle Einteilung:

- Mild: Vorübergehendes oder leichtes Unbehagen; keine Einschränkung der Aktivität; keine medizinische Intervention/Therapie erforderlich.
- Moderat: Leichte bis mäßige Einschränkung der Aktivität – möglicherweise ist etwas Hilfe erforderlich; keine oder minimale medizinische Intervention/Therapie erforderlich.
- Schwer: Starke Einschränkung der Aktivität, in der Regel ist etwas Hilfe erforderlich; medizinische Intervention/Therapie erforderlich; Krankenhausaufenthalt möglich.

UE wurden ab Baseline bis zu 1 Monat nach der letzten Dosis der Studienmedikation (Sicherheitsnachbeobachtung) für Patientinnen und Patienten erhoben, die nicht an der offenen, einarmigen Extensionsstudie teilnahmen.

Als schwerwiegendes unerwünschtes Ereignis (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;
- wichtige medizinische Ereignisse, die Studienteilnehmende auf andere Weise gefährdeten und ein Eingreifen erforderten, um eines der obigen aufgeführten Ereignisse zu verhindern.

Im Studienprotokoll wurden UE von besonderem Interesse als Sicherheitsendpunkte genannt; die Definition, welche Ereignisse zu welchem UE von besonderem Interesse zusammengefasst werden, erfolgte allerdings erst im Studienbericht und in Modul 4 des Nutzendossiers. Sie wurden aus den potentiellen, substanzklassenspezifischen Risiken, die für Somatotropin-haltige Produkte identifiziert wurden, ausgewählt:

- Reaktionen an der Injektionsstelle: „Schwere Reaktionen an der Injektionsstelle“ (High-Level Terms (HLT), alle Pfade) und ausgewählte Preferred Term (PT) „Lipoatrophie und Hautdystrophie“.

- Immunogenität (Standardised MedDRA Query (SMQ)): „Anti-Somatrogon neutralisierende Antikörper vorhanden“, Anaphylaktische Reaktion (SMQ, enger Bereich), „Angioödem“ (SMQ, enger Bereich), „Hypersensibilität“ (SMQ, enger Bereich).
- Beeinträchtigung des Glukosestoffwechsels: „Hyperglykämie / Neuauftreten eines Diabetes mellitus“ (SMQ, enger Bereich).
- Beeinträchtigung der Schilddrüsenfunktion: „Schilddrüsenfunktionsstörung“ (SMQ, breiter und enger Bereich).
- Kortisolveränderungen: HLT für „Hypofunktionen der Nebennierenrinde“.

Im Studienbericht und in Modul 4 wurden folgende andere nicht-schwerwiegende UE von besonderem Interesse hinzugefügt, die während der Sicherheitsprüfung identifiziert wurden, um eine Vergleichbarkeit zwischen der vorliegenden Studie und der Dosisfindungsstudie CP-4-004 zu gewährleisten:

- Arthralgie
- Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen (speziell: Schmerz in einer Extremität)
- Erhöhte Proteinkinase (speziell Kreatinphosphokinase im Blut erhöht)
- Ödem

Zudem wurden nach Rückmeldung der U. S. Food and Drug Administration (FDA) post hoc noch folgende UE von besonderem Interesse hinzugefügt:

- Pankreatitis: „Akute Pankreatitis“ (SMQ, breit).
- Epiphysenerkrankungen: „Epiphysiolyse“ (Lowest Level Term (LLT)), „Verrutschte Oberschenkelkapitalepiphyse“ (LLT).

Im Studienbericht und in Modul 4 des Nutzendossiers wurden noch folgende, nicht a priori definierte UE von besonderem Interesse ausgewertet:

- Neoplasien: „Gutartige, bösartige und nicht spezifizierte Neubildungen (einschließlich Zysten und Polypen)“ (Systemorganklasse (SOC)).
- Intrakranielle Blutung und intrakranielles Aneurysma: „Hämorrhagische Gefäßerkrankungen des Zentralnervensystems“ (SMQ, enger Bereich), „Erkrankungen im Zusammenhang mit Blutungen des Zentralnervensystems und zerebrovaskulären Unfällen“ (SMQ, enger Bereich).
- Intrakranielle Hypertonie: „Erhöhter intrakranieller Druck und Hydrozephalus“ (High Level Group Terms (HLGT), alle Pfade) und ausgewählte PT („Liquordruck erhöht“).

Die Bewertung der lokalen Verträglichkeit erfolgte durch Untersuchung der Injektionsstellen (durch die Prüffärztin / den Prüfarzt, wenn bei der Studienvsiste (außer beim Screening) eine Reaktion auftrat) sowie anhand von anamnestischen Daten und Aufzeichnungen im Patiententagebuch. Die Beobachtungen lokaler Reaktionen an der Injektionsstelle wurden im eCRF dokumentiert. Die Patientinnen und Patienten sowie Eltern/Erziehungsberechtigten wurden darin geschult, mindestens einmal wöchentlich jede Reaktion an der Injektionsstelle in ihr Tagebuch einzutragen. Laut Studienbericht erfolgte die Datenerhebung im Somatrogon-Arm im Anschluss an die wöchentliche Injektion und im Genotropin-Arm ebenfalls einmal wöchentlich.

Eine abnorme Reaktion an der Injektionsstelle war definiert als:

- Reaktion an der Injektionsstelle, die zum Zeitpunkt der Studienvsiste oder bei einem Telefonanruf gemeldet oder beobachtet wurde und von mittlerer bis schwerer Intensität war.
- Reaktion an der Injektionsstelle, die zwischen der letzten und der jetzigen Studienvsiste aufgetreten war oder zum Zeitpunkt der Studienvsiste verblieb und ärztlich behandelt

werden musste, oder Reaktion an der Injektionsstelle, die auf eine frühere Injektion zurückzuführen war.

- Jede andere Reaktion an der Injektionsstelle, die nach dem Urteil der Prüffärztin / des Prüffarztes als abnormal angesehen wurde und nicht zu den üblichen Reaktionen bei subkutanen Injektionen gehörte.
- Schmerzwert ≥ 4 (wie im Patiententagebuch angegeben).

Wenn eine Reaktion an der Injektionsstelle die oben definierten Kriterien für "abnormal" erfüllte, wurde sie als UE bewertet. Schmerzen an der Injektionsstelle wurden mithilfe der Pain Rating Scale auf einer Skala von 0 (kein Schmerz) bis 5 (schlimmster Schmerz) im Patiententagebuch erfasst. Zudem wurden alle Patientinnen und Patienten sowie Eltern/Erziehungsberechtigten während der Studienvisiten zu möglichen Schmerzen an der Injektionsstelle befragt.

Bewertung

Die Operationalisierung ist im Wesentlichen nachvollziehbar. Im Studienprotokoll wurden UE von besonderem Interesse als Sicherheitsendpunkte genannt; die Definition, welche Ereignisse zu welchen UE von besonderem Interesse zusammengefasst werden, erfolgte allerdings erst im Studienbericht und in Modul 4 des Nutzendossiers. Daher werden in der Nutzenbewertung nur die UE von besonderem Interesse dargestellt, die als Sicherheitsendpunkt im Studienprotokoll präspezifiziert waren und jene, die post hoc von den Zulassungsbehörden gefordert wurden.

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“.

Das Auftreten von UE ist patientenrelevant. Die Patientenrelevanz von Laborparametern ist unklar.

Validität

Die Erhebung der UE wird als valide bewertet.

2.3.5 Übersicht der Erhebungszeitpunkte

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

Tabelle 6: Erhebungszeitpunkte der berücksichtigten Endpunkte in der Studie CP-4-006

Studienvisite \ Endpunkt	Screening ¹⁾	Base-line (Tag 1)	Monat 3	Monat 6	Monat 9	Monat 12 ²⁾	Sicherheitsnachbeobachtung ³⁾⁴⁾
Todesfälle		kontinuierlich ⁵⁾					
Körpergröße (z-Score)	x	x	x	x	x	x	
QoLISSY		x				x	
Unerwünschte Ereignisse		kontinuierlich ⁵⁾⁶⁾					

¹⁾ Die Screeningphase konnte um 4 Wochen verlängert werden, wenn sich die Studienvisite einer Patientin / eines Patienten aufgrund einer harmlosen Erkrankung oder technischer Probleme verzögerte (maximal 12 Wochen).

²⁾ Bei Patientinnen und Patienten, die die Studie vorzeitig abbrachen, wurden dieselben Erhebungen durchgeführt, die für Monat 12 geplant waren (es sollte mindestens eine Messung der Körpergröße zu Monat 12 erfolgen). Zudem sollten alle Anstrengungen unternommen werden zur Durchführung der Sicherheitsnachbeobachtung.

³⁾ Einmonatige Sicherheitsnachbeobachtung bei Studienteilnehmenden, die die 12-monatige vergleichende Studie vorzeitig abbrachen oder die nicht an der offenen einarmigen Extensionsstudie teilnahmen.

⁴⁾ Telefonischer Kontakt nur für Patientinnen und Patienten, die nicht an der offenen einarmigen Extensionsstudie teilnahmen.

⁵⁾ Zu jeder Studienvisite ab Baseline erhoben.

⁶⁾ Reaktionen an der Injektionsstelle / Schmerzen an der Injektionsstelle wurden von den Studienteilnehmenden sowie Eltern/Erziehungsberechtigten mindestens einmal wöchentlich im Patiententagebuch erfasst.

Abkürzungen: QoLISSY: Quality of Life in Short Stature Youth.

2.4 Statistische Methoden

Die letzte SAP-Version (4.0) wurde am 14.06.2019 finalisiert.

Analysepopulationen

Im SAP sind folgende Analysepopulationen definiert:

- Full Analysis Set (FAS): Alle randomisierten Patientinnen und Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Patientinnen und Patienten werden nach ihrer randomisierten Behandlungsgruppe ausgewertet.
- Sicherheitspopulation: Alle Patientinnen und Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Patientinnen und Patienten werden nach ihrer aktuell verabreichten Studienmedikation ausgewertet.

Post hoc wurde für die Auswertung des QoLISSY folgende Analysepopulation definiert:

- PRO (Patient-Reported Outcome) Analysis Set: Alle Patientinnen und Patienten aus dem FAS, die mindestens zu Baseline den QoLISSY-Fragebogen ausgefüllt hatten.

Datenschnitte

Es wurde keine Interimsanalyse durchgeführt. Vorliegend ist der Datenschnitt nach Ende der vergleichenden Studie zu Monat 12 (23.08.2019).

Präspezifizierung geplanter und durchgeführter Analysen

Im SAP waren zur Bestimmung der Mittelwertdifferenzen zwischen den beiden Behandlungsgruppen für die Körpergröße (z-Score) (nach 12 Monaten) Kovarianzanalysen (ANCOVA) unter Berücksichtigung der Stratifizierungsfaktoren der Randomisierung geplant.

Der Endpunkt „QoLISSY“ und die Endpunkte der Sicherheit waren zur deskriptiven Auswertung vorgesehen. Post hoc wurde eine Responderanalyse für eine Verbesserung in der Gesamtskala des QoLISSY zu Monat 12 mit einer Relevanzschwelle von 15 % der Skalenspannweite vorgelegt. Dies entspricht einer Zunahme auf der Skala im Vergleich zum Baseline-Wert von ≥ 15 Punkten. Für die Sicherheitsendpunkte wurden post hoc stratifizierte relative Risiken (RR) mit 95%-Konfidenzintervall (KI) mittels Logit-Methode und Nullzellenkorrektur berechnet. P-Werte wurden mittels stratifizierten Cochran-Mantel-Haenszel-Test berechnet.

Fehlende Werte und Daten-Imputation

Laut SAP sollten fehlende Werte des Endpunkts „Körpergröße (z-Score)“ über multiple Imputationen ersetzt werden, unter der Annahme, dass die Werte nicht zufällig fehlen (Missing Not At Random (MNAR)). Post hoc wurden fehlende Werte im QoLISSY mittels „Missing At Random“ (MAR)-Annahme imputiert. Es konnte in den Studienunterlagen keine Begründung zu den Missing-Annahmen identifiziert werden.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 7: Verzerrungspotential der Studie CP-4-006

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
CP-4-006	Ja ¹⁾	Ja ²⁾	Nein ³⁾	Nein ³⁾	Nein	Nein	Hoch

¹⁾ Patientinnen und Patienten wurden 1:1 auf die Behandlungsarme (Somatrogen:Genotropin) randomisiert. Die Randomisierung erfolgte in Blöcken (Blockrandomisierung), jeweils stratifiziert nach „Region“ (Westeuropa, Israel, Australien, Neuseeland, Kanada, USA vs. Mittel- und Osteuropa, Griechenland, Türkei, Lateinamerika, Asien (mit Ausnahme von Indien und Vietnam) vs. Indien, Vietnam), „Höchste gemessene GH-Konzentration beim Screening“ (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml) und „Chronologisches Alter“ (≥ 3 bis ≤ 7 vs. > 7 Jahre).

²⁾ Zuteilung der Studienteilnehmenden auf die Studienarme erfolgte über ein IRT.

³⁾ Patientinnen und Patienten sowie Studienpersonal waren aufgrund der unterschiedlichen Verabreichung des Studienmedikaments (wöchentliche vs. tägliche Injektion) unverblindet. Die fehlende Verblindung führt zu einem hohen Verzerrungspotential.

Abkürzungen: GH: Wachstumshormon; IRT: Interaktives Web Response Technologiesystem.

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

Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studie CP-4-006

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Todesfälle	Nein ¹⁾	Ja	Nein	Nein	Niedrig ²⁾
Körpergröße (z-Score)	Nein ¹⁾	Ja	Nein	Nein	Niedrig ³⁾
QoLISSY	Nein ¹⁾	Unklar ⁴⁾	Nein	Nein ⁵⁾	Hoch ⁶⁾
Unerwünschte Ereignisse	Nein ¹⁾	Ja ⁷⁾	Unklar ⁸⁾	Nein	Hoch ⁶⁾

¹⁾ Patientinnen und Patienten sowie Studienpersonal waren aufgrund der unterschiedlichen Verabreichung des Studienmedikaments (wöchentliche vs. tägliche Injektion) unverblindet.

²⁾ Trotz fehlender Verblindung kein hohes Verzerrungspotential, da es sich beim Endpunkt „Todesfälle“ um einen objektiv erhobenen Endpunkt handelt.

³⁾ Trotz fehlender Verblindung kein hohes Verzerrungspotential, da die Messungen durch geschultes Personal mithilfe eines kalibrierten Wand-Stadiometers durchgeführt wurden und der Mittelwert der Messungen für die Analysen verwendet wurde. Eine Verzerrung durch die subjektive Erhebung kann weitestgehend ausgeschlossen werden.

⁴⁾ Der QoLISSY-Fragebogen wurde nur in Ländern eingesetzt, für die eine validierte Übersetzung vorlag, sodass sich die Auswertungen auf N = 54 im Somatrogen-Arm und N = 63 im Genotropin-Arm (PRO Analysis Set) beschränken. Da keine Angaben zu Rücklaufquoten von den jeweiligen Ländern vorliegen, in denen der QoLISSY eingesetzt wurde, ist unklar, ob sich die berichtete Rücklaufquote von $\geq 90\%$ in beiden Behandlungsarmen auf alle Personen bezieht, die den Fragebogen hätten beantworten können.

⁵⁾ In den post hoc durchgeführten Analysen wurden die Stratifizierungsfaktoren der Randomisierung aus Konvergenzgründen nicht berücksichtigt.

⁶⁾ Aufgrund fehlender Verblindung hohes Verzerrungspotential, da es sich um einen subjektiv erhobenen Endpunkt handelt.

⁷⁾ Auswertung erfolgte in der Sicherheitspopulation.

⁸⁾ UE von besonderem Interesse wurden im Studienprotokoll als Sicherheitsendpunkte genannt; die Definition, welche Ereignisse zu welchen UE von besonderem Interesse zusammengefasst werden, erfolgte allerdings erst im Studienbericht und in Modul 4 des Nutzendossiers.

Abkürzungen: ITT: Intention to Treat; PRO: Patient-Reported Outcome; QoLISSY: Quality of Life in Short Stature Youth; UE: Unerwünschtes Ereignis.

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

In die Studie CP-4-006 wurden von den 536 gescreenten Personen insgesamt 228 Patientinnen und Patienten im Verhältnis 1:1 stratifiziert randomisiert auf die Behandlungsarme (Somatrogen:Genotropin) zugeteilt. 4 Personen (N = 3 im Somatrogen-Arm, N = 1 im Genotropin-Arm) erhielten keine Dosis des Studienmedikaments, sodass das FAS und die Sicherheitspopulation 224 Personen umfasste (N = 109 Somatrogen, N = 115 Genotropin). 1 Person brach die Studie aufgrund eines UE im Somatrogen-Arm ab und 1 Person im Genotropin-Arm aufgrund des Entzugs der Einverständniserklärung (Tabelle 9). Nach der 12-monatigen vergleichenden Behandlungsdauer gingen ca. 95 % der Studienteilnehmenden in beiden Behandlungsarmen in die einarmige, offene Extensionsstudie über. Die mediane Behandlungsdauer mit 364 Tagen im Somatrogen-Arm und 358 Tagen im Genotropin-Arm sowie die mediane Beobachtungsdauer mit 371 und 370 Tagen waren vergleichbar zwischen den Behandlungsarmen. Bei 12 Personen im Somatrogen-Arm und keiner Person im Genotropin-Arm wurde die Dosis während der Studie reduziert. Angaben zu Dosisunterbrechungen oder -erhöhungen konnten nicht identifiziert werden.

Tabelle 9: Allgemeine Angaben zur Studie CP-4-006

Studie CP-4-006 Allgemeine Angaben	Somatrogen n (%)	Genotropin n (%)
FAS ¹⁾	109 (100)	115 (100)
Sicherheitspopulation ¹⁾	109 (100)	115 (100)
Abbruch der Einnahme der Studienmedikation	0 (0)	0 (0)
Abbruch der Studie aufgrund von:		
UE	1 (0,9)	1 (0,9)
Lost to Follow-up	1 (0,9)	0 (0)
Tod	0 (0)	0 (0)
Entzug der Einverständniserklärung	0 (0)	1 (0,9)
Übergang in einarmige, offene Extensionsphase	104 (95,4)	108 (93,9)
Mediane Behandlungsdauer ²⁾ , Tage (min; max)	364 (63; 413)	358 (78; 382)
Mediane Beobachtungsdauer ³⁾ , Tage (min; max)	371 (355; 414)	370 (349; 404)
Dosisanpassungen		
Unterbrechung	k. A.	k. A.
Erhöhung	k. A.	k. A.
Reduktion aufgrund von IGF-1 (z-Score) > 2	12 (11,0)	0 (0)

¹⁾ Definition, siehe Kapitel 2.4 Statistische Methoden.

²⁾ Behandlungsdauer Somatrogen = Datum der letzten Dosis - Datum der ersten Dosis + 7; Behandlungsdauer Genotropin = Datum der letzten Dosis - Datum der ersten Dosis + 1.

³⁾ Die Beobachtungsdauer bezieht sich auf den Endpunkt „Körpergröße“ (z-Score).

Abkürzungen: FAS: Full Analysis Set; IGF-1: Insulinähnlicher Wachstumsfaktor 1; k. A.: keine Angabe; UE: Unerwünschtes Ereignis.

Die demographischen und krankheitsspezifischen Patientencharakteristika sind in beiden Behandlungsarmen weitestgehend vergleichbar (Tabelle 10). Insgesamt nahmen mehr Kinder im Alter > 7 Jahre (ca. 60 %) in beiden Behandlungsarmen sowie mehr Jungen (75,2 % im Somatrogen-Arm, 68,7 % im Genotropin-Arm) an der Studie teil. Es ist unklar, warum nicht alle Studienteilnehmenden zu Baseline eine endokrine Erkrankung sowie einen Wachstumshormonmangel aufwiesen. Im Somatrogen-Arm lag der Anteil bei 92,7 bzw. 87,2 % und im Genotropin-Arm bei 89,6 und 88,7 %.

Tabelle 10: Charakterisierung der Studienpopulation der Studie CP-4-006; Sicherheitspopulation

Studie CP-4-006 Charakterisierung der Studienpopulation¹⁾	Somatrogon N = 109	Genotropin N = 115
<i>Alter (Jahre)</i> MW (SD) Median (min; max)	7,8 (2,7) 7,9 (3,0; 12,0)	7,6 (2,4) 7,8 (3,1; 11,9)
<i>Altersgruppe (Jahre)²⁾, n (%)</i> ≥ 3 bis ≤ 7 Jahre > 7 Jahre	43 (39,4) 66 (60,6)	47 (40,9) 68 (59,1)
<i>Geschlecht, n (%)</i> männlich weiblich	82 (75,2) 27 (24,8)	79 (68,7) 36 (31,3)
<i>Abstammung (genetisch), n (%)</i> Weiß Asiatisch Schwarz/Afroamerikanisch Amerikanischer Ureinwohner / Alaska Pazifische Inselbewohner andere	81 (74,3) 24 (22,0) 0 (0) 1 (0,9) 0 (0) 3 (2,8)	86 (74,8) 21 (18,3) 2 (1,7) 0 (0) 1 (0,9) 5 (4,3)
<i>Region²⁾³⁾⁴⁾, n (%)</i> Region 1 Region 2 Region 3	45 (41,3) 52 (47,7) 12 (11,0)	50 (43,5) 51 (44,4) 14 (12,2)
<i>Höchste gemessene GH-Konzentration im Stimulationstest²⁾, n (%)</i> ≤ 3 ng/ml > 3 bis ≤ 7 ng/ml > 7 ng/ml	22 (20,2) 53 (48,6) 34 (31,2)	21 (18,3) 56 (48,7) 38 (33,0)
<i>Höchste gemessene GH-Konzentration im Stimulationstest (ng/ml)</i> MW (SD) Median (min; max)	5,5 (2,8) 5,9 (0,1; 9,9)	5,8 (2,6) 6,2 (0,1; 9,9)
<i>Körpergröße (cm)</i> MW (SD) Median (min; max)	110,0 (15,3) 109,8 (75,1; 139,6)	109,9 (13,9) 110,5 (77,1; 144,3)
<i>Körpergröße (z-Score)</i> MW (SD) Median (min; max)	-2,9 (1,3) -2,5 (-7,5; -1,1)	-2,8 (1,3) -2,6 (-10,0; -0,5)
<i>Knochenalter⁵⁾ (Jahre)</i> MW (SD) Median (min; max)	5,5 (2,7) 5,0 (1,0; 11,0)	5,2 (2,5) 5,0 (1,3; 11,0)
<i>Vorherige Einnahme von Medikamenten, n (%)</i>	48 (44,0)	47 (40,9)
<i>Krankheitsdauer in Jahre</i> MW (SD) Median (min; max)	k. A. k. A.	k. A. k. A.
Vorerkrankungen nach MedDRA-Systemorganklasse / Preferred Term, n (%)		
<i>Endokrine Erkrankungen</i> Wachstumshormonmangel	101 (92,7) 95 (87,2)	103 (89,6) 102 (88,7)

- ¹⁾ Angaben aus dem Studienbericht und Modul 4 des Nutzendossiers.
- ²⁾ Stratifizierungsfaktoren der Randomisierung.
- ³⁾ Region 1: Westeuropa, Israel, Griechenland, Australien, Neuseeland, Kanada, USA; Region 2: Mittel- und Osteuropa, Türkei, Lateinamerika, Asien (außer Indien und Vietnam); Region 3: Indien, Vietnam.
- ⁴⁾ Bei einer verdeckten Datenüberprüfung wurde festgestellt, dass im Studienprotokoll und im SAP Griechenland fälschlicherweise Region 2 zugeordnet worden war. Die Regionen wurden in den Tabellen und Listen korrigiert und Griechenland wurde Region 1 zugeordnet.
- ⁵⁾ Knochenalter nach der Atlasmethode von Greulich & Pyle.

Abkürzungen: GH: Wachstumshormon; k. A.: keine Angabe; MedDRA: Medical Dictionary for Regulatory Activities; MW: Mittelwert; SAP: Statistischer Analyseplan; SD: Standardabweichung.

Begleitmedikation

Die Angaben zu den Begleitmedikationen mit mindestens 20 % in einem der beiden Behandlungsarme sind in Tabelle 11 für die Sicherheitspopulation abgebildet. Es zeigte sich ein Unterschied in der Begleittherapie von mehr als 10 % zwischen den Studienarmen: Antihistaminika zur systemischen Anwendung wurden häufiger im Somatrogen-Arm (32,1 %) als im Genotropin-Arm (21,7 %) eingenommen.

Tabelle 11: Begleitmedikation mit einer Häufigkeit von ≥ 20 % in beiden Behandlungsarmen in der Studie CP-4-006; Sicherheitspopulation

Studie CP-4-006 Begleitmedikation mit einer Häufigkeit von ≥ 20 % ATC-Klassifikation	Somatrogen N = 109 n (%)	Genotropin N = 115 n (%)
Erhalt mindestens einer Begleitmedikation	87 (79,8)	88 (76,5)
Analgetika	27 (24,8)	32 (27,8)
Paracetamol	26 (23,9)	29 (25,2)
Antibiotika zur systemischen Anwendung	34 (31,2)	32 (27,8)
Antihistaminika zur systemischen Anwendung	35 (32,1)	25 (21,7)
Antiphlogistika und Antirheumatika	35 (32,1)	39 (33,9)
Ibuprofen	34 (31,2)	37 (32,2)

Abkürzungen: ATC: Anatomisch-Therapeutisch-Chemisch.

3.2 Mortalität

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

3.3 Morbidität

Körpergröße (z-Score)

Die Veränderung der Körpergröße wurde während der 12-monatigen Studiendauer als z-Score ermittelt. Die mediane Beobachtungsdauer für die Körpergröße betrug 371 Tage (Spannweite: 355–414 Tage) im Somatrogen-Arm und 370 Tage (Spannweite: 349–404 Tage) im Genotropin-Arm. Bezüglich der Körpergröße befinden sich die untersuchten Kinder im gesamten Untersuchungszeitraum im Mittel unterhalb des Mittelwerts der Referenzpopulation (Tabelle 12).

Im Vergleich zu Baseline erreichten die Patientinnen und Patienten des Somatrogen-Arms zu Monat 12 eine Veränderung des z-Scores um 0,92 im Least Squares (LS) Mean, die des

Genotropin-Arms um 0,87. Es zeigte sich zwischen den Behandlungsarmen kein statistisch signifikanter Unterschied (LS-Mean-Differenz [95%-KI]: 0,05 [-0,06; 0,16]; $p = 0,388$).

Post hoc durchgeführte Sensitivitätsanalysen mit verschiedenen multiplen Imputationen und ohne Imputation bestätigen die Robustheit des Ergebnisses. Die Subgruppenanalysen zeigten keine statistisch signifikanten Interaktionstests.

Tabelle 12: Veränderung der Körpergröße (z-Score) zu Monat 12 in der Studie CP-4-006; FAS

Studie CP-4-006 Körpergröße (z-Score)	Somatrogon N = 109	Genotropin N = 115
Baseline n (%) MW (SD)	109 (100) -2,9 (1,3)	115 (100) -2,8 (1,3)
Monat 12 n (%) MW (SD)	108 (99,1) -2,0 (1,1)	113 ¹⁾ (98,3) -1,9 (1,1)
Veränderung von Baseline zu Monat 12 n (%) LS Mean [95%-KI]	109 (100) 0,92 [0,82; 1,02]	115 (100) 0,87 [0,78; 0,97]
Gruppenunterschied der Veränderung LS-Mean-Differenz [95%-KI]; p-Wert ²⁾	0,05 [-0,06; 0,16]; 0,388	

¹⁾ Eine Person im Genotropin-Arm schloss die 12-monatige vergleichende Studie zwar ab, hatte aber keine Erhebung der Körpergröße zu Monat 12. Daher ist die Anzahl der Studienteilnehmenden zu Monat 12 N = 113.

²⁾ ANCOVA-Modell schließt Behandlung, Altersgruppe, die höchste gemessene GH-Konzentration im Stimulationstest, Region, Geschlecht und Baseline-Körpergröße (z-Score) als Kovariablen ein. Fehlende Werte wurden mithilfe multipler Imputation ersetzt (FCS-Option). Es bestehen Unklarheiten, ob wie a priori im SAP festgelegt eine Imputation mit MNAR-Annahme durchgeführt wurde oder wie in Modul 4 des Nutzendossiers beschrieben eine MAR-Annahme verwendet wurde. Im Somatrogon-Arm wurden zu Monat 12 für 1 Person und im Genotropin-Arm für 2 Personen fehlende Werte imputiert.

Abkürzungen: ANCOVA: Kovarianzanalyse; FAS: Full Analysis Set; FCS: Fully Conditional Specification; GH: Wachstumshormon; KI: Konfidenzintervall; LS: Least Squares; MAR: Missing At Random; MNAR: Missing Not At Random; MW: Mittelwert; SAP: Statistischer Analyseplan; SD: Standardabweichung.

3.4 Lebensqualität

QoLISSY

Da der QoLISSY-Fragebogen nur in den Ländern eingesetzt wurde, für die eine validierte Übersetzung vorlag, beschränken sich die Auswertungen auf N = 54 im Somatrogon-Arm und N = 63 im Genotropin-Arm. Es liegen nur gepoolte Responderanalysen für die selbstberichteten Angaben der Kinder ≥ 7 Jahre und die elternberichteten Angaben für die Kinder < 7 Jahre im QoLISSY vor. Die mediane Beobachtungsdauer für den QoLISSY betrug im Somatrogon-Arm 370 Tage (Spannweite: 15–414 Tage), im Genotropin-Arm 370 Tage (Spannweite: 6–403 Tage).

Zu Monat 12 zeigte sich für 34,7 % der Patientinnen und Patienten im Somatrogon-Arm und für 23,7 % im Genotropin-Arm eine Verbesserung um mindestens 15 Punkte in der Gesamtskala des QoLISSY (Tabelle 13). Der Unterschied war nicht statistisch signifikant ((RR) [95%-KI]: 1,43 [0,81; 2,50]; $p = 0,218$).

Post hoc durchgeführte Sensitivitätsanalysen mit verschiedenen multiplen Imputationen und ohne Imputation bestätigen die Robustheit des Ergebnisses. Die Subgruppenanalysen zeigten keine statistisch signifikanten Interaktionstests.

Tabelle 13: Ergebnisse zu Responderanalysen für eine Verbesserung im QoLISSY (Gesamtskala) von ≥ 15 Punkten in der Studie CP-4-006; PRO Analysis Set

Studie CP-4-006 QoLISSY-Gesamtskala¹⁾	Somatrogen N = 54²⁾³⁾	Genotropin N = 63²⁾³⁾
Baseline n (%) ⁴⁾	54 (100)	63 (100)
MW (SD)	61,6 (20,6)	58,1 (20,0)
Median (min; max)	64,4 (0; 93,4)	59,0 (11,8; 93,8)
Monat 12 n (%) ⁵⁾	49 (90,7)	59 (93,7)
MW (SD)	72,9 (17,5)	66,7 (20,9)
Median (min; max)	76,7 (12,2; 100)	72,6 (11,1; 96,5)
Verbesserung zu Monat 12 um ≥ 15 Punkte		
n/N (%)	17/49 (34,7)	14/59 (23,7)
RR [95%-KI] ⁶⁾ ; p-Wert	1,43 [0,81; 2,50]; 0,218	

¹⁾ Skala von 0 bis 100; höhere Werte bedeuten eine bessere Lebensqualität.

²⁾ PRO Analysis Set: Definition, siehe Kapitel 2.4 Statistische Methoden.

³⁾ Der QoLISSY-Fragebogen wurde nur in den folgenden Ländern eingesetzt, für die eine validierte Übersetzung vorliegt: USA, Australien, Neuseeland, Belarus, Russland, Ukraine, Großbritannien, Spanien.

⁴⁾ Umfasst die selbstberichtete Kinderversion des QoLISSY für Studienteilnehmende ≥ 7 Jahre (N = 35 im Somatrogen-Arm, N = 35 im Genotropin-Arm) und den Eltern-Proxy-Fragebogen für Studienteilnehmende < 7 Jahre (N = 19 im Somatrogen-Arm, N = 28 im Genotropin-Arm).

⁵⁾ Umfasst die selbstberichtete Kinderversion des QoLISSY für Studienteilnehmende ≥ 7 Jahre (N = 32 im Somatrogen-Arm, N = 33 im Genotropin-Arm) und den Eltern-Proxy-Fragebogen für Studienteilnehmende < 7 Jahre (N = 17 im Somatrogen-Arm, N = 26 im Genotropin-Arm).

⁶⁾ Post hoc berechnet: RR mit 95%-KI wurde durch ein generalisiertes lineares Modell mit einer Binomialverteilungsannahme und mit natürlichem Logarithmus als Link-Funktion errechnet. Es wurde für den Baseline-Wert adjustiert, auf weitere Stratifizierungsfaktoren wurde aus Konvergenzgründen verzichtet. Fehlende Werte wurden mithilfe multipler Imputation ersetzt (FCS-Option und MAR-Annahme). Es ist unklar, ob Personen, für die Werte fehlten (N = 5 im Somatrogen-Arm, N = 4 im Genotropin-Arm) als Responder oder Non-Responder in die Analyse eingegangen sind. Bei einer Imputation als Non-Responder wäre jedoch nicht von einer abweichenden Gesamtaussage der Analyse auszugehen.

Abkürzungen: FCS: Fully Conditional Specification; KI: Konfidenzintervall; MAR: Missing At Random; MW: Mittelwert; PRO: Patient-Reported Outcome; QoLISSY: Quality of Life in Short Stature Youth; RR: Relatives Risiko; SD: Standardabweichung.

3.5 Sicherheit

Die mediane Beobachtungsdauer für die Sicherheitspopulation in der Studie CP-4-006 betrug im Somatrogen-Arm 373 Tage (Spannweite: 95–414 Tage) und im Genotropin-Arm 371 Tage (Spannweite: 94–589 Tage).

Die zusammenfassende Darstellung der UE ist in Tabelle 14 abgebildet. 87,3 % der Studienteilnehmenden im Somatrogen-Arm und 84,3 % im Genotropin-Arm erlitten im Studienverlauf ein UE jeglichen Schweregrads. Insgesamt traten nur wenige schwere UE (8,3% im Somatrogen-Arm, 5,2 % im Genotropin-Arm) und SUE (2,8 % im Somatrogen-Arm, 1,7 % im Genotropin-Arm) auf. 1 Person im Somatrogen-Arm brach die Studienmedikation und die Studie aufgrund von UE in der SOC „Reaktionen an der Injektionsstelle“ im PT „Erythem an der Injektionsstelle“ und „Verhärtung der Injektionsstelle“ ab. Es zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen. Die Subgruppenanalysen zeigten keine statistisch signifikanten Interaktionstests.

Tabelle 14: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE in der Studie CP-4-006; Sicherheitspopulation

Studie CP-4-006 Zusammenfassung der UE	Somatrogon N = 109 n (%)	Genotropin N = 115 n (%)	RR [95%-KI] ¹⁾ p-Wert ²⁾
UE	95 (87,2)	97 (84,3)	-
Schwere UE	9 (8,3)	6 (5,2)	1,91 [0,79; 4,58] 0,213
SUE	3 (2,8)	2 (1,7)	1,99 [0,45; 8,72] 0,229
UE, das zum Abbruch der Studienmedikation führte	1 (0,9)	0 (0)	6,00 [0,31; 116,61] 0,127

¹⁾ Post hoc berechnet: Stratifiziertes RR mit 95%-KI mittels Logit-Methode mit Nullzellenkorrektur berechnet.

²⁾ Cochran-Mantel-Haenszel-Test: Stratifiziert nach „Alter“ (≥ 3 bis ≤ 7 vs. > 7 Jahre), „Region“ (Region 1 vs. Region 2 vs. Region 3), „Höchste gemessene GH-Konzentration im Stimulationstest“ (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml) und „Geschlecht“ (männlich vs. weiblich).

Abkürzungen: GH: Wachstumshormon; KI: Konfidenzintervall; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Unerwünschte Ereignisse

UE, die bei ≥ 10 % der Patientinnen und Patienten in mindestens einem Studienarm auftraten, sind in Tabelle 15 dargestellt. Es zeigte sich für die SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (RR = 1,48 [95%-KI: 1,10; 1,99]; p = 0,004) und dazugehörigem PT „Schmerzen an der Injektionsstelle“ (RR [95%-KI] = 1,37 [1,02; 1,85]; p = 0,021) sowie für die SOC „Augenerkrankungen“ (RR [95%-KI] = 2,97 [0,98; 9,05]; p = 0,010) jeweils ein statistisch signifikanter Unterschied zuungunsten von Somatrogon im Vergleich zu Genotropin. In der SOC „Untersuchungen“ zeigte sich ein statistisch signifikanter Unterschied zugunsten von Somatrogon im Vergleich zu Genotropin (RR [95%-KI] = 0,48 [0,25; 0,91]; p = 0,004).

Tabelle 15: UE mit Inzidenz ≥ 10 % oder bei ≥ 10 Patientinnen und Patienten und mit einer Inzidenz von ≥ 1 % in einem Studienarm nach MedDRA-Systemorganklassen und Preferred Terms in der Studie CP-4-006; Sicherheitspopulation

Studie CP-4-006 UE mit Inzidenz ≥ 10 %	Somatrogon N = 109 n (%)	Genotropin N = 115 n (%)	Somatrogon vs. Genotropin RR [95%-KI] ¹⁾ p-Wert ²⁾
MedDRA-Systemorganklasse Preferred Term	n (%)	n (%)	
Erkrankungen des Blutes und des Lymphsystems	12 (11,0)	15 (13,0)	0,84 [0,44; 1,61] 0,549
Erkrankungen des Gastrointestinaltrakts	20 (18,3)	23 (20,0)	0,81 [0,49; 1,36] 0,525
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	54 (49,5)	38 (33,0)	1,48 [1,10; 1,99] 0,004
Schmerzen an der Injektionsstelle	43 (39,4)	29 (25,2)	1,37 [1,02; 1,85] 0,021
Fieber	18 (16,5)	16 (13,9)	1,42 [0,85; 2,36] 0,453

Studie CP-4-006 UE mit Inzidenz $\geq 10\%$	Somatrogon N = 109	Genotropin N = 115	Somatrogon vs. Genotropin
<i>MedDRA-Systemorganklasse Preferred Term</i>	<i>n (%)</i>	<i>n (%)</i>	<i>RR [95%-KI]¹⁾ p-Wert²⁾</i>
Infektionen und parasitäre Erkrankungen	59 (54,1)	56 (48,7)	1,17 [0,91; 1,50] 0,222
Nasopharyngitis	25 (22,9)	29 (25,2)	0,80 [0,51; 1,27] 0,694
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	14 (12,8)	9 (7,8)	1,22 [0,63; 2,35] 0,231
Untersuchungen	8 (7,3)	22 (19,1)	0,48 [0,25; 0,91] 0,004
Stoffwechsel- und Ernährungsstörungen	11 (10,1)	8 (7,0)	1,36 [0,64; 2,86] 0,332
Skelettmuskulatur-, Bindegewebs-, und Knochenerkrankungen	14 (12,8)	15 (13,0)	0,80 [0,42; 1,54] 0,579
Erkrankungen des Nervensystems	22 (20,2)	26 (22,6)	1,17 [0,79; 1,73] 0,744
Kopfschmerzen	18 (16,5)	25 (21,7)	1,07 [0,71; 1,62] 0,392
Erkrankungen der Atemwege, des Brustraums und Mediastinums	24 (22,0)	21 (18,3)	1,16 [0,72; 1,87] 0,268
Erkrankungen der Haut und des Unterhautgewebes	14 (12,8)	14 (12,2)	1,03 [0,49; 2,15] 0,922
Augenerkrankungen	10 (9,2)	1 (0,9)	2,97 [0,98; 9,05] 0,010

¹⁾ Post hoc berechnet: Stratifiziertes RR mit 95%-KI mittels Logit-Methode mit Nullzellenkorrektur berechnet.

²⁾ Cochran-Mantel-Haenszel-Test: Stratifiziert nach „Alter“ (≥ 3 bis ≤ 7 vs. > 7 Jahre), „Region“ (Region 1 vs. Region 2 vs. Region 3), „Höchste gemessene GH-Konzentration im Stimulationstest“ (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml) und „Geschlecht“ (männlich vs. weiblich).

Abkürzungen: GH: Wachstumshormon; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

Schwere unerwünschte Ereignisse

In der Studie CP-4-006 lag die Inzidenz schwerer UE auf Ebene der SOC und PT nur in der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ über 5 %. Im Somatrogon-Arm erlitten 7 Personen (6,4 %) und im Genotropin-Arm 3 Personen (2,6 %) ein schweres UE. Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen (RR [95%-KI] = 1,93 [0,72; 5,21], p = 0,141).

Schwerwiegende unerwünschte Ereignisse

Die Inzidenz schwerwiegender UE auf Ebene der SOC und PT war in beiden Behandlungsarmen unter 5 %.

UE von besonderem Interesse

UE von besonderem Interesse (jeglicher Schweregrad) sind in Tabelle 16 dargestellt. Es zeigten sich statistisch signifikante Unterschiede zuungunsten von Somatrogen für „Reaktionen an der Injektionsstelle“ (RR [95%-KI] = 1,43 [1,06; 1,93], $p = 0,004$) und „Immunogenität“ (RR [95%-KI] = 1,78 [0,91; 3,51], $p = 0,034$). Schwere UE von besonderem Interesse traten bei 7 Personen (6,4 %) im Somatrogen-Arm und 3 Personen (2,6 %) im Genotropin-Arm für „Reaktionen an der Injektionsstelle“ (RR [95%-KI] = 1,93 [0,72; 5,21], $p = 0,141$) und bei jeweils 1 Person (0,9 %) im Somatrogen-Arm und keiner Person im Genotropin-Arm für „Immunogenität“ (RR [95%-KI] = 3,00 [0,19; 47,96], $p = 0,317$) bzw. „Pankreatitis“ auf; für Pankreatitis wurden vom pU keine stratifizierte RR berichtet (unstratifiziertes RR [95%-KI] = 3,16 [0,13; 76,84], $p = 0,487$). Es zeigten sich für die drei UE von besonderem Interesse keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen. Es traten keine schwerwiegenden UE von besonderem Interesse auf.

Für das UE von besonderem Interesse „Pankreatitis“ (jeglicher Schweregrad) zeigte sich ein statistisch signifikanter Interaktionstest für das Subgruppenmerkmal „Geschlecht“ ($p = 0,039$). Es zeigte sich in beiden Subgruppen (Jungen und Mädchen) kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen (Jungen: RR [95%-KI] = 1,01 [0,46; 2,23], $p = 0,212$; Mädchen: RR [95%-KI] = 1,47 [0,50; 4,32], $p = 0,293$).

Tabelle 16: UE von besonderem Interesse (jeglicher Schweregrad) in der Studie CP-4-006; Sicherheitspopulation

Studie CP-4-006 UE von besonderem Interesse (unabhängig vom Schweregrad)	Somatrogen N = 109 n (%)	Genotropin N = 115 n (%)	Somatrogen vs. Genotropin RR [95%-KI] ¹⁾ p-Wert ²⁾
Reaktionen an der Injektionsstelle ³⁾	47 (43,1)	29 (25,2)	1,43 [1,06; 1,93] 0,004
Immunogenität ³⁾	20 (18,3)	9 (7,8)	1,78 [0,91; 3,51] 0,034
Beeinträchtigung des Glukosestoffwechsels ³⁾	0 (0)	3 (2,6)	0,46 [0,09; 2,40] 0,162
Beeinträchtigung der Schilddrüsenfunktion ³⁾	10 (9,2)	11 (9,6)	1,21 [0,64; 2,30] 0,983
Kortisolveränderungen ³⁾	0 (0)	1 (0,9)	0,33 [0,02; 7,45] 0,317
Pankreatitis ⁴⁾	12 (11,0)	14 (12,2)	1,15 [0,61; 2,18] 0,608
Epiphysenerkrankungen ⁴⁾	0 (0)	0 (0)	n. b.

¹⁾ Post hoc berechnet: Stratifiziertes RR mit 95%-KI mittels Logit-Methode mit Nullzellenkorrektur berechnet.

²⁾ Cochran-Mantel-Haenszel-Test: Stratifiziert nach „Alter“ (≥ 3 bis ≤ 7 vs. > 7 Jahre), „Region“ (Region 1 vs. Region 2 vs. Region 3), „Höchste gemessene GH-Konzentration im Stimulationstest“ (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml) und „Geschlecht“ (männlich vs. weiblich).

³⁾ Im Studienprotokoll als Sicherheitsendpunkte genannt; die Definition, welche Ereignisse zu welchen UE von besonderem Interesse zusammengefasst werden, erfolgte allerdings erst im Studienbericht und in Modul 4 des Nutzendossiers.

⁴⁾ Post hoc von den Zulassungsbehörden gefordert.

Abkürzungen: GH: Wachstumshormon; KI: Konfidenzintervall; n. b.: nicht berechenbar; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Somatrogen

Somatrogen (Ngenla®) wird angewendet zur Behandlung von Kindern und Jugendlichen ab einem Alter von 3 Jahren mit Wachstumsstörung durch unzureichende Ausschüttung von Wachstumshormon. Gemäß Fachinformation beträgt die empfohlene Dosierung 0,66 mg/kg Körpergewicht, verabreicht einmal wöchentlich als subkutane Injektion [12]. Die Behandlung mit Somatrogen soll bei Anzeichen für einen Schluss der Epiphysenfugen und wenn die jährliche Wachstumsgeschwindigkeit unter 2 cm/Jahr fällt oder das Knochenalter > 14 Jahre bei Mädchen oder > 16 Jahre bei Jungen beträgt beendet werden [12].

In der Studie CP-4-006 wurden 224 Kinder und Jugendliche ab einem Alter von 3 Jahren mit einem diagnostizierten GHD eingeschlossen, die zuvor keine Therapie mit rekombinanten humanen Wachstumshormonen (r-hGH) erhalten hatten, eine verminderte Körpergröße und Wachstumsgeschwindigkeit aufwiesen und deren Knochenalter (< 10 Jahre für Mädchen, < 11 Jahre für Jungen) nicht älter als das chronologische Alter (< 11 Jahre für Mädchen, < 12 Jahre für Jungen) war. Das maximale Knochenalter in beiden Behandlungsarmen betrug 11 Jahre (siehe Tabelle 10), sodass eine Aussage zu älteren Kindern und Jugendlichen, deren Knochenalter ≥ 10 bis ≤ 14 Jahre (Mädchen) bzw. > 11 bis ≤ 16 Jahre (Jungen) beträgt, nicht möglich ist.

In die Studie wurde keine Person aus Deutschland eingeschlossen. Die Studie wurde weltweit in 21 Ländern durchgeführt, darunter Länder in Europa, Nordamerika sowie Australien. Es liegen keine Informationen vor, wie viele Personen aus welchen Ländern stammten. Jedoch waren rund zwei Drittel der Patientinnen und Patienten kaukasischer Abstammung. Es ist daher von einer Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

4.2 Design und Methodik der Studie

Die Studie CP-4-006 (C0311009) ist eine offene, randomisierte, aktiv kontrollierte, multizentrische Phase-III-Studie im Parallelgruppendesign zur Untersuchung der Wirksamkeit und Sicherheit von wöchentlich verabreichtem Somatrogen im Vergleich zu täglich verabreichtem Genotropin bei pädiatrischen Patientinnen und Patienten mit Wachstumsstörungen aufgrund einer unzureichenden Wachstumshormon-Sekretion. Die Studie gliederte sich in eine Screeningphase, eine 12-monatige vergleichende Behandlungsphase und eine langfristige, einarmige, offene Extensionsphase, in der alle Studienteilnehmenden mit Somatrogen behandelt wurden. Für Studienteilnehmende, die die 12-monatige vergleichende Behandlungsphase vorzeitig abbrechen oder die nicht an der offenen einarmigen Extensionsstudie teilnahmen, war eine Nachbeobachtung der Sicherheit über weitere 4 Wochen vorgesehen. Der in der Nutzenbewertung herangezogene Datenschnitt vom 23.08.2019 erfolgte, nachdem alle Patientinnen und Patienten die 12-monatige vergleichende Behandlungsphase abgeschlossen hatten. Der pU hat keine Daten für die einarmige, offene Extensionsphase vorgelegt.

Insgesamt wurden 228 Patientinnen und Patienten im Verhältnis 1:1 randomisiert auf die Behandlungsarme (Somatrogen:Genotropin) zugeteilt. 4 Personen (N = 3 im Somatrogen-Arm, N = 1 im Genotropin-Arm) erhielten keine Dosis des Studienmedikaments, sodass das FAS und die Sicherheitspopulation 224 Personen umfasste (N = 109 Somatrogen, N = 115 Genotropin). Die Randomisierung erfolgte in Blöcken jeweils stratifiziert nach „Region“ (Westeuropa, Israel, Australien, Neuseeland, Kanada, USA vs. Mittel- und Osteuropa, Griechenland, Türkei, Lateinamerika, Asien (mit Ausnahme von Indien und Vietnam) vs. Indien, Vietnam), „Höchste gemessene Wachstumshormon(GH)-Konzentration beim Screening“ (≤ 3 vs. > 3 bis

≤ 7 vs. >7 bis ≤ 10 ng/ml) und „Chronologisches Alter“ (≥ 3 bis ≤ 7 vs. > 7 Jahre). Bei einer verdeckten Datenüberprüfung wurde festgestellt, dass im Studienprotokoll und im SAP Griechenland fälschlicherweise der Region 2 zugeordnet worden war; die Regionen wurden in den Ergebnistabellen korrigiert und Griechenland wurde Region 1 zugeordnet. Es ist unklar wie viele Personen dies betraf. Tabelle 10 ist zu entnehmen, dass sich die Anzahl der Personen für Region 1 und Region 2 zwischen den Behandlungsarmen nur geringfügig unterscheidet, sodass von keiner wesentlichen Verzerrung bei der Randomisierung ausgegangen wird.

Nach der 12-monatigen vergleichenden Behandlungsdauer gingen 104 Studienteilnehmende (95,4 %) im Somatrogen-Arm und 108 (93,9 %) im Genotropin-Arm in die einarmige, offene Extensionsstudie über. Jeweils 1 Person pro Behandlungsarm brach die Studie vorzeitig ab, sodass 4 Personen im Somatrogen-Arm und 6 Personen im Genotropin-Arm die 12-monatige Studie abschlossen und nicht an der offenen Extensionsstudie teilnahmen.

Die mediane Behandlungsdauer mit 364 Tagen im Somatrogen-Arm und 358 Tagen im Genotropin-Arm sowie die mediane Beobachtungsdauer mit 371 und 370 Tagen waren vergleichbar zwischen den Behandlungsarmen. Bei 12 Personen im Somatrogen-Arm und keiner Person im Genotropin-Arm wurde die Dosis während der Studie reduziert. Angaben zu Dosisunterbrechungen oder -erhöhungen konnten nicht identifiziert werden. Die Patientencharakteristika zu Baseline waren in beiden Behandlungsarmen weitestgehend vergleichbar (Tabelle 10). Insgesamt nahmen mehr Kinder im Alter > 7 Jahre (ca. 60 %) in beiden Behandlungsarmen sowie mehr Jungen (75,2 % im Somatrogen-Arm, 68,7 % im Genotropin-Arm) an der Studie teil. Laut Einschlusskriterien wurden Kinder und Jugendliche mit entweder einem isolierten GHD oder einem GHD als Teil eines multiplen Hypophysenhormon-Mangels eingeschlossen. Es ist daher unklar, warum nicht alle Studienteilnehmenden zu Baseline eine endokrine Erkrankung sowie einen Wachstumshormonmangel aufwiesen (siehe Tabelle 10). Im Somatrogen-Arm lag der Anteil bei 92,7 bzw. 87,2 % und im Genotropin-Arm bei 89,6 und 88,7 %. In der Studie CP-4-006 wurde ein Cut-off von ≤ 10 ng/ml in der höchsten gemessenen GH-Konzentration in zwei verschiedenen GH-Stimulationstests als Einschlusskriterium festgelegt, der laut pU der gängigen klinischen Praxis entspricht. Gemäß aktueller S2e-Leitlinie wird ein Cut-off von < 8 ng/ml in der höchsten gemessenen GH-Konzentration in zwei GH-Stimulationstests zur Diagnostik eines Wachstumshormonmangels im Kindes- und Jugendalter empfohlen [5]. Dieser Cut-off setzt die Messung mit einem Assay voraus, der den Standard 98/574 (1 mg = 3 IU) für rekombinantes Wachstumshormon als Kalibrator verwendet [5].

Das Verzerrungspotential der offenen, randomisierten, aktiv kontrollierten Studie CP-4-006 wird aufgrund der fehlenden Verblindung als hoch eingestuft.

4.3 Mortalität

Todesfälle wurden im Rahmen der Sicherheitserhebung dokumentiert. Im Verlauf der Studie CP-4-006 ist keine Person verstorben.

4.4 Morbidität

In der Endpunktkategorie „Morbidität“ wurde „Körpergröße (z-Score)“ als patientenrelevanter Endpunkt für die Nutzenbewertung herangezogen.

Der primäre Endpunkt „Jährliche Wachstumsgeschwindigkeit“ wurde hingegen aufgrund fehlender Patientenrelevanz nicht für die Nutzenbewertung herangezogen. Der Endpunkt liefert keine über die Körpergröße hinausgehende, für die Nutzenbewertung relevanten Informationen. Die Ergebnisse des primären Endpunkts wurden in Analogie zum Nutzenbewertungsverfahren zu Vosoritid [7] im Anhang dargestellt.

Körpergröße (z-Score)

Die Veränderung der Körpergröße wurde als z-Score ermittelt. Fehlende Werte wurden mittels multipler Imputation ersetzt. Es bestehen jedoch Unklarheiten, ob fehlende Werte wie a priori im SAP festgelegt mittels MNAR-Annahme imputiert wurden oder wie in Modul 4 des Nutzen dossiers beschrieben eine MAR-Annahme verwendet wurde. Da im Somatrogen-Arm zu Monat 12 nur für 1 Person und im Genotropin-Arm für 2 Personen fehlende Werte imputiert wurden, wird davon ausgegangen, dass sich die Ergebnisse nicht wesentlich unterscheiden.

Die mediane Beobachtungsdauer für die Körpergröße war vergleichbar zwischen den Behandlungsarmen (371 Tage im Somatrogen-Arm, 370 Tage im Genotropin-Arm). Bezüglich der Körpergröße befinden sich die untersuchten Kinder im gesamten Untersuchungszeitraum im Mittel unterhalb des Mittelwerts der Referenzpopulation (Tabelle 12). Im Vergleich zu Baseline nahm die Körpergröße (z-Score) zu Monat 12 sowohl bei den Patientinnen und Patienten des Somatrogen-Arms als auch des Genotropin-Arms zu. Es zeigte sich zwischen den Behandlungsarmen kein statistisch signifikanter Unterschied (LS-Mean-Differenz [95%-KI]: 0,05 [-0,06; 0,16]; $p = 0,388$). Post hoc durchgeführte Sensitivitätsanalysen mit verschiedenen multiplen Imputationen und ohne Imputation bestätigen die Robustheit des Ergebnisses. Die Subgruppenanalysen zeigten keine statistisch signifikanten Interaktionstests.

Das Verzerrungspotential für den Endpunkt „Körpergröße (z-Score)“ wird trotz fehlender Verblindung als niedrig eingestuft. Da die Messungen durch geschultes Personal mithilfe eines kalibrierten Wand-Stadiometers durchgeführt wurden und der Mittelwert der Messungen für die Analysen verwendet wurde, kann eine Verzerrung durch die subjektive Erhebung weitestgehend ausgeschlossen werden.

Zusammenfassende Einschätzung zur Morbidität

Für den Endpunkt „Körpergröße (z-Score)“ zeigte sich bei niedrigem Verzerrungspotential kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

4.5 Lebensqualität

In der Endpunktkategorie „Lebensqualität“ wurde „QoLISSY“ als patientenrelevanter Endpunkt für die Nutzenbewertung herangezogen. Es wurden die selbstberichteten Angaben der Kinder und Jugendlichen ≥ 7 Jahre und die elternberichteten Angaben für die Kinder von 3 bis 7 Jahren berücksichtigt. In der Studie CP-4-006 füllten bereits 7-jährige Kinder den Fragebogen selbst aus, obwohl eine direkte Befragung laut Entwicklung erst ab 8 Jahren vorgesehen ist [3]. Der pU begründet diese Abweichung damit, dass im Rahmen der Studie Kindern in diesem Alter zugetraut wurde, eine valide Selbstbewertung abzugeben. Aus den Studienunterlagen geht nicht hervor, wie viele Kinder dies betraf. Laut SAP konnten die Eltern beim Ausfüllen des Fragebogens helfen, sodass dieses Vorgehen in der vorliegenden Konstellation als akzeptabel eingeschätzt wird. Zudem wurde der QoLISSY-Fragebogen nur in den Ländern

eingesetzt, für die eine validierte Übersetzung vorlag. Dadurch beschränken sich die Auswertungen auf 54 Patientinnen und Patienten im Somatrogen-Arm und 63 Patientinnen und Patienten im Genotropin-Arm (PRO Analysis Set). Da keine Informationen vorliegen, wie viele Personen in welchen Ländern in die Studie eingeschlossen wurden, ist dementsprechend unklar, für wie viele Personen ein validierter Fragebogen vorlag und ob sich die berichtete Rücklaufquote von $\geq 90\%$ in beiden Behandlungsarmen auf alle Personen bezieht, die den Fragebogen hätten beantworten können. Der pU legt eine Responderanalyse für eine Verbesserung im QoLISSY von ≥ 15 Punkten im PRO Analysis Set vor. Es wurde nur eine gepoolte Responderanalyse für die selbstberichteten Angaben der Kinder und Jugendlichen und die elternberichteten Angaben im QoLISSY vorgelegt. In den post hoc durchgeführten Analysen wurden die Stratifizierungsfaktoren der Randomisierung aus Konvergenzgründen nicht berücksichtigt. Fehlende Werte im QoLISSY wurden unter der MAR-Annahme durch multiple Imputation ersetzt. Eine Begründung für diese Annahme konnte in den Studienunterlagen nicht identifiziert werden. Es ist unklar, ob Personen, für die Werte fehlten ($N = 5$ im Somatrogen-Arm, $N = 4$ im Genotropin-Arm) als Responder oder Non-Responder in die Analyse eingegangen sind. Bei einer Imputation als Non-Responder wäre jedoch nicht von einer abweichenden Gesamtaussage der Analyse auszugehen.

Die mediane Beobachtungsdauer für den QoLISSY war mit jeweils 370 Tagen in den Behandlungsarmen vergleichbar. Zu Monat 12 zeigte sich für 34,7 % der Patientinnen und Patienten im Somatrogen-Arm und für 23,7 % im Genotropin-Arm eine Verbesserung um mindestens 15 Punkte in der Gesamtskala des QoLISSY (Tabelle 13). Der Unterschied zwischen den Behandlungsgruppen war nicht statistisch signifikant (RR [95%-KI]: 1,43 [0,81; 2,50]; $p = 0,218$). Post hoc durchgeführte Sensitivitätsanalysen mit verschiedenen multiplen Imputationen und ohne Imputation bestätigen die Robustheit des Ergebnisses. Die Subgruppenanalysen zeigten keine statistisch signifikanten Interaktionstests.

Das Verzerrungspotential für den Endpunkt „QoLISSY“ wird aufgrund der fehlenden Verblindung als hoch eingestuft, da es sich um einen subjektiv erhobenen Endpunkt handelt.

Zusammenfassende Einschätzung zur Lebensqualität

Für den Endpunkt „QoLISSY“ zeigte sich bei hohem Verzerrungspotential kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

4.6 Sicherheit

Die mediane Beobachtungsdauer für die Sicherheitspopulation in der Studie CP-4-006 war mit 373 Tagen im Somatrogen-Arm und 371 Tagen im Genotropin-Arm vergleichbar. Die monatige Sicherheitsnachbeobachtung fand nur für Patientinnen und Patienten statt, die nicht an der Extensionsstudie teilnahmen oder die Studie vorzeitig abbrachen. Dies betraf lediglich 12 Personen.

Innerhalb der 12-monatigen Behandlungsphase erlitten 87,3 % der Studienteilnehmenden im Somatrogen-Arm und 84,3 % im Genotropin-Arm ein UE jeglichen Schweregrads. Insgesamt traten nur wenige schwere UE (8,3% im Somatrogen-Arm, 5,2 % im Genotropin-Arm) und SUE (2,8 % im Somatrogen-Arm, 1,7 % im Genotropin-Arm) auf. 1 Person im Somatrogen-Arm brach die Studienmedikation und die Studie aufgrund von einem UE ab. Es zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen.

Hinsichtlich UE jeglichen Schweregrads zeigten sich statistisch signifikante Unterschiede zuungunsten von Somatrogen im Vergleich zu Genotropin für die SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (RR [95%-KI] = 1,48 [1,10; 1,99]; $p = 0,004$) und

dazugehörigem PT „Schmerzen an der Injektionsstelle“ (RR [95%-KI] = 1,37 [1,02; 1,85]; p = 0,021) sowie für die SOC „Augenerkrankungen“ (RR [95%-KI] = 2,97 [0,98; 9,05]; p = 0,010). Im SOC „Untersuchungen“ zeigte sich ein statistisch signifikanter Unterschied zugunsten von Somatrogen im Vergleich zu Genotropin (RR [95%-KI] = 0,48 [0,25; 0,91]; p = 0,004).

Die Inzidenz schwerer UE lag nur in der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ über 5 %. Im Somatrogen-Arm erlitten 7 Personen (6,4 %) und im Genotropin-Arm 3 Personen (2,6 %) ein schweres UE. Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen. Der Schweregrad der UE wurde anhand einer studienindividuellen Einteilung vorgenommen. Die Inzidenz der SUE auf Ebene der SOC und PT war in beiden Behandlungsarmen unter 5 %.

UE von besonderem Interesse wurden im Studienprotokoll als Sicherheitsendpunkte genannt, die Definition, welche Ereignisse zu welchen UE von besonderem Interesse zusammengefasst werden, erfolgte allerdings erst im Studienbericht und in Modul 4 des Nutzendossiers. Es zeigten sich statistisch signifikante Unterschiede zuungunsten von Somatrogen für die UE von besonderem Interesse (jeglicher Schweregrad) „Reaktionen an der Injektionsstelle“ (RR [95%-KI] = 1,43 [1,06; 1,93], p = 0,004) und „Immunogenität“ (RR [95%-KI] = 1,78 [0,91; 3,51], p = 0,034). Die Ereignisse, die für das UE von besonderem Interesse „Reaktionen an der Injektionsstelle“ berücksichtigt wurden, überschneiden sich zum Großteil mit Ereignissen, die im PT „Schmerzen an der Injektionsstelle“ aufgetreten sind. Schwere UE von besonderem Interesse traten für „Reaktionen an der Injektionsstelle“, „Immunogenität“ und „Pankreatitis“ auf. Für Pankreatitis wurden vom pU keine stratifizierten RR berichtet. Es zeigten sich für die drei UE von besonderem Interesse keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen. Es traten keine SUE von besonderem Interesse auf. Ausschließlich für das UE von besonderem Interesse „Pankreatitis“ zeigte sich eine statistisch signifikante Interaktion für das Subgruppenmerkmal „Geschlecht“.

Die Somatrogen-Injektionen erfolgten wöchentlich, die Genotropin-Injektionen hingegen täglich. Dies ist bei der Interpretation des UE von besonderem Interesse „Reaktionen an der Injektionsstelle“ bzw. des PT „Schmerzen an der Injektionsstelle“ zu berücksichtigen, da diese Daten laut Studienbericht im Somatrogen-Arm im Anschluss an die wöchentliche Injektion und im Genotropin ebenfalls einmal wöchentlich mit dem Patiententagebuch erhoben wurden.

Laut EPAR war die wöchentliche Verabreichung von Somatrogen im Allgemeinen gut verträglich und mit Genotropin vergleichbar [6]. Die wichtigsten Sicherheitsbedenken sind schmerzhafte Reaktionen an der Injektionsstelle und Immunogenitätsreaktionen. Die Sicherheit von langwirksamem Somatrogen (der Mechanismus, durch den die GH-Wirkung verlängert wird, und die Dauer der Verlängerung) wird in einer Sicherheitsstudie nach der Markteinführung (PASS) weiter untersucht, um die langfristige Sicherheit von Somatrogen bei pädiatrischen Patientinnen und Patienten in Europa zu bewerten [6].

Das Verzerrungspotential für den Endpunkt „Unerwünschte Ereignisse“ wird aufgrund der fehlenden Verblindung als hoch eingestuft, da es sich um einen subjektiv erhobenen Endpunkt handelt.

Zusammenfassende Einschätzung zur Sicherheit

Insgesamt traten nur wenige schwere UE und SUE in beiden Behandlungsarmen auf. Für schwere UE, SUE und UE, die zum Abbruch der Studienmedikation führten, zeigte sich bei hohem Verzerrungspotential kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Für die UE von besonderem Interesse (jeglicher Schweregrad) „Reaktionen an der Injektionsstelle“ und „Immunogenität“ zeigten sich statistisch signifikante Unterschiede zuungunsten von Somatrogen.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Somatrogon (Ngenla®) ist zugelassen für Kinder und Jugendliche ab einem Alter von 3 Jahren mit Wachstumsstörung durch unzureichende Ausschüttung von Wachstumshormon. Die Nutzenbewertung von Somatrogon basiert auf der zulassungsbegründenden Studie CP-4-006. Es handelt sich bei der Studie CP-4-006 um eine offene, randomisierte, aktiv kontrollierte, multizentrische Phase-III-Studie im Parallelgruppendesign zur Untersuchung der Wirksamkeit und Sicherheit von wöchentlich verabreichtem Somatrogon im Vergleich zu täglich verabreichtem Genotropin bei pädiatrischen Patientinnen und Patienten mit Wachstumsstörungen aufgrund einer unzureichenden Wachstumshormon-Sekretion.

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 Somatrogon 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.

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

Tabelle 17: Zusammenfassende Darstellung der Ergebnisse der Studie CP-4-006

Studie CP-4-006 Darstellung der Ergebnisse	Somatrogon N = 109			Genotropin N = 115			Somatrogon vs. Genotropin	Effekt
	N	Personen mit Ereignis, n (%)	Veränderung Baseline zu Monat 12 LS Mean [95%-KI]	N	Personen mit Ereignis, n (%)	Veränderung Baseline zu Monat 12 LS Mean [95%-KI]		
Mortalität								
	N	Personen mit Ereignis, n (%)		N	Personen mit Ereignis, n (%)		Effekt- schätzer	
Todesfälle	109	Keine Ereignisse aufgetreten		115	Keine Ereignisse aufgetreten		n. b.	-
Morbidität								
	N ¹⁾²⁾	Baseline MW (SD)	Veränderung Baseline zu Monat 12 LS Mean [95%-KI]	N ¹⁾²⁾	Baseline MW (SD)	Veränderung Baseline zu Monat 12 LS Mean [95%-KI]	LS-Mean- Differenz [95%-KI]; p-Wert ³⁾	
Körpergröße (z-Score)	109	-2,9 (1,3)	0,92 [0,82; 1,02]	115	-2,8 (1,3)	0,87 [0,78; 0,97]	0,05 [-0,06; 0,16]; 0,388	↔
Lebensqualität								
	N ⁴⁻⁶⁾	Personen mit Ereignis, n (%)		N ⁴⁻⁶⁾	Personen mit Ereignis, n (%)		RR [95%-KI] ⁷⁾⁸⁾ ; p-Wert	
QoLISSY- Gesamtskala ⁹⁾	49	17 (34,7)		59	14 (23,7)		1,43 [0,81; 2,50]; 0,218	↔

Studie CP-4-006 Darstellung der Ergebnisse	Somatrogen N = 109		Genotropin N = 115		Somatrogen vs. Genotropin	Effekt
Sicherheit¹⁰⁾						
	N	Personen mit Ereignis, n (%)	N	Personen mit Ereignis, n (%)	RR [95%-KI] ¹¹⁾ ; p-Wert ¹²⁾	
Schwere UE	109	9 (8,3)	115	6 (5,2)	1,91 [0,79; 4,58]; 0,213	↔
SUE	109	3 (2,8)	115	2 (1,7)	1,99 [0,45; 8,72]; 0,229	↔
UE, das zum Abbruch der Studienmedi- kation führte	109	1 (0,9)	115	0 (0)	6,00 [0,31; 116,61]; 0,127	↔

¹⁾ Ergebnisse für die Baseline-Werte liegen für 109 Personen im Somatrogen-Arm und 115 im Genotropin-Arm vor, für Monat 12 liegen für 108 Personen im Somatrogen-Arm und 113 im Genotropin-Arm Werte vor (FAS).

²⁾ Für die Veränderungen von Baseline zu Monat 12 und den Gruppenunterschied der Veränderung wurden fehlende Werte mithilfe multipler Imputation ersetzt (FCS-Option). Es bestehen Unklarheiten, ob wie a priori im SAP festgelegt eine Imputation mit MNAR-Annahme durchgeführt wurde oder wie in Modul 4 des Nutzen-dossiers beschrieben eine MAR-Annahme verwendet wurde. Im Somatrogen-Arm wurden zu Monat 12 für 1 Person und im Genotropin-Arm für 2 Personen fehlende Werte imputiert.

³⁾ ANCOVA-Modell schließt „Behandlung“, „Altersgruppe“, „Höchste gemessene GH-Konzentration im Stimulationstest“, „Region“, „Geschlecht“ und „Baseline-Körpergröße (z-Score)“ als Kovariablen ein.

⁴⁾ Der QoLISSY-Fragebogen wurde nur in den folgenden Ländern eingesetzt, für die eine validierte Übersetzung vorliegt: USA, Australien, Neuseeland, Belarus, Russland, Ukraine, Großbritannien, Spanien (PRO Analysis Set).

⁵⁾ Ergebnisse für die Baseline-Werte liegen für die selbstberichtete Kinderversion des QoLISSY für Studienteilnehmende ≥ 7 Jahre (N = 35 im Somatrogen-Arm, N = 35 im Genotropin-Arm) und den Eltern-Proxy-Fragebogen für Studienteilnehmende < 7 Jahre (N = 19 im Somatrogen-Arm, N = 28 im Genotropin-Arm) vor.

⁶⁾ Ergebnisse für Monat 12 liegen für die selbstberichtete Kinderversion des QoLISSY für Studienteilnehmende ≥ 7 Jahre (N = 32 im Somatrogen-Arm, N = 33 im Genotropin-Arm) und den Eltern-Proxy-Fragebogen für Studienteilnehmende < 7 Jahre (N = 17 im Somatrogen-Arm, N = 26 im Genotropin-Arm) vor.

⁷⁾ Responderanalyse für eine Verbesserung zu Monat 12 im QoLISSY von ≥ 15 Punkten.

⁸⁾ Post hoc berechnet: RR mit 95%-KI wurde durch ein generalisiertes lineares Modell mit einer Binomialverteilungsannahme und mit natürlichem Logarithmus als Link-Funktion berechnet. Es wurde für den Baseline-Wert adjustiert, auf weitere Stratifizierungsfaktoren wurde aus Konvergenzgründen verzichtet. Fehlende Werte wurden mithilfe multipler Imputation ersetzt (FCS-Option und MAR-Annahme).

⁹⁾ Skala von 0 bis 100; höhere Werte bedeuten eine bessere Lebensqualität.

¹⁰⁾ Weitere Ergebnisse zu UE (nach Organsystemen und Einzelereignissen) sind in Ergebniskapitel 3.5 Sicherheit dargestellt.

¹¹⁾ Post hoc berechnet: Stratifiziertes RR mit 95%-KI mittels Logit-Methode mit Nullzellenkorrektur berechnet.

¹²⁾ Cochran-Mantel-Haenszel-Test: Stratifiziert nach „Alter“ (≥ 3 bis ≤ 7 vs. > 7 Jahre), „Region“ (Region 1 vs. Region 2 vs. Region 3), „Höchste gemessene GH-Konzentration im Stimulationstest“ (≤ 3 vs. > 3 bis ≤ 7 vs. > 7 bis ≤ 10 ng/ml) und „Geschlecht“ (männlich vs. weiblich).

Abkürzungen: ANCOVA: Kovarianzanalyse; FAS: Full Analysis Set; FCS: Fully Conditional Specification; GH: Wachstumshormon; KI: Konfidenzintervall; LS: Least Squares; MAR: Missing At Random; MNAR: Missing Not At Random; MW: Mittelwert; n. b.: nicht berechenbar; PRO: Patient-Reported Outcome; QoLISSY: Quality of Life in Short Stature Youth; RR: Relatives Risiko; SAP: Statistischer Analyseplan; SD: Standardabweichung; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. **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 and quality of life outcomes* 2019;17(1):49.
2. **Bullinger M, Quitmann J, Power M, Herdman M, Mimoun E, DeBusk K, et al.** Assessing the quality of life of health-referred children and adolescents with short stature: development and psychometric testing of the QoLISSY instrument. *Health and quality of life outcomes* 2013;11(1):76.
3. **Chaplin J, European QoLISSY Group.** Quality of life in short stature youth: the QoLISSY questionnaire user's manual. Berlin: Pabst Science Publishers; 2013.
4. **Deal C, Steelman J, Vlachopapadopoulou E, Stawerska R, Silverman LA, Phillip M, et al.** Efficacy and safety of weekly somatrogen vs daily somatropin in children with growth hormone deficiency: a phase 3 study. *J Clin Endocrinol Metab* 2022 [Online ahead of print].
5. **Deutsche Gesellschaft für Kinderendokrinologie und -diabetologie (DGKED).** Diagnostik des Wachstumshormonmangels im Kindes- und Jugendalter; S2e-Leitlinie, Langfassung [online]. AWMF-Registernummer 174-002. Berlin (GER): Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF); 2022. [Zugriff: 03.05.2022]. URL: https://www.awmf.org/uploads/tx_szleitlinien/174-002l_S2e_Diagnostik-des-Wachstumshormonmangels-im-Kindes-und-Jugendalter_2022-04.pdf.
6. **European Medicines Agency (EMA).** Ngenla (Somatrogen): European public assessment report EMEA/H/C/005633/0000 [online]. 16.12.2021. Amsterdam (NED): EMA. [Zugriff: 03.05.2022]. URL: https://www.ema.europa.eu/en/documents/assessment-report/ngenla-epar-public-assessment-report_en.pdf.
7. **Gemeinsamer Bundesausschuss (G-BA).** Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII - Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V): Vosoritid (Achondroplasie, ≥ 2 Jahre), vom 18. März 2022 [online]. Berlin (GER): G-BA; 2022. [Zugriff: 09.06.2022]. URL: https://www.g-ba.de/downloads/40-268-8358/2022-03-18_AM-RL-XII_Vosoritid_D-737_TrG.pdf.
8. **Kuczmarski RJ, Ogden CL, Guo SS.** 2000 CDC growth charts for the United States: methods and development. *Vital and health statistics* [online]. 2002;11(246):1-190. URL: https://www.cdc.gov/nchs/data/series/sr_11/sr11_246.pdf.
9. **OPKO Biologics.** A phase 3, open-label, randomized, multicenter, 12 months, efficacy and safety study of weekly MOD-4023 compared to daily Genotropin therapy in pre-pubertal children with growth hormone deficiency (CP4006): statistical analysis plan [unveröffentlicht]. 2019.
10. **OPKO Biologics.** A phase 3, open-label, randomized, multicenter, 12 months, efficacy and safety study of weekly MOD-4023 compared to daily Genotropin therapy in pre-pubertal children with growth hormone deficiency (CP4006): study protocol [unveröffentlicht]. 2018.

11. **OPKO Biologics, Pfizer.** A phase 3, open-label, randomized, multicenter, 12 months, efficacy and safety study of weekly MOD-4023 compared to daily Genotropin therapy in pre-pubertal children with growth hormone deficiency (CP4006): clinical study report [unveröffentlicht]. 2020.
12. **Pfizer Europe.** Ngenla 24 mg Injektionslösung im Fertigpen, Ngenla 60 mg Injektionslösung im Fertigpen [online]. 02.2022. [Zugriff: 03.05.2022]. URL: <https://www.fachinfo.de/>.
13. **Pfizer Europe.** Studie CP4006: Zusatzanalysen Efficacy [unveröffentlicht]. 2021.
14. **Pfizer Europe.** Studie CP4006: Zusatzanalysen PRO [unveröffentlicht]. 2021.
15. **Pfizer Europe.** Studie CP4006: Zusatzanalysen Safety [unveröffentlicht]. 2021.
16. **Pfizer Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Somatrogen (Ngenla), Kinder und Jugendliche ab dem Alter von 3 Jahren mit Wachstumsstörungen aufgrund einer unzureichenden Wachstumshormon-Sekretion; Modul 4A Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 29.03.2022.

Anhang

Tabelle 18: Veränderung der Jährlichen Wachstumsgeschwindigkeit zu Monat 12 in der Studie CP-4-006; FAS

Studie CP-4-006 Jährliche Wachstumsgeschwindigkeit	Somatrogon N = 109	Genotropin N = 115
Baseline n (%) MW (SD)	k. A. k. A.	k. A. k. A.
Monat 12 n (%) MW (SD)	108 (99,1) 10,2 (2,4)	113 (98,3) 9,7 (2,5)
Veränderung von Baseline zu Monat 12 n (%) LS Mean [95%-KI]	109 (100) 10,10 [9,58; 10,63]	115 (100) 9,78 [9,29; 10,26]
Gruppenunterschied der Veränderung LS-Mean-Differenz [95%-KI]; p-Wert ¹⁾	0,33 [-0,24; 0,89]; 0,259	

¹⁾ ANCOVA-Modell schließt „Behandlung“, „Altersgruppe“, „Höchste gemessene GH-Konzentration im Stimulationstest“, „Region“, „Geschlecht“ und „Baseline-Körpergröße (z-Score)“ als Kovariablen ein. Fehlende Werte wurden mithilfe multipler Imputation ersetzt (FCS-Option und MNAR-Annahme) und die imputierten Werte in der Somatrogon-Gruppe um die Nicht-Unterlegenheitsgrenze von 1,8 cm/Jahr reduziert.

Abkürzungen: ANCOVA: Kovarianzanalyse; FAS: Full Analysis Set; FCS: Fully Conditional Specification; GH: Wachstumshormon; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MNAR: Missing Not At Random; MW: Mittelwert; SD: Standardabweichung.