

Dokumentvorlage, Version vom 18.04.2013

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Burosumab (Crysvita[®])

Kyowa Kirin GmbH

Modul 4 A

X-chromosomale Hypophosphatämie (XLH)

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 23.09.2019

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	8
Abkürzungsverzeichnis	9
4 Modul 4 – allgemeine Informationen	13
4.1 Zusammenfassung der Inhalte von Modul 4	14
4.2 Methodik	25
4.2.1 Fragestellung	25
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	26
4.2.3 Informationsbeschaffung	31
4.2.3.1 Studien des pharmazeutischen Unternehmers	31
4.2.3.2 Bibliografische Literaturrecherche	31
4.2.3.3 Suche in Studienregistern	32
4.2.3.4 Selektion relevanter Studien	33
4.2.4 Bewertung der Aussagekraft der Nachweise	34
4.2.5 Informationssynthese und -analyse	36
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	36
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien	36
4.2.5.3 Meta-Analysen	56
4.2.5.4 Sensitivitätsanalysen	57
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren	59
4.2.5.6 Indirekte Vergleiche	63
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	65
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel	65
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	65
4.3.1.1.1 Studien des pharmazeutischen Unternehmers	65
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche	66
4.3.1.1.3 Studien aus der Suche in Studienregistern	69
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel	70
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	71
4.3.1.2.1 Studiendesign und Studienpopulationen	71
4.3.1.2.2 Verzerrungspotenzial auf Studienebene	77
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien	79
4.3.1.3.1 Ergebnisse patientenrelevanter Endpunkte – RCT	80
4.3.1.3.1.1 Rachitissymptomatik – RCT	81
4.3.1.3.1.2 Körpergröße (Stehhöhe/Liegegröße) – RCT	90
4.3.1.3.1.3 Gehfähigkeit – RCT	95
4.3.1.3.1.4 Serumphosphat – RCT	101
4.3.1.3.1.5 Schmerzerhebung und funktionelle Behinderung – RCT	106

4.3.1.3.1.6	Lebensqualität – RCT	113
4.3.1.3.1.7	Sicherheit – RCT	119
4.3.1.3.2	Subgruppenanalysen – RCT	135
4.3.1.3.2.1	Überblick über Ergebnisse der Interaktionstests aus Subgruppenanalysen	136
4.3.1.3.2.2	Ergebnisse für Subgruppenanalysen mit positivem Interaktionstest ($p < 0,05$)	139
4.3.1.3.2.3	Ergebnisse für präspezifizierten Subgruppenanalysen	145
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien	170
4.3.2	Weitere Unterlagen	175
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	175
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	175
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche	176
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	176
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	176
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	179
4.3.2.2	Nicht randomisierte vergleichende Studien	179
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	179
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien	180
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	180
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien	180
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	182
4.3.2.3	Weitere Untersuchungen	182
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	182
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	183
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	183
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	183
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	184
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	184
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens	185
4.4.1	Beurteilung der Aussagekraft der Nachweise	185
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß	186
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	198
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	199
4.5.1	Begründung für die Vorlage indirekter Vergleiche	199
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen	199
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	199
4.5.4	Verwendung von Surrogatendpunkten	200
4.6	Liste der eingeschlossenen Studien	201
4.7	Referenzliste	202

Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche	207
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....	211
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....	214
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....	215
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT	219
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	235

Tabellenverzeichnis

	Seite
Tabelle 4-1: Zusammenfassung des Ausmaßes des Zusatznutzens per Endpunkt.....	17
Tabelle 4-2: Zusammenfassung des Ausmaßes des Zusatznutzens für die Endpunktkategorie Lebensqualität	21
Tabelle 4-3: Patientenrelevante Endpunkte für die Bewertung von Burosumab in der Therapie der XLH bei Kindern im Alter von 1-17 Jahren	28
Tabelle 4-4: Einschlusskriterien für RCT mit Burosumab als Intervention.....	30
Tabelle 4-5: Ausschlusskriterien für RCT mit Burosumab als Intervention.....	30
Tabelle 4-6: Kriterien für die Berechnung der Interaktionstests in Anlehnung an das Methodenpapier 5.0 des IQWiG	60
Tabelle 4-7: Übersicht der für die Endpunkte der Studie UX023-CL301 durchgeführten Subgruppenanalysen.....	62
Tabelle 4-8: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	66
Tabelle 4-9: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	66
Tabelle 4-10: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	70
Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	71
Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	72
Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	74
Tabelle 4-15: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	78
Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	79
Tabelle 4-17: Operationalisierung von Rachitissymptomatik.....	81
Tabelle 4-18: Bewertung des Verzerrungspotenzials für Rachitissymptomatik in RCT mit dem zu bewertenden Arzneimittel.....	82
Tabelle 4-19: Ergebnisse für <i>Radiographic Global Impression of Change (RGI-C) Total Score</i> aus RCT mit dem zu bewertenden Arzneimittel	83
Tabelle 4-20: Ergebnisse für <i>Radiographic Global Impression of Change (RGI-C) - Responderanalyse</i> aus RCT mit dem zu bewertenden Arzneimittel	85
Tabelle 4-21: Ergebnisse für <i>Rickets Severity Score (RSS) – Total Score</i> aus RCT mit dem zu bewertenden Arzneimittel.....	87

Tabelle 4-22: Operationalisierung von Körpergröße	90
Tabelle 4-23: Bewertung des Verzerrungspotenzials für Körpergröße in RCT mit dem zu bewertenden Arzneimittel	90
Tabelle 4-24: Ergebnisse für Körpergröße (Stehhöhe/Liegeänge) aus RCT mit dem zu bewertenden Arzneimittel	92
Tabelle 4-25: Operationalisierung von Gehfähigkeit.....	95
Tabelle 4-26: Bewertung des Verzerrungspotenzials für Gehfähigkeit in RCT mit dem zu bewertenden Arzneimittel	96
Tabelle 4-27: Ergebnisse für Gehfähigkeit aus RCT mit dem zu bewertenden Arzneimittel..	97
Tabelle 4-28: Operationalisierung von Serumphosphat.....	101
Tabelle 4-29: Bewertung des Verzerrungspotenzials für Serumphosphat in RCT mit dem zu bewertenden Arzneimittel	101
Tabelle 4-30: Ergebnisse für Serumphosphat aus RCT mit dem zu bewertenden Arzneimittel.....	103
Tabelle 4-31: Operationalisierung von Schmerzerhebung.....	106
Tabelle 4-32: Bewertung des Verzerrungspotenzials für Schmerzerhebung und funktionelle Behinderung in RCT mit dem zu bewertenden Arzneimittel	107
Tabelle 4-33: Ergebnisse für Schmerzerhebung und funktionelle Behinderung mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel	108
Tabelle 4-34: Ergebnisse für Schmerzerhebung mittels FPS-R aus RCT mit dem zu bewertenden Arzneimittel	111
Tabelle 4-35: Operationalisierung von Lebensqualität	113
Tabelle 4-36: Bewertung des Verzerrungspotenzials für Lebensqualität> in RCT mit dem zu bewertenden Arzneimittel	113
Tabelle 4-37: Ergebnisse für Lebensqualität (SF-10) aus RCT mit dem zu bewertenden Arzneimittel.....	115
Tabelle 4-38: Operationalisierung von Sicherheit	119
Tabelle 4-39: Bewertung des Verzerrungspotenzials für Sicherheit in RCT mit dem zu bewertenden Arzneimittel	121
Tabelle 4-40: Ergebnisse für Unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse, die zum Therapieabbruch führen und Todesfälle aus RCT mit dem zu bewertenden Arzneimittel - Effektschätzer	122
Tabelle 4-41: Ergebnisse für Unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse, die zum Therapieabbruch führen und Todesfälle aus RCT mit dem zu bewertenden Arzneimittel	124
Tabelle 4-42: Ergebnisse für unerwünschten Ereignisse nach System Organ Class (SOC) und Preferred Term (PT) aus RCT mit dem zu bewertenden Arzneimittel	126
Tabelle 4-43: Zusammenfassung der zahnmedizinischen Untersuchungen	132
Tabelle 4-44: Ergebnisse der dentalen unerwünschten Ereignisse	133
Tabelle 4-45: Subgruppenanalysen für die Endpunkte der Studie UX023-CL301.....	137

Tabelle 4-46: Subgruppenanalysen – Ergebnisse für FPS-R aus RCT mit dem zu bewertenden Arzneimittel	140
Tabelle 4-47: Subgruppenanalysen – Ergebnisse für Serumphosphat – Prozentsatz erreichter Normwert aus RCT mit dem zu bewertenden Arzneimittel	143
Tabelle 4-48: Subgruppenanalyse des Endpunktes RGI-C Total Score	146
Tabelle 4-49: Subgruppenanalyse des Endpunktes RSS Total Score	148
Tabelle 4-50: Subgruppenanalyse des Endpunktes Responderanalyse des RGI-C.....	151
Tabelle 4-51: Subgruppenanalyse des Endpunktes Körpergröße	152
Tabelle 4-52: Subgruppenanalyse des Endpunktes Gehfähigkeit	156
Tabelle 4-53: Subgruppenanalyse des Endpunktes Serumphosphat	161
Tabelle 4-54: Subgruppenanalyse des Endpunktes PROMIS	164
Tabelle 4-55: Subgruppenanalyse des Endpunktes FPS-R	168
Tabelle 4-56: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	176
Tabelle 4-57: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	177
Tabelle 4-58: Operationalisierung von <Endpunkt xxx>.....	177
Tabelle 4-59: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	178
Tabelle 4-60: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	178
Tabelle 4-61: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	180
Tabelle 4-62: Operationalisierung von <Endpunkt xxx>.....	181
Tabelle 4-63: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	181
Tabelle 4-64: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	183
Tabelle 4-65: Zusammenfassung der Ergebnisse zur Bestimmung des Ausmaßes des Zusatznutzens von Burosumab für Patienten mit XLH	188
Tabelle 4-66: Übersicht der unerwünschten Ereignisse nach PT ($\geq 20\%$) in beiden Behandlungsarmen der Studie UX023-CL301 zu Woche 40 und Woche 64	195
Tabelle 4-67: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	199
Tabelle 4-68: Liste der zur Bewertung des Zusatznutzens eingeschlossenen Studien	201
Tabelle 4-69: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren in EMBASE.....	208
Tabelle 4-70: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren in MEDLINE	209
Tabelle 4-71: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren in Cochrane.....	210

Tabelle 4-72: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in clinicaltrials.gov	211
Tabelle 4-73: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EU Clinical Trials Register (EU-CTR)	212
Tabelle 4-74: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in WHO International Clinical Trials Registry Platform (WHO ICTRP)	212
Tabelle 4-75: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in Pharmnet.Bund	212
Tabelle 4-76: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EMA <i>clinical data</i>	213
Tabelle 4-77: Ausgeschlossene Studien_ ClincialTrials.gov	215
Tabelle 4-78: Ausgeschlossene Studien_ EU Clinical Trials Register (EU-CTR)	216
Tabelle 4-79: Ausgeschlossene Studien_ WHO International Clinical Trials Registry Platform (WHO ICTRP)	217
Tabelle 4-80 (Anhang): Studiendesign und -methodik für Studie UX023-CL301	219
Tabelle 4-81 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UX023-CL301	236

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	68
Abbildung 4-2: Studiendesign der UX023-CL301-Studie	76
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	89
Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	94
Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	100
Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	104
Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	112
Abbildung 8: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	117
Abbildung 9: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	134

Abkürzungsverzeichnis

Abkürzung	Bedeutung
1,25(OH) ₂ -Vitamin D ₃	Aktives Vitamin D
6MWT	6-Minuten-Gehdistanz Test (engl.: <i>6-Minute Walk Test</i>)
ALP	Alkalische Phosphatase
AMNOG	Arzneimittelmarktneuordnungsgesetz
ANCOVA	Kovarianzanalyse (engl.: <i>analysis of covariance</i>)
ARR	Absolute Risikoreduktion (engl. <i>absolute risk reduction</i>)
ATS	<i>American Thoracic Society</i>
BMI	Body-Mass-Index
B – P	Baseline zu Post-Baseline
CDC	<i>Center of disease control and prevention</i>
CHQ	<i>Child Health Questionnaire</i>
Cm	Zentimeter
COMP	<i>Committee of Orphan Medicinal Products</i>
CONSORT	<i>Consolidated Standards of Reporting Trials</i>
CSR	Studienbericht (engl. <i>Clinical study report</i>)
CTCAE	<i>Common Terminology Criteria for Adverse Events</i>
DGKED	Deutsche Gesellschaft für Kinderendokrinologie und -diabetologie
DIMDI	Deutsches Institut für Medizinische Dokumentation
ECHO	Echokardiogramm
EDC system	<i>Electronic data capture system</i>
EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm
EMA	Europäische Arzneimittelagentur (engl.: <i>European Medicines Agency</i>)
EU-CTR	<i>EU Clinical Trials Registry</i>
EU	Europäische Union
FAS	<i>Full Analysis Set</i>
FDA	<i>Food Drug Administration</i>
FGF23	Fibroblasten-Wachstumsfaktor 23

	(engl.: <i>Fibroblast growth factor 23</i>)
FPS-R	<i>Face Pain Scale-Revised</i>
G-BA	Gemeinsamer Bundesausschuss
GEE	<i>Generalized estimation equation</i>
GFR	Glomeruläre Filtrationsrate
H‘g	Hedges‘g
IgG1	Immunglobulin G1
iPTH	Intakt-Parathormon
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	<i>Intention to treat</i>
IWRS	<i>Interactive Web Response System</i>
KI	Konfidenzintervall
KRN23	Synonym für Burosumab
LOCF	<i>Last Observation Carried Forward</i>
LS Mean	<i>Least Squares Means</i>
MAR	Behandlung fehlender Werte (engl.: <i>missing at random</i>)
MCMC	Markov Chain Monte Carlo
MeSH	<i>Medical Subject Headings</i>
MI	<i>Multiple Imputation</i>
MID	Minimal wichtiger Unterschied (engl. <i>minimal important difference</i>)
MTC	<i>Mixed Treatment Comparison</i>
MW	Mittelwert
MWD	Mittelwertdifferenz
NaPi-2	Natrium-Phosphat-Kotransporter
NIH	<i>National Institutes of Health</i>
NST CTCAE	<i>National Cancer Institute Common Terminology Criteria for Adverse Events</i>
OR	<i>Odds ratio</i>
PEI	Paul-Ehrlich-Institut
<i>PHEX</i>	Phosphat regulierendes Gen mit Homologie zu Endopeptidasen auf dem X-Chromosom
PHS	<i>Physical Health Summary Score</i>

PODCI	<i>Pediatric Outcomes Data Collection Instrument</i>
POSNA	<i>Pediatric Orthopedic Society of North America</i>
PROMIS	<i>Patient Reported Outcomes Measurement Information System</i>
PSS	<i>Psychosocial Summary Score</i>
PT	<i>Preferred Term</i>
PTH	Parathormon
Q2W	2-wöchiges Behandlungsregime [(<i>quaque</i>) every two weeks]
RCT	Randomisierte, kontrollierte Studie (engl.: <i>randomized controlled trial</i>)
RGI-C	<i>Radiographic Global Impression of Change</i>
ROW	Rest der Welt (engl. <i>Rest of the World</i>)
RR	<i>Risk ratio</i>
RSS	<i>Rickets Severity Score</i>
s.c.	Subkutan
SAP	Statistischer Analyseplan (engl.: <i>Statistical Analysis Plan</i>)
SD	Standardabweichung (engl. <i>standard deviation</i>)
SE	Standardfehler (engl. <i>standard error</i>)
SF-10	<i>Short form 10</i>
SF-36	<i>Short form 36</i>
SGB	Sozialgesetzbuch
SOC	Systemorganklasse (engl.: <i>System Organ Class</i>)
SOP	<i>Standard operating procedure</i>
SUE	Schwerwiegendes unerwünschtes Ereignis
TmP	Maximale tubuläre Rückresorption von Phosphat
UE	Unerwünschtes Ereignis
ULN	Untere Grenze des Normbereichs (engl.: <i>upper limit normal</i>)
USA oder US	Vereinigte Staaten von Amerika (engl.: <i>The United States of America</i>)
VerfO	Verfahrensordnung
WHO	<i>World Health Organization</i>

WHO ICTRP	<i>WHO International Clinical Trials Registry Platform</i>
XLH	X-chromosomale Hypophosphatämie (engl.: <i>X-linked hypophosphatemia</i>)
zVT	Zweckmäßige Vergleichstherapie

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.1)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

Wie ist das Ausmaß des Zusatznutzens mit Burosumab (Crysvita®) zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Wachstumsphase mit X-chromosomaler Hypophosphatämie (XLH) und röntgenologischem Nachweis einer Knochenerkrankung hinsichtlich patientenrelevanter Endpunkte im Sinne des § 35a SGB V (Sozialgesetzbuch)?

Zielsetzung

Ziel des vorliegenden Dossiers ist die Darstellung des Ausmaßes des Zusatznutzens von Burosumab (Crysvita®) in der zugelassenen Formulierung und Dosierung bei der X-chromosomalen Hypophosphatämie bei Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischem Nachweis einer Knochenerkrankung. Bei der XLH handelt es sich um eine seltene Erkrankung, die schwerwiegende Behinderungen zur Folge hat.

Burosumab wurde bereits am 04. Oktober 2018 vom Gemeinsamen Bundesausschuss (G-BA) bewertet. Im damaligen Dossier wurden die beiden pivotalen und einarmigen Studien UX023-CL201 und UX023-CL205, sowie ein historischer Vergleich (UX023-CL002), dargestellt. Während des laufenden Verfahrens reichte Kyowa Kirin die zu diesem Zeitpunkt neu verfügbaren 40-Wochen-Daten der randomisierten und aktiv kontrollierten Phase-III-Studie UX023-CL301 ein. Die Ergebnisse belegten eindeutig den Zusatznutzen von Burosumab im Vergleich zu einer weniger wirksamen Therapie, die Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D. Da die UX023-CL301-Studie noch laufend war, wurde der Beschluss des G-BA vom 04. Oktober 2018 befristet, um Einblicke in die 64-Wochen-Daten der Studie zu gewährleisten. Im Beschluss wurde ein nicht quantifizierbarer Zusatznutzen vergeben, den die Patientenvertretung „nur schweren Herzens“ aufgrund der Befristung zustimmte, da Sie in den Endpunkten Serumphosphat und Rachitissymptomatik einen beträchtlichen Zusatznutzen sehe.

Laut Anwendungsgebiet ist Burosumab zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischem Nachweis einer Knochenerkrankung angezeigt [1]. Zur vereinfachten Darstellung wird im weiteren Verlauf des Dossiers die Zielpopulation als „Kinder im Alter von 1-17 Jahren“ bezeichnet. Dabei ist zu berücksichtigen, dass sich das Skelett in der Wachstumsphase befinden muss.

Die X-chromosomale Hypophosphatämie stellt die häufigste erbliche Form der Rachitis dar. Sie ist gekennzeichnet durch eine Störung der renalen Phosphat-Rückresorption, was zur einer gestörten Knochenmineralisierung sowie defekter Epiphysen (Wachstumsfugen) führt. XLH wird durch eine X-chromosomal lokalisierte Funktionsverlustmutation (meist in der Region Xp 22.1-Xp 22.2) des *PHEX*-Gens (phosphatregulierendes Gen mit Homologie zu Endopeptidasen auf dem X-Chromosom) verursacht [2, 3].

Burosumab ist ein rekombinanter, vollständig humaner, monoklonaler Antikörper (IgG1, Immunglobulin G1), der an den Fibroblasten-Wachstumsfaktor 23 (FGF23) bindet und dadurch dessen Aktivität hemmt. Durch die Hemmung von FGF23 stellt Burosumab die tubuläre Rückresorption von Phosphat aus der Niere wieder her. Die Produktion von 1,25(OH)₂-Vitamin D₃ wird gesteigert, was wiederum die Darmabsorption von Calcium und Phosphat erhöht. Gemeinsam verbessern diese Maßnahmen den Serumphosphat-Wert und die Knochenmineralisierung. Burosumab greift direkt in den Signalweg ein, hemmt die pathologisch erhöhte Phosphatausscheidung und stellt die Regulation der Mineralisierung des gesamten Skelettes wieder her. Burosumab ist die einzige pathophysiologisch kausale Therapiemöglichkeit für Kinder im Alter von 1-17 Jahren und deckt den medizinischen Bedarf [1].

Datenquellen

Zu Bewertung des Zusatznutzens wurde die für die Zulassung von Burosumab relevante Studie UX023-CL301 zugrunde gelegt.

Bei der Studie UX023-CL301 handelt es sich um eine multizentrische, randomisierte und aktiv kontrollierte Studie der Phase III im Parallelgruppendesign, welche die Wirksamkeit und Sicherheit von Burosumab im Vergleich zu einer Substitutionstherapie, die aus oralem Phosphat und aktivem Vitamin D besteht, abfragt. Dabei werden Patienten mit einer X-chromosomalen Hypophosphatämie im Alter von 1-12 Jahren über einen Zeitraum von 64 Wochen kontrolliert untersucht. Die sich hieran anschließende freiwillige Extensionsphase mit einer Burosumab-Behandlung ist in den Ländern Europa, USA, Kanada und Australien bis zu Woche 140 möglich.

Ein-/Ausschlusskriterien für Studien

Eingeschlossen wurden Studien im RCT-Design, sofern sie ganz oder teilweise in der oben aufgeführten Indikation von Burosumab durchgeführt wurden, sofern Burosumab in der zugelassenen Dosis von 0,8 mg/kg im Abstand von zwei Wochen subkutan verabreicht wurde und sofern sie mindestens einen patientenrelevanten Endpunkt beinhalteten.

Ausgeschlossen wurden Studien, die die hier genannten Einschlusskriterien nicht erfüllten, für die keine Vollpublikation oder kein Studienregistereintrag mit primären Ergebnissen vorlag oder bei denen es sich um eine Mehrfachpublikation handelte. Die Ein- und Ausschlusskriterien sind detailliert in Abschnitt 4.2.3 dargestellt.

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Zur Bewertung der Aussagekraft der Ergebnisse wurden die Verzerrungsaspekte für jede der eingeschlossenen Studien sowohl auf Endpunktebene als auch auf Studienebene einzeln beschrieben und bewertet.

Dabei entspricht die Studie UX023-CL301 dem höchsten Evidenzgrad der Klassifikation des IQWiG (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen). Darüber hinaus weist UX023-CL301 auf Studienebene ein niedriges Verzerrungspotenzial auf. Das Verzerrungspotenzial auf Endpunktebene wurde unterschiedlich bewertet. Da die Studie aufgrund der unterschiedlichen Applikationen im jeweiligen Studienarm unverblindet ist, konnte bei Endpunkten wie der Messung der Rachitissymptomatik nur der Auswerter/die Auswerter verblindet werden. Ebenso stellt die Serumphosphat-Konzentration durch die Messung am Gerät ein geringes Verzerrungspotenzial dar. Sowohl bei der Messung des Endpunktes 6-Minuten-Gehdistanz Test (6MWT) als auch der Erhebung der Fragebögen, die z.T. von den Eltern oder vom klinischen Personal ausgefüllt werden mussten, kann eine Verzerrung nicht vollständig ausgeschlossen werden [4-6]. Jedoch wurden einheitliche Schulungen durchgeführt, um die Restunsicherheit möglichst gering zu halten [7-9].

In Anhang 4-F bzw. in dem Abschnitt 4.3.1.2.2 (Verzerrungspotenzial auf Studienebene) sowie in dem Abschnitt 4.3.1.3.1 (Verzerrung auf Endpunktebene) ist die Beurteilung des Verzerrungspotenzials detailliert dargestellt.

Die Ergebnisse bezüglich der Wirksamkeit und Sicherheit von Burosumab werden im Abschnitt 4.3.1.3.1 ausführlich auf Basis klinisch- und patientenrelevanter Endpunkte dargestellt. Als Grundlage dienen der Studienbericht sowie weitere Analysen, die speziell für das vorliegende Dossier durchgeführt wurden.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die Therapieergebnisse einer Burosumab-Behandlung des familiären Phosphatmangels hervorgerufen durch einen FGF23-Defekt bei Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase basieren auf einer offenen, randomisierten und aktiv kontrollierten Phase-III-Studie UX023-CL301. Die Randomisierung erfolgte 1:1 in zwei Behandlungsgruppen. Die aktive Kontrolle erfolgte mit einer Substitutionstherapie bestehend aus oralem Phosphat und aktivem Vitamin D.

Mortalität

Mortalität ist ein patientenrelevanter Endpunkt, der jedoch nicht für die XLH-Therapie geeignet ist. Es handelt sich um eine schwere, chronische Erkrankung des Knochenskeletts und des Stoffwechsels, die erwartungsgemäß im Verlauf einer Studiendauer von 40 – 64 Wochen nicht zum Tode führt. Die Mortalität wurde in der Zulassungsstudien UX023-CL301 durch die

Datenerhebung der schwerwiegenden unerwünschten Ereignisse berücksichtigt. Es trat kein Todesfall während der Studiendauer auf.

Morbidität

Die Bewertung von Burosumab erfolgte anhand der multizentrischen, offenen, randomisierten und aktiv kontrollierten zulassungsrelevanten Studie UX023-CL301. Tabelle 4-1 zeigt in einer Übersicht die zugrundeliegenden Ergebnisse der Endpunktkategorie Morbidität der Zulassungsstudie UX023-CL301.

Tabelle 4-1: Zusammenfassung des Ausmaßes des Zusatznutzens per Endpunkt

Patientenrelevante Endpunkte	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 40	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 64
Rachitissymptomatik		
<i>Radiographic Global Impression of Change (RGI-C) Score</i>		
LsMean^a	1,14	1,02
[95% KI]	[0,83; 1,45]	[0,72; 1,33]
p-Wert	< 0,0001	< 0,0001
Hedges'g^b [95% KI]	H ^g = 1,846 [1,246; 2,446]	H ^g = 1,600 [1,023; 2,178]
<i>Rickets Severity Score (RSS)</i>		
LsMean^a	-1,34	-1,21
[95% KI]	[-1,74; -0,94]	[-1,59; -0,83]
p-Wert	< 0,0001	< 0,0001
Hedges'g^b [95% KI]	H ^g = -1,399 [-1,964; -0,833]	H ^g = -1,119 [-1,710; -0,529]
Körpergröße		
<i>Stehhöhe/Liegelänge [Z-Score]</i>		
LsMean^a	0,12	0,14
[95% KI]	[0,01; 0,24]	[0,00; 0,29]
p-Wert	0,0408	0,0490
Hedges'g^b [95% KI]	H ^g = 0,537 [0,021; 1,053]	H ^g = 0,490 [-0,025; 1,005]
Gefähigkeit^c		
<i>6-Minuten-Gehdistanz Test^c (6MWT) [in Meter]</i>		
LsMean^a	38,30	45,55
[95% KI]	[-2,38; 78,98]	[2,09; 89,02]
p-Wert	0,0650	0,0399
Hedges'g^b [95% KI]	-	H ^g = 0,679 [-0,038; 1,396]

Patientenrelevante Endpunkte	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 40	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 64
<i>6-Minuten-Gehdistanz Test^c (6MWT) [in Prozent]</i>		
LsMean^a	6,19	7,27
[95% KI]	[-0,68; 13,06]	[0,01; 14,52]
p-Wert	0,0775	0,0496
Hedges'g^b [95% KI]	-	H'g = 0,703 [-0,015; 1,422]
Biochemischer Parameter		
<i>Serumphosphat-Konzentration [mg/dl]</i>		
LsMean^a	0,77	0,74
[95% KI]	[0,60; 0,94]	[0,58; 0,91]
p-Wert	< 0,0001	< 0,0001
Hedges'g^b [95% KI]	H'g = 2,047 [1,427; 2,667]	H'g = 1,981 [1,368; 2,595]
Schmerzerhebung		
<i>Patient Reported Outcomes Measurement Information System (PROMIS) Pediatric Pain Interference Domain Score</i>		
LsMean^a	-4,91	-2,26
[95% KI]	[-9,19; -0,63]	[-6,61; 2,09]
p-Wert	0,0245	0,3091
Hedges'g^b [95% KI]	H'g = -0,697 [-1,390; -0,008]	-
<i>Patient Reported Outcomes Measurement Information System (PROMIS) Physical Function Mobility Domain Score</i>		
LsMean^a	2,68	1,90
[95% KI]	[-0,53; 5,88]	[-1,80; 5,59]
p-Wert	0,1014	0,3145
Hedges'g^b [95% KI]	-	-
<i>Patient Reported Outcomes Measurement Information System (PROMIS) Fatigue Domain Score</i>		
LsMean^a	-3,20	-1,08
[95% KI]	[-7,79; 1,39]	[-6,21; 4,06]
p-Wert	0,1724	0,6810
Hedges'g^b [95% KI]	-	-
<i>Face Pain Scale – Revised (FPS-R)</i>		
LsMean^a	-0,00	0,05
[95% KI]	[-0,80; 0,79]	[-0,58; 0,68]
p-Wert	0,9905	0,8786
Hedges'g^b [95% KI]	-	-
(B-P): Behandlungsdifferenz betrachtet von Baseline zu Post-Baseline (Woche 40 bzw. 64) der Burosumab-Gruppe mit der aktiven Kontrollgruppe (orales Phosphat und aktives Vitamin D)		
[95 %-KI]: 95 %-Konfidenzintervall		
^a : LsMean: Statistische Methode für den jeweiligen Endpunkt siehe Abschnitt 4.3.1.3.1		
^b : Hedges'g-Berechnung erfolgt nur, falls p < 0,05		
^c : Die korrigierte Analyse des 6MWT wurde dargestellt, da zwei Patienten einen fehlerhaften Datensatz aufwiesen (zu Baseline und zu Woche 40). Eine detaillierte Beschreibung erfolgt im Abschnitt 4.3.1.3.1.3.		

Verbesserung der Rachitissymptomatik

Erstmals konnte in der Indikation X-chromosomale Hypophosphatämie durch die RCT-Studie UX023-CL301 anhand einer aktiven Kontrolle die Heilung der Rachitissymptomatik durch eine Burosumab-Behandlung gezeigt werden. Außerdem konnten die ebenfalls eindrucksvollen Ergebnisse der beiden einarmigen und pivotalen Studien UX023-CL201 und UX023-CL205 mit der RCT UX023-CL301 bestätigt werden. Diese zeigen, dass Burosumab deutlich besser gegenüber einer schwach wirksamen Therapie, der Substitutionstherapie, anstatt eines Vergleichs zu Placebo, ist. Die Burosumab-Behandlung führte zu einer signifikanten Erhöhung der RGI-C-Werte und zur Reduktion der RSS-Werte bereits nach 40 Wochen. Die adjustierte mittlere Änderung des primären Endpunktes RGI-C von Baseline zu Woche 40 lag bei 1,92 in der Burosumab-Gruppe und 0,77 in der aktiven Kontrollgruppe. Der Behandlungsunterschied von 1,14 ist sowohl statistisch signifikant als auch klinisch relevant. Dieser Unterschied war zu Woche 64 ebenfalls signifikant sowie klinisch relevant (Differenz: 1,02; $p < 0,0001$).

Die Auswertung der RSS-Methode zeigte eine stärkere Reduktion des Schweregrads der Rachitissymptomatik mithilfe der Burosumab-Behandlung. Der durchschnittliche RSS Total Score sank um 64 % von Baseline zu Woche 40 in der Burosumab-Gruppe, von 3,17 bis 1,13, im Vergleich zu einem Rückgang in der aktiven Kontrollgruppe von 23 %, von 3,19 bis 2,47. Die Behandlungsdifferenz der beiden Gruppen zu Woche 64 war mit -1,21 statistisch signifikant ($p < 0,0001$) und klinisch relevant.

Verbesserung der Körpergröße

Sowohl in der Burosumab-Gruppe als auch in der aktiven Kontrollgruppe konnte eine Wachstumsbeschleunigung zu Woche 40 und Woche 64 beobachtet werden. Das stark verlangsamte Wachstum in der aktiven Kontrollgruppe unterstützt die in der Literatur auffindbaren Ergebnisse. Hier wird berichtet, dass die Substitutionstherapie bei XLH-Patienten nur begrenzte Auswirkungen auf das Wachstum hat und folglich den Kleinwuchs, die Mobilität und die Lebensqualität der Patienten nicht verbessert [2, 10, 11].

Die Analysen der Körpergröße zeigen eine Veränderung des Wachstums in Richtung Normalisierung durch die Burosumab-Behandlung im Vergleich zur aktiven Kontrolle. Der Unterschied in den Behandlungsgruppen bei Veränderung des Z-Scores war zu Woche 40 und zu Woche 64 statistisch signifikant ($p_{40} = 0,0408$, $p_{64} = 0,0490$; GEE-Modell).

Verbesserung der Gehfähigkeit

Die Ergebnisse des 6MWT (Gesamtentfernung in Meter und als Prozentsatz der vorhergesagten Gehdistanz basierend auf normativen Werten) in der UX023-CL301 zeigten eine deutliche Verbesserung in der Burosumab-Gruppe, allerdings nicht in der aktiven Kontrollgruppe. Bei der Interpretation der Ergebnisse ist jedoch zu berücksichtigen, dass die Behandlungsgruppen

abweichende Baseline-Werte aufweisen. Weiterhin erfolgten zwei Analysen, da zwei Patienten der Burosumab-Gruppe den 6MWT nicht abgeschlossen hatten (1-mal zu Baseline und 1-mal zu Woche 40). Aus diesem Grund wurde in Abschnitt 4.3.1.3.1.3 sowohl die Primäranalyse mit allen Patienten als auch die Analyse exklusive der zwei Patienten in der Burosumab-Gruppe dargestellt und diskutiert.

Veränderung der Serumphosphat-Konzentration

Burosumab ist die erste, kausale Therapie zur Behandlung der XLH, die die zugrunde liegende Pathophysiologie der FGF23-induzierten Hypophosphatämie verbessert. In der aktiv kontrollierten Studie konnte eine schnelle Normalisierung der Serumphosphat-Konzentration in der Burosumab-Gruppe erreicht werden, während bei der aktiven Kontrollgruppe nur ein kleiner Anstieg der Serumphosphat-Konzentration beobachtet wurde, welcher nach einer Auswaschphase jedoch ausschließlich auf der oralen Einnahme von Phosphat beruht. Die bisherige Therapie der XLH bestand aus einer Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D, welche seit über 50 Jahren praktiziert wurde. Sowohl Carpenter et al., Linglart et al. als auch die *Key Opinion Leader* der mündlichen Anhörung [12] berichten, dass trotz der hohen Dosierung von Phosphat-Präparaten zu keiner Zeit eine Normalisierung der Serumphosphat-Konzentration erreicht werden könne. Hingegen bestehe ein hohes Risiko einer Folgeerkrankung durch die hochdosierte Behandlung mit Phosphat. Mit der UX023-CL301 konnte nun eindeutig belegt werden, dass die Burosumab-Behandlung zielgerichtet in den Phosphatstoffwechsel eingreift, eine Normalisierung des Serumphosphats erzielt wird und dabei keine Belastung der Nieren erfolgt.

Verringerung der Schmerzen und funktionellen Behinderung

In der aktiv kontrollierten Burosumab-Studie UX023-CL301 wurde der PROMIS-Fragebogen zur Ermittlung des Schmerzempfindens, der Körperfunktion und des Ermüdungsempfindens bei Patienten ab einem Alter von 5 Jahren verwendet. Dabei zeigte sich zu Beginn, dass die Studienteilnehmer im Gegensatz zu gesunden Kindern (standardisierter Mittelwert von 50) eine deutlich beeinträchtigte Körperfunktion aufweisen. Zu Baseline wurde in der Burosumab-Gruppe über mehr Schmerzen als in der aktiven Kontrollgruppe berichtet. Dies änderte sich zu Woche 40. Hier wurde weniger Schmerz in der Burosumab-Gruppe berichtet. Der Unterschied zwischen den Behandlungsgruppen war zu Woche 40 statistisch signifikant, jedoch nicht mehr zu Woche 64. Der FPS-R Schmerzfragebogen, der ebenfalls in der UX023-CL301 erhoben wurde, zeigte keine merklichen Veränderungen innerhalb der Behandlungsgruppen. Jedoch wurde bereits zu Baseline kein akuter Schmerz in den Behandlungsgruppen berichtet.

Lebensqualität

In der Studie UX023-CL301 erfolgte die Erhebung der Lebensqualität anhand des pädiatrischen SF-10-Fragebogens. Dieser ist zur Beurteilung körperlicher und psychosozialer Funktionen bei Kindern ab einem Alter von 5 Jahren geeignet. Die Werte beider Behandlungsgruppen in der psychosozialen Komponente (PSS) blieben während der Studiendauer unverändert. Die zweite Komponente des Fragebogens „körperliche Funktion“ (PHS) verbesserte sich in der Burosumab-Gruppe deutlicher als in der aktiven Kontrollgruppe. Die Tabelle 4-2 zeigt in einer Übersicht die zugrundeliegenden Ergebnisse der Endpunktkategorie Lebensqualität der Zulassungsstudie UX023-CL301.

Tabelle 4-2: Zusammenfassung des Ausmaßes des Zusatznutzens für die Endpunktkategorie Lebensqualität

Patientenrelevante Endpunkte	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 40	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 64
Lebensqualität		
<i>Short Form (SF)-10 Health Survey</i> <i>Physical Health Score (PHS-10)</i>		
LsMean^a	4,35	5,51
[95% KI]	[-1,15; 9,85]	[-0,17; 11,18]
p-Wert	0,1213	0,0571
Hedges'g^b [95% KI]	-	-
<i>Short Form (SF)-10 Health Survey</i> <i>Psychosocial Summary Score (PSS-10)</i>		
LsMean^a	2,24	-0,44
[95% KI]	[-1,86; 6,35]	[-4,59; 3,70]
p-Wert	0,2841	0,8337
Hedges'g^b [95% KI]	-	-
(B-P): Behandlungsdifferenz betrachtet von Baseline zu Post-Baseline (Woche 40 bzw. 64) der Burosumab-Gruppe mit der aktiven Kontrollgruppe (orales Phosphat und aktives Vitamin D)		
[95 %-KI]: 95 %-Konfidenzintervall		
^a : LsMean: Statistische Methode für den jeweiligen Endpunkt siehe Abschnitt 4.3.1.3.1		
^b : Hedges'g-Berechnung erfolgt nur, falls $p < 0,05$		

Diese Ergebnisse decken sich mit den Morbiditätsendpunkten PROMIS *Physical Function Mobility* und 6MWT und deuten darauf hin, dass die Burosumab-Behandlung einen höheren Grad an Verbesserungen der Morbidität und der damit einhergehenden Verbesserung der Lebensqualität erzielt.

Sicherheit

Insgesamt erweist sich Burosumab als wirksames, gut verträgliches und sicheres Arzneimittel zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischen Nachweis einer Knochenerkrankung. In der aktiv kontrollierten Phase-III-Studie ergaben sich aus der Bewertung von unerwünschten Ereignissen (UEs) keine signifikanten Sicherheitsbedenken. Es wurde weder ein Therapie-/Studienabbruch noch ein Todesfall aufgrund von UEs berichtet.

Die häufigsten UEs waren Kopfschmerzen, Husten, Pyrexie, Nasopharyngitis, Erbrechen und Durchfall. Diese genannten UEs waren leicht oder moderat ausgeprägt. Als typisch in einer pädiatrischen Population werden UEs angesehen wie Husten, Nasopharyngitis, Erbrechen, Infektionen der oberen Atemwege und Pyrexie. Schmerzen in Extremitäten und Arthralgien treten häufig bei Patienten mit XLH auf. In der aktiv kontrollierten Studie UX023-CL301 traten in beiden Behandlungsgruppen ebenfalls die häufig berichteten UEs auf, die in den einarmigen pivotalen Phase-II-Studien (UX023-CL201 und UX023-CL205) beobachtet wurden. Mit Ausnahme der Reaktion an der Injektionsstelle waren diese UEs typisch für eine pädiatrische Population oder ein häufiges Symptom der XLH-Erkrankung.

Die Auswertung der UEs ergab keine signifikanten Sicherheitsbedenken (Ausnahme: Therapie-assoziierte UEs). Während der Studie wurde kein Kind aufgrund von unerwünschten Ereignissen oder aus einem anderen Grund aus der Studie ausgeschlossen. Kein Studienteilnehmer verstarb. Die meisten Studienteilnehmer hatten mindestens ein unerwünschtes Ereignis (Burosumab-Gruppe: 100 %; Aktive Kontrollgruppe: 84 %). Bei 7 Patienten (4 in der Burosumab-Gruppe; 3 in der aktiven Kontrollgruppe) traten unerwünschte Ereignisse Grad 3 auf. Keines der UEs Grad 3 war schweren Ausmaßes außer das schwerwiegende UE (SUE) Kraniosynostose in der aktiven Kontrollgruppe. Keines der UEs Grad 3 wurde vom Prüfarzt als Therapie-assoziiert angesehen, außer das UE Grad 3 Arthralgie in der Burosumab-Gruppe. Als Reaktion auf diese Therapie-assoziierten UEs des Grad 3 wurden keine Änderungen in der Verabreichung von Studienmedikamenten vorgenommen. Die restlichen UEs waren mild oder moderat (Grad 1 oder 2) und es wurden keine lebensbedrohlichen UEs Grad 4 gemeldet.

Die häufigsten Therapie-assoziierten UEs in der Studie UX023-CL301 waren Reaktionen an der Injektionsstelle, welche nur leicht ausgeprägt und gut zu behandeln waren. Eine Reaktion an der Injektionsstelle ist eine bekannte und zu erwartende Folge einer subkutanen Verabreichung von Proteintherapeutika (Antikörpertherapeutika). Die Ausprägung dieser UEs an der Injektionsstelle waren in den meisten Fällen mild oder moderat (UX023-CL301: 1 Patient mit Hautausschlag an der Injektionsstelle (Grad 2: moderat)) und wurden von den Prüfarzten als allgemeine lokale Reizung und nicht als schwere Überempfindlichkeitsreaktion angesehen. Die Dauer der meisten Reaktionen an der Injektionsstelle betrug etwa ein bis zwei Tage.

In der Burosumab-Studie wurde kein Ereignis berichtet, bei dem die Serumphosphat-Konzentration über der Norm lag [3,2 bis 6,1 mg/dl]. Ultraschalluntersuchungen der Nieren

waren in den meisten Fällen in beiden Behandlungsarmen der Studie UX023-CL301 bei einem Untersuchungszeitraum von 64 Wochen unauffällig (Burosumab-Gruppe: 90 %, Aktive Kontrollgruppe: 78 %). Ebenfalls wurden keine klinisch bedeutsamen Erhöhungen von Nephrokalzinose-Werten beobachtet. Hervorzuheben ist, dass unter der Therapie mit Burosumab die Folgeerkrankungen wie eine Nephrokalzinose und ein sekundärer Hyperparathyreoidismus, die häufig durch eine Substitutionstherapie auftreten, ausbleibt.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Burosumab ist das erste Arzneimittel, das für Kinder mit X-chromosomaler Hypophosphatämie entwickelt wurde und das als Antikörpertherapie an der Ursache der Erkrankung ansetzt, indem es das Protein FGF23 bindet, seine renale Aktivität hemmt und somit die tubuläre Rückresorption von Phosphat aus dem Primärharn wiederherstellt [1].

Der medizinische Zusatznutzen von Burosumab wird eindrucksvoll in der randomisierten und aktiv kontrollierten Studie gezeigt. Ebenfalls wurden die Ergebnisse der im ersten Dossier dargestellten Studien UX023-CL201 und UX023-CL205 bestätigt. In der Gesamtschau der Endpunkte Morbidität, Lebensqualität und Sicherheit wird für Burosumab zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH insgesamt ein beträchtlicher Zusatznutzen abgeleitet. Es handelt sich um eine bisher nicht erreichte deutliche Verbesserung der Erkrankung insbesondere basierend auf einer Abschwächung der schwerwiegenden Symptomatik (schwere Rachitis, verringerte Wachstumsgeschwindigkeit/Kleinwuchs, stark eingeschränkte Gehfähigkeit, erhebliche Funktionsstörungen, einhergehende starke Schmerzen, geringere Lebensqualität). Die Sicherheit zeigt keine auffälligen unerwünschten Ereignisse und ein deutlich besseres Nebenwirkungsprofil im Vergleich zur Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D, die schwerwiegende Nebenwirkungen wie eine Nephrokalzinose und ein sekundärer Hyperparathyreoidismus zur Folge haben kann. Burosumab stellt die erste und einzige kausale Therapie zur Behandlung der XLH dar und setzt dort an, wo die Erkrankung entsteht. Aufgrund der raschen Normalisierung der Mineral- und Vitamin-Dysregulation und der damit verbundenen normalisierten Knochenmineralisierung kommt es zur Rückbildung skelettaler Veränderungen der Knochen und Gelenkschmerzen. Auch der Zahnstatus sollte sich so im Laufe der Zeit verbessern. Die Markteinführung und das vorherige Arzneimittel-Härtefallprogramm beim Paul-Ehrlich-Institut zeigte die fehlende therapeutische Alternative zur Behandlung der XLH und ebenso, dass Burosumab nun eine erhebliche Versorgungslücke geschlossen hat. Insgesamt wird durch die oben aufgeführten Gründe ein beträchtlicher Zusatznutzen für Burosumab zur Behandlung der XLH abgeleitet.

Fazit

Die für die vorliegende Nutzenbewertung herangezogene Studie UX023-CL301 besteht auf Studien- und Endpunktebene ein niedriges Verzerrungspotenzial und somit eine hohe Aussagekraft. Die Ergebnisse der Studien sind zudem vollständig auf den deutschen Versorgungskontext übertragbar (vgl. Abschnitt 4.3.1.2.2, Abschnitt 4.3.1.3.1 und Anhang 4-F). In den Endpunkten, aus denen sich ein Zusatznutzen ableiten lässt, liegen homogene Ergebnisse vor. Hinweise, die eine Herabstufung der Wahrscheinlichkeit des Zusatznutzens für einzelne Endpunkte notwendig erscheinen lassen, bestehen indes keine. Aus den Ergebnissen der UX023-CL301-Studie lässt sich somit ein **Beleg** für einen Zusatznutzen ableiten.

Die Ergebnisse der Studie UX023-CL301 belegen eine ausgeprägte Wirksamkeit von Burosumab zur pathologisch kausalen Therapie der X-chromosomalen Hypophosphatämie bei Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase. Insbesondere führt die Anwendung von Burosumab bei den Endpunkten Gehdistanz, Wachstum, Serumphosphat und Rachitissymptomatik zu einer beträchtlichen Verbesserung bei den Patienten. Damit ermöglicht Burosumab den Patienten eine langfristige Freiheit von Knochen- und Gelenkschmerzen, die wiederum durch Deformationen der Extremitäten hervorgerufen werden. Ebenso von schwerwiegenden Symptomen, die durch die XLH selbst aber auch durch eine Substitutionstherapie auftreten können. Hier ist vor allem die Gefahr einer Nephrokalzinose und eines sekundären Hyperparathyreoidismus zu nennen.

In der Gesamtschau der einzelnen Endpunktkategorien ist eine nachhaltige und bisher gegenüber der Substitutionstherapie nicht erreichte erhebliche Verbesserung des therapielevanten Nutzens festzustellen, die insbesondere auf der Verbesserung der Rachitissymptomatik, Schmerzreduktion, Verbesserung der Mobilität und damit auf der langfristigen Freiheit von Symptomen der XLH und Verbesserung der Lebensqualität beruht.

Entsprechend § 5 Abs. 7 ist für Burosumab zur Behandlung der XLH bei Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit röntgenologischen Nachweis einer Knochenerkrankung im Vergleich zur Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D damit von einem **beträchtlichen Zusatznutzen** auszugehen.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Das vorliegende Dossier bezieht sich auf den Nachweis der Wirksamkeit und Sicherheit des Wirkstoffs Burosumab (Crysvita®), welcher zur Behandlung von Kindern im Alter von 1-17 Jahren mit X-chromosomaler Hypophosphatämie (XLH) angezeigt ist. Ziel des vorliegenden Dossiers ist die Darstellung des medizinischen Zusatznutzens von Burosumab gegenüber der bisherigen Substitutionstherapie im zugelassenen Anwendungsgebiet, d.h. Kinder ab 1 Jahr und Jugendliche in der Skelettwachstumsphase mit XLH und röntgenologischem Nachweis einer Knochenerkrankung.

Patientenpopulation, Intervention und zweckmäßige Vergleichstherapie (zVT) werden im Folgenden näher erläutert, ebenso die patientenrelevanten Endpunkte im Anwendungsgebiet und der vorgelegte Studientyp.

Die im vorliegenden Dossier zu beantwortende Frage lautet:

Wie ist das Ausmaß des Zusatznutzens mit Burosumab (Crysvita®) zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit X-chromosomaler Hypophosphatämie (XLH) und röntgenologischem Nachweis einer Knochenerkrankung hinsichtlich patientenrelevanter Endpunkte im Sinne des § 35a SGB V?

Die **Patientenpopulation** umfasst laut Anwendungsgebiet pädiatrische Patienten im Alter von 1 bis einschließlich 17 Jahren (< 18 Jahre) in der Skelettwachstumsphase mit XLH und röntgenologischem Nachweis einer Knochenkrankung.

Die **Intervention** ist Burosumab, ein rekombinanter, vollständig humaner monoklonaler Antikörper (IgG1), der den Fibroblasten-Wachstumsfaktor 23 (FGF23) bindet und dadurch dessen biologische Aktivität hemmt [1].

Im Beratungsgespräch vom 08. November 2017 hatte der G-BA Phosphatsubstitution als **zweckmäßige Vergleichstherapie** festgelegt. Weiterhin wurde erläutert, dass der G-BA davon ausgehe, dass im Rahmen einer Phosphatsubstitutionstherapie zusätzlich mit aktiven Vitamin D therapiert wird, um die Folgeerkrankung einer Phosphatsubstitution (z. B. sekundärer Hyperparathyreoidismus) zu vermeiden.

Zur Beurteilung des Ausmaßes des Zusatznutzens wurden die **patientenrelevanten Endpunkte** im Anwendungsgebiet XLH betrachtet. Diese umfassen Morbidität, Lebensqualität und Sicherheit. Die Mortalität wird in der randomisierten und aktiv kontrollierten Studie UX023-CL301 durch die Datenerhebung der schwerwiegenden unerwünschten Ereignisse berücksichtigt.

Die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention liefern im Allgemeinen randomisierte, kontrollierte Studien (RCTs) (**Studientyp**). Das vorliegende Dossier verwendet ausschließlich Daten aus einer methodischen und der jeweiligen Fragestellung angemessenen RCT. Dieses Vorgehen ermöglicht Aussagen zum medizinischen Zusatznutzen mit hoher Ergebnissicherheit.

Bei der hier dargestellten Studie UX023-CL301 handelt es sich um eine randomisierte, offene und aktiv kontrollierte Studie der Phase III von Burosumab bei Kindern im Alter von 1-12 Jahren mit XLH. Die im Dossier vom 15. April 2018 dargestellten zulassungsrelevanten, einarmigen Phase-II-Studien (UX023-CL201 und UX023-CL205) wurden für die bedingte Zulassung (*conditional marketing authorisation*) herangezogen. Während des laufenden Arzneimittelmarktneuordnungsgesetz (AMNOG)-Verfahrens konnte Kyowa Kirin die Ergebnisse des Datenschnitts nach 40 Wochen der Studie UX023-CL301 zusammen mit der Stellungnahme dem G-BA vorlegen. Im Folgenden werden die jetzt verfügbaren 40- und 64-Wochen-Daten der Studie UX023-CL301 dargestellt, um die bereits eindrucksvollen Ergebnisse der Phase-II-Studien zu bestätigen.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann

beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Population

Die Population der eingeschlossenen Studie musste der im vorherigen Abschnitt 4.2.1 beschriebenen Definition der Zielpopulation entsprechen. In die Nutzenbewertung wurden daher nur Studien eingeschlossen, die mindestens die folgende Studienpopulation untersuchten:

Patienten im Alter von 1-17 Jahren mit X-chromosomaler Hypophosphatämie, die sich im skelettalen Wachstum befinden. Weiter müssen röntgenologische Hinweise auf eine Knochenerkrankung vorliegen.

Intervention

Für die Bewertung des Ausmaßes des Zusatznutzens von Burosumab werden nur Studien eingeschlossen, die Burosumab als Intervention benennen.

Die empfohlene Anfangsdosis von Burosumab beträgt 0,4 mg/kg Körpergewicht und die normale Erhaltungsdosis beträgt 0,8 mg/kg Körpergewicht. Burosumab wird alle zwei Wochen subkutan verabreicht. Die Maximaldosis liegt bei 90 mg. Alle Dosen sollen auf 10 mg genau abgemessen werden [1].

Zweckmäßige Vergleichstherapie

Die seit über 50 Jahre historisch gewachsene Behandlung von Kindern mit XLH vor Burosumab war nach Angaben der Leitlinie „Hereditäre hypophosphatämische Rachitiden“ der Deutschen Gesellschaft für Kinderendokrinologie und -diabetologie (DGKED) e.V. eine Phosphatsubstitution. Um die durch die Phosphatsubstitution bedingte Folgeerkrankung eines sekundären Hyperparathyreoidismus zu vermeiden, empfiehlt die Leitlinie zusätzlich eine Substitution mit aktivem Vitamin D [13, 14]. Im Anwendungsgebiet XLH existiert neben Burosumab keine zugelassene kausale Therapie. Der hohe medizinische Bedarf und die fehlenden therapeutischen Alternativen vor der Einführung von Burosumab wurden mit der Durchführung eines Arzneimittel-Härtefallprogramms durch das Paul-Ehrlich-Institut bestätigt.

Die zweckmäßige Vergleichstherapie im Anwendungsgebiet von Burosumab ist eine Phosphatsubstitution. Nach dem Beratungsgespräch verdeutlichte der G-BA zudem, dass eine

zusätzliche Therapie mit aktivem Vitamin D erfolgen sollte, um die Nebenwirkung der Substitutionstherapie zu vermeiden.

Endpunkte

Zur Beurteilung des Ausmaßes des Zusatznutzens werden für die Erkrankung patientenrelevante Endpunkte herangezogen. Diese umfassen Morbidität, Lebensqualität und Sicherheit. In der folgenden Tabelle 4-3 sind die für die Darstellung von Burosumab in der pädiatrischen Population relevanten Endpunkte dargestellt. Begründungen für die Relevanz und Validität der einzelnen Endpunkte erfolgen im Abschnitt 4.3. Aufgrund des Alters gibt es Abweichungen bei der Erhebung der patientenrelevanten Endpunkte der pädiatrischen XLH-Patienten (z. B. 6MWT und PROMIS-Fragebogen)

Tabelle 4-3: Patientenrelevante Endpunkte für die Bewertung von Burosumab in der Therapie der XLH bei Kindern im Alter von 1-17 Jahren

Endpunktkategorie und Ausprägung	Beispiele für Erhebung/Messinstrument
Mortalität	
Im Rahmen der Sicherheit erhoben	-
Morbidität	
Rachitissymptomatik	<ul style="list-style-type: none"> • <i>Radiographic Global Impression of Change (RGI-C) Score</i> • <i>Rickets Severity Score (RSS)</i>
Körpergröße	<ul style="list-style-type: none"> • Stehhöhe/Liegegröße [Z-Score]
Gehfähigkeit	<ul style="list-style-type: none"> • <i>6-Minute Walk Test (6MWT)</i> <ul style="list-style-type: none"> ○ absolut ○ relativ
Serumphosphat	<ul style="list-style-type: none"> • Serumphosphat [mg/dl]
Schmerzerhebung und funktionelle Behinderung	<ul style="list-style-type: none"> • <i>Patient-Reported Outcomes Measurement Information System</i> <ul style="list-style-type: none"> ○ <i>Pediatric Pain Interference domain scores</i> ○ <i>Physical Function Mobility domain scores</i> ○ <i>Fatigue domain scores</i> • <i>Face Pain Scale – Revised (FPS-R)</i>
Lebensqualität	
Spezifischer Lebensqualitätsfragebogen für Kinder	<ul style="list-style-type: none"> • <i>Short Form (SF)-10 Health Survey</i>
Sicherheit	
Unerwünschte Ereignisse (UE)	Gesamtrate von Patienten mit Ereignis
Schwerwiegende unerwünschte Ereignisse (SUE)	Gesamtrate von Patienten mit Ereignis
Unerwünschte Ereignisse (UE) mit CTCAE-Grad 3 oder höher: Grad 3: Unterbrechung von täglichen	Gesamtrate von Patienten mit Ereignis

Aktivitäten des Patienten (Therapie/Behandlung erforderlich) Grad 4: Sterberisiko Grad 5: Tod	
Unerwünschten Ereignisse, die zum Studienabbruch führten	Gesamtrate von Patienten mit Ereignis
Unerwünschte Ereignisse von besonderem Interesse	Veränderungen der Vitalfunktionen und des Elektrokardiogramms (EKG) im Laufe der Zeit
CTCAE: <i>Common Terminology Criteria for Adverse Events</i>	

Studiendesign

Für die Nutzenbewertung von Burosumab wurden ausschließlich RCTs herangezogen, die einen direkten Vergleich mit der zVT „Phosphatsubstitution“ ermöglichen.

Studiendauer

Es wird keine Einschränkung bezüglich der Studiendauer vorgenommen, wobei in der Vergangenheit Nutzenbewertungen für chronische Erkrankungen auf Grundlage von Studien mit einer Dauer von mindestens 24 Wochen (6 Monate) durchgeführt wurden (siehe z. B. Tragende Gründe Secukinumab: „Bei einer chronischen Erkrankung, wie der Plaque-Psoriasis, wird jedoch eine Studiendauer von mindestens 24 Wochen als notwendig erachtet.“ [15]. Unterstützt wird dies durch die ICH-E1-Leitlinie der Europäischen Arzneimittelagentur (EMA), die eine Untersuchung von mindestens 6 Monaten bei chronischen Erkrankungen empfiehlt [16].

Ergebnisdarstellung

Für die Nutzenbewertung wurden Studienberichte, Vollpublikationen sowie – soweit möglich – in Studienregistereinträgen veröffentlichte Ergebnisse berücksichtigt. Konferenzbeiträge wurden nicht berücksichtigt. Publikationen, die nicht in deutscher oder englischer Sprache vorlagen, wurden ebenfalls ausgeschlossen.

Ein- und Ausschlusskriterien für randomisierte kontrollierte Studien mit dem zu bewertenden Arzneimittel

Die bibliografische Literaturrecherche sowie die Suche nach RCT in Studienregistern werden entsprechend den Vorgaben des G-BA durchgeführt (siehe Abschnitt 4.2.3.2 und 4.2.3.3).

Im Folgenden sind die Ein- und Ausschlusskriterien für randomisierte, kontrollierte Studien mit Burosumab als Intervention dargestellt (Tabelle 4-4 und Tabelle 4-5).

Tabelle 4-4: Einschlusskriterien für RCT mit Burosumab als Intervention

Einschlusskriterien		
E1	Patientenpopulation	<ul style="list-style-type: none"> • Patienten mit X-chromosomaler Hypophosphatämie (XLH) • Kinder im Alter von 1-17 Jahren
E2	Intervention	<ul style="list-style-type: none"> • Burosumab
E3	Komparator	<ul style="list-style-type: none"> • Phosphatsubstitution (in Kombination mit aktivem Vitamin D)
E4	Endpunkte	<ul style="list-style-type: none"> • Erhebung mindestens eines patientenrelevanten Endpunktes, wie zum Beispiel Morbidität (z. B. Stehhöhe/Liegelänge und Schweregrad der Rachitis), Lebensqualität oder unerwünschte Ereignisse
E5	Studien- und Publikationstyp	<ul style="list-style-type: none"> • Randomisierte kontrollierte Studien (RCT) • Volltextpublikation oder Studienbericht • Veröffentlichung in deutscher oder englischer Sprache
E6	Studiendauer	<ul style="list-style-type: none"> • Ohne Einschränkung

Tabelle 4-5: Ausschlusskriterien für RCT mit Burosumab als Intervention

Ausschlusskriterien		
A1	Patientenpopulation nicht entsprechend E1	<ul style="list-style-type: none"> • Andere Formen der Rachitis (z. B. Vitamin-D-Mangel, Mangelernährung) • Erwachsene (≥ 18 Jahre) • Säuglinge/Kleinkinder (< 1 Jahr)
A2	Intervention nicht entsprechend E2	<ul style="list-style-type: none"> • Ausschließlich andere Wirkstoffe als Burosumab
A3	Komparator	<ul style="list-style-type: none"> • Placebo
A4	Endpunkte nicht entsprechend E4	<ul style="list-style-type: none"> • Studien mit ausschließlich Surrogat-Parametern
A5	Studien- und Publikationstyp nicht entsprechend E5	<ul style="list-style-type: none"> • Keine randomisierte, kontrollierte Studie • Nur Abstract vorliegend • Review-Artikel, Surveys, Notes, Letter, Editorial, Errata, Konferenzbeiträge (Conference abstracts), Poster • Publikation nur in einer anderen Sprache als Deutsch oder Englisch • Keine Volltextpublikation mit primären Studiendaten, Studienberichte oder ausführlicher Ergebnisbericht aus einem Studienregister verfügbar
A6	Art der Information	<ul style="list-style-type: none"> • Bereits identifizierte Publikation ohne neue inhaltliche Informationen (Dublette)

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt. Für MEDLINE (inklusive inprocess & other non-indexed citations) und EMBASE erfolgte die Suche über die Oberfläche von Ovid, für die Cochrane-Datenbank wurde die Cochrane-eigene Suchoberfläche (Cochrane Central Register of Controlled Trials (Clinical Trails) verwendet. Die Suchstrategien wurden unter Verwendung eines Pools von verwandten Suchbegriffen in Bezug auf die Einschlusskriterien der jeweiligen Recherche für jede Datenbank angepasst. Die Suchstrategien beinhalten jeweils eine Kombination von Freitext und Begriffen der Medical Subject Headings (MeSH) bzw. verwandter Konzepte in anderen Datenbanken. Da aufgrund der erst im letzten Jahr erlangten Zulassung von Burosumab nur mit einer geringen Anzahl an Studien mit diesem Wirkstoff zu rechnen war und um die Sensitivität der Recherche nicht unnötig einzuschränken, wurde die Suchstrategie in einem mehrstufigen Verfahren verfeinert. Erster Schritt war die Suche mit Suchbegriffen für die Intervention in Verbindung mit einem RCT-Filter. Nur wenn bei diesem Schritt eine größere Anzahl möglicher Studien identifiziert worden wäre, wäre die Suchstrategie in den nächsten Schritten mit Suchbegriffen für die Vergleichstherapie und für die Indikation verfeinert worden. Dies war jedoch nicht der Fall. Aufgrund der allgemein niedrigen Trefferzahlen wurden a priori keine weiteren Einschränkungen in der Suchstrategie, wie z. B. die Beschränkung auf Humanstudien, vorgenommen. Die verwendeten Suchstrategien und Suchbegriffe sind in Anhang 4-A detailliert beschrieben und dokumentiert.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov) (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm>) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Neben der systematischen Suche in medizinischen Datenbanken wird auch eine Suche in öffentlich zugänglichen Studienregistern durchgeführt. Das Ziel dieser Suche ist die Identifikation aller klinischen Studien mit Burosumab bei Kindern im Alter von 1-17 Jahren mit XLH, die zur Bewertung des Ausmaßes des Zusatznutzens herangezogen werden können. Die Suche in den Registern gilt der Identifizierung aller abgeschlossenen, abgebrochenen und laufenden Studien.

Die Suche wird in den oben angeführten Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov), *WHO International Clinical Trials Registry Plattform* (WHO ICTRP), *EU Clinical Trials Registry* (EU-CTR), *EMA Clinical Data* und Klinische Prüfungen PharmNet.Bund durchgeführt. Die detaillierte Suchstrategie ist für jedes Studienregister im Anhang 4-B getrennt dargestellt.

Die Suche wurde vom 09.-16.07.2019 durchgeführt; es wurden keine generellen Einschränkungen bei den genannten Recherchen vorgenommen.

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Literaturrecherche wurde von zwei Personen unabhängig voneinander durchgeführt. Zunächst werden die Titel und Abstracts der gefundenen Treffer gesichtet und anhand der vordefinierten in Abschnitt 4.2.2 dargelegten Kriterien selektiert. Im Weiteren werden die verbleibenden Studien im Volltext betrachtet, tabellarisch aufgelistet und entsprechend der Ein- und Ausschlusskriterien ausgewählt. Abweichende Beurteilungen werden im abschließenden Konsensverfahren durch beide Personen, oder durch Hinzuziehen einer dritten unabhängigen Person, geklärt und entsprechend dokumentiert.

Die Ergebnisse der Studienselektion aus der bibliographischen Literaturrecherche sind im Abschnitt 4.3.1.1.2 dargestellt.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Validität der in diesem Dossier präsentierten Evidenz wurde in Übereinstimmung mit den oben beschriebenen Anforderungen und den Kriterien in Anhang 4-F bewertet. Dazu wurden die Verzerrungsaspekte der Ergebnisse der eingeschlossenen Studie beschrieben. Die Bewertung der einzelnen Aspekte wurde für jeden Endpunkt mithilfe des Bewertungsbogens aus Anhang 4-F dokumentiert. Die auf Studienebene berücksichtigten Faktoren waren insbesondere: Erstellung einer Randomisierungsreihenfolge, Geheimhaltung der Gruppenzuweisung, Verblindung von Patient und Prüfarzt, Vergleichbarkeit der Eigenschaften der Studienarme bei Studienbeginn, ergebnisorientierte oder selektive Berichterstattung und Berücksichtigung von weiteren für die Prognose relevanten Faktoren. Die auf Endpunktebene berücksichtigten Faktoren waren insbesondere: Verblindung der Endpunkterheber, Umsetzung des *Intention-to-treat* (ITT)-Prinzips, ergebnisorientierte oder selektive Berichterstattung und andere Faktoren, die zur Verzerrung der Ergebnisse beitragen könnten.

Das Verzerrungspotenzial wurde entweder als niedrig oder als hoch eingestuft. Wenn eine Verzerrung der relevanten Ergebnisse mit hoher Wahrscheinlichkeit ausgeschlossen werden konnte, wurde dies als niedriges Verzerrungspotenzial angesehen. Wenn von der Verzerrungsquelle eine relevante Beeinflussung der Ergebnisse zu erwarten war, wurde dies als hohes Verzerrungspotenzial angesehen. Dieser Fall musste auch angenommen werden, wenn keine ausreichenden Informationen zur Bewertung des Verzerrungspotenzials vorlagen (z. B. wenn kein Studienbericht, sondern ausschließlich Publikationen vorlagen). Lagen hingegen zusätzlich Unterlagen – z. B. Informationen von Zulassungsbehörden – vor, wurden diese in die Bewertung des Verzerrungspotenzials einbezogen. Die Verzerrungspotenzialbewertung wurde zur Bewertung der Aussagekraft der Nachweise verwendet und führte nicht zum Ausschluss von Daten.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Design und Methodik der eingeschlossenen Studie werden in den Abschnitten 4.3.1.2 und 4.3.1.3 sowie den zugehörigen Anhängen dokumentiert. Die Bewertung folgt den *Consolidated Standards of Reporting Trials* (CONSORT)-Vorgaben (Items 2b bis 14) sowie dem CONSORT-Flow-Chart. Bei der eingeschlossenen Studie handelt es sich um eine RCT, somit entspricht die Beurteilung und Darstellung des Designs und der Methodik der Studien den Vorgaben. Nicht randomisierte Interventionsstudien und epidemiologische Beobachtungsstudien werden nicht zur Bewertung des Zusatznutzens von Burosumab herangezogen. Die Ergebnisse der Burosumab Studien der Evidenzstufe IV wurden bereits im ersten AMNOG-Verfahren D-349 vom 15. April 2018 dargestellt (UX023-CL201, UX023-CL205 und die historische Kontrollstudie mit einer Substitutionstherapie UX023-CL002).

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten

¹ Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.

² Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Publ Health* 2004; 94(3): 361-366.

³ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

(Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Patientencharakteristika

Im Rahmen der Studie UX023-CL301 wurden demographische und krankheitsspezifische Charakteristika erhoben, um die in den Studien eingeschlossenen Patienten zu beschreiben. Hierbei wurden folgende Parameter gewählt:

Demographische Parameter

- Alter
- Geschlecht (Subgruppen: XLH-Patienten < 5 Jahre und \geq 5 Jahre)
- Rasse und Ethnie
- Gewicht, *body mass index* (BMI) und Größe

Krankheitsspezifische Parameter

- Vorbehandlung mit einer Substitutionstherapie und deren Initiierung und Dauer
- Radiographische Nachweise einer Rachitis (\geq 2,0 RSS)
- *PHEX*-Mutation

Stratifizierungscharakteristika der Studie

- Baseline RSS Total Score: \leq 2,5 (Low RSS) vs. $>$ 2,5 (High RSS)
- Alter: < 5 Jahre und \geq 5 Jahre
- Region: Japan vs. Rest der Welt

Patientenrelevante Endpunkte

Im Folgenden werden die Ergebnisse der randomisierten und aktiv kontrollierten Phase-III-Studie UX023-CL301 von Baseline zu Woche 40 und 64 dargestellt. Für die frühe Nutzenbewertung von Burosumab zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit X-chromosomaler Hypophosphatämie (XLH) und

röntgenologischem Nachweis einer Knochenerkrankung wurden Endpunkte bezüglich der Morbidität, Lebensqualität und Sicherheit untersucht. Die Mortalität wurde im Rahmen der Sicherheit erhoben. Folgende patientenrelevante Endpunkte werden betrachtet und dargestellt:

- **Morbidität**

- Rachitissymptomatik (RGI-C und RSS)
- Körpergröße [Z-Score]
- Gehfähigkeit: 6-Minuten-Gehdistanz Test (6MWT) [in Meter und Prozent]
- Serumphosphat-Konzentration [mg/dl]
- Schmerzerhebung und funktionelle Behinderung:
 - *Patient-Reported Outcomes Measurement Information System (PROMIS): Pediatric Pain Interference, Physical Function Mobility and Fatigue domain scores*
 - *Faces Pain Scale - Revised (FPS-R)*
- Zahnärztliche Untersuchungen

- **Lebensqualität**

- **Sicherheit**

- Gesamtrate
- Unerwünschte Ereignisse (UEs)
- Schwerwiegende UEs (SUEs)
- UEs, die zum Studienabbruch führten
- Todesfälle

Mortalität

Mortalität ist ein patientenrelevanter Endpunkt, der jedoch für die XLH kein relevanter Endpunkt ist, da es sich um eine schwere, chronische Erkrankung des Knochenskeletts und des Stoffwechsels handelt. Burosumab ist ein Wirkstoff, der keinen direkten Einfluss auf die Mortalität hat. In der Studie UX023-CL301 wird die Mortalität durch die Datenerhebung der schwerwiegenden unerwünschten Ereignisse berücksichtigt.

Morbidität

Rachitissymptomatik (RGI-C, RSS)

Seit Jahrzehnten gilt die Radiologie als Standard für die Diagnostik und Beurteilung der Rachitissymptomatik. Folglich wurde die therapiebedingte Verbesserung der Rachitissymptomatik anhand von Röntgenbildaufnahmen unter Anwendung des *Rickets Severity Scores* (RSS) und des *Radiographic Global Impression of Change* (RGI-C)-Scores zu verschiedenen Erhebungszeitpunkten in der Studie UX023-CL301 verwendet [17-21].

Während der RSS eine quantitative Methode zur Beurteilung des Schweregrads der Rachitissymptomatik zu jeweils einem Messzeitpunkt darstellt, dient der RGI-C-Score ergänzend als Methode zur qualitativen Beurteilung zu mehreren Zeitpunkten. Durch die Erhebung der beiden Endpunkte kann sowohl eine quantitative als auch qualitative Bewertung der Veränderung der Rachitissymptomatik unter der Burosumab-Behandlung und Substitutionstherapie erfolgen. Unter Berücksichtigung der Studie von Thacher et al. (2019), konnte ein Zusammenhang zwischen einem erhöhten RSS und einer Beeinträchtigung auf biochemischer, klinischer sowie funktionaler Ebene bei Kindern mit XLH festgestellt werden. Diesbezüglich ist die verminderte Leistungsfähigkeit der Patienten mit erhöhtem RSS beim patientenrelevanten Endpunkt 6MWT explizit zu nennen. Aufgrund dessen sind der RSS und RGI-C als patientenrelevant einzustufen [17, 22].

Rickets Severity Score (RSS)

Der RSS ist eine radiographische Bewertungsmethode, welche zur Beurteilung des Schweregrads einer ernährungsbedingten Rachitis in den Handgelenken und Knien basierend auf dem Grad der metaphysären Unschärfe, Konkavität und dem Anteil der betroffenen Wachstumsfuge validiert wurde. Der RSS-Gesamtscore setzt sich aus dem Handgelenkscore mit max. 4 Punkten und dem Kniescore mit max. 6 Punkten zusammen. Bei letzterem ist ein Multiplikator-Faktor von 0,5 zu beachten, wenn ≤ 1 Kondylus oder Plateau betroffen ist [17, 23]. Trotz ähnlicher radiologischer Anomalien variieren die ernährungsbedingte und XLH-bedingte Rachitis stark in ihren Schweregraden. So ist der RSS bei XLH-Patienten in der Regel deutlich niedriger. Dies beruht auf den Unterschieden der radiologischen Bilder zwischen beiden Krankheitsbildern, durch die die Scoring-Methodik beeinflusst wird. Bei XLH-Röntgenbildaufnahmen werden nur bestimmte Abschnitte (z. B. Konkavität an den Metaphysen der Ulna) bewertet, die zu geringeren Scoring-Ergebnissen führen. Insbesondere am Knie ist die Punktzahl bei der XLH durch radiologische Unterschiede in den Röntgenbildaufnahmen gesenkt. So ist bei Kindern mit XLH eine komplette Aufhellung sowie eine Verbreiterung der Epiphysen an beiden Kondylen selten zu beobachten, was zu einem Multiplikator-Faktor von 0,5 führt und die Punktzahl senkt [17].

Bei der ernährungsbedingten Rachitis ist aus der Addition der Punktzahlen der analysierten Knochenveränderungen in Handgelenk und Knie ein Gesamtscore von 0-10 Punkten bei einer Skalierung von 0,5 möglich, wobei ein hoher Wert eine schwere Rachitissymptomatik anzeigt. Der übliche Bereich vom Gesamtscore in XLH dagegen liegt zwischen 0 und 6,5 und spiegelt nur den epiphysären und distalen metaphysären Anteil der Skelettanomalien wider, die bei betroffenen Kindern sehr häufig auftreten. Dieser Unterschied im üblichen Gesamtscore der beiden Krankheitsbilder ist bei der Bewertung des Schweregrads zu berücksichtigen. So würde beispielsweise ein Gesamtscore von 5 bei einer ernährungsbedingten Rachitis als milder interpretiert werden können, während dieser Gesamtscore bei der XLH-bedingten Rachitissymptomatik bereits im oberen Bereich des üblichen Gesamtscores liegen würde [17, 22].

Die Beurteilung aller Röntgenbildaufnahmen wurde von einem zentralen, unabhängigen Auswerter, Thomas Thacher, dem Entwickler des RSS, vorgenommen. Um das Risiko eines Bias zu minimieren, erfolgte die Bewertung nach einer vordefinierten Methodik, in der Thacher zuvor geschult wurde (*Biomedical Systems, Independent Review Manual* [8, 9]). Die bilateralen

Röntgenbildaufnahmen wurden einzeln in einer randomisierten Reihenfolge präsentiert, ohne dass ein Vergleich zu einer weiteren Röntgenbildaufnahme gegeben war. Somit war es möglich, den Auswerter für den Zeitpunkt der Röntgenbildaufnahme und den Behandlungsstatus inklusive Informationen zu Patientenalter/-geschlecht, Dosierungsregime oder Patientenhistorie zu verblinden. Es wurde die Seite mit der stärksten Symptomausprägung zur Auswertung herangezogen. Demnach stellt der Gesamtscore den schlechtesten Wert der beiden Körperseiten dar. Jede Bewertung wurde direkt in ein elektronisches Datenerfassungssystem (*EDC system – electronic data capture system*) eingetragen und elektronisch an eine zentrale Bildgebungseinrichtung übermittelt. So konnte der Auswerter die Punktzahlen nachträglich nicht mehr abrufen bzw. korrigieren [17, 20, 22].

Radiographic Global Impression of Change (RGI-C)

Die Ratingskala RGI-C ist ein anerkannter Bewertungsmaßstab zur qualitativen Bewertung der Verbesserung oder der Verschlechterung skelettaler Manifestationen bei der Rachitis. Die RGI-C-Skala wurde von drei pädiatrischen Radiologen mit Fachkenntnissen in der Skelettdysplasie entwickelt, die sich mit der Untersuchung von Handgelenks- und Knieaufnahmen von Kindern mit XLH befassten [17, 23]. Dieser wurde ergänzend zum RSS erhoben, um Veränderungen des Schweregrads der Rachitissymptomatik und Deformationen im Zeitverlauf zu bewerten. Dadurch ist eine weitere Einschätzung des Erfolgs der Burosumab-Behandlung gegeben. Die RGI-C-Skala bewertet dieselben radiographischen Merkmale der Rachitissymptomatik wie auch die RSS-Skala. Der Unterschied liegt darin, dass die RGI-C-Methodik einem Auswerter ermöglicht, alle auf einer Röntgenbildaufnahme sichtbaren Anomalien aus einer umfassenden Liste auszuwählen, statt wie bei der RSS-Methodik auf ein hierarchisches Punktesystem beschränkt zu sein. Beim RSS wird ein Schweregrad anhand einzelner Handgelenk- und Knie-Röntgenbildaufnahmen von einem einzigen Zeitpunkt abgeleitet. Die RGI-C-Methodik dagegen ermöglicht die Beurteilung der Veränderung des radiologischen Erscheinungsbilds durch einen direkten Vergleich von einem früheren Bild (Baseline Bild A) und einem Bild zu einem späteren Zeitpunkt (Post-Baseline Bild B). Diese zwei Bilder werden dem Auswerter zeitgleich präsentiert [17, 20].

Der Auswerter wird zunächst aufgefordert, das Vorhandensein oder Fehlen einer Reihe von XLH-bezogener Anomalien in Bild A zu bewerten. Zu diesen XLH-bezogenen Anomalien zählen:

- Metaphysäre Aufhellung
- Metaphysäre/epiphysäre Trennung
- Metaphysäre Unschärfe
- Metaphysäre konkave Becherung

Anschließend erfolgt die Beurteilung der Veränderung jeder einzelnen Anomalie mithilfe der 7-Punkte-Ordinal-RGI-C-Skala basierend auf dem Vergleich mit Bild B.

- **+3:** sehr deutliche Verbesserung/(nahezu) vollständige Heilung der Rachitissymptomatik
- **+2:** erhebliche Verbesserung/wesentliche Heilung der Rachitissymptomatik
- **+1:** minimale Verbesserung der Rachitissymptomatik
- **0:** keine Veränderung
- **-1:** minimale Verschlechterung der Rachitissymptomatik
- **-2:** mittelstarke Verschlechterung der Rachitissymptomatik
- **-3:** starke Verschlechterung der Rachitissymptomatik

Neben den so ermittelten RGI-C-Scores für die Knie- und die Handgelenke, wird der Auswerter abschließend aufgefordert, einen Gesamtscore zu definieren. Dieser beruht nicht auf Berechnungen der Einzelscores, sondern auf dem generellen Eindruck des Auswerter bezüglich der Veränderungen der Röntgenbildaufnahmen [20].

Die Beurteilung aller Röntgenbilder wurde von drei unabhängigen, pädiatrischen Radiologen vorgenommen. Die Bewertung erfolgte nach einer vordefinierten Methodik, in welcher die Auswerter zuvor geschult wurden (*Biomedical Systems, Independent Review Manual*). Die Auswerter waren bezüglich Dosierung, Dosierungsregime, Befolgen des Studienprotokolls, Behandlungsdauer und Vorbehandlung mit einer Substitutionstherapie verblindet. Des Weiteren erfolgte die Risikominimierung eines Bias durch das randomisierte Mischen der Röntgenbildaufnahmen verschiedener Aufnahmezeitpunkte der unterschiedlichen Behandlungsgruppen [20].

Wie bei der RSS-Methodik wurde jede Bewertung direkt in ein elektronisches Datenerfassungssystem (*EDC system – electronic data capture system*) eingetragen und elektronisch an eine zentrale Bildgebungseinrichtung übermittelt. Den Auswertern war es nicht möglich die Scores nachträglich abzurufen bzw. zu korrigieren. Abschließend wurden die RGI-C-Scores für Knie, Handgelenke und gesamt als Mittelwerte aus den unabhängigen Scores der einzelnen Auswerter berechnet [17, 20, 23].

Die Erhebungszeitpunkte der Röntgenbildaufnahmen der Knie- und Handgelenke zur Ermittlung des RSS und RGI-C-Scores waren zu Screeningvisite, zu Woche 40 und 64. Die Röntgenbilder der langgestreckten Beine der stehenden Patienten wurden zu Baseline, Woche 40 und 64 aufgenommen [20].

Sowohl beim RSS als auch beim RGI-C-Score wurden die Auswerter für die Standardisierung des Bewertungsprozesses mit einem *Independent Review Manual* in der Methodik geschult (*Biomedical Systems, Independent Review Manual*). Weiterhin erfolgte wie beschrieben eine Verblindung der Auswerter sowie eine Randomisierung der Reihenfolge der Röntgenbildaufnahmen, um das Risiko eines Bias zu minimieren. Zur Gewährleistung der standardisierten Erstellung der Röntgenbildaufnahmen in jedem Studienzentrum wurde zudem eine Charta entwickelt. Ferner zeigte die Untersuchung der Intra- und Interrater-Reliabilität, dass der RSS und der RGI-C-Score verlässliche Ergebnisse unabhängig von der Auswerter-Anzahl liefern [17, 20, 22, 23]. Somit lässt sich konstatieren, dass die Datenvalidität für beide Scoring-Methoden gegeben ist.

Körpergröße (Stehhöhe/Liege­länge)

Bei der XLH handelt es sich um eine schwere und chronisch bedingte Erkrankung des Knochenskeletts, welche durch Kleinwuchs und Knochendeformationen gekennzeichnet ist. Ein verzögertes Wachstum sowie Gedeihstörungen sind charakteristisch für die XLH-Patienten. Folglich sind diese Patienten sowohl bei ihren alltäglichen Aktivitäten als auch in ihrer Lebensqualität eingeschränkt. Das Ziel der medikamentösen Therapie ist das Erlangen eines altersbezogenen normalen Wachstums und einer daraus resultierenden Verbesserung der Lebensqualität [13, 20]. In der Vergangenheit wurden bereits durch den G-BA anthropometrische Parameter, insbesondere bei Kindern mit charakteristischen, krankheitsbedingten Wachstumsstörungen, als patientenrelevante Morbiditätsparameter eingeschätzt [24]. Aufgrund dessen gilt die Verbesserung des Wachstums als ein patientenrelevanter Endpunkt.

Die Körpergröße in Zentimeter [cm] wurde bei Patienten ≥ 2 Jahre durch Messung der Stehhöhe mit einem Radiometer ermittelt. Bei Patienten < 2 Jahre bzw. bei Kindern, welche nicht fähig oder nicht bereit waren, für die Messung zu stehen, wurde die Liege­länge bestimmt [25]. Um Messfehler zu minimieren, wurden je Erhebungszeitpunkt drei Messungen durchgeführt und aus diesen Daten ein Mittelwert berechnet. Während der kontrollierten Behandlungsphase wurden die Daten zu Baseline, zu Woche 24, 40 und 64 erhoben. In der Extensionsphase liegen die Erhebungszeitpunkte zu Woche 76, 88, 100, 112, 124 und Woche 140. Im Gegensatz zu den anderen Endpunkten erfolgte die erste Datenauswertung zu Woche 64. Die Körpergröße wurde als alters- und geschlechtsadjustierte Z-Scores berichtet. Der Z-Score beschreibt die Anzahl der Standardabweichungen jedes einzelnen Messwertes von den alters- und geschlechtsadjustierten normalen Durchschnittswerten [26]. Sowohl die Vergleichsdaten der Z-Scores als auch die Perzentile beruhen dabei auf Wachstumscharts von einer Stichprobe gesunder Kinder aus den USA, veröffentlicht vom *Center of disease control and prevention (CDC)* [20, 21, 27].

Bei der Körpergröße als auch beim Wachstum handelt es sich um Parameter, welche von sehr vielen Faktoren beeinflusst werden. Hierzu gehören sowohl die in den Wachstumscharts berücksichtigten Variablen Geschlecht und Alter als auch ethnische Unterschiede sowie Unterschiede im Zeitpunkt der Datenerhebungen. Letztere können insbesondere aufgrund von sozioökonomischen Veränderungen im Zeitverlauf Einfluss auf das Größenwachstum nehmen [26]. Angesichts des relativ hohen Patienten-Anteils von 82 % mit kaukasischem Ursprung (50/61) und 51 % Amerikanern (31/61) sowie der länderspezifischen komplexen Variabilität des Parameters stellen die Vergleichsdaten der Z-Scores von der amerikanischen gesunden Population des CDC das Verfahren mit der größten zur Verfügung stehenden Validität dar. Ergänzend ist zu erwähnen, dass kein deutscher Studienstandort in der RCT UX023-CL301 eingeschlossen war [20]. Ferner wurden in dem früheren Nutzenbewertungsverfahren des Wirkstoffs Asfotase alfa die Z-Scores auf Datengrundlage des CDC vom G-BA als plausibel anerkannt [24, 28]. Weiterhin bestätigt die vergleichende Studie von Rosario et al., dass die Wachstumscharts der CDC für den Vergleich mit deutschen Wachstumskurven geeignet sind [29]. Daraus folgt, dass länderspezifische Z-Scores nicht zwingend erforderlich sein müssen,

sondern viel mehr die Herkunft der Patientenpopulation entscheidend ist. Folglich ist die Datenvalidität gewährleistet.

Gehfähigkeit (6MWT)

Gangstörungen sind häufig ein klinisches Merkmal bei pädiatrischen Patienten mit XLH. Zurückzuführen ist diese Störung auf eine Reihe von Folgeerscheinungen der XLH, wie die Knochendeformation, die Knochen- und Gelenkschmerzen sowie die eingeschränkte Muskelfunktion aufgrund der Hypophosphatämie. Die Knochendeformation in den unteren Extremitäten und die Achsenfehlstellung, die mit der XLH assoziiert sind, verursachen oftmals ein „watschelndes“ und anormales Gangbild. Die Muskel- und Skelettauswirkungen der Hypophosphatämie können verminderte Kraft, Gleichgewichts-, Koordinationsstörungen und erhöhte Steifigkeit bewirken, die die Gehfähigkeit und andere grobmotorische Funktionen beeinflussen. Diese Symptome führen neben dem im Alter mit XLH assoziierten erhöhten Arthrose- und Enthesiopathie-Risiko zu starken Einschränkungen der körperlichen Bewältigung des Alltags [30]. Eine vom Patienten bewusst wahrnehmbare Verbesserung der körperlichen Leistungsfähigkeit unter Burosumab-Behandlung hat demnach einen beträchtlichen Anstieg der Lebensqualität zur Folge.

Zur Ermittlung der Verbesserung der körperlichen Leistungsfähigkeit wird der Endpunkt 6-Minuten-Gehdistanz Test (6MWT) herangezogen. Dieser ist als Messinstrument zur Bestimmung der körperlichen Leistungsfähigkeit anerkannt. Der 6MWT untersucht eine Gehstrecke, die ein Patient in 6 Minuten zurücklegen kann. Unter differenzierter Betrachtung folgender, ermittelter Werte lässt sich der 6MWT als patientenrelevanter Endpunkt definieren: Der absolute Wert der Gehdistanz zu Baseline, der absolute Wert zum Zeitpunkt eines zuvor festgelegten Zeitpunkts nach Therapiebeginn (Post-Baseline) und der Veränderung der 6MWT von Baseline zu Post-Baseline. Ferner wird der Prozentsatz der vorhergesagten Normalwerte anhand von veröffentlichten normativen Daten basierend auf Alter, Geschlecht und Größe berechnet. Dies dient zur Ergänzung der absoluten Werte, da aufgrund der bisherigen, geringen Studienlage unklar ist, ob relative oder absolute Werte eine bessere klinische Einschätzung zulassen [21, 30-32]. Laut ATS (*American Thoracic Society*) ist die Angabe der absoluten Werte zu empfehlen [31]. Die absoluten Werte der Veränderung der 6MWT zwischen Baseline und Post-Baseline zeigen dem Patienten die veränderte Gehfähigkeit unter Burosumab-Behandlung auf. Insbesondere den Kindern hilft der spielerische Anreiz der Softring-Technik (siehe unten), sich der positiven Veränderung ihrer körperlichen Leistungsfähigkeit bewusst zu werden. Folglich handelt es sich beim 6MWT um einen patientenrelevanten Endpunkt.

Während der kontrollierten Behandlungsphase erfolgt die Datenerhebung für Übungszwecke zum Zeitpunkt der Screeningvisite und anschließend zu Baseline, Woche 24, 40, 64 sowie bei vorzeitigem Studienabbruch innerhalb von 3 Monaten. Mit Ausnahme der nach Woche 64 abgeschlossenen Studien aus Japan und Korea sind weitere Erhebungszeitpunkte in Woche 88, 112 und 140 gegeben (Extensionsphase) [20]. Das Mindestalter der pädiatrischen Patientenpopulation betrug am Tag der Screeningvisite 5 Jahre. Zudem sind zur

Gewährleistung der Patientensicherheit kardiale Ausschlusskriterien definiert. Hierzu gehört eine Ruheherzfrequenz von $<120/\text{min}$ und ein Blutdruck von $<180/100 \text{ mmHg}$, um das Auftreten von Arrhythmien sowie negativen, kardiovaskulären Ereignissen zu vermeiden [25, 31].

Nach standardisierter Erklärung des Testverfahrens durch das geschulte klinische Personal geht der Patient unter standardisierter, minütiger Zeitangabe eine Rundstrecke um zwei Pylonen. Bei Erschöpfung kann sich der Patient jederzeit zum Ausruhen auf einen bereitgestellten Stuhl setzen und individuell entscheiden, wenn er weitergehen möchte. Über Rundenanzahl und mittels eines Maßbandes wurde die Gehstrecke in Metern gemessen. Zudem wurde der gesamte Test zur Beurteilung der Geh-Qualität per Video aufgezeichnet. Zusätzlich wurden nach dem Test die vom Patienten berichteten Symptome notiert. Die Anforderung an die Gehstrecke ist ein Flur mit einem flachen, harten Boden und einer Länge von 30 m sowie einer Breite von 2,5 m. Zur Standardisierung der Strecke werden Pylonen im Abstand von 20 m aufgestellt, sodass eine absolvierte Runde 40 m umfasst [25].

Um eine möglichst hohe Validität zu erhalten, werden die Kriterien der *ATS Guidelines* als qualitätssichernde Maßnahme berücksichtigt [25, 31]. Aufgrund dessen erfolgten auf Basis der *ATS Guidelines* an allen teilnehmenden Studienzentren persönliche Schulungen des klinischen Personals durch Ultragenyx. So wird das klinische Personal angewiesen, die im *Clinical Evaluator Manual* beschriebenen Standardanweisungen (*standard operating procedures - SOPs*) zu beachten [25, 31, 32]. Hierzu gehören beispielsweise die standardisierten Formulierungen für die Kommunikation mit dem Patienten. Erst nach erfolgreichem Abschluss der Schulung und Erhalt des Schulungszertifikats ist das klinische Personal berechtigt, den 6MWT zu leiten [7, 25, 33]. Zudem beträgt die Anzahl des klinischen Personals in der Regel lediglich zwei Personen pro Standort, um die Messvariabilität während der gesamten Studie zu minimieren, und die Validität der Daten trotz unterschiedlicher Studienorte zu gewährleisten.

Folgende Änderungen der *ATS Guidelines* wurden wegen der pädiatrischen Patientenpopulation und der XLH-bedingten Symptomatik vorgenommen: Aufgrund des unsicheren Ganges der XLH-Patienten führt die Messung der Gehstrecke mittels Messrades zu einem potenziell erhöhten Sturzrisiko der Patienten. Folglich wurde auf ein Messrad verzichtet, um eine freie Bewegung der oberen Extremitäten und eine damit verbundene Gleichgewichts-anpassung zu ermöglichen. Zur Vorbeugung einer möglichen Motivationsverzerrung durch das fehlende Messrad, erhielten die Patienten die Möglichkeit an jeder Biegung Softringe zu sammeln sowie fallen zu lassen. Dementsprechend ist mit der Softring-Technik für die Kinder ein spielerischer Anreiz geschaffen, die 6 Minuten kontinuierlich zu gehen, um möglichst viele Ringe zu sammeln [25, 31]. Trotz dieser Modifikationen der *ATS Guidelines* lässt sich konstatieren, dass die Datenvalidität aufgrund des geschulten klinischen Personals gewährleistet ist.

Serumphosphat

Eine Beeinträchtigung des Wachstums ist eines der vorherrschenden Merkmale von Kindern mit XLH. Es wurde gezeigt, dass das Wachstum der Beine und des Rumpfes bei XLH entkoppelt ist und mit dem Phosphatstoffwechsel zusammenhängt [34]. Durch die *PHEX*-Mutation wird die Aktivität einer Endopeptidase verringert, die normalerweise einen phosphaturischen Faktor (FGF23) inaktiviert [35]. Die dadurch erhöhte FGF23-Konzentration hemmt den für die Phosphat-Rückresorption wichtigen Natrium-Phosphat-Kotransporter (NaPi-2). Somit haben XLH-Patienten stark verminderte Serumphosphatwerte.

Der Indikator für eine erfolgreiche Therapie mit Burosumab ist die Serumphosphat-Konzentration, die als Zielparamester gilt. Das Erreichen einer normal-physiologischen Serumphosphat-Konzentration bei einer Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D ist nicht möglich. Dies stellt einen großen Vorteil bei der Behandlung mit Burosumab gegenüber der Substitutionstherapie dar, da die starke Strahlenbelastung, die durch die häufigen Röntgenaufnahmen entsteht, entfällt [2, 3, 35].

Schmerzerhebung und funktionelle Behinderung

Patient-Reported Outcomes Measurement Information System (PROMIS): Pediatric Pain Interference, Physical Function Mobility and Fatigue Domain Score

Schmerzen und funktionelle Behinderungen sind für die an XLH erkrankten Patienten laut Fachexperten am meisten belastend. Die funktionellen Einschränkungen und chronischen Schmerzen sind Folge einer XLH-Erkrankung. Sowohl die Behinderung als auch die starken Knochen- und Gliederschmerzen sind eine bedeutende Belastung für betroffene Patienten. Schmerzen und daraus resultierende Behinderungen haben enorme Auswirkungen auf die Aktivität des täglichen Lebens und führen somit zu erheblichen Einschränkungen der Lebensqualität und oft zur Erschöpfung der XLH-Patienten. Daher ist die Reduktion von Schmerzen und eingeschränkter Mobilität ein patientenrelevanter Endpunkt. Dieser Endpunkt wurde in der Studie UX023-CL301 anhand der *Patient Reported Outcomes Measurement Information System* (PROMIS) pädiatrischen Schmerzskala gemessen [32].

Mithilfe der Erkenntnisse aus der pivotalen Studie UX023-CL201 und der engen Zusammenarbeit mit der *Food Drug Administration* (FDA) wurde abweichend zur Studie UX023-CL201, die Endpunktbestimmung „funktionelle Behinderung und Schmerz“ in der RCT UX023-CL301 mit dem PROMIS-Fragebogen anstatt des *Pediatric Orthopaedic Society of North America - Pediatric Outcomes Data Collection Instrument* (POSNA-PODCI) durchgeführt. Hiermit wurde der Empfehlung der FDA gefolgt, die die Ansicht vertrat, dass der PROMIS-Fragebogen den Gesundheitszustand der XLH-Patienten ab einem Alter von 5 Jahren eindeutiger erfasse [32].

Der PROMIS-Fragebogen wurde von dem *National Institutes of Health* (NIH) entwickelt und verwendet spezifische Domänen zur Beurteilung des Wohlbefindens der Patienten [36]. Der

Ansatz basiert darauf, dass gesundheitliche Merkmale wie Schmerz und körperliche Funktionen nicht für eine bestimmte Erkrankung spezifisch sind. Das PROMIS-System besteht aus einem umfangreichen Verzeichnis mit Fragen, die nach Relevanz ausgewählt und für die Erstellung eines krankheitsspezifischen Fragebogens verwendet werden können. Das pädiatrische PROMIS-Verzeichnis ist in zahlreiche Domänen aus den drei Hauptmodulen physische, mentale und soziale Gesundheit aufgebaut. Auf Grundlage der bekannten Hauptsymptome der XLH-Patienten aus der pivotalen Studie UX023-CL201 wurde in der RCT UX023-CL301 eine Auswahl von 22 Items aus den Domänen Schmerz (bestehend aus 20 Items), Mobilität (bestehend aus 24 Items) und Fatigue (bestehend aus 25 Items) abgefragt [32]. Die Beantwortung der Fragebögen in Papierform erfolgte zur Screeningvisite, zu Baseline, zu Woche 24, 40, 64 und in der Extensionsphase zu Woche 88, 112 sowie 140. Nach Vervollständigung wurden die Antworten in das PROMIS *Online Scoring System* hochgeladen [20, 21]. Aus der Summe der Item-Scores wurden die Domänen-Scores als T-Werte zusammengefasst. Der Referenz-T-Score beträgt als standardisierter Mittelwert 50 mit einer Standardabweichung von 10. Positive Effekte führen bei der Domäne Mobilität zu einem Anstieg und bei den Domänen Schmerz und Fatigue zu einer Abnahme des Scores verglichen zur Referenz. Es erfolgte die Betrachtung der Veränderungen der T-Werte im Studienverlauf im Vergleich zu den Werten zu Baseline [20]. Es wurden alle Patienten mit einem Alter von ≥ 5 Jahren in die Befragung aufgenommen. Dabei wurden zwei Fragebogen-Varianten verwendet: die Selbstbericht-Version [6] und die Eltern-Version [5]. Die Auswahl der korrekten Versionen war abhängig von dem Alter der Patienten zum Zeitpunkt der Screeningvisite. Für Kinder im Alter von 5 bis einschließlich 7 Jahren zur Screeningvisite, erfolgte die Beantwortung der Fragen der Eltern-Version durch einen Erziehungsberechtigten. Wenn möglich wurde dieser im Laufe der Studie immer vom selben Erziehungsberechtigten ausgefüllt. Kinder ≥ 8 Jahre zur Screeningvisite erhielten die Selbstbericht-Version, die eigenständig von den Kindern ausgefüllt werden konnte. Für Kinder mit XLH, die während der Studie 8 Jahre alt wurden, war es nicht gestattet, von der Eltern- in die Selbstbericht-Version zu wechseln. Bei Kindern ≥ 8 Jahre, die mit dem Ausfüllen der Selbstbericht-Version Schwierigkeiten hatten, gab es zwei Möglichkeiten. Entweder es konnte der Selbstbericht-Fragebogen per Interview durch ein geschultes Personal ausgefüllt werden oder es wurde wie bei den unter 8-Jährigen vorgegangen, so dass die Eltern-Version von einem Erziehungsberechtigten ausgefüllt wurde. Um eine hohe Validität zu gewährleisten, musste folglich bei jeder weiteren Visite die zuvor festgelegte Messpraxis befolgt werden. Die Entscheidung, ob ein Interview der Selbstbericht-Version durch ein geschultes Klinikpersonal erfolgt oder die Eltern-Version von einem Erziehungsberechtigten ausgefüllt wird, oblag dem Prüfarzt und musste von diesem ausführlich begründet und dokumentiert werden [19, 20].

Wie bereits erwähnt, erhielten die Patienten oder Erziehungsberechtigten zur Selbst- bzw. Fremdeinschätzung eine statische Fragebogenversion aus den Domänen Schmerz, Mobilität und Fatigue. Die Auswahl der Items erfolgte einerseits auf Grundlage einer ausführlichen Literaturrecherche zur XLH, um das XLH-spezifische Erkrankungsbild und die Symptomatik spezifisch abzufragen. Andererseits wurden die Items so gewählt, dass eine Vergleichbarkeit mit der pivotalen Studie UX023-CL201, in der der Endpunkt „Schmerz und funktionelle Behinderung“ mit dem Messinstrument POSNA-PODCI erhoben wurde, gewährleistet war.

Für die Items der Selbstbericht-Version für Kinder ab einem Alter von 8 Jahren wurde zusätzlich auf eine gute Verständlichkeit geachtet. Aus den oben genannten drei Gründen, konnten nicht alle Items der jeweiligen Domänen, die in der Shortform der PROMIS-Entwickler aufgeführt sind, ausgewählt und abgefragt werden. Die Versionen zur Selbst- und Fremdeinschätzung waren bis auf eine zusätzliche Frage zur Schmerzintensität in der Selbstbericht-Version inhaltlich identisch. Insgesamt bestehen die Fragebögen aus 22 Items aufgeteilt in 10 Items Mobilität, 8 Items Fatigue und 4 Items Schmerz. Alle Items waren nach einem 5-Punkte-Likert-Skala Typ, mit einer Punkteskala von 0 bis 4, aufgebaut. Die Items der Domänen Schmerz und Fatigue umfassten die Antwortmöglichkeiten: nie, fast nie, manchmal, oft und fast immer. Die Antwortmöglichkeiten der Domäne Mobilität waren: ohne Schwierigkeiten, mit etwas Schwierigkeiten, mit einigen Schwierigkeiten, mit viel Schwierigkeiten und nicht in der Lage, es zu tun. Der Bezugszeitraum war bei allen Items jeweils auf die letzten 7 Tage definiert. Abweichend zu den 22 Items ist die Frage (23) zur Schmerzintensität in der Selbstbericht-Version zu betrachten. Diese Frage forderte das Kind auf, seine Schmerzen im Durchschnitt der letzten 7 Tage anhand einer 10-Punkte-Likert-Skala von 0= kein Schmerz bis 10= stärkster Schmerz, den man sich vorstellen kann, zu bewerten [32].

In der Studie UX023-CL301 wurde die Version 2 des PROMIS-Fragebogens verwendet, obwohl ursprünglich die Verwendung der Version 1 geplant war. Auf Anraten der PROMIS-Entwickler wurde die aktualisierte Version des Fragebogens (Version 2) verwendet. Dies wurde in einer späteren Novellierung des Protokolls dokumentiert. Der Unterschied von Version 2 zu Version 1 ist eine Ergänzung der Itembänke um weitere Items. Die Domäne Fatigue wurde von ursprünglich 23 auf 25 Items (je 10 Shortform-Items) und die Domäne Schmerz von 13 auf 20 Items (je 8 Shortform-Items) erweitert [37-39]. Die Domäne Mobilität dagegen wurde um ein weiteres Item auf 24 Items in Version 2 ergänzt. Wobei hier eine Reduktion der angegeben Shortform-Items von 10 auf 8 durchgeführt wurde [37, 40].

Anzumerken ist, dass ausschließlich die Version 2 des PROMIS in der UX023-CL301-Studie verwendet wurde. Demnach bestehen durch die genannten Versions-Modifikationen keine Auswirkungen auf die Untersuchung in der Studie UX023-CL301 [20, 32].

Es waren keine formalen Einweisungen für Patienten oder ihre Erziehungsberechtigten nötig, um den Fragebogen zu verwenden. Ein Benutzerhandbuch (*User/Scoring manual*) mit weiterführenden Informationen zum PROMIS-Verzeichnis, Anleitungen zur Scoringmethodik und statistischen Charakteristiken ist gegeben. Es liegt keine weitere Dokumentation und damit kein Hinweis in Bezug auf die Unterstützung der Mitarbeit von klinischem Personal für das Ausfüllen der Selbstbericht-Version oder die Verwendung der Eltern-Version bei Kindern im Alter ≥ 8 Jahren vor [20, 32]. Die Validierung des PROMIS-Fragebogens auf die pädiatrische XLH-Patientenpopulation erfolgte durch die Interviewstudie UX7575A, in der durch eine kognitive Nachbesprechung die Qualität der Fragen und Antwortmöglichkeiten aus Perspektive der Patienten und ihrer Erziehungsberechtigten ermittelt werden sollte. Inhaltlich umfassten die Interviews Erfahrungen, Symptome und funktionelle Einschränkungen der Patienten mit dem Krankheitsbild der XLH. Ferner wurde über das Verständnis der Fragen, der

Antwortmöglichkeiten und die Relevanz der Fragen in Bezug auf die XLH gesprochen. Die Ergebnisse der Studie deuten auf ein überwiegend sehr gutes Verständnis der Fragen sowie der Antwortmöglichkeiten in den drei Domänen hin, sodass sich die Operationalisierung der Interviews mit möglicher Hilfestellung durch das Klinikpersonal nicht auf die Antworten der Versuchspersonen ausgewirkt haben sollte. Als negativer Kritikpunkt wurde lediglich der für einige Patienten geringe Bezugszeitraum von 7 Tagen genannt. Hierbei wurde sich allerdings an der Empfehlung der FDA orientiert, welche Items, die den aktuellen Gesundheitszustand abfragen, präferieren. Außerdem konnte die Relevanz der Items bezüglich der Hauptsymptome der XLH von den Teilnehmern bestätigt werden, was auf eine valide Auswahl der Items hindeutet [41]. Ein weiterer Nachweis für die Validität der generierten Daten mittels des PROMIS-Fragebogens zeigt eine Validierungsstudie, die die grundsätzliche Eignung der Anwendung der Items aus dem pädiatrischen PROMIS-Verzeichnis bei Kindern untersucht. In dieser wurden 88 kognitive Interviews mit 77 Kindern im Alter von 8-17 Jahren zur Beurteilung der Items des pädiatrischen PROMIS-Verzeichnisses durchgeführt. Die Studie zeigte auch bei den 8-Jährigen gute Ergebnisse hinsichtlich der Verständlichkeit der Fragen, Antwortmöglichkeiten und des Betrachtungszeitraums. Bei auffälligen Verständnisproblemen konnten auf Grundlage der Studienergebnisse Formulierungsanpassungen vorgenommen werden [42]. Demnach ist die Validität der erstellten Fragebögen mithilfe des pädiatrischen PROMIS-Verzeichnisses gegeben.

Faces Pain Scale – Revised (FPS-R)

Um die Intensität des akuten Schmerzes der Kinder zu messen, wurde die überarbeitete Gesichtsschmerzskala (FPS-R) verwendet. Diese Skala ist ein selbstberichtendes Instrument, das speziell für Kinder entwickelt wurde. Diese wurde für die Verwendung bei Kindern im Alter von 5 bis 16 Jahren validiert [43]. Dabei wird Schmerzempfinden und somit die Schmerzintensität auf einer Skala von 0 bis 10 graphisch dargestellt, wobei Gesichter mit Werten von 0, 2, 4, 6, 8 und 10 verwendet werden (0 = Kein Schmerz; 10 = stärkster Schmerz). Die Schmerzintensität der Patienten wurde von Kindern ab 5 Jahren selbst berichtet [32]. In der Studie UX023-CL301 wurde das Kind aufgefordert, auf ein Gesicht zu zeigen, das am besten die Intensität des aktuellen Schmerzes abbildet [32].

Angewendet wurde die FPS-R bei allen Kindern, die zur Screeningvisite mindestens 5 Jahre alt waren (15 Kinder in der Burosumab-Gruppe und 20 Kinder in der Kontrollgruppe) [19-21]. Es wurde nicht wie im CAO-Manual [32] berichtet, nur Kinder im Alter zwischen 5 und 7 Jahren befragt. Der Fragebogen wurde - wie im Studienprotokoll und -bericht beschrieben – ebenfalls bei Kindern in der Altersgruppe 8- bis 12-Jährigen eingesetzt. Der FPS-R wurde zu Baseline, zu Woche 24, 40 und 64 oder zum vorzeitigen Abbruch abgefragt.

Eines der Hauptmerkmale der XLH sind die starken Knochen- und Gelenkschmerzen, welche sich im weiteren Verlauf der Erkrankung durch das geschädigte Skelett, die Knochendeformationen und den Gelenkverschleiß intensivieren. Zudem können zahnassoziierte Anomalien aufgrund von Defekten im Dentin Schmerzen verursachen [2, 3,

10]. Da der FPS-R die akute Schmerzintensität des Kindes abfragt und somit Auskunft über Verbesserung des Gesundheitszustandes gibt, handelt es sich hierbei um einen patientenrelevanten Endpunkt. Synergistisch mit der Domäne Schmerz (*Pain Interference*) des PROMIS-Fragebogens kann vollumfänglich die Kategorie Schmerz sowohl akut als auch chronisch in der Studie UX023-CL301 abgefragt werden.

Zahnmedizinische Untersuchungen

Zahnassoziierte Anomalien sind ein häufiges klinisches Bild der XLH. Durch den gestörten Mineral- und Knochenstoffwechsels kommt es zu Defekten im Dentin, die wiederum zu gestörter Zahnentwicklung mit Zahnwurzelabszessen, Hohlräumen und abnormaler Zahnschmelzbildung führen. Lichtmikroskopische Untersuchungen zeigen, dass außer endogener Zahnwurzelfehlbildungen auch an der Zahnwurzel Entzündungen durch das Eindringen von Mikroorganismen in Mikrospalten oder in pathologisch veränderten Mikrostrukturen des Zahnschmelzes entstehen können, ohne dass kariöse Läsionen sichtbar sind. Durch den Schmelzverlust und das Ausbleiben des Speichelflusses entsteht ein erhöhtes Risiko für Infektionen im Mundbereich [2, 3, 10, 44]

Dieser Endpunkt erfasst die zahnärztlichen Anomalien zu Baseline und nach Verabreichung der Studienmedikation, sodass eine Verbesserung des Gesundheitszustandes dargestellt werden kann. Daher handelt es sich bei der zahnärztlichen Untersuchung um einen patientenrelevanten Endpunkt.

In der Studie erfolgte eine deskriptive Darstellung der zahnärztlichen Untersuchungen. Dabei wurde die Anzahl der zahnmedizinischen Ereignisse Karies, verzögertes Zahnen, Zahnschmelzhypoplasie, Abszesse und Gingivitis, welche zu Baseline vorhanden waren und nach Verabreichung der Studienmedikation auftraten, berichtet. Hierzu wurde aktiv zu Baseline, Woche 4, 8, 16, 24, 32, 40, 52, 64 und Woche 74 das Kind während der Visite befragt, ob die oben genannten zahnmedizinischen Ereignisse eingetreten sind.

Lebensqualität (SF-10)

Die Erhebung der Lebensqualität ist als patientenrelevanter Endpunkt im Sinne der Verfahrensordnung (VerfO) anzusehen. Dies wurde auch durch den G-BA im Beratungsgespräch zu Burosumab bestätigt. Der G-BA forderte im Beratungsgespräch im Anwendungsgebiet mindestens einen generischen Lebensqualitätsfragebogen zu verwenden und zählte beispielsweise den validierten und anerkannten Short Form-36 (SF-36) auf. Dieses Messinstrument findet jedoch Anwendung bei erwachsenen Patienten und ist für eine pädiatrische Population weder geeignet noch validiert. Im Anwendungsgebiet von Burosumab handelt es sich um Kinder im Alter von 1 Jahr und Jugendlichen, so dass der SF-36 kein geeigneter Fragebogen für die Patientenpopulation von Burosumab ist. Der Gesundheitsfragebogen SF-10 ist ein von der Pflegekraft/Erziehungsberechtigten ausgefüllter

Fragebogen zur Beurteilung körperlicher und psychosozialer Funktionen kranker und gesunder Kinder im Alter von 5-18 Jahren. Aus verschiedenen Zulassungsverfahren geht hervor, dass die europäische Zulassungsbehörde aus den Ergebnissen des SF-10 ein positives Nutzen-Risiko-Verhältnis herleitet [45, 46]. Dieser Fragebogen leitet sich vom SF-36 ab und wurde speziell für Kinder entwickelt [47]. Die Items wurden aus dem *Child Health Questionnaire* (CHQ) adaptiert und bezogen sich auf die letzten 4 Wochen.

Hierbei werden zwei Komponenten generiert:

- Physische/körperliche Gesamtpunktzahl (PHS-10): Items 1, 2 (bestehend aus 2a und 2b durch die Abfrage der Einschränkungen von zwei unterschiedlichen Aktivitäten), 3 und 5
- Psychosoziale Gesamtpunktzahl (PSS-10): Items 4, 6 und 9

Die PHS-10 Komponente beinhaltet vier Skalen, welche die körperliche Funktion, die soziale Funktion aufgrund körperlicher Gesundheit, die generelle Gesundheit und die körperlichen Schmerzen abfragt. Die PSS-10 besteht ebenfalls aus vier Skalen. Dabei wird das soziale, emotionale Verhalten, das Selbstwertgefühl, die mentale Gesundheit und das allgemeine Verhalten abgefragt. Zusätzlich gibt es zwei weitere Skalen, die an die Eltern des betroffenen Kindes gerichtet sind. Hierbei wird die eigens zu gestaltende Zeit sowie die elterliche Belastung bewertet [4].

Die Skalenwerte wurden so standardisiert, dass eine Punktzahl von 50 dem Durchschnittswert einer Stichprobe von 2006 entspricht. Skalenwerte sind auf einen Mittelwert von 50 und eine Standardabweichung von 10 in der kombinierten US-Allgemeinbevölkerung aus gesunden, chronisch kranken Kindern und Kindern mit Behinderung standardisiert. Die Bewertung des SF-10 erfolgt nach der in Saris-Baglana et al. (2007) beschriebenen Methode. Höhere globale Punktzahlen sind mit einer besseren Lebensqualität verbunden [48].

Der Fragebogen wurde zu Baseline, zu Woche 40, 64, 88 und 160 oder bei vorzeitigem Abbruch ausgefüllt.

Die Lebensqualität von Kindern mit XLH ist durch die stark eingeschränkte Gehfähigkeit, die erheblichen Funktionsstörungen, und die einhergehenden starken Schmerzen deutlich vermindert. Erstmals wurde die Lebensqualität von XLH-Patienten kontrolliert in der UX023-CL301 erhoben. Durch die Burosumab-Behandlung konnte eine Verbesserung der Lebensqualität von XLH-Patienten beobachtet werden. Alleine der Wechsel einer alle 3-4-stündigen oralen Phosphat- und Vitamin-D-Einnahme auf eine subkutane Injektion im zweiwöchigen Abstand im häuslichen Umfeld wird von den betroffenen Kindern und den betreuenden Personen als eine enorme Verbesserung der Lebensqualität empfunden, da die Applikation einen normalen Tagesablauf ermöglicht. Weiterhin ist dies begründet durch eine Reduktion der Schmerzen und eine deutliche Verbesserung der Mobilität.

Aus diesem Grund ist die Erhebung der Lebensqualität mittels SF-10 ein patientenrelevanter Endpunkt.

Sicherheit

Ein unerwünschtes Ereignis (UE) ist definiert als jegliche unerwünschte medizinische Erscheinung bei einem Patienten, unabhängig davon, ob das Ereignis in kausalem Zusammenhang mit der Verabreichung der Studienmedikation steht.

Alle UEs wurden gemäß MedDRA (Version 18.1) von dem Zeitpunkt aufgezeichnet, an dem der Patient oder dessen Erziehungsberechtigter die Einwilligung unterzeichnet hat bis zum Zeitpunkt 12 Wochen nach der letzten Dosis des Studienmedikaments. Auch nach Amendment 1 des Studienprotokolls UX023-CL301 wurden die Kinder, die anderweitig (nicht über die Studie UX023-CL301) Burosumab erhielten, dennoch nach Auflagen der Zulassungsbehörde beobachtet. Dies betrifft Patienten, die nicht in die Extensionsphase aufgenommen werden konnten (Kinder aus Japan oder Korea). Bei jedem Besuch wurden die Patienten zu neuen oder laufenden unerwünschten Ereignissen seit dem letzten Besuch befragt. Die Beurteilung der unerwünschten Ereignisse erfolgte bei jedem Besuch des Studienzentrums. Der Prüfarzt beurteilt die Schwere, die Kausalität (Zusammenhang mit der Studienmedikation) und die Bedeutung jedes UEs.

Schwere der UEs

Der Prüfarzt schätzt die Schwere aller UEs gemäß der *National Cancer Institute Common Terminology Criteria for Adverse Events (NCCTCAE)* ein, die wie folgt definiert sind:

- Mild: Jedes UE, das normalerweise vorübergehend ist und nur minimale Behandlung oder therapeutisches Eingreifen benötigt. Das Ereignis beeinflusst normale tägliche Aktivitäten generell nicht.
- Moderat: Jedes UE, das normalerweise mit zusätzlicher spezifischer therapeutischer Intervention verringert wird. Das Ereignis beeinflusst normale tägliche Aktivitäten, verursacht Unwohlsein, aber stellt kein signifikantes oder dauerhaftes Risiko für die Studienteilnehmer dar.
- Schwerwiegend (Stufe 3): Jedes UE, das normale tägliche Aktivitäten unterbricht oder das den klinischen Zustand signifikant beeinflusst oder intensives therapeutisches Eingreifen notwendig macht.
- Lebensbedrohlich (Grad 4): Ereignisse, bei denen der Teilnehmer ein Risiko zu versterben hat.
- Tod (Grad 5): Ereignisse, die zum Tod führen.

Kausalität der UEs

Der Prüfarzt bewertete die potenzielle Beziehung der UEs der Studie in Abhängigkeit des Auftretens durch die Medikation. Hierzu dienen folgende Kategorien:

- Definitiv nicht zusammenhängend: Diese Kategorie bezieht sich auf UEs, die eindeutig nicht mit der Studienmedikation/dem Verfahren zusammenhängen und bei denen eine andere plausible Ursache für das Ereignis zu sehen ist und/oder eine klinisch plausible zeitliche Abfolge mit dem Beginn des Ereignisses und der Exposition gegenüber der Studienmedikation inkonsistent ist und/oder eine kausale Beziehung als biologisch unplausibel angesehen werden kann.
- Wahrscheinlich nicht zusammenhängend: Diese Kategorie gilt für UEs, die wahrscheinlich durch andere Ursachen (z. B. Begleitmedikation(en), Begleiterkrankung(en)) aufgetreten

sind. Das heißt, eine alternative Erklärung war wahrscheinlicher oder die zeitliche Beziehung legen nahe, dass ein kausaler Zusammenhang besteht.

- Eventuell zusammenhängend: Diese Kategorie bezieht sich auf UEs, die möglicherweise mit der Studienmedikation/dem Verfahren zusammenhängen könnten. Der zeitliche Zusammenhang des klinischen Ereignisses mit der Verabreichung der Studienmedikation und der unerwartet auftretenden Reaktion könnte leicht durch andere Faktoren verursacht worden sein.
- Wahrscheinlich zusammenhängend: Der zeitliche Zusammenhang des klinischen Ereignisses mit der Verabreichung der Studienmedikation macht einen kausalen Zusammenhang denkbar und nicht unwahrscheinlich. Andere Medikamente, therapeutische Interventionen und Basiszustände liefern keine ausreichende Erklärung für das beobachtete Ereignis.
- Definitiv zusammenhängend: Begründeter, zeitlicher Zusammenhang des klinischen Ereignisses mit der Verabreichung der Studienmedikation UND keine sinnvolle Erklärung des Ereignisses durch andere Faktoren (z. B. der klinische Zustand der Person, Begleittherapie und/oder andere Interventionen).
- Therapie-assoziierte UEs waren Ereignisse, die eventuell, wahrscheinlich oder definitiv zusammenhängend kategorisiert wurden.

Schwerwiegendes unerwünschtes Ereignis (SUE)

Ein schwerwiegendes unerwünschtes Ereignis ist jegliches UE, das folgendes ist oder zu folgendem führt:

- Tod
- unmittelbar lebensbedrohend
- benötigte stationäre Hospitalisierung oder Verlängerung einer bereits vorhandenen Hospitalisierung
- persistierende oder signifikante Unfähigkeit oder wesentliche Störung der Fähigkeit normale Lebensfunktionen zu erfüllen
- kongenitale Anomalie/Geburtsfehler

Ein wichtiges medizinisches Ereignis, das – basierend auf angemessener medizinischer Einschätzung – den Patienten gefährdet und medizinisches oder chirurgisches Eingreifen benötigen könnte, um eines der oben gelisteten Ereignisse zu verhindern.

Üblicherweise erfolgt in der frühen Nutzenbewertung die Bewertung eines Arzneimittels bezüglich seiner Nebenwirkungen über die unerwünschten Ereignisse, die innerhalb der relevanten Studien im zeitlichen Zusammenhang mit der Anwendung der Studienmedikation aufgetreten sind. Diesem Vorgehen wird auch in der vorliegenden Nutzenbewertung gefolgt. Generell ist hierbei zu beachten, dass der zeitliche Zusammenhang zwischen Anwendung der Studienmedikation und Auftreten eines UE nicht zwangsläufig mit einem kausalen Zusammenhang zwischen beiden gleichzusetzen ist.

Entsprechend den Vorgaben des G-BA werden im vorliegenden Nutzendossier UEs und SUEs sowohl als Gesamtraten als auch aufgeteilt nach Systemorganklasse (SOC) bzw. *Preferred*

Term (PT) dargestellt. Da durch den Endpunkt Sicherheit eine Verringerung von Nebenwirkungen beurteilt werden kann, handelt es sich um einen patientenrelevanten Endpunkt.

Expositionszeiten

Behandlungsphase (Woche 0-64)

Burosumab-Gruppe

Alle Kinder der Burosumab-Gruppe absolvierten die 64 Wochen der Behandlungsphase und wurden vollständig in die Analysen einbezogen. Die durchschnittliche Behandlungsdauer der Exposition gegenüber Burosumab betrug 15,13 Monate. Die Dauer der Exposition lag zwischen 14,7 und 15,3 Monaten. Die Gesamtdauer der Exposition gegenüber Burosumab während der 64-wöchigen Behandlungsphase betrug 36,56 Patientenjahre. Während der ersten 64 Wochen der Studie erhielten alle Kinder, die in die Burosumab-Gruppe randomisiert wurden, alle geplanten Dosen von Burosumab, mit Ausnahme von 3 der 29 (10 %) Kinder, die jeweils eine der 32 (3 %) geplanten Dosen verpassten. Die meisten Kinder in der Burosumab-Gruppe (21/29) erhielten die zugewiesene Dosis von 0,8 mg/kg, 8 Kinder hatten dabei eine Dosiserhöhung von Burosumab auf 1,2 mg/kg basierend auf protokolldefinierten Kriterien, die mit dieser Dosis bis Woche 64 konstant weiterbehandelt wurden [20].

Aktive Kontrollgruppe (Substitutionstherapie)

Alle Kinder in der aktiven Kontrollgruppe absolvierten die 64 Wochen Behandlungsphase und gingen somit vollständig in die Analyse mit ein. Die mittlere Expositionsdauer der aktiven Kontrollgruppe betrug 15,11 Monate. Die verabreichte orale Phosphatdosis der Substitutionstherapie lag durchschnittlich innerhalb der empfohlenen Werte (20-60mg/kg/Tag verteilt auf 3-5 Dosen/Tag) [11]. 10 der 32 Kinder hatten keine Dosisanpassung, 3 Kinder eine Dosisreduktion und 19 Kinder eine Dosiserhöhung während der Studiendauer. Von den 19 Kindern mit Dosiserhöhungen hatten 13 eine Dosiserhöhung von über 20 % zu Baseline bei einer meist viermal täglichen Einnahme.

Die durchschnittlichen täglichen Dosen von aktivem Vitamin D lagen im Allgemeinen im empfohlenen Bereich (Alfacalcidol 40 bis 60 ng/kg/Tag oder Calcitriol 20 bis 30 ng/kg/Tag unterteilt in zwei oder drei Dosen/Tag) [10, 11]. 22 der 32 Kinder erhielten Calcitriol. 9 dieser Kinder hatten keine Dosisanpassung, 2 hatten eine Dosisreduktion und 11 bekamen eine Dosiserhöhung. 9 der 32 Kinder erhielten Alfacalcidol. 4 Kinder hatten während der Studie keine Dosisanpassung, 1 Kind eine Dosisreduktion und 4 eine Erhöhung der Gesamtdosis. Die 13 der 15 Kinder mit einer Vitamin-D-Dosiserhöhung hatten eine Dosiserhöhung von über 20 % zu Baseline bei einer zweimal täglichen Einnahme [20].

Compliance während der Behandlungsphase

Bezogen auf die Gesamtzahl der geplanten Dosierungstage während der Behandlungsphase (448 Tage) betrug die Compliance der Patienten in der aktiven Kontrollgruppe (N=32) > 95 %. 19 der Kinder (60 %) waren zu 100 % compliant zu dem verschriebenen Dosierungsregime der Substitutionstherapie. 5 Kinder verpassten 1 Tag, 5 Kinder verpassten 2 Tage, 1 Kind verpasste 5 Tage, 1 Kind verpasste 6 Tage und 1 Kind verpasste 7 Tage (nicht aufeinanderfolgend) der Medikation mit oralem Phosphat und/oder aktivem Vitamin D.

Die maximale Zeitspanne der ausgesetzten Medikation eines Patienten (7 Tage) entspricht 1,6 % (7/448) der geplanten Dosiertage. Die Hauptgründe für das Auslassen der Medikation waren Nebenwirkungen (wie leichtes Erbrechen, leichte Nasopharyngitis, Ohrenentzündung, akute virale Gastritis und Gastroenteritis), kein Zugang zur Medikation und Non-Compliance.

Extensionsphase (Woche 66 – 140)

Die Patienten in der EU und in den Ländern Kanada, USA und Australien konnten freiwillig an der Extensionsphase (ab Woche 66 bis maximal Woche 140) der Studie UX023-CL301 teilnehmen. In dieser Extensionsphase wurden alle Patienten mit Burosumab behandelt.

Aktive Kontrollgruppe → Extensionsphase Burosumab:

Vor Beginn der Extensionsphase erfolgte für die Kontrollgruppe eine Auswaschphase von 14 Tagen. Im Anschluss (ab Woche 66) erhielten diese Patienten die Studienmedikation Burosumab. Insgesamt konnten aus den oben genannten Ländern maximal 27 Patienten (32 Kinder der aktiven Kontrollgruppe abzüglich 5 Kinder aus Japan/Korea) in die Extensionsphase aufgenommen werden. An der Extensionsphase nahmen 26 Kinder aus der aktiven Kontrollgruppe teil, da ein Kind kommerzielles Burosumab erhielt. Von diesen 26 Kindern vollendeten 4 die Datenerhebung in Woche 64, erhielten jedoch danach nicht ihre erste Behandlung mit Burosumab.

Bei den 22 Kinder, die mit der Extensionsphase ihre erste Burosumab-Behandlung erhielten, betrug die Exposition gegenüber Burosumab 2,60 Monate. Alle Kinder erhielten alle geplanten Injektionen von Burosumab.

Burosumab-Gruppe → Extensionsgruppe Burosumab:

Aus der Burosumab-Gruppe konnten 27 Patienten (29 Kinder aus der Burosumab-Gruppe abzüglich 2 Kinder aus Japan/Korea) an der Extensionsphase teilnehmen. 2 Kinder entschieden sich dazu nicht an der Extensionsphase teilzunehmen und kommerzielles Burosumab zu erhalten. Somit wurden 25 Kinder aus der Burosumab-Gruppe in die Extensionsphase ausgenommen.

Die durchschnittliche gesamte Expositionsdauer gegenüber Burosumab in der Behandlungsphase und Extensionsphase betrug somit 17,50 Monate und lag zwischen 15,0 und 21,3 Monate bei 32 – 93 Injektionen Burosumab. Nach Woche 64 gab es keine Dosisanpassungen [20].

Verwendete statistische Methode

Auswertungspopulationen Full Analysis Set (FAS)

Die FAS-Population war definiert als alle randomisierten Patienten, die eine Injektion der Studienmedikation erhalten hatten. Die Zuordnung zu den Behandlungsgruppen bei Auswertungen innerhalb der FAS-Population folgte der Zuordnung der Patienten während der Randomisierung. Die FAS-Population wurde für die Analyse sämtlicher Wirksamkeitsendpunkte herangezogen.

Auswertungsmethoden

Die Ergebnisse der relevanten Studie werden durch adäquate Effektschätzer mit zugehörigem 95%-KI und p-Werten geeigneter statistischer Tests zu einem Signifikanzniveau von 5% dargestellt.

Obwohl sich durch die Vielzahl an *post hoc* durchgeführten statistischen Tests ein multiples Testproblem ergibt, wurde keine Adjustierung des Signifikanzniveaus vorgenommen. Bei den im Dossier aufgeführten Endpunkten handelt es sich sowohl um binäre als auch um kontinuierliche Endpunkte. Diese wurden nach der prä-definierten Berechnungssystematik wie im *Statistical Analysis Plan* der Studie UX023-CL301 dargestellt ausgewertet [21].

Umgang mit fehlenden Werten

In der Studie UX023-CL301 wurden im Allgemeinen fehlende Daten - sofern nicht anders angegeben - als fehlend behandelt. Wenn eine Änderung gegenüber der Baseline vorgenommen wurde, werden nur Patienten mit einer Baseline und mindestens einer Post-Baseline-Messung in die Analyse einbezogen. Für wiederholende Messungen werden bei der *generalized estimated equation* (GEE)-Analyse die Modellparameter gleichzeitig unter Verwendung aller beobachteten Daten geschätzt. Eine Sensitivitätsanalyse kann unter Verwendung von gewichteter GEE durchgeführt werden, indem die inverse wahrscheinlichkeitsgewichtete Methode implementiert wird, um Studienabbrecher („dropouts“) unter der fehlenden zufälligen (MAR) Annahme zu berücksichtigen [20, 21].

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁴ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁵ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität⁶ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

⁴ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

⁵ Deeks JJ, Higgins JPT, Altman DG. *Analysing data and undertaking meta-analyses*. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

⁶ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. *Measuring inconsistency in meta-analyses*. *BMJ* 2003;327(7414):557-560.

Nicht zutreffend.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

In der Studie UX023-CL301 wurde das *Full-Analysis-Set* (FAS) zur Auswertung herangezogen. Dieses ist definiert als alle Patienten, die randomisiert wurden und mindestens eine Studienmedikation erhalten haben ($N_{\text{(Burosumab)}} = 29$; $N_{\text{(aktive Kontrolle)}} = 32$). Da alle randomisierten Patienten mindestens eine Dosis der geplanten Studienmedikation erhalten haben, ist die FAS in diesem Fall eine adäquate Umsetzung des ITT-Prinzips.

Lediglich die Endpunkte, bei denen ein Mindestalter erforderlich war, bilden eine Ausnahme. Der 6MWT, der PROMIS-, der FPS-R- und der SF-10-Fragebogen konnten nur bei Kindern ab einem Alter von 5 Jahren erhoben werden. Für diese Endpunkte wurde das FAS zur Auswertung herangezogen, welches definiert ist als alle FAS-Patienten, die zu Baseline ein Alter von mindestens 5 Jahren aufwiesen ($N_{\text{(Burosumab)}} = 15$; $N_{\text{(aktive Kontrolle)}} = 20$).

In der Studie UX023-CL301 wurden im Allgemeinen fehlende Daten - sofern nicht anders angegeben - als fehlend behandelt. Sensitivitätsanalysen wurden für die Ergebnisse der Primäranalyse der Endpunkte RSS und RGI-C durchgeführt. Diese beinhalteten [20]:

- *Last Observation Carried Forward* (LOCF): Die letzte nicht fehlende Post-Baseline wurde zur Auswertung für den entsprechenden Endpunkt herangezogen. Wenn keine Post-Baseline verfügbar war, wurde die Baseline für die Bewertung verwendet. Wenn keine Baseline verfügbar war, wurde die LOCF-Methode nicht durchgeführt.
- *Multiple imputation* (MI): Das PROC MI-Verfahren in SAS wurde verwendet, um fehlende Werte mithilfe eines *Markov Chain Monte Carlo* (MCMC)-Ansatzes 100 Mal zu berechnen. Der kalkulatorische Datensatz wird mit dem ANCOVA-Modell zu Woche 40 oder dem GEE-Modell zu Woche 64 analysiert. Die Ergebnisse werden mit

dem PROC MIANALYZE-Verfahren in SAS kombiniert. Der Baseline RSS Total Score, die Behandlungsgruppe und der Baseline-Alter-Stratifizierungsfaktor wurden als Einflussvariablen in der Imputation verwendet. Andere Variablen wie beispielsweise die Veränderung der alkalischen Phosphatase (ALP) zu Baseline wurden als Einflussvariable berücksichtigt.

Sensitivitätsanalysen

In der Studie UX023-CL301 wurde zur Überprüfung der Robustheit der Endpunkte RGI-C, RSS und 6MWT eine Sensitivitätsanalyse durchgeführt.

Für den Endpunkt 6MWT wurde eine Sensitivitätsanalyse mit Patienten durchgeführt, die den Gehetest zu jedem Messzeitpunkt abgeschlossen haben.

Darüber hinaus wurde für die 40-Wochen-Analyse des Endpunktes RGI-C eine Sensitivitätsanalyse mit Patienten durchgeführt, bei der alle Kinder mit orthopädischem Eingriff (vor der 40-Wochen-Visite) ausgeschlossen wurden.

Bei der 64-Wochen-Sensitivitätsanalyse wurden bei Patienten, die sich erstmalig zwischen Woche 40 und 64 einem orthopädischen Eingriff unterziehen mussten, die RGI-C-Daten der Woche 40 zur Berechnung der Sensitivität verwendet.

Eine weitere Sensitivitätsanalyse wurde durchgeführt, um die Robustheit der Analyseergebnisse gegenüber fehlender Daten zu bewerten. Von einem Kind fehlten die Knie- und Gesamtwerte des RSS zu Woche 40. Der RSS-Kniewert zu Woche 40 wurde nicht bewertet, da die Oberschenkelwachstumsplatten bereits verschmolzen waren. Der Baseline-Kniewert wurde imputiert, um den Kniewert der Woche 40 zu berechnen und somit den Gesamtwert der Woche 40 abzuleiten. Die Ergebnisse dieser LOCF-Analyse waren denen der Primäranalyse sehr ähnlich. Alle Röntgenbildaufnahmen wurden für die RGI-C-Bewertungen in Woche 64 einbezogen.

Die Ergebnisse der Sensitivitätsanalysen des Endpunktes Rachitissymptomatik zeigten konsistente Ergebnisse zur Hauptanalyse, weshalb auf die Darstellung verzichtet wurde. Dabei waren die Patientenzahlen der Analysen der Endpunkte Rachitissymptomatik identisch zur Gesamtpopulation der Hauptanalyse. Die 6MWT-Sensitivitätsanalyse wird im Abschnitt 4.3.1.3.1.3 dargestellt und diskutiert. Hier wurde die Sensitivitätsanalyse exklusive zwei Patienten in der Burosumab-Gruppe durchgeführt, da diese 1-mal zu Baseline und 1-mal zu Woche 40 den 6MWT nicht abgeschlossen hatten (13 Patienten in der Burosumab-Gruppe; 20 Patienten in der aktiven Kontrollgruppe).

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Subgruppenanalyse in der Studie UX023-CL301

Gemäß den Anforderungen des Nutzendossiers sind zur Bestimmung möglicher Effektmodifikatoren des medizinischen Zusatznutzens a priori präspezifizierte Subgruppenanalysen, zu berichten.

Mehrstufiges Vorgehen der Interpretation der Subgruppenanalyse

Erster Schritt zur Überprüfung ist der statistische Nachweis unterschiedlicher Effekte zwischen verschiedenen Subgruppen mithilfe eines Interaktions- oder Homogenitätstests. Damit das

Ergebnis der Subgruppenanalyse sinnvoll interpretierbar ist, muss einerseits eine ausreichende Anzahl an Patienten in den Subgruppen eingeschlossen und andererseits genügend Ereignisse aufgetreten sein. In Anlehnung an das IQWiG-Methodenpapier 5.0 wird der Interaktionstest daher nur dann interpretiert, wenn folgende Kriterien erfüllt sind [49]:

Tabelle 4-6: Kriterien für die Berechnung der Interaktionstests in Anlehnung an das Methodenpapier 5.0 des IQWiG

	Burosumab		Aktive Kontrolle	
	Anzahl Patienten	Anzahl Events	Anzahl Patienten	Anzahl Events
Kategorie 1	n ₁	e ₁	n ₂	e ₂
Kategorie 2	n ₃	e ₃	n ₄	e ₄
e: Anzahl von Ereignissen n: Anzahl von Patienten				

- Es müssen mindestens 10 Patienten in jeder Kategorie enthalten sein: $n_1 + n_2 \geq 10$ und $n_3 + n_4 \geq 10$.
- Zusätzlich gilt bei binären Subgruppen, dass in mindestens einer Kategorie mindestens 10 Ereignisse enthalten sein müssen: $e_1 + e_2 \geq 10$ oder $e_3 + e_4 \geq 10$.

Die Ergebnisse der Interaktionstests werden für alle relevanten Endpunkte im Abschnitt 4.3.1.3.2.1 und getrennt nach Endpunkten tabellarisch im Abschnitt 4.3.1.3.2.2 dargestellt. Für Subgruppenanalysen, bei denen der p-Wert des Interaktionstests $\geq 0,05$ ist, wird von einem homogenen Effekt über alle Subgruppen ausgegangen. In diesem Fall ist ausschließlich das Ergebnis der Gesamtstudienpopulation relevant, auf eine Darstellung der Ergebnisse der Subgruppenanalyse im Dossier wird verzichtet. Zur formalen Vollständigkeit und um eine vollständige Nutzenbewertung bei abweichender Entscheidung des G-BA zu ermöglichen, wurden die Ergebnisse dieser Subgruppenanalysen (auch der *post hoc* Analysen) dem Dossier dennoch als beigelegt.

Ein zum Niveau $\alpha=0,05$ signifikantes Ergebnis im Interaktionstest wird zunächst als statistischer Beleg für einen Unterschied zwischen den Subgruppen gewertet, jedoch ist zu beachten, dass - wie oben beschrieben - mindestens 10 Patienten in jedem Behandlungsarm enthalten sein müssen. Bei folgenden Endpunkten ist dieses Kriterium im jeweiligen Behandlungsarm nicht gegeben [Patientenzahl (N) < 10]:

- Region:
 - Japan: $N_{\text{(Burosumab)}} = 2$; $N_{\text{(aktive Kontrolle)}} = 3$
 - Rest der Welt: $N_{\text{(Burosumab)}} = 29$; $N_{\text{(aktive Kontrolle)}} = 27$
- PHEX-Mutation:
 - Positiv: $N_{\text{(Burosumab)}} = 27$; $N_{\text{(aktive Kontrolle)}} = 28$
 - Nicht-Positiv: $N_{\text{(Burosumab)}} = 2$; $N_{\text{(aktive Kontrolle)}} = 4$

Bei diesen Subgruppenanalysen, auch bei signifikantem Ergebnis im Interaktionstest, kann durch die Nicht-Einhaltung der Mindestanforderungen des Methodenpapiers 5.0 keine valide Interpretation erfolgen, weshalb hier nur das Ergebnis der Gesamtstudienpopulation zu berücksichtigen ist.

Für die Morbiditätsendpunkte Rachitissymptomatik (RGI-C, RSS) und Veränderung des Wachstums wurden folgende Subgruppenanalysen a priori definiert:

- Krankheitsschweregrad: Baseline RSS Total Score ($\leq 2,5$: Geringe Rachitis-Schwere vs. $> 2,5$: Hohe Rachitis-Schwere)
- Alter (< 5 Jahre vs. ≥ 5 Jahre)
- Geschlecht (weiblich vs. männlich)
- *PHEX*-Mutation (positive vs. nicht-positiv)
- Region (Japan vs. Rest der Welt)

Post hoc Subgruppenanalysen wurden für die Endpunkte PROMIS, FPS-R, SF-10 und 6MWT – sofern nicht vollständig im *Clinical Study Report* der UX023-CL301-Studie [20] vorhanden – nachträglich durchgeführt, um die Vorgaben der Dossievorlage zu erfüllen und eine einheitliche Analyse aller Endpunkte transparent darzustellen. In der nachfolgenden

Tabelle 4-7 erfolgt eine Darstellung, welche Subgruppenanalysen nachträglich von Kyowa Kirin durchgeführt wurden.

Tabelle 4-7: Übersicht der für die Endpunkte der Studie UX023-CL301 durchgeführten Subgruppenanalysen

UX023-CL301								
Endpunkt-Kategorie	Morbidität							Sicherheit
Subgruppen	Rachitis-symptomatik	Wachstum	6MWT	Serum-phosphat	PROMIS	FPS-R	SF-10	
Rachitis-Schwere (≤ 2,5 vs. > 2,5)	x	x	x	x	x	x	o	x
Geschlecht (weiblich vs. männlich)	x	x	x	o	o	o	o	x
Alter (< 5 Jahre vs. ≥ 5 Jahre)	x	x	o	x	o	o	o	x
Region (Japan vs. Rest der Welt)	x	x	o	x	o	o	o	x
<i>PHEX</i> -Mutation (positiv vs. nicht-positiv)	x	x	o	o	o	o	o	x

6MWT: 6-Minute Walk Test; PROMIS: Patient Reported Outcomes Measurement Information System; FPS-R: Face Pain Scale - Revised; SF-10: Short form - 10
x: Subgruppe a priori definiert
o: Subgruppe nicht a priori definiert. Subgruppenanalysen wurden *post hoc* durchgeführt.

Darstellung der Subgruppenanalysen im Dossier

Zunächst wurden mittels Interaktionstests mögliche Effektunterschiede zwischen den Kategorien einer Subgruppe bewertet. Diese sind für jeden Endpunkt gesammelt in einer Übersichtstabelle im Abschnitt 4.3.1.3.2.1 dargestellt. Eine bedeutsame Heterogenität wurde ab einem Interaktions-p-Wert < 0,05 angenommen und als Beleg für Effektmodifikation gewertet. Lag ein Beleg für eine Interaktion in einem Endpunkt vor, wurde die Subgruppenanalyse für die entsprechende Population in Abschnitt 4.3.1.3.2.2 dargestellt und diskutiert. Die präspezifizierten Subgruppen der jeweiligen Endpunkte werden trotz negativem Interaktionstest in Abschnitt 4.3.1.3.2.3 dargestellt.

Bei der Durchführung einer großen Anzahl statistischer Tests ist zu beachten, dass eine gewisse Anzahl der durchgeführten Tests aufgrund der vorgegebenen Irrtumswahrscheinlichkeit zu falsch-positiven Ergebnissen führen. Ein positiver Interaktionstest ($p < 0,05$) wird hier daher nur als Ausgangspunkt zur Überprüfung einer möglichen Effektmodifikation herangezogen. Liegt ein positiver Interaktionstest vor, werden daher zusätzlich folgende Kriterien geprüft.

- Besteht eine medizinisch, biologisch oder physiologisch plausible Erklärung für die mögliche Effektmodifikation?
- Liegen konsistente Ergebnisse zur möglichen Effektmodifikation über mehrere Endpunkte vor?

Außerdem ist zu beachten, dass aufgrund des von der EMA erteilten *Orphan-Drug*-Status und der damit im Zusammenhang stehenden geringen Anzahl der Studienteilnehmer in der Studie eine „geringe Power“ besteht. Dadurch können möglicherweise tatsächlich vorliegende Effekte nicht mit signifikanten Ergebnissen belegt werden. Außerdem handelt es sich bei einigen Subgruppen um teilweise *post hoc* durchgeführte Analysen, die nicht methodisch korrekt nach dem im Methodenpapier des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) einer Hypothesenprüfung unterzogen werden konnten.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethoden dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁷. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹.

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in

⁷ Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

⁸ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

⁹ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹⁰ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17(3): 279-301.

¹¹ B. Schöttker, D. Lüthmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

Frage kommen, vorliegen. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Nicht zutreffend.

¹² Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*
- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-8: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
UX023-CL301 ^a	ja	ja	abgeschlossen	64 Wochen ^b	<ul style="list-style-type: none"> • <u>Intervention:</u> Burosumab • <u>Aktive Kontrolle:</u> Orales Phosphat und aktives Vitamin D
<p>^a: Die primäre Wirksamkeits- und Sicherheits-Analyse nach 40 Wochen ist abgeschlossen und wurde im ersten Verfahren (D-349) mit der schriftlichen Stellungnahme präsentiert. Die sich für die Zusatzanalyse daran anschließenden aktiv kontrollierten weiteren 24 Wochen sind ebenfalls abgeschlossen und werden im vorliegenden Dossier dargestellt.</p> <p>^b: Die Angabe bezieht sich auf die offene, randomisierte und aktiv kontrollierte Phase, welche aus insgesamt 64 Wochen besteht. Die sich daran anschließende offene Extensionsphase konnte bis Woche 140 erweitert werden.</p>					

Geben Sie an, welchen Stand die Information in Tabelle 4-8 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Aktualisierung der Angaben des Dossiers erfolgte am 10.07.2019.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-8 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-9: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
Nicht zutreffend.	

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie

viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

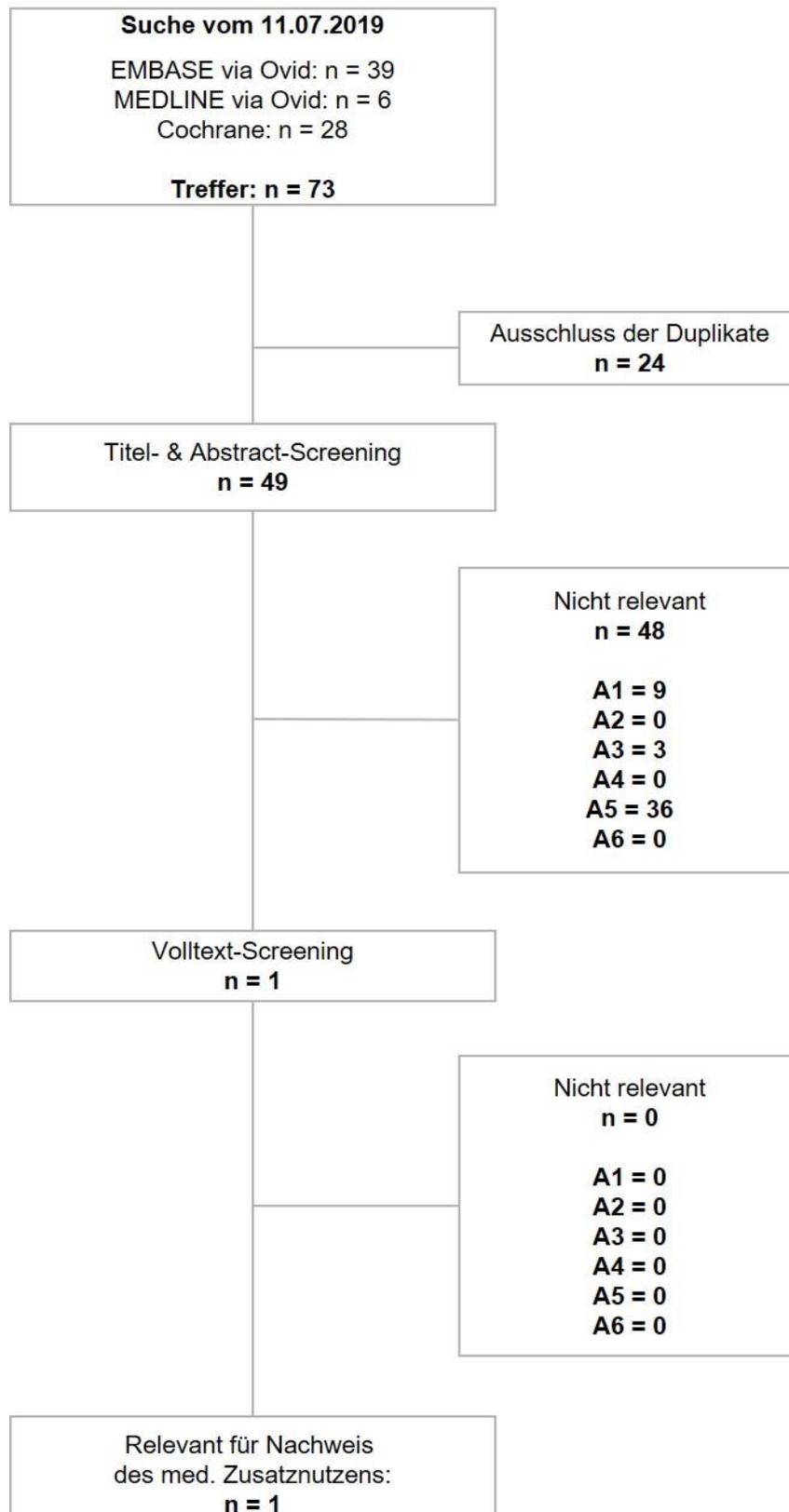


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die Recherche in bibliographischen Datenbanken wurde am 11.07.2019 durchgeführt. Insgesamt ergab diese Recherche 73 Treffer von denen 24 als Duplikate ausgeschlossen wurden. Die verbleibenden 49 Treffer wurden anhand von *Titel* und *Abstract* mit den in 4.2.3.4 genannten Kriterien bewertet. 48 konnten so ausgeschlossen werden. Eine Publikation wurde als relevant für die Neubewertung des medizinischen Zusatznutzens von Burosumab aufgrund neuer medizinischer Evidenz erachtet und eingeschlossen. Diese bezieht sich auf die pivotale Phase-III-Studie UX023-CL301 mit Datenschnitt zu Woche 64 und kann somit als Evidenz für das vorliegende Anwendungsgebiet dienen [50].

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-8) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-10: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
UX023-CL301 NCT02915705	Clinicaltrials.gov NCT02915705 [51] EU Clinical Trials Registry [52] WHO ICTRP 2016-000600-29 [53] Pharmnet.bund [54] EMA <i>clinical data</i> ^b	ja	ja [50]	abgeschlossen
^a Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. ^b Die Suche unter EMA <i>clinical data</i> am 16.07.2019 ergab keine Treffer.				

Geben Sie an, welchen Stand die Information in Tabelle 4-10 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information: 16.07.2019

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie

eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
Nicht zutreffend.						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
UX023-CL301	ja	ja	nein	ja [20]	ja [51-54]	ja [50]
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.1 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
UX023-CL301	Randomisierte, aktiv kontrollierte, offene und multizentrische Phase-III-Studie	XLH-Patienten im Alter von 1-12 Jahren	Burosumab subkutan alle 2 Wochen Initialdosis	64 Wochen gefolgt von einer freiwilligen Extensionsphase bis zu Woche 140	Australien, Kanada, Dänemark, Italien, Japan, Korea, Spanien, Schweden, United Kingdom, United States, Deutschland und Irland <u>Studienbeginn</u> 09/2016 <u>Ende Primär-Analyse:</u> 02/2018 <u>Voraus-sichtliches Studienende:</u> 06/2019	<p>Primärer Endpunkt: Evaluation der Wirksamkeit von Burosumab auf die Rachitis-Schwere bei Kindern mit XLH anhand des RGI-C Total Scores</p> <p>Sekundäre Endpunkte:</p> <ul style="list-style-type: none"> • Biochemische Parameter (Serumphosphat-Konzentration, 1,25(OH)₂-Vitamin D₃, ALP, TmP/GFR • 6MWT • Körpergröße • Schmerzerhebung: PROMIS Schmerz, Mobilität und Fatigue und FPS-R <p>Explorative Endpunkte: Kinderspezifischer Gesundheitsfragebogen <i>Short Form 10</i> (SF-10)</p>

Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Aktive Kontrolle (orales Phosphat und aktives Vitamin D)	Burosumab	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
UX023-CL301	<p>Behandlungsphase (Woche 0 – 64): Mehrere tägliche Dosen von oralem Phosphat und aktivem Vitamin D. Aufgrund der Variabilität der Dosen und Einnahmezeiten erfolgt eine Patientenindividuelle Gabe nach Ermessen des Arztes. Zu Verfügung gestellte Leitlinien: EU-Leitlinie Orales Phosphat 45 – 70 mg/KG/Tag in 3–4 geteilten täglichen Dosen und Alfacalcidol 1 – 2 µg/Tag 1 x täglich US-Leitlinie Orales Phosphat 20 – 40 mg/KG/Tag in 2–5 geteilten täglichen Dosen und Calcitriol 20 – 30 ng/KG/Tag in 2–3 geteilten täglichen Dosen [10, 11] Die Behandlung mit Alfacalcidol oder Calcitriol erfolgt basierend auf klinischen Parametern</p>	<p>Behandlungsphase (Woche 0 – 64): Initialdosis Burosumab von 0,8 mg/kg KG alle 2 Wochen subkutan Die Dosis kann auf 1,2 mg/kg KG basierend auf den Nüchtern-Serumphosphatwert erhöht werden.</p>	<p>Alle Studienteilnehmer erhielten vor der Studienteilnahme eine Substitutionstherapie. Die am häufigsten verwendeten Phosphatpräparate waren zu 34 % Natriumphosphat, 23 % Kaliumphosphat, 20 % elementares Phosphat und 15 % Natriumkaliumphosphat. Das aktive Vitamin D wurde in Form von Calcitriol zu 82 % und in Form von Alfacalcidol zu 21 % verwendet. Studienteilnehmer, die der Burosumab-Gruppe zugeteilt wurden, durften kein orales Phosphat und aktives Vitamin D während der Behandlungsphase einnehmen. Verboten wurden folgende Medikamente für beide Behandlungsarme: Aluminiumhydroxy-Antazida, systemisches Kortison, Acetazolamid, Thiazide, Wachstumshormone, PTH Suppressoren, monoklonale Antikörper-Therapien</p>
<p>KG: Körpergewicht; kg: Kilogramm; mg: Milligramm; ng: Nanogramm; EU: Europäische Union; US: United States; PTH: Parathormon</p>			

Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie	UX023-CL301	
	Aktive Kontrolle (N=32)	Burosumab (N=29)
Gruppe		
N (FAS)		
Allgemeine Charakteristika		
Alter, MW (SD)	6,34 (3,24)	5,83 (3,43)
Weibliches Geschlecht, n (%)	18 (56,3)	16 (55,2)
Männliches Geschlecht, n (%)	14 (43,8)	13 (44,8)
Rasse Asiaten, n (%)	6 (18,8)	2 (6,9)
Rasse Weiße, n (%)	25 (78,1)	25 (86,2)
Rasse Andere, n (%)	1 (3,1)	2 (6,9)
Ethnizität hispanisch/lateinamerikanisch; n (%)	3 (9,4)	3 (10,3)
Ethnizität nicht-hispanisch/lateinamerikanisch; n (%)	29 (90,6)	26 (89,7)
Allgemeine Baseline-Charakteristika		
Gewicht [kg] MW (SD)	21,55 (8,910)	19,59 (8,984)
BMI [kg/m ²] MW (SD)	18,22 (2,157)	18,00 (2,450) ^b
Größe [Z-Score] MW (SD)	-2,05 (0,868)	-2,32 (1,167) ^b
Krankheitsbezogene Baseline-Charakteristika		
RSS Total Score, MW (SD)	3,19 (1,141)	3,17 (0,975)
Serumphosphat [mg/dl], MW (SD)	2,30 (0,257)	2,42 (0,244)
Serum 1,25(OH) ₂ -Vitamin D ₃ [pg/ml], MW (SD)	40,18 (14,886)	46,00 (20,060)
Alkalische Phosphatase [U/L], MW (SD)	523,4 (154,42)	510,8 (124,90)
Substitutionstherapie ^a (SOC) erhalten, n (%)	32 (100)	29 (100)
Alter bei Initiierung der SOC [Jahre], MW (SD)	2,07 (2,032)	2,61 (2,555)
Dauer der konventionellen Therapie [Jahre], MW (SD)	4,31 (3,031)	3,32 (3,118)
Stratifizierungscharakteristika (zu Randomisierung)		
Alter in Jahren; n (%)		
< 5 Jahre	12 (37,5)	14 (48,3)
≥ 5 Jahre	20 (62,5)	15 (51,7)
RSS Total Score; n (%)		
≤ 2,5	12 (37,5)	10 (34,5)
> 2,5	20 (62,5)	19 (65,5)
Region; n (%)		

Studie	UX023-CL301	
	Aktive Kontrolle (N=32)	Burosumab (N=29)
Gruppe		
N (FAS)		
Japan	3 (9,4)	2 (6,9)
ROW	29 (90,6)	27 (93,1)

MW: Mittelwert; SD: *standard deviation* (Standardabweichung); n: Anzahl der Patienten des FAS;
 ROW: *rest of the world* (Rest der Welt); BMI: Body-Mass-Index

^a Die Substitutionstherapie besteht aus einer oralen Gabe von Phosphat und aktivem Vitamin D

^b Die Baseline-Werte der Burosumab-Gruppe beziehen sich auf N=28

Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Studiendesign der Studie UX023-CL301

Bei der Phase-III-Studie UX023-CL301 handelt es sich um eine randomisierte, offene und aktiv kontrollierte Studie an der insgesamt 16 Studienzentren in weltweit 7 Ländern (4 in den USA, 3 in Kanada, 3 in Japan, 2 in Großbritannien, 1 in Australien, 1 in Korea und 1 in Schweden) teilgenommen haben. Es wurden 61 Patienten mit XLH im Alter von 1-12 Jahren eingeschlossen.

Die Studie besteht aus einer Screeningphase, Wash-out-Phase, Behandlungsphase und einer freiwilligen Extensionsphase.

Vor der Screeningphase und der Randomisierung in die jeweilige Behandlungsgruppe unterzogen sich alle Studienteilnehmer einer Wash-out-Phase, in der die Kinder die Behandlung mit oralem Phosphat und aktivem Vitamin D für 7 Tage abbrachen.

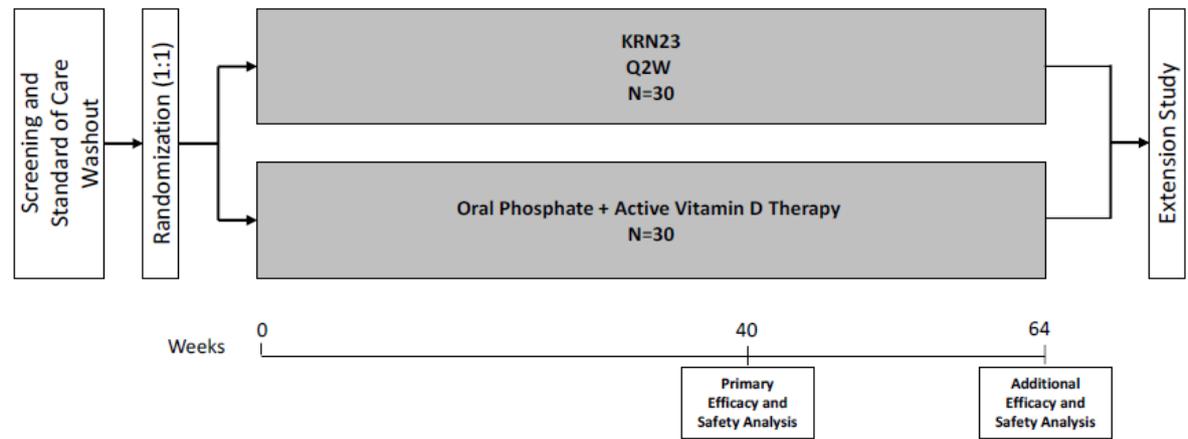
Figure 2.1: UX023-CL301 Study Schema

Abbildung 4-2: Studiendesign der UX023-CL301-Studie

Für Patienten, die zu der Burosumab-Behandlung randomisiert wurden, blieb eine Phosphat und Vitamin-D-Substitution ausgeschlossen. Diese erhielten eine Burosumab-Anfangsdosis von 0,8 mg/kg KG subkutan im 2-wöchigen Abstand. Die Dosis konnte basierend auf dem Nüchtern-Serumphosphatwert auf 1,2 mg/kg KG erhöht werden.

Patienten, die der aktiven Kontrollgruppe zugeteilt wurden, erhielten eine Patienten-individuelle Gabe von oralem Phosphat und aktivem Vitamin D nach Ermessen des Arztes. Die EU- und US-Leitlinie für die Phosphat- und Vitamin-D-Substitution wurde den behandelnden Ärzten zur Verfügung gestellt [10, 11].

Zur Analyse der Behandlungsphase (Woche 0-64) erfolgten zwei Datenschnitte (Woche 40 und 64). Die Extensionsphase (Woche 64-140) galt nur für Patienten in Europa, den USA, Kanada und Australien (ausgenommen waren Patienten aus Japan und Korea). Das Ziel der Extensionsphase bestand darin, die Behandlung mit Burosumab fortzusetzen und weiterhin langfristige Daten zur Sicherheit und Wirksamkeit zu sammeln. Alle Studienteilnehmer konnten nach der 64-wöchigen Behandlungsphase somit für weitere 76 Wochen mit Burosumab behandelt werden.

Patientencharakteristika der Studienpopulation

Diagnose

In der Zulassungsstudie wurden Patienten im Alter von 1 bis einschließlich 12 Jahren mit XLH eingeschlossen. Einschlusskriterium war ein RSS Total Score von mindestens 2, während eine

gesicherte *PHEX*-Mutation nicht notwendig war. Hier reichte ein Nachweis eines direkten Familienmitglieds mit einer X-chromosomalen Vererbung. Hierbei wurden 90,2 % Patienten mit einer gesicherten *PHEX*-Mutation, 4,9 % mit einer möglichen pathogenen *PHEX*-Variante, 3,3 % mit einer ungesicherten Variante und 3,1 % mit einer nicht-positiven *PHEX*-Mutation identifiziert. In den beiden Behandlungsgruppen war die Krankheitsgeschichte der Patienten sehr ähnlich (Tabelle 4-14). Typische biochemische Parameter wie Serumphosphat, Kreatinin und $1,25(\text{OH})_2$ -Vitamin D_3 mussten über-/ oder unterhalb der altersspezifischen Normgrenze sein.

Baseline-Charakteristika

Die Patienten wurden nach einem Alter unter und über 5 Jahren stratifiziert, wobei 12 % der aktiven Kontrollgruppe und 14 % der Burosumab-Gruppe jünger als 5 Jahre waren. Der Grenzwert der Rachitis-Schwere, ausgedrückt durch den RSS Total Score, war im Vergleich zur Phase-II-Zulassungsstudie UX023-CL201 mit 2,5 höher angesetzt (UX023-CL201: Baseline RSS Total Score Grenzwert 1,5). Der Anteil der schweren Rachitissymptomatik war in beiden Gruppen höher als eine leichte Symptomatik (Aktive Kontrolle: 62,5 %; Burosumab: 65,5 %). Der mittlere RSS Total Score-Wert lag bei 3,19 bzw. 3,17. Alle Patienten erhielten oder haben eine konventionelle Therapie mit oralem Phosphat und aktivem Vitamin D erhalten, die im Schnitt im Alter von 2 Jahren initiiert und für ungefähr 3-4 Jahre durchgeführt wurde.

Zusammenfassend lässt sich sagen, dass die Patientencharakteristika, der Behandlungsgruppen sehr vergleichbar waren. Diese wurde zusätzlich zu der Randomisierung durch eine Stratifizierung nach Alter und Rachitis-Schwere sichergestellt.

Übertragbarkeit auf den deutschen Versorgungskontext

Da die Mehrheit der Patienten kaukasischer Herkunft war und die XLH eine genetisch bedingte Erkrankung ist, welche unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen. Daher wird die Studie UX023-CL301 auf den deutschen Versorgungskontext als übertragbar eingeschätzt.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-15: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
UX023-CL301	ja	nein ^a	nein	nein	ja ^b	ja	niedrig
^a Bei der Studie UX023-CL301 handelt es sich um eine randomisierte, offene und aktiv kontrollierte Studie. ^b Die behandelnden Personen wie Ärzte und Pfleger waren nicht verblindet. Die Auswerter der bildgebenden Verfahren zur Ermittlung des <i>RSS Total Scores</i> und <i>RGI-C-Scores</i> waren verblindet.							

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern, problematisch wäre. Die Randomisierung der Patienten wurde 1:1 mittels *Interactive Web Response System* (IWRS) durchgeführt. Die Randomisierung wurde nach Rachitis-Schwere, Alter und Region stratifiziert. Patienten und klinisches Personal waren nicht verblindet. Ausschließlich die Auswerter der Röntgenbildaufnahmen, die den *RSS Total Score* und *RGI-C Score* ermittelten, waren verblindet. Es handelt sich um eine ergebnisunabhängige Berichterstattung. Die Fragebögen wurden von den Patienten selbst, ihren Erziehungsberechtigten oder mit Unterstützung vom klinischen Personal ausgefüllt, biochemische Parameter wurden objektiv erhoben und die Auswertung der Röntgenbilder wurde verblindet durchgeführt. Daher weisen die in der Studie UX023-CL301 berichteten Endpunkte ein geringes Verzerrungspotenzial auf.

Es werden die Ergebnisse des *Full-Analysis-Subsets* (FAS) dargestellt. Ausnahme stellen die Endpunkte 6MWT, die PROMIS-Fragebögen, FPS-R und der SF-10 dar. Hier wurden nur Patienten der FAS-Population herangezogen, die 5 Jahre oder älter waren, da diese Endpunkterhebungen altersabhängig waren. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugewiesene Medikation (aktive Kontrolle oder Burosumab) erhalten haben und mindestens eine Post-Baseline-Bewertung hatten. Da alle randomisierten Patienten eine zugewiesene Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des *Intension-To-Treat* (ITT)-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial der Studie UX023-CL301 somit insgesamt als niedrig eingestuft.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Endpunktkategorie	Endpunkt	UX023-CL301
Mortalität	Im Rahmen der Sicherheit erhoben	- ^a
Morbidität	<i>Radiographic Global Impression of Change (RGI-C)</i>	X
	RGI-C-Responder	X
	<i>Rickets Severity Score (RSS)</i>	X
	Körpergröße: Stehhöhe/Liegelänge [Z-Score]	X
	6-Minuten-Gehdistanz Test (6MWT) [m und Prozentsatz]	X
	Serumphosphat [mg/dl]	X
	<i>Patient-Reported Outcomes Measurement Information System (PROMIS): Pediatric Pain Interference, Physical Function Mobility und Fatigue Domain Score</i>	X
	<i>Faces Pain Scale – Revised (FPS-R)</i>	X
Gesundheitsbezogene Lebensqualität	<i>Short Form 10 (SF-10)</i>	X
Sicherheit	Unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche wegen unerwünschter Ereignisse und Todesfälle	X
X: Endpunkt wurde erhoben. – Endpunkt wurde nicht erhoben. ^a Todesfälle wurden im Rahmen der unerwünschten Ereignisse erhoben.		

4.3.1.3.1 Ergebnisse patientenrelevanter Endpunkte – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

4.3.1.3.1.1 Rachitissymptomatik – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-17: Operationalisierung von Rachitissymptomatik

Studie	Operationalisierung
Radiographic Global Impression of Change (RGI-C) Score	
UX023-CL301	Röntgenbildaufnahmen des Skeletts, die nach den Protokollanforderungen angefertigt wurden, wurden von drei zentralen pädiatrischen Radiologen bewertet. Die Radiologen wurden darauf trainiert, RGI-C-Bewertungen durchzuführen, um einen Konsens über die Terminologie zu erzielen, die zur Beschreibung von XLH-bezogenen radiologischen Anomalien und zur Etablierung einer Interrater-Reliabilität verwendet wurde. Sie waren bezüglich Dosierung, Dosierungsregime, Befolgen des Studienprotokolls, Behandlungsdauer und Vorbehandlung mit einer Substitutionstherapie verblindet. Die Bewertungen wurden unabhängig unter Verwendung eines EDC-Systems durchgeführt, wobei die Bewerter keine Gelegenheit hatten, Bilder zu diskutieren oder Bewertungen zu vergleichen. Bewertungen konnten von den Radiologen nach der Einreichung nicht abgerufen oder geändert werden. Röntgenbildpaare wurden in randomisierter Reihenfolge zur Überprüfung vorgelegt.
Rickets Severity Score (RSS)	
UX023-CL301	<p>Röntgenbildaufnahmen des Skeletts, die gemäß den Protokollanforderungen angefertigt wurden, wurden individuell von Thomas Thacher bewertet, der als einziger zentraler unabhängiger Auswerter für die Burosumab-Studien (UX023-CL201, UX023-CL205 und UX023-CL301) zuständig war. Für die Röntgenbildaufnahmen während der UX023-CL301-Studie wurde Thacher über den Studienbesuch, bei dem die Röntgenbildaufnahme erfolgte, Dosis, Einhaltung des Studienprotokolls und Dauer der Behandlung verblindet. Für die Röntgenbildaufnahmen von UX023-CL301-Patienten vor der Aufnahme in die Studie erfolgte die Verblindung Thachers für den Behandlungsstatus mit SOC-Therapie, SOC-Behandlungsdauer und Einhaltung der SOC-Therapie. Jede von Thacher durchgeführte Bewertung wird zum Zeitpunkt des Ratings in ein EDC-System eingegeben und elektronisch an eine zentrale Bildgebungsanlage übermittelt. Die Ergebnisse können nach der Einreichung von Thacher nicht aus dem System abgerufen oder geändert werden.</p> <p>Der RSS ist eine radiographische Bewertungsmethode, die entwickelt wurde, um die Schwere der Rachitissymptomatik an den Handgelenken und an den Knien basierend auf den metaphysären Aufhellungen, Unschärfen und metaphysären konkaven Becherungen und den Anteil der betroffenen Wachstumsplatten zu bewerten.</p> <p>Die Baseline RSS-Gesamtpunktzahl wird als Summe der RSS Knie Punktzahl bei der Screeningvisite und der RSS Handgelenk Punktzahl zu Baseline berechnet. Die Röntgenbildaufnahmen des Knies wurden bereits bei der Screeningvisite durchgeführt, um die Teilnahmeberechtigung zu bestätigen.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-18: Bewertung des Verzerrungspotenzials für Rachitissymptomatik in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	ja	ja	niedrig
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Für den Endpunkt „Rachitissymptomatik“ in der Studie UX023-CL301 wird ein geringes Verzerrungspotenzial angenommen. Der Endpunkt RGI-C wurde von drei verblindeten pädiatrischen Radiologen ausgewertet, während bei dem Endpunkt RSS ausschließlich eine Person (Thacher) verblindet die Auswertung aller Röntgenbilder durchgeführt hat.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern problematisch wäre. Es liegt eine ergebnisunabhängige Berichterstattung vor. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Rachitissymptomatik“ in der Studie UX023-CL301 somit als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-19: Ergebnisse für *Radiographic Global Impression of Change (RGI-C) Total Score* aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	Radiographic Global Impression of Change (RGI-C) Total Score					
	Änderung von Baseline zu Woche 40 n MWD (SE)	Änderung von Baseline zu Woche 40 n LsMean ^a (SE)	Änderung von Baseline zu Woche 64 n MWD (SE)	Änderung von Baseline zu Woche 64 n LsMean ^a (SE)	Behandlungsdifferenz von Baseline zu Woche 40 LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	Behandlungsdifferenz von Baseline zu Woche 64 LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Gesamtrate						
Aktive Kontrolle ^b (N=32)	32 0,79 (0,125)	32 0,77 (0,107)	32 1,04 (0,134)	32 1,03 (0,136)	1,14 [0,83; 1,45] < 0,0001	1,02 [0,72; 1,33] < 0,0001
Burosumab (N=29)	29 1,92 (0,080)	29 1,92 (0,110)	29 2,06 (0,074)	29 2,06 (0,072)	H'g = 1,846 [1,246; 2,446]	H'g = 1,600 [1,023; 2,178]
MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)						
^a : Berechnung mit ANCOVA-Modell zu Woche 40 und <i>generalized estimation equation</i> (GEE)-Modell zu Woche 64.						
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						
ANCOVA: Behandlungsgruppe, Baseline-Alter-Stratifizierungsfaktor als unabhängige Variable und Baseline RSS als kontinuierliche Kovariate.						
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-Alter-Stratifizierungsfaktor als Faktoren und Baseline RSS als kontinuierliche Kovariate.						

Die Burosumab-Behandlung führte im Vergleich zur aktiven Kontrolle (orales Phosphat und aktives Vitamin D) bereits nach 40 Wochen zu einer deutlich stärkeren Heilung der Rachitissymptomatik. Bei der Betrachtung der Gesamtpopulation der Analyse des Endpunkts RGI-C Global Score betrug die adjustierte mittlere Änderung (LsMean) nach 40 Wochen in der Burosumab-Gruppe +1,92 (+2: erhebliche Verbesserung der Rachitissymptomatik) und +0,77 in der aktiven Kontrollgruppe, was eine statistisch signifikante und auch klinisch relevante Behandlungsdifferenz von +1,14 ergibt ($p < 0,0001$; ANCOVA-Modell). Zu Woche 64 zeigten beide Behandlungsgruppen eine Verbesserung der Rachitissymptomatik, jedoch mit deutlicher Differenz zu Baseline (Burosumab-Gruppe: +2,06; Aktive Kontrollgruppe: +1,03). Die Behandlungsdifferenz von +1,02 zu Woche 64 ($p < 0,0001$; GEE-Modell) unterstützt die Überlegenheit von Burosumab, die sich bereits in Woche 40 deutlich abzeichnete. Burosumab heilt die Rachitissymptomatik im Vergleich zu einer wirksamen Behandlung (Substitutionstherapie: Orales Phosphat und aktives Vitamin D) und nicht im Vergleich zu einer unwirksamen Placebo-Kontrollgruppe. Der Unterschied ist zu beiden Datenschnitten klinisch signifikant.

Erstmalig konnte die Studie UX023-CL301 mit der aktiven Kontrolle (orales Phosphat und aktives Vitamin D) die „Heilung“ der Rachitissymptomatik prospektiv zeigen und die Ergebnisse der Burosumab-Studien des ersten AMNOG-Verfahrens von Burosumab (UX023-CL201 und UX023-CL205 im Vergleich zur historischen Kontrollstudie UX023-CL002) bestätigen.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-20: Ergebnisse für Radiographic Global Impression of Change (RGI-C) -Responderanalyse aus RCT mit dem zu bewertenden Arzneimittel

Radiographic Global Impression of Change (RGI-C) – Responderanalyse								
Population (N)	Woche 40				Woche 64			
	Patienten n (%)	OR ^a [95% KI] p-Wert	RR ^b [95% KI] p-Wert	ARR ^b [95% KI] p-Wert	Patienten n (%)	OR [95% KI] p-Wert	RR ^b [95% KI] p-Wert	ARR ^b [95% KI] p-Wert
Anteil der Patienten mit einem RGI-C \geq +2,0								
Aktive Kontrolle ^c (32)	2 (6,3)	39,139 [7,238; 211,656]	11,586 [2,972; 45,168]	0,662 [0,479; 0,845]	6 (18,8)	34,075 [5,627; 206,340]	4,598 [2,203; 9,596]	0,675 [0,490; 0,859]
Burosumab (29)	21 (72,4)	< 0,0001	< 0,0001	< 0,0001	25 (86,2)	0,0002	< 0,0001	< 0,0001
<p>N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Responder; %: Anteil von N; RSS: <i>Rickets Severity Score</i>; [95% KI]: 95%-iges Konfidenzintervall; OR: <i>Odds Ratio</i>; RR: <i>Risk Ratio</i> (relatives Risiko); ARR: <i>Absolute Risk Reduction</i> (absolute Risikoreduktion) RGI-C-Responder ist definiert als die Patienten mit einem RGI-C \geq +2,0</p> <p>^a: Das logistische Regressionsmodell beinhaltet die Behandlungsgruppe und den Baseline-Alter-Stratifizierungsfaktor als unabhängige Variablen, den Baseline RSS Total Score als kontinuierliche Kovariable. Die geschätzte Responder-Rate, das 95 %-KI, Odds Ratio und der zweiseitiger p-Wert stammen aus dem logistischen Regressionsmodell.</p> <p>^b: Berechnung Effektmodifikatoren (RR, ARR) erfolgte post-hoc</p> <p>^c: Aktive Kontrolle: orales Phosphat und aktives Vitamin D</p>								

Eine erhebliche Verbesserung der „Heilung“ der Rachitissymptomatik ($\geq +2$), wie durch den Endpunkt RGI-C Score gezeigt, bedeutet in diesem Zusammenhang eine Verbesserung der radiologischen Anomalien und impliziert nicht, dass eine vollständige Heilung der XLH beobachtet wurde.

Eine erhebliche „Heilung“ der Rachitissymptomatik wurde zu Woche 40 bei 72,4 % in der Burosumab-Gruppe und bei 6,3 % in der aktiven Kontrollgruppe beobachtet (Odds Ratio [95% KI]: 39,139 [7,238; 211,656], $p < 0,0001$; logistisches Regressionsmodell). Dieses Ergebnis konnte zu Woche 64 mit 67,4 % Behandlungsdifferenz bestätigt werden. Hier konnte ein Odds Ratio von 34,075 (95%-KI: [5,627; 206,340]; $p < 0,0002$) ermittelt werden.

Tabelle 4-21: Ergebnisse für Rickets Severity Score (RSS) – Total Score aus RCT mit dem zu bewertenden Arzneimittel

Intervention		Rickets Severity Score (RSS) - Total Score					
(N)	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n	n	n	n	n	LsMean ^a [95% KI]	LsMean ^a [95% KI]
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	p-Wert falls p < 0,05: Hedges'g [95% KI]	p-Wert falls p < 0,05: Hedges'g [95% KI]
Gesamtrate							
Aktive	32	32	32	32	32		
Kontrolle ^b (N=32)	3,19 (1,141)	2,47 (1,092)	2,17 (0,947)	-0,72 (0,182) -0,71 (0,138)	-1,02 (0,192) -1,02 (0,151)	-1,34 [-1,74; -0,94] < 0,0001	-1,21 [-1,59; -0,83] < 0,0001
Burosumab (N=29)	3,17 (0,975)	1,13 (0,715)	0,95 (0,724)	-2,04 (0,143) -2,04 (0,145)	-2,22 (0,166) -2,23 (0,117)	H'g = -1,399 [-1,964; -0,833]	H'g = -1,119 [-1,710; -0,529]
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)							
^a : Berechnung mit ANCOVA-Modell zu Woche 40 und <i>generalized estimation equation</i> (GEE)-Modell zu Woche 64							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
ANCOVA: Behandlungsgruppe, Baseline-Alter-Stratifizierungsfaktor als unabhängige Variable und Baseline RSS als kontinuierliche Kovariate							
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-Alter-Stratifizierungsfaktor als Faktoren und Baseline RSS als kontinuierliche Kovariate							

Die Abnahme des Schweregrads der Rachitissymptomatik in beiden Behandlungsgruppen wird durch die Veränderung des RSS Total Score von Baseline zu Woche 40 und 64 veranschaulicht. In der Burosumab-Gruppe ist im Vergleich zur aktiven Kontrollgruppe eine deutliche Verbesserung des RSS zu erkennen. Bei Baseline lag der Mittelwert der Burosumab-Gruppen bei 3,17, nach 40 Wochen reduzierte sich der RSS auf 1,13 und in Woche 64 auf 0,95. Die adjustierte mittlere Änderung betrug somit zu Woche 40 -2,04 (ANCOVA-Modell) und -2,23 (GEE-Modell) zu Woche 64. In der aktiven Kontrollgruppe sank der RSS Total Score gegenüber der Baseline um 23 % zu Woche 40 und 32 % zu Woche 64, dabei betrug die adjustierte mittlere Änderung -0,71 zu Woche 40 (ANCOVA-Modell) und -1,02 zu Woche 64 (GEE-Modell). Zu beiden Datenschnitten war der Rückgang des RSS Total Scores in der Burosumab-Gruppe mindestens doppelt so hoch wie in der aktiven Kontrollgruppe. Dabei war der Behandlungsunterschied zwischen den Gruppen sowohl zu Woche 40 als auch zu Woche 64 statistisch signifikant und klinisch relevant (LsMean₄₀: -1,34, $p < 0,0001$, ANCOVA-Modell; LsMean₆₄: -1,21 $p < 0,0001$, GEE-Modell).

Alle Kinder in der Burosumab-Gruppe hatten eine Verbesserung der Rachitis-Schwere von Baseline zu Woche 64 von ≥ 1 im RSS Total Score. Im Gegensatz dazu hat nur die Hälfte (16/32) der Kinder in der aktiven Kontrollgruppe eine Verbesserung der Rachitis-Schwere ≥ 1 gezeigt.

Sensitivitätsanalyse des Endpunktes RSS

Eine Sensitivitätsanalyse wurde durchgeführt, um die Robustheit des primären Analyseergebnisses gegenüber fehlenden Daten zu bewerten. Von einem Kind fehlten der Knie- und Gesamtscore des RSS zu Woche 40. Der RSS-Kniewert zu Woche 40 wurde nicht bewertet, da die Oberschenkelwachstumsplatten bereits verschmolzen waren. Der Baseline-Kniewert wurde imputiert, um den Kniewert der Woche 40 zu ersetzen und somit den Gesamtwert der Woche 40 ableiten zu können. Die Ergebnisse dieser LOCF-Analyse waren denen der Primäranalyse sehr ähnlich. Alle Röntgenbildaufnahmen wurden für die RGI-C-Bewertungen in Woche 64 einbezogen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

4.3.1.3.1.2 Körpergröße (Stehhöhe/Liege­länge) – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-22: Operationalisierung von Körpergröße

Studie	Operationalisierung
UX023-CL301	<p>Das Wachstum wurde durch Änderungen der Stehhöhe gemessen mittels Radiometer vor und nach der Behandlung bestimmt. Bei Patienten unter 2 Jahren oder bei Personen, die nicht in der Lage oder nicht bereit sind, für die Messung zu stehen, wurde die Liege­länge ermittelt.</p> <p>Bei der ersten Screeningvisite wurde die Stehhöhe/Liege­länge gemessen, um die Eignung der Teilnahme an der Studie zu bestätigen. Um das Wachstum während der Behandlung unter Burosumab zu beurteilen, wurde die Stehhöhe/Liege­länge von einem Physiotherapeuten zu Baseline, Woche 24, 40 und 64 erhoben, wobei die erste Datenauswertung zu Woche 64 stattfand. Zu jedem Zeitpunkt erfolgte eine Dreifachmessung, womit ein Mittelwert gebildet wurde.</p> <p>Das Wachstum wurde unter Verwendung von alters- und geschlechtsangepassten Z-Scores bewertet.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-23: Bewertung des Verzerrungspotenzials für Körpergröße in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	ja	ja	niedrig
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie

erforderlich sind, vor allem bei Kindern problematisch wäre. Es liegt eine ergebnisunabhängige Berichterstattung vor. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Körpergröße“ in der Studie UX023-CL301 somit als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-24: Ergebnisse für Körpergröße (Stehhöhe/Liegeänge) aus RCT mit dem zu bewertenden Arzneimittel

Intervention	Körpergröße (Stehhöhe/Liegeänge) [Z-Score]							
	(N)	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n	n	n	n	n	n	LsMean ^a [95% KI]	LsMean ^a [95% KI]
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)		p-Wert falls p < 0,05: Hedges'g [95% KI]	p-Wert falls p < 0,05: Hedges'g [95% KI]
Gesamtrate								
Aktive Kontrolle ^b (N=32)	32 -2,05 (0,868)	32 -2,02 (0,849)	32 -2,03 (0,829)	32 0,03 (0,036) 0,03 (0,031)	32 0,02 (0,039) 0,02 (0,035)		0,12 [0,01; 0,24]	0,14 [0,00; 0,29]
Burosumab (N=29)	28 -2,32 (1,167)	29 -2,12 (1,222)	29 -2,11 (1,111)	28 0,16 (0,051) 0,16 (0,052)	28 0,17 (0,071) 0,17 (0,066)		0,0408 H'g = 0,537 [0,021; 1,053]	0,0490 H'g = 0,490 [-0,025; 1,005]
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)								
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.								
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D								
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Alter, Baseline- Stehhöhe/Liegeänge (Z-Scores) als kontinuierliche Kovariate								

Die XLH ist eine schwere, chronische Erkrankung des Knochenskeletts, die durch Kleinwuchs und Beindeforamationen gekennzeichnet ist. Ein verzögertes Wachstum und Gedeihstörungen sind charakteristisch für XLH-Patienten. Folglich sind diese Patienten sowohl bei den alltäglichen Aktivitäten als auch in ihrer Lebensqualität eingeschränkt. Das Ziel der medikamentösen Therapie ist das Erlangen eines altersbezogenen normalen Wachstums. Das Wachstum wurde in der Studie UX023-CL301 über die Änderung der Stehhöhe/Liegegröße durch den Z-Score beurteilt. Die primäre Analyse für den Endpunkt Wachstum ist Woche 64.

Sowohl in der Burosumab-Gruppe als auch in der aktiven Kontrollgruppe konnte ein Wachstum zu Woche 40 und Woche 64 beobachtet werden. Das dennoch stark verzögerte Wachstum in der aktiven Kontrollgruppe unterstützt die in der Literatur auffindbaren Ergebnisse. Carpenter et al. [10] und Linglart et al. [11] berichten, dass die Substitutionstherapie bei XLH-Patienten nur begrenzte Auswirkungen auf das Wachstum und folglich ebenfalls die Lebensqualität der Patienten hat.

Die Stehhöhe zu Baseline gemessen anhand des Z-Scores ist in den Behandlungsgruppen unterschiedlich. Das GEE-Modell berücksichtigt jedoch durch den Baseline Z-Score als kontinuierliche Kovariable das Ungleichgewicht. Zu Baseline lagen die mittleren Z-Scores der Stehhöhe/Liegegröße in der Burosumab-Gruppe bei -2,32 und in der aktiven Kontrollgruppe bei -2,05.

Zu Woche 40 erhöhte sich der Z-Score der mittleren Stehhöhe/Liegegröße in der Burosumab-Gruppe auf -2,12 und zeigte in Woche 64 einen Wert von -2,11. In der aktiven Kontrollgruppe veränderte sich der Z-Score der mittleren Stehhöhe/Liegegröße zu Woche 40 auf -2,02 und blieb ebenfalls bis Woche 64 nahezu konstant bei -2,03. Der Unterschied in den Behandlungsgruppen bei Veränderung des Z-Scores war zu Woche 40 mit 0,12 ($p = 0,0408$, GEE-Modell) und zu Woche 64 mit 0,14 ($p = 0,0490$, GEE-Modell) statistisch signifikant.

Sowohl bei der Körpergröße als auch beim Wachstum handelt es sich um Parameter, welche von sehr vielen Faktoren beeinflusst werden. Hierzu gehören sowohl die in den Wachstumscharts berücksichtigten Variablen Geschlecht und Alter als auch ethnische Unterschiede sowie Unterschiede im Zeitpunkt der Datenerhebungen. Letztere können insbesondere aufgrund von sozioökonomischen Veränderungen im Zeitverlauf Einfluss auf das Größenwachstum nehmen [26]. Angesichts des relativ hohen Patienten-Anteils von 82 % mit kaukasischem Ursprung (50/61) und 51 % Amerikanern (31/61) sowie der länderspezifischen komplexen Variabilität des Parameters stellen die Vergleichsdaten der Z-Scores von der amerikanischen gesunden Population des CDC das Verfahren mit der größten zur Verfügung stehenden Validität dar. Ergänzend ist zu erwähnen, dass kein deutscher Studienstandort in der RCT UX023-CL301 eingeschlossen war [20]. Weiterhin bestätigt die vergleichende Studie von Rosario et al., dass die Wachstumscharts der CDC für den Vergleich mit deutschen Wachstumskurven geeignet sind [29]. Daraus folgt, dass länderspezifische Z-Scores nicht zwingend erforderlich sein müssen, sondern viel mehr die Herkunft der Patientenpopulation entscheidend ist. Folglich ist die Datenvalidität gewährleistet.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

4.3.1.3.1.3 Gehfähigkeit – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-25: Operationalisierung von Gehfähigkeit

Studie	Operationalisierung
UX023-CL301	<p>Der 6-Minuten-Gehdistanz Test (6MWT) wurde von einem geschulten klinischen Personal in Übereinstimmung mit allgemeinen Prinzipien durchgeführt, die in den Richtlinien der <i>American Thoracic Society</i> (ATS) festgelegt sind. Erst nach Erhalt eines Schulungszertifikats bezüglich der ATS-Richtlinien war es dem klinischen Personal gestattet, den Test nach SOP zu leiten. Die Patienten wurden angewiesen, einen festgelegten Weg sechs Minuten zu gehen. Die zurückgelegte Strecke nach sechs Minuten wurde in Metern aufgezeichnet [32].</p> <p>Abweichend von den ATS-Richtlinien erfolgte die Messung der Gehstrecke aufgrund eines erhöhten Sicherheitsrisikos nicht mithilfe eines Messrads, sondern über die Zählung der Rundenzahl und eines Maßbandes. Da XLH-Patienten starke Rotationsfehlbildungen der unteren Extremitäten aufweisen und dadurch einen unsicheren Gang haben bzw. häufig stolpern, galt das Schieben eines Messrads entlang der Laufstrecke als potenzielles Risiko. Um eine repräsentative Gehbewegung zu erfassen, wurde darauf geachtet, dass eine freie Bewegung der oberen Extremitäten für das „Ausbalancieren“ möglich war und somit eine natürliche Gehbewegung/Gleichgewichtsanpassung erfasst werden konnte. Um die mögliche Motivation zum Weitergehen durch das Schieben des Rades der Kinder nachzuahmen, wurde ebenfalls ein Anreiz für die Kinder geschaffen, indem die Kinder alternativ an jeder Biegung, ohne anzuhalten, Softringe in Armbandgröße sammeln und fallen lassen konnten. Die Verwendung der Ringe ermutigte die Kinder, kontinuierlich für 6 Minuten zu gehen und nach Ende des Tests die Anzahl der gesammelten Ringe zu zählen. Dieser Test wurde von den Kindern gut angenommen und gleicht mit der Softring-Technik die Motivationsverzerrung, die durch den Verzicht auf ein Rad entstanden sein könnte, aus.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-26: Bewertung des Verzerrungspotenzials für Gehfähigkeit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	ja	ja	niedrig
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern problematisch wäre. Es liegt eine ergebnisunabhängige Berichterstattung vor. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Gehfähigkeit“ in der Studie UX023-CL301 somit als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-27: Ergebnisse für Gehfähigkeit aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	6-Minuten-Gehdistanz Test (6MWT) in Metern (Baseline Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n MW (SD)	n MWD (SE) LsMean ^a (SE)	n MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Gesamtrate – Messung in Meter							
Aktive Kontrolle ^b (N=20)	20 450,50 (106,432)	20 457,40 (95,571)	20 481,30 (113,057)	20 6,90 (12,767) 11,85 (12,981)	20 30,80 (18,106) 35,75 (16,711)	38,03 [-2,28; 78,35] 0,0644	43,20 [2,33; 84,07] 0,0383 H'g = 0,811 [0,115; 1,507]
Burosumab (N=15)	15 365,93 (118,083)	15 434,73 (79,342)	15 463,80 (81,987)	15 68,80 (24,829) 49,88 (16,298)	15 97,87 (19,339) 78,95 (11,410)		
Gesamtrate – Prozentsatz							
Aktive Kontrolle ^b (N=20)	20 76,20 (14,838)	20 75,28 (14,198)	20 78,13 (17,527)	20 -0,92 (2,080) 0,52 (2,205)	20 1,94 (2,935) 3,38 (2,839)	5,50 [-1,12; 12,12] 0,1035	6,23 [-0,70; 13,15] 0,0781
Burosumab (N=15)	15 62,13 (18,629)	15 71,85 (11,931)	15 75,43 (12,040)	15 9,72 (4,301) 6,02 (2,711)	15 13,30 (3,319) 9,61 (1,934)		
Korrigierte Analyse^c							
Gesamtrate – Messung in Meter							

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention	6-Minuten-Gehdistanz Test (6MWT) in Metern (Baseline Alter der Kinder ≥ 5 Jahren)						
	(N)	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40
	n	n	n	n	n	LsMean ^a [95% KI]	LsMean ^a [95% KI]
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	p-Wert falls p < 0,05: Hedges'g [95% KI]	p-Wert falls p < 0,05: Hedges'g [95% KI]
Aktive Kontrolle ^b (N=20)	20 450,50 (106,432)	20 457,40 (95,571)	20 481,30 (113,057)	20 6,90 (12,767) 5,38 (12,656)	20 30,80 (18,106) 29,28 (16,834)	38,30 [-2,38; 78,98]	45,55 [2,09; 89,02]
Burosumab (N=15)	13 385,00 (86,157)	13 435,31 (85,605)	13 466,46 (81,717)	13 50,31 (14,331) 43,68 (15,262)	13 81,46 (13,778) 74,83 (12,513)	0,0650 -	0,0399 H'g = 0,679 [-0,038; 1,396]
Gesamtrate – Prozentsatz							
Aktive Kontrolle ^b (N=20)	20 76,20 (14,838)	20 75,28 (14,198)	20 78,13 (17,527)	20 -0,92 (2,080) -0,97 (2,077)	20 1,94 (2,935) 1,88 (2,789)	6,19 [-0,68; 13,06]	7,27 [0,01; 14,52]
Burosumab (N=15)	13 65,07 (12,120)	13 71,60 (12,832)	13 75,53 (11,582)	13 6,53 (2,452) 5,22 (2,634)	13 10,46 (2,259) 9,15 (2,056)	0,0775 -	0,0496 H'g = 0,703 [-0,015; 1,422]
<p>MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Ricketts Severity Score</i>; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)</p> <p>^a: Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.</p> <p>^b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D</p> <p>^c: Bei dieser Analyse wurden 2 Patienten nicht berücksichtigt, da diese einen unvollständigen Datensatz 1-mal zu Baseline und 1-mal zu Woche 40 aufweisen.</p> <p>GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Baseline-6MWT als Kovariate</p>							

Gangstörungen sind ein häufiges klinisches Merkmal bei pädiatrischen Patienten mit XLH. Diese treten durch die Kombination aus Knochendeformationen, Knochen- und Gelenkschmerzen, sowie muskuläre Beschwerden auf. Der 6-Minuten-Gehdistanz Test (6MWT) ist ein Indikator für die Mobilität und die körperliche Funktionsfähigkeit. Dieser Test konnte erst bei Kindern ≥ 5 Jahren zur Screeningvisite durchgeführt werden. Die Gehfähigkeit wurde als die gesamte zurückgelegte Strecke in Meter und als Prozentsatz der vorhergesagten Werte auf Grundlage normativer Daten (basierend auf Alter, Geschlecht und Größe [55]) innerhalb der 6 Minuten ermittelt. Eine erhöhte Gehfähigkeit konnte sowohl als absoluter als auch als relativer Wert in der Burosumab-Gruppe, jedoch nicht in der aktiven Kontrollgruppe von Baseline zu Woche 40 beobachtet werden. Zu Woche 64 zeigte die Burosumab-Gruppe im Vergleich zur aktiven Kontrollgruppe eine signifikante Verbesserung in der Gehdistanz gemessen innerhalb der 6 Minuten.

Die Analyse der Gehfähigkeit wurde durch die unterschiedlichen Baseline-Werte der Behandlungsgruppen des 6MWT gestört, da die Burosumab-Gruppe zu Baseline bereits eine deutlich stärkere Beeinträchtigung als die aktive Kontrollgruppe zeigte. Zwei pädiatrische Patienten der Burosumab-Gruppe (ein Kind zu Baseline und ein Kind zu Woche 40) wurden aus der Gesamtanalyse ausgeschlossen, da diese den 6-Minuten-Gehdistanz Test nicht abgeschlossen haben (siehe auch 4.2.5.4). In Tabelle 4-27 sind als Ergänzung neben den Ergebnissen der korrigierten Analyse ohne Berücksichtigung der zwei genannten Patienten, die Ergebnisse der ursprünglichen Analyse (inklusive der zwei Patienten) dargestellt.

Die mittlere Distanz in der Burosumab- und aktiven Kontrollgruppe betrug zu Baseline 385,00 m bzw. 450,50 m. Die mittleren relativen Distanzwerte betragen daher 65,07 % bzw. 76,20 %.

Die Gehdistanz verbesserte sich zu Woche 40 in der Burosumab-Gruppe um 50,31 m und in der aktiven Kontrollgruppe um 6,90 m. Die Burosumab-Gruppe zeigte im vorhergesagten Prozentsatz eine mittlere Gehfähigkeitsverbesserung von 6,53 %, während die aktive Kontrollgruppe nahezu unverändert blieb mit -0,92 %. Der Unterschied zwischen den Behandlungsgruppen von 6,19 %, ausgedrückt als Prozentsatz der vorhergesagten Werte, war statistisch nicht signifikant ($p_{40} = 0,0775$, GEE-Modell).

In Woche 64 stieg der Mittelwert in der Burosumab-Gruppe um 21,26 % von 385,00 m auf 466,46 m, was eine adjustierte mittlere Änderung von 74,83 m ist. Hingegen konnte die aktive Kontrollgruppe nur eine Verbesserung um 6,84 % von 450,51 m auf 481,30 m erzielen. Nach der Kontrolle der Baseline-Differenzen mithilfe des GEE-Modells, zeigte die Behandlungsdifferenz eine signifikante Verbesserung der zurückgelegten Wegstrecke im 6MWT zugunsten der Burosumab-Gruppe ($LsMean_{64}$: 45,55 m; $p = 0,0399$; GEE-Modell). Ähnliche Ergebnisse wurden in der Auswertung der relativen 6MWT zu Woche 64 erzielt. Dabei zeigte die Burosumab-Gruppe eine mittlere Gehfähigkeitsverbesserung von 65,07 % zu Baseline auf 75,53 % zu Woche 64. Die mittlere Änderung betrug somit 10,46 %, während die aktive Kontrollgruppe nur eine mittlere Änderung von 1,94 % erreichen konnte.

Nach der Kontrolle der Baseline-Differenz anhand des GEE-Modells zeigte die Burosumab-Gruppe eine signifikant größere Zunahme der zurückgelegten relativen 6MWT im Vergleich zur aktiven Kontrollgruppe (LsMean₆₄: 7,27; p = 0,0496; GEE-Modell).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

4.3.1.3.1.4 Serumphosphat – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-28: Operationalisierung von Serumphosphat

Studie	Operationalisierung
UX023-CL301	Den nüchternen Patienten wurden Blutproben zu Baseline sowie zu den geforderten Post-Baseline-Zeitpunkten entnommen. Die Post-Baseline Proben wurden den Patienten, die Burosumab injiziert bekamen, vor der Verabreichung von Burosumab abgenommen und in einem Zentrallabor ausgewertet.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-29: Bewertung des Verzerrungspotenzials für Serumphosphat in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	ja	ja	niedrig
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern problematisch wäre. Es liegt eine ergebnisunabhängige Berichterstattung vor. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Serumphosphat“ in der Studie UX023-CL301 somit als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-30: Ergebnisse für Serumphosphat aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	Serumphosphat [mg/dl]						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n MW (SD)	n MWD (SE) LsMean ^a (SE)	n MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Gesamtrate							
Aktive Kontrolle ^b (N=32)	32 2,30 (0,257)	32 2,55 (0,291)	32 2,56 (0,300)	32 0,25 (0,056) 0,23 (0,058)	32 0,26 (0,056) 0,24 (0,058)	0,77 [0,60; 0,94] < 0,0001 H'g = 2,047 [1,427; 2,667]	0,74 [0,58; 0,91] < 0,0001 H'g = 1,981 [1,368; 2,595]
Burosumab (N=29)	29 2,42 (0,244)	29 3,38 (0,374)	29 3,36 (0,365)	29 0,96 (0,066) 1,00 (0,062)	29 0,94 (0,065) 0,98 (0,061)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)							
^a : Berechnung mit ANCOVA-Modell zu Woche 40 und zu Woche 64							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
ANCOVA: Behandlungsgruppe, Baseline-Alter-Stratifizierungsfaktor als unabhängige Variable und Baseline RSS als kontinuierliche Kovariate							

Die chronische Hypophosphatämie bei Patienten mit XLH führt zu einer Rachitis und zu einer Osteomalazie, den zwei wichtigsten pathologischen Merkmalen der XLH. Daher besteht das Ziel der Burosumab-Therapie darin, die Serumphosphat-Konzentration auf einen normalen oder nahezu normalen Wert innerhalb des Normbereichs [3,2 mg/dl – 6,1 mg/dl] zu erhöhen. Die Patienten unterbrachen 7 Tage vor der Randomisierung die orale Einnahme von Phosphat und aktiven Vitamin D. Alle Patienten, die in die Studie aufgenommen wurden, mussten einen hypophosphatämischen Status unter 3 mg/dl aufweisen.

Die Burosumab-Behandlung erhöhte die mittlere Serumphosphat-Konzentration schnell auf die untere Grenze des Normbereichs (LLN: 3,2 mg/dl). Die mittlere Änderung des Serumphosphats von Baseline zu Woche 40 betrug in der Burosumab-Gruppe 0,96 mg/dl und lag in Woche 40 bei 3,38 mg/dl, welche innerhalb des altersspezifischen Normbereichs liegt. Diese Normalisierung der Serumphosphat-Konzentration konnte konstant gehalten werden und lag in Woche 64 bei 3,36 mg/dl. In der aktiven Kontrollgruppe konnte nur eine leichte Erhöhung der Serumphosphat-Konzentration gezeigt werden. Dieser lag jedoch zu allen gemessenen Zeitpunkten deutlich unter LLN. Zu Woche 40 konnte in der aktiven Kontrollgruppe eine Serumphosphat-Konzentration von 2,55 mg/dl erreicht werden, welche in Woche 64 nahezu unverändert bei 2,56 mg/dl lag. In der Kontrollgruppe konnte eine adjustierte Veränderung von 0,23 mg/dl zu Woche 64 gezeigt werden, hingegen konnte die Burosumab-Gruppe eine adjustierte Änderung von 1,00 mg/dl von Baseline zu Woche 64 erzielen.

Die Veränderung der Serumphosphat-Konzentration in der Burosumab-Gruppe von Baseline zu Woche 40 und 64 (0,77 mg/dl; 0,74 mg/dl) war statistisch signifikant höher als bei der aktiven Kontrollgruppe (mit jeweils $p < 0,0001$; ANCOVA). Die Behandlungsdifferenz zu Woche 40 und 64 war sowohl statistisch signifikant als auch klinisch relevant.

Ein theoretisches potenzielles Sicherheitsrisiko für Burosumab stellt die Hyperphosphatämie dar. Kein Studienteilnehmer hatte während der Studie eine Serumphosphat-Konzentration oberhalb des Normbereichs. Ebenfalls wurde kein unerwünschtes Ereignis mit *preferred terms* (PT) „Hyperphosphatämie“ berichtet.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

4.3.1.3.1.5 Schmerzerhebung und funktionelle Behinderung– RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-31: Operationalisierung von Schmerzerhebung

Studie	Operationalisierung
PROMIS Pain Interference, Physical Function und Fatigue Domain Score	
UX023-CL301	<p>Die pädiatrische PROMIS-Fragebogen wurde von nationalen Gesundheitsinstituten entwickelt und verwendet spezifische Domänen zur Beurteilung des Wohlbefindens der Patienten [3]. Die Domänen wurden so gewählt, dass Gesundheitsmerkmale wie Schmerz und körperliche Funktion nicht für eine bestimmte Krankheit spezifisch sind. Der PROMIS enthält eine Reihe von Fragen zu den Domänen Schmerz, Körperfunktionsmobilität und Fatigue.</p> <p>Es wurde ausschließlich die Version 2.0 des PROMIS-Fragebogens in der Studie UX023-VL301 angewandt. Der Fragebogen eignet sich bei Kindern ab 5 Jahren. Für Kinder von 5 bis einschließlich 7 Jahren wurde eine Version des Fragebogens (Eltern-Version) verwendet, bei dem die Eltern/Erziehungsberechtigten die Fragen beantworten mussten. Selbst wenn das Kind während der Studiendauer 8 Jahre alt wurde, erfolgte konstant die Beantwortung der Eltern-Version durch die Eltern. Kinder ab 8 Jahren erhielten eine Version, die eigenständig von den Kindern ausgefüllt werden konnte (Selbstbericht-Version). Hatten Kinder ab 8 Jahren Schwierigkeiten beim Ausfüllen der Selbstbericht-Version, konnte der Prüfarzt zustimmen, dass der Fragebogen von einem geschulten klinischen Personal per Interview ausgefüllt wird oder dass ein Erziehungsberechtigter die Eltern-Version ausfüllt. Der Grund musste sorgfältig dokumentiert werden und es musste während der gesamten Studie eine einheitliche Verwendung und Durchführung der Fragebogenversionen beibehalten werden [32]. Die Validierung des PROMIS-Fragebogens auf die pädiatrische XLH-Patientenpopulation erfolgte durch die Interviewstudie UX7575A, in der durch eine kognitive Nachbesprechung mit den Patienten die Validität gezeigt werden konnte.</p>
Face Pain Scale – Revised (FPS-R)	
UX023-CL301	<p>Die überarbeitete Gesichtsschmerzskala (FPS-R) ist ein selbstberichtendes Maß zur Ermittlung des akuten Schmerzempfindens, das speziell für Kinder entwickelt wurde [4]. Dieses wurde für die Verwendung bei Kindern im Alter von 5 bis 16 Jahren validiert. Dabei wird Schmerzempfinden auf einer Skala von 0 bis 10 graphisch dargestellt, wobei Gesichter mit Werten von 0, 2, 4, 6, 8 und 10 verwendet werden (0 = Kein Schmerz; 10 = stärkster Schmerz). Die Schmerzintensität der Patienten wurde von Kindern ab 5 Jahren selbst berichtet [32]. Es erfolgte die Befragung aller Patienten im Alter von 5-12 Jahren.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-32: Bewertung des Verzerrungspotenzials für Schmerzerhebung und funktionelle Behinderung in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	unklar	ja	niedrig
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern problematisch wäre. Für die PROMIS Selbstbericht-Version und dem FPS-R liegt eine ergebnisunabhängige Berichterstattung vor. Eine Ausnahme ist die Eltern-Version des PROMIS-Fragebogens. Hier liegt keine objektive Einschätzung der Kindersymptomatik vor, da diese auf der subjektiven Wahrnehmung der Eltern beruht. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Schmerzerhebung“ in der Studie UX023-CL301 somit als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-33: Ergebnisse für Schmerzerhebung und funktionelle Behinderung mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder \geq 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n	n	n	n	n	LsMean ^a [95% KI] p-Wert falls $p < 0,05$: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls $p < 0,05$: Hedges'g [95% KI]
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)		
PROMIS Pediatric Pain Interference Domain Score							
Gesamtrate							
Aktive Kontrolle ^b (N=20)	20 49,9 (12,05)	20 50,4 (9,51)	20 49,4 (9,52)	20 0,5 (2,01) -0,31 (1,526)	20 -0,5 (1,60) -1,29 (1,267)	-4,91 [-9,19; -0,63] 0,0245	-2,26 [-6,61; 2,09] 0,3091
Burosumab (N=15)	15 53,1 (10,95)	15 47,6 (9,84)	15 49,3 (8,07)	15 -5,5 (1,78) -5,23 (1,707)	15 -3,8 (2,59) -3,55 (1,873)	H'g = -0,697 [-1,390; -0,008]	-
PROMIS Pediatric Physical Function Mobility Domain Score							
Gesamtrate							
Aktive Kontrolle ^b (N=20)	20 45,5 (9,86)	20 45,5 (9,71)	20 46,3 (9,63)	20 0,0 (1,01) 0,09 (0,966)	20 0,9 (0,99) 0,92 (0,962)	2,68 [-0,53; 5,88] 0,1014	1,90 [-1,80; 5,59] 0,3145
Burosumab (N=15)	15 45,2 (9,05)	15 47,9 (8,32)	15 47,9 (9,24)	15 2,7 (1,49) 2,76 (1,338)	15 2,8 (1,80) 2,82 (1,648)	-	-

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention	Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)						
	(N)	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
PROMIS Pediatric Fatigue Domain Score							
Gesamtrate							
Aktive Kontrolle ^b (N=20)	20 47,0 (13,70)	20 46,6 (10,73)	20 45,0 (11,17)	20 -0,5 (2,21) -0,97 (1,761)	20 -2,1 (1,87) -2,57 (1,547)	-3,20 [-7,79; 1,39] 0,1724	-1,08 [-6,21; 4,06] 0,6810
Burosumab (N=15)	15 48,8 (9,60)	15 44,7 (10,49)	15 45,2 (10,69)	15 -4,1 (1,50) -4,17 (1,673)	15 -3,6 (2,32) -3,65 (2,119)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)							
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Baseline-PROMIS als Kovariate							

Der pädiatrische *Patient-Reported Outcomes Measurement Information System* (PROMIS) Fragebogen mit den Domänen Schmerz, Mobilität und Fatigue konnte bei Patienten ab einem Alter von 5 Jahren angewendet werden. Die Domänen zeigen Scores mit einem Mittelwert von 50 und einer Standardabweichung (SD) von 10. Höhere Scores und somit positive Effekte werden durch einen Anstieg der Mobilität und einen Rückgang der Schmerz- und Fatigue-Domäne angezeigt. Minimal klinisch wichtige Unterschiede (MIDs) für den PROMIS-Skalenwert werden auf 2,4 bis 3,5 geschätzt [56].

Zu Baseline lagen die Werte für die drei untersuchten Domänen innerhalb von 1 SD des standardisierten Mittelwerts von 50. Die Mobilitätsdomäne (PROMIS *Physical Function Mobility Domain Score*) zeigte die größte Beeinträchtigung. Hier lagen die Werte bei 45,2 in der Burosumab-Gruppe und 45,5 in der aktiven Kontrollgruppe.

In der Schmerzdomäne (PROMIS *Pediatric Pain Interference Domain Score*) nahmen die Werte in der Burosumab-Gruppe zu Woche 40 von 53,1 auf 47,6 ab und nahmen zu Woche 64 wieder auf 49,3 zu. In der aktiven Kontrollgruppe sank der Baseline-Wert um 0,31 zu Woche 40 und um 1,29 zu Woche 64. Zu Baseline wurde in der Burosumab-Gruppe über mehr Schmerzen als in der aktiven Kontrollgruppe berichtet. Dies änderte sich zu Woche 40. Hier wurde in der Burosumab-Gruppe im Vergleich zur aktiven Kontrollgruppe über weniger Schmerz berichtet. Der Unterschied von -4,91 zwischen den Behandlungsgruppen in der Schmerzdomäne war statistisch signifikant ($p = 0,0245$; GEE-Modell) und war größer als die geschätzte MID von 2,4 bis 3,5 [56]. Zu Woche 64 konnte die Behandlungsdifferenz von -2,26 dagegen keinen signifikanten Vorteil zugunsten der Burosumab-Gruppe zeigen.

Die mittleren Werte der Mobilitätsdomäne (PROMIS *Physical Function Mobility Domain Score*) stiegen in der Burosumab-Gruppe von 45,2 zu Baseline auf 47,9 zu Woche 40 an und blieben zu Woche 64 unverändert. Die Werte der aktiven Kontrollgruppe blieben zu Woche 40 mit 45,5 unverändert. In Woche 64 konnte nur eine mittlere Änderung von 0,9 zu Baseline erzielt werden. Der Unterschied zwischen den Behandlungsgruppen zu Woche 40 und 64 in der Mobilitätsdomäne war statistisch nicht signifikant. Die Größe der minimal wichtigen Differenz war jedoch ähnlich der geschätzten MID.

In der Fatigue-Domäne (PROMIS *Fatigue Domain Score*) nahmen in der Burosumab-Gruppe die Werte von 48,8 zu Baseline auf 44,7 zu Woche 40 ab. Zu Woche 64 konnte eine Reduktion von 3,6 zu Baseline erzielt werden. Die aktive Kontrollgruppe konnte hingegen zu Woche 64 eine Reduktion von 2,1 auf 45,0 zeigen. Der Unterschied zwischen den Behandlungsgruppen in der Fatigue-Domäne war zu keiner Zeit statistisch signifikant, jedoch war die Größe der Differenz ähnlich der geschätzten MID.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-34: Ergebnisse für Schmerzerhebung mittels FPS-R aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder ≥ 5 Jahre)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n	n	n	n	n	LsMean ^a [95% KI]	LsMean ^a [95% KI]
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	p-Wert falls p < 0,05: Hedges'g [95% KI]	p-Wert falls p < 0,05: Hedges'g [95% KI]
Gesamtrate							
Aktive Kontrolle ^b (N=20)	20 0,7 (1,17)	20 0,6 (1,60)	19 0,5 (1,12)	20 -0,1 (0,34) 0,04 (0,327)	19 -0,1 (0,32) -0,01 (0,234)	-0,00 [-0,80; 0,79] 0,9905	0,05 [-0,58; 0,68] 0,8786
Burosumab (N=15)	15 0,4 (1,12)	15 0,5 (1,19)	15 0,5 (0,92)	15 0,1 (0,36) 0,04 (0,312)	15 0,1 (0,41) 0,04 (0,270)	-	-

MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: *Rickets Severity Score*; [95% KI]: 95%-iges Konfidenzintervall; SD: *standard deviation* (Standardabweichung); SE: *standard error* (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)

^a: Berechnung mit *generalized estimation equation* (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.

^b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Baseline-FPS-R als Kovariate

Die Schmerzintensität wurde von Kindern ab einem Alter von ≥ 5 Jahren zur Screeningvisite unter Verwendung des pädiatrischen Fragebogens (*Faces Pain Scale – Revised* (FPS-R)) selbst berichtet. Der FPS-R ist eine Bewertung der aktuellen Schmerzintensität im Ruhezustand und nicht während einer körperlichen Aktivität. Skalenwerte liegen im Bereich von 0 bis 10 (0 = Kein Schmerz; 10 = stärkster Schmerz), wobei nur ganze Zahlen verwendet werden dürfen.

Die meisten Studienteilnehmer in beiden Behandlungsgruppen berichteten von keinem FPS-R-Schmerz zu Baseline und bei allen Visiten. Der FPS-R-Wert betrug 0,4 (Burosumab) bzw. 0,7 (aktive Kontrolle) zu Baseline und 0,5 bzw. 0,6 zu Woche 40. Zu Woche 64 zeigten beide Behandlungsgruppen eine Änderung zu Baseline von 0,1 bzw. -0,1. Es konnte kein signifikanter Unterschied zwischen den Behandlungsgruppen von Baseline zu Woche 40 bzw. 64 beobachtet werden ($p_{40} = 0,9905$ und $p_{64} = 0,8786$; GEE-Modell).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

4.3.1.3.1.6 Lebensqualität – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-35: Operationalisierung von Lebensqualität

Studie	Operationalisierung
UX023-CL301	<p>Der <i>Short Form 10</i> (SF-10) für Gesundheitsbefragungen ist ein von der Pflegekraft/Erziehungsberechtigten ausgefüllter Fragebogen zur Beurteilung körperlicher und psychosozialer Funktionen von kranken und gesunden Kindern im Alter von 5 – 18 Jahren [47]. Die 10 Items wurden aus dem <i>Child Health Questionnaire</i> (CHQ) adaptiert.</p> <p>Hierbei werden zwei Komponenten generiert:</p> <ul style="list-style-type: none"> • Physische Gesamtpunktzahl (PHS-10) • Psychosoziale Gesamtpunktzahl (PSS-10) <p>Der PHS-10 besteht aus 4 Skalen, während der PSS-10 aus 4 Skalen + 2 Skalen, die an die Eltern des Patienten gerichtet sind, aufgebaut ist.</p> <p>Die Skalenwerte wurden so zentriert, dass eine Punktzahl von 50 der durchschnittlichen Punktzahl in einer umfassenden Stichprobe von 2006 entspricht (eine Kombination aus Proben aus der gesunden Allgemeinbevölkerung, Kindern mit Behinderungen sowie chronischen Erkrankungen). Skalenwerte sind auf einen Mittelwert von 50 und eine Standardabweichung von 10 in der kombinierten US-Allgemeinbevölkerung standardisiert. Die Bewertung des SF-10 erfolgt nach der in Saris-Baglama et al. beschriebenen Methode. Höhere globale Punktzahlen sind mit einer besseren Lebensqualität verbunden.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-36: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	nein	ja	hoch
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern problematisch wäre. Es liegt eine ergebnisunabhängige Berichterstattung vor, da diese auf der subjektiven Wahrnehmung der Eltern beruht. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden. Da der Fragebogen nicht selbstständig von den Kindern ausgefüllt wurde, erfolgte ausschließlich eine Fremdeinschätzung durch die Eltern. Somit liegt keine objektive Einschätzung der Lebensqualität der Patienten vor, da diese auf der rein subjektiven Wahrnehmung der Eltern beruht.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Lebensqualität“ in der Studie UX023-CL301 somit als hoch eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-37: Ergebnisse für Lebensqualität (SF-10) aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	Gesundheitsspezifischer Fragebogen – SF-10 (Baseline-Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Physical Health Summary Score (PHS-10)							
Gesamtrate							
Aktive Kontrolle ^b (N=20)	20 40,74 (15,298)	20 42,27 (12,510)	20 41,07 (15,093)	20 1,53 (2,731) 1,69 (2,247)	20 0,33 (2,418) 0,48 (2,346)	4,35 [-1,15; 9,85] 0,1213	5,51 [-0,17; 11,18] 0,0571
Burosumab (N=15)	15 40,03 (10,069)	15 46,16 (9,904)	15 46,12 (9,844)	15 6,13 (2,047) 6,03 (1,946)	15 6,08 (2,187) 5,99 (2,006)		
Psychosocial Summary Score (PSS-10)							
Gesamtrate							
Aktive Kontrolle ^b (N=20)	20 52,79 (9,401)	20 51,85 (9,330)	20 53,95 (8,457)	20 -0,94 (1,519) -0,58 (1,413)	20 1,16 (1,395) 1,52 (1,229)	2,24 [-1,86; 6,35] 0,2841	-0,44 [-4,59; 3,70] 0,8337
Burosumab (N=15)	15 50,76 (9,652)	15 52,66 (9,225)	15 52,06 (9,137)	15 1,90 (1,738) 1,67 (1,640)	15 1,31 (2,112) 1,07 (1,799)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: *Rickets Severity Score*; [95% KI]: 95%-iges Konfidenzintervall; SD: *standard deviation* (Standardabweichung); SE: *standard error* (Standardfehler); H^g: Hedges'g (wurden post-hoc berechnet)

^a: Die Berechnung der LsMean mit dem *generalized estimation equation* (GEE)-Modell erfolgte post hoc. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.

^b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Baseline-Alter, Standhöhe/Liegelänge [Z-Score] als Kovariate

Die körperliche und psychosoziale gesundheitsbezogene Lebensqualität wurde anhand des für Kinder entwickelten SF-10-Fragebogen bewertet. Der mittlere *Physical Health Summary Score* (PHS-10) betrug bei der Burosumab-Gruppe zu Baseline 40,03 und bei der aktiven Kontrollgruppe 40,74, was auf eine signifikante Beeinträchtigung der körperlichen Funktion im Vergleich zur gesunden US-Allgemeinbevölkerung hinweist. Die mittleren PHS-10-Werte stiegen in der Burosumab-Gruppe in Woche 40 auf 46,16 und zeigten eine adjustierte mittlere Veränderung von 6,03, die bis zu Woche 64 mit 5,99 nahezu unverändert blieb. In der aktiven Kontrollgruppe stiegen die mittleren PHS-10-Werte in Woche 40 auf 42,27, was eine adjustierte Veränderung von 1,69 bedeutet. Zu Woche 64 war dieser Vorteil bei einer mittleren Änderung zu Baseline mit lediglich 0,48 nicht mehr eindeutig. Die Verbesserung der Lebensqualität zwischen den Behandlungsgruppen war statistisch nicht signifikant ($p_{40} = 0,1213$; $p_{64} = 0,0571$). Diese Ergebnisse des PHS-10 deuten auf eine größere Verbesserung in der Burosumab-Gruppe hin. Dieser Trend konnte ebenfalls durch den statistisch nicht signifikanten Endpunkt PROMIS *Physical Function Mobility* und den signifikanten Ergebnissen des 6MWT gezeigt werden.

Bei Baseline lagen die mittleren Werte für den *Psychosocial Summary Score* (PSS-10) nahe dem standardisierten Mittelwert von 50 (Burosumab-Gruppe: 50,76; aktive Kontrollgruppe: 52,79), was keine Beeinträchtigung der psychosozialen Funktion im Vergleich zur US-Allgemeinbevölkerung zeigt. In der Burosumab-Gruppe stiegen die mittleren PSS-10-Werte in Woche 40 auf 52,66 und in Woche 64 auf 52,06. Die mittleren PSS-10-Werte der aktiven Kontrollgruppe sanken in Woche 40 auf 51,85 und stiegen wieder in Woche 64 auf 53,95. Weder zu Woche 40 noch zu Woche 64 lag ein statistisch signifikanter Behandlungsunterschied zwischen den Behandlungsgruppen vor ($p_{40} = 0,2841$; $p_{64} = 0,8337$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 8: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren

Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

4.3.1.3.1.7 Sicherheit – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-38: Operationalisierung von Sicherheit

Studie	Operationalisierung
UX023-CL301	<p>Unerwünschte Ereignisse (UEs)</p> <p>Ein unerwünschtes Ereignis (UE) ist jegliche unerwünschte medizinische Erscheinung bei einem Patienten, unabhängig davon, ob das Ereignis in kausalem Zusammenhang mit der Verabreichung der Studienmedikation steht.</p> <p>Alle UEs wurden von dem Zeitpunkt ab, an dem der Patient oder dessen Erziehungsberechtigter die Einwilligung unterzeichnet hat bis zu 12 Wochen nach der letzten Dosis des Studienmedikaments aufgezeichnet. Bei jedem Besuch wurden die Patienten zu neuen oder laufenden UEs seit dem letzten Besuch befragt. Die Beurteilung der UEs erfolgte bei jedem Besuch der Prüfstelle.</p> <p>Der Prüfarzt beurteilt die Schwere, die Kausalität (Zusammenhang mit der Studienmedikation) und die Bedeutung jedes UE.</p> <p><i>Schwere der UEs</i></p> <p>Der Prüfarzt schätzt die Schwere aller UEs gemäß der <i>National Cancer Institute Common Terminology Criteria for Adverse Events (NST CTCAE)</i> ein, die wie folgt definiert sind:</p> <ul style="list-style-type: none"> • Mild: Jedes UE, das normalerweise vorübergehend ist und nur minimale Behandlung oder therapeutisches Eingreifen benötigt. Das Ereignis beeinflusst normale tägliche Aktivitäten generell nicht. • Moderat: Jedes UE, das normalerweise mit zusätzlicher spezifischer therapeutischer Intervention verringert wird. Das Ereignis beeinflusst normale tägliche Aktivitäten, verursacht Unwohlsein, aber stellt kein signifikantes oder dauerhaftes Risiko für die Studienteilnehmer dar. • Schwerwiegend (Stufe 3): Jedes UE, das normale tägliche Aktivitäten unterbricht oder das den klinischen Zustand signifikant beeinflusst oder das intensives therapeutisches Eingreifen notwendig macht. • Lebensbedrohlich (Grad 4): Ereignisse, bei denen der Studienteilnehmer ein Risiko zu versterben hat. • Tod (Grad 5): Ereignisse, die zum Tod führen. <p><i>Kausalität der UEs</i></p> <p>Der Prüfarzt bewertete die potenzielle Beziehung der UEs der Studie in Abhängigkeit des Auftretens durch die Medikation. Hierzu dienen folgende Kategorien:</p> <ul style="list-style-type: none"> • Definitiv nicht zusammenhängend: Diese Kategorie bezieht sich auf UEs, die eindeutig nicht mit der Studienmedikation/dem Verfahren zusammenhängen und bei denen eine andere plausible Ursache für das Ereignis zu sehen ist und/oder eine klinisch plausible zeitliche Abfolge mit dem Beginn des Ereignisses und der Exposition gegenüber der Studienmedikation inkonsistent ist und/oder eine kausale Beziehung als biologisch unplausibel angesehen werden kann. • Wahrscheinlich nicht zusammenhängend: Diese Kategorie gilt für UEs, die wahrscheinlich durch andere Ursachen (z. B. Begleitmedikation(en), Begleiterkrankung(en)) aufgetreten sind. Das heißt, eine alternative Erklärung war wahrscheinlicher oder die zeitliche Beziehung legen nahe, dass kein kausaler Zusammenhang zur Studienmedikation besteht. • Eventuell zusammenhängend: Diese Kategorie bezieht sich auf UEs, die

möglicherweise mit der Studienmedikation/dem Verfahren zusammenhängen könnten. Der zeitliche Zusammenhang des klinischen Ereignisses mit der Verabreichung der Studienmedikation und der unerwartet auftretenden Reaktion könnte leicht durch andere Faktoren verursacht worden sein.

- Wahrscheinlich zusammenhängend: Der zeitliche Zusammenhang des klinischen Ereignisses mit der Verabreichung der Studienmedikation macht einen kausalen Zusammenhang denkbar und nicht unwahrscheinlich. Andere Medikamente, therapeutische Interventionen und Basiszustände liefern keine ausreichende Erklärung für das beobachtete Ereignis.
- Definitiv zusammenhängend: Begründeter, zeitlicher Zusammenhang des klinischen Ereignisses mit der Verabreichung der Studienmedikation UND keine sinnvolle Erklärung des Ereignisses durch andere Faktoren (z. B. der klinische Zustand der Person, Begleittherapie und/oder andere Interventionen).

Therapie-assoziierte UEs waren Ereignisse, die eventuell, wahrscheinlich oder definitiv zusammenhängend kategorisiert wurden.

Schwerwiegendes unerwünschtes Ereignis (SUE)

Ein schwerwiegendes unerwünschtes Ereignis ist jegliches UE, das folgendes ist oder zu folgendem führt:

- Tod
- unmittelbar lebensbedrohend
- benötigte stationäre Hospitalisierung oder Verlängerung einer bereits vorhandenen Hospitalisierung
- persistierende oder signifikante Unfähigkeit oder wesentliche Störung der Fähigkeit normale Lebensfunktionen zu erfüllen
- kongenitale Anomalie/Geburtsfehler

Ein wichtiges medizinisches Ereignis, das – basierend auf angemessener medizinischer Einschätzung – den Patienten gefährdet und medizinisches oder chirurgisches Eingreifen benötigen könnte, um eines der oben gelisteten Ereignisse zu verhindern.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-39: Bewertung des Verzerrungspotenzials für Sicherheit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UX023-CL301	niedrig	nein ^a	ja	ja	ja	niedrig
^a Bei der UX023-CL301 handelt es sich um eine offene Phase-III-Studie.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Es handelt sich bei der Studie UX023-CL301 um eine randomisierte, offene und aktiv kontrollierte Phase-III-Studie. Das offene Design wurde gewählt, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, vor allem bei Kindern problematisch wäre. Es liegt eine ergebnisunabhängige Berichterstattung vor. Die FAS-Population ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS alle randomisierten Patienten. Daher wird die FAS als eine adäquate Umsetzung des ITT-Prinzips verstanden.

Abschließend wird das Verzerrungspotenzial für den Endpunkt „Sicherheit“ in der Studie UX023-CL301 somit als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-40: Ergebnisse für Unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse, die zum Therapieabbruch führen und Todesfälle aus RCT mit dem zu bewertenden Arzneimittel - Effektschätzer

UX023-CL301 Population (Safety)	Patienten n (%)	OR^a [95% KI] p-Wert	RR^a [95% KI] p-Wert	ARR^a [95% KI] p-Wert
Baseline zu Woche 40				
Gesamtrate UEs				
Aktive Kontrolle ^b (32)	26 (81,3)	6,121 [0,652; 57,447]	1,188 [0,992; 1,423]	0,153 [0,002; 0,304]
Burosumab (29)	28 (96,6)	0,1107	0,1064	0,1064
Gesamtrate SUEs				
Aktive Kontrolle ^b (32)	1 (3,1)	3,473 [0,320; 37,691]	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]
Burosumab (29)	3 (10,3)	0,3002	0,3385	0,3385
UEs Grad 3, 4 und 5				
Aktive Kontrolle ^b (32)	2 (6,3)	2,832 [0,440; 18,213]	2,207 [0,436; 11,166]	0,075 [-0,076; 0,226]
Burosumab (29)	4 (13,8)	0,2673	0,4107	0,4107
Studienabbruch durch UEs				
Aktive Kontrolle ^b (32)	0 (0)	-	-	-
Burosumab (29)	0 (0)	-	-	-
Therapieabbruch durch UEs				
Aktive Kontrolle ^b (32)	0 (0)	-	-	-
Burosumab (29)	0 (0)	-	-	-
Tod durch UEs				
Aktive Kontrolle ^b (32)	0 (0)	-	-	-
Burosumab (29)	0 (0)	-	-	-
Baseline zu Woche 64				
Gesamtrate UEs				
Aktive Kontrolle ^b (32)	27 (84,4)	n.b.	1,185 [1,021; 1,376]	0,156 [0,030; 0,282]
Burosumab (29)	29 (100,0)		0,0538	0,0538
Gesamtrate SUEs				
Aktive Kontrolle ^b (32)	3 (9,4)	1,133 [0,197; 6,520]	1,115 [0,207; 6,018]	0,010 [-0,140; 0,160]
Burosumab (29)	3 (10,3)	0,8868	1,0000	1,0000

UX023-CL301 Population (Safety)	Patienten n (%)	OR^a [95% KI] p-Wert	RR^a [95% KI] p-Wert	ARR^a [95% KI] p-Wert
UEs Grad 3, 4 und 5				
Aktive Kontrolle ^b (32)	3 (9,4)	1,808 [0,338; 9,664]	1,547 [0,316; 7,582]	0,044 [-0,117; 0,205]
Burosumab (29)	4 (13,8)	0,4824	0,6988	0,6988
Studienabbruch durch UEs				
Aktive Kontrolle ^b (32)	0 (0)	-	-	-
Burosumab (29)	0 (0)	-	-	-
Therapieabbruch durch UEs				
Aktive Kontrolle ^b (32)	0 (0)	-	-	-
Burosumab (29)	0 (0)	-	-	-
Tod durch UEs				
Aktive Kontrolle ^b (32)	0 (0)	-	-	-
Burosumab (29)	0 (0)	-	-	-
<p>N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; UEs: unerwünschte Ereignisse; SUEs: schwerwiegende unerwünschte Ereignisse; [95% KI]: 95%-iges Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio (relatives Risiko); ARR: Absolute Risk Reduction (Absolute Risikoreduktion); n.b.: nicht berechenbar</p> <p>^a: Berechnung Effektmodifikatoren (OR, RR, ARR) erfolgte post-hoc</p> <p>^b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin</p>				

Tabelle 4-41: Ergebnisse für Unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse, die zum Therapieabbruch führen und Todesfälle aus RCT mit dem zu bewertenden Arzneimittel

UX023-CL301 Population (N) (Safety)	Woche 40 Patienten n (%)	Woche 64 Patienten n (%)	Extensionsphase^a Patienten n (%)
Gesamtrate UEs			
Aktive Kontrolle ^b (32)	26 (81,3)	27 (84,4)	13 (59,1)
Burosumab (29)	28 (96,6)	29 (100,0)	29 (100,0)
Gesamtrate SUEs			
Aktive Kontrolle ^b (32)	1 (3,1)	3 (9,4)	0 (0,0)
Burosumab (29)	3 (10,3)	3 (10,3)	4 (13,8)
Therapie-assoziierte UEs			
Aktive Kontrolle ^b (32)	4 (12,5)	7 (21,9)	13 (59,1)
Burosumab (29)	17 (58,6)	17 (58,6)	17 (58,6)
Therapie-assoziierte SUEs			
Aktive Kontrolle ^b (32)	0 (0)	0 (0)	0 (0)
Burosumab (29)	0 (0)	0 (0)	1 (3,4)
UEs Grad 3, 4 und 5			
Aktive Kontrolle ^b (32)	2 (6,3)	3 (9,4)	1 (4,5)
Burosumab (29)	4 (13,8)	4 (13,8)	4 (13,8)
Studienabbruch durch UEs			
Aktive Kontrolle ^b (32)	0 (0)	0 (0)	0 (0)
Burosumab (29)	0 (0)	0 (0)	0 (0)
Therapieabbruch durch UEs			
Aktive Kontrolle ^b (32)	0 (0)	0 (0)	0 (0)
Burosumab (29)	0 (0)	0 (0)	0 (0)
Tod durch UEs			
Aktive Kontrolle ^b (32)	0 (0)	0 (0)	0 (0)
Burosumab (29)	0 (0)	0 (0)	0 (0)
N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; UEs: unerwünschte Ereignisse; SUEs: schwerwiegende unerwünschte Ereignisse; [95% KI]: 95%-iges Konfidenzintervall;			
^a Extensionsphase: Eine Burosumab-Behandlung für alle Studienteilnehmer (Ausnahme: Patienten aus Japan und Korea) war ab Woche 66 bis zu Woche 140 möglich.			
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin			

Die Auswertung der unerwünschten Ereignisse (UEs) ergab keine signifikanten Sicherheitsbedenken (Ausnahme: Therapie-assoziierte UEs). Während der Studie wurde kein Kind aufgrund von UEs oder aus einem anderen Grund aus der Studie ausgeschlossen. Kein Studienteilnehmer verstarb. Die meisten Studienteilnehmer hatten mindestens ein UE (Burosumab: 100 %; Aktive Kontrolle: 84 %). Bei 7 Patienten ($n_{\text{Burosumab}} = 4$; $n_{\text{aktive Kontrolle}} = 3$) traten UEs Grad 3 auf. Die restlichen UEs waren mild oder moderat (Grad 1 oder 2) und es wurden keine lebensbedrohlichen UEs Grad 4 gemeldet.

Unerwünschte Ereignisse, die durch den Prüfarzt als möglicherweise mit der Studienmedikation in Zusammenhang gebracht werden konnten, sind bei 59 % der Kinder in der Burosumab-Gruppe und 22 % der Kinder in der aktiven Kontrollgruppe aufgetreten. Der Großteil der therapie-assoziierten UEs waren Reaktionen an der Injektionsstelle.

Extensionsphase (Woche 66 bis 140)

Die Bewertung der Extensionsphase ergab keine Sicherheitsbedenken. Kein Patient wurde aufgrund von UEs aus der Studie ausgeschlossen. Es wurde kein Todesfall berichtet. 59 % der Patienten, die von der Substitutionstherapie auf die Burosumab-Behandlung wechselten, berichteten von einem UE. Die meisten dieser Ereignisse waren ebenfalls Reaktionen an der Injektionsstelle. In der Burosumab-Burosumab-Gruppe trat während der Extensionsphase ein schweres unerwünschtes Ereignis auf. Dabei handelte es sich um ein bilaterales Papillenödem, welches vom Prüfarzt als schwer eingestuft wurde und möglicherweise mit der Burosumab-Behandlung in Verbindung gebracht werden kann. Zu beachten ist jedoch, dass dieses Kind bereits eine Kraniosynostose vor Studienbeginn erlitt, die mit intrakraniellm Druck und ophthalmologischen Komplikationen verbunden war und das Papillenödem dadurch bedingt sein könnte. Während der Extensionsphase waren alle UEs mild oder moderat mit Ausnahme von zwei UEs Grad 3. Diese traten bei den beiden Patienten auf, die während der Behandlungsphase bereits schwere UEs berichteten. Eines war das oben genannte bilaterale Papillenödem. Bei dem zweiten schweren UE handelte es sich um einen Zahnabszess, welcher in der Substitutions-Burosumab-Gruppe (während der Burosumab-Behandlung) berichtet wurde. Dieser Zahnabszess wurde nicht als schwerwiegend oder therapie-assoziiert eingestuft und konnte innerhalb von zwei Tagen mit einer Zahnextraktion behoben werden.

Tabelle 4-42: Ergebnisse für unerwünschten Ereignisse nach System Organ Class (SOC) und Preferred Term (PT) aus RCT mit dem zu bewertenden Arzneimittel

Studie UX023-CL301	Baseline zu Woche 40		Baseline zu Woche 64		Extensionsphase	
	Aktive Kontrolle ^a	Burosumab	Aktive Kontrolle ^a	Burosumab	Akt. Kontrolle ^a → Burosumab	Burosumab → Burosumab
	(N=32) n (%)	(N=29) n (%)	(N=32) n (%)	(N=29) n (%)	(N=22) n (%)	(N=29) n (%)
Organklassen (System Organ Class)						
Gesamtrate UE	26 (81,3)	28 (96,6)	27 (84,4)	29 (100,0)	13 (59,1)	29 (100,0)
Infektionen und parasitäre Erkrankungen	23 (71,9)	22 (75,9)	24 (75,0)	23 (79,3)	3 (13,6)	23 (79,3)
Nasopharyngitis	13 (40,6)	10 (34,5)	14 (43,8)	11 (37,9)	1 (4,5)	12 (41,4)
Infektionen der oberen Atemwege	2 (6,3)	3 (10,3)	3 (9,4)	3 (10,3)	0 (0,0)	3 (10,3)
Ohrinfektion	3 (9,4)	2 (6,9)	3 (9,4)	2 (6,9)	0 (0,0)	2 (6,9)
Zahnabszess	3 (9,4)	7 (24,1)	3 (9,4)	8 (27,6)	2 (9,1)	8 (27,6)
Gastroenteritis	2 (6,3)	2 (6,9)	3 (9,4)	2 (6,9)	0 (0,0)	2 (6,9)
Virale Infektionen der oberen Atemwege	2 (6,3)	3 (10,3)	2 (6,3)	0 (0,0)	0 (0,0)	0 (0,0)
Influenza	6 (18,8)	4 (13,8)	6 (18,8)	4 (13,8)	0 (0,0)	4 (13,8)
Virale Infektionen	1 (3,1)	2 (6,9)	1 (3,1)	2 (6,9)	0 (0,0)	2 (6,9)
Rhinitis	2 (6,3)	3 (10,3)	2 (6,3)	2 (6,9)	0 (0,0)	2 (6,9)
Erkrankungen des Gastrointestinaltrakts	15 (46,9)	20 (69,0)	17 (53,1)	23 (79,3)	2 (9,1)	25 (86,2)
Erbrechen	8 (25,0)	11 (37,9)	8 (25,0)	12 (41,4)	2 (9,1)	13 (44,8)
Durchfall	2 (6,3)	7 (24,1)	2 (6,3)	7 (24,1)	0 (0,0)	7 (24,1)
Zahnschmerz	1 (3,1)	4 (13,8)	1 (3,1)	4 (13,8)	0 (0,0)	5 (17,2)
Zahnen	0 (0,0)	3 (10,3)	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)
Karies	2 (6,3)	5 (17,2)	2 (6,3)	9 (31,0)	0 (0,0)	10 (34,5)
Übelkeit	0 (0,0)	2 (6,9)	1 (3,1)	3 (10,3)	1 (4,5)	4 (13,8)

Studie UX023-CL301	Baseline zu Woche 40		Baseline zu Woche 64		Extensionsphase	
	Aktive Kontrolle ^a	Burosumab	Aktive Kontrolle ^a	Burosumab	Akt. Kontrolle ^a → Burosumab	Burosumab → Burosumab
	(N=32) n (%)	(N=29) n (%)	(N=32) n (%)	(N=29) n (%)	(N=22) n (%)	(N=29) n (%)
Abdominale Beschwerden	2 (6,3)	2 (6,9)	2 (6,3)	2 (6,9)	1 (4,5)	2 (6,9)
Schmerzen im Oberbauch	3 (9,4)	3 (10,3)	3 (9,4)	3 (10,3)	0 (0,0)	4 (13,8)
Verstopfung	0 (0,0)	2 (6,9)	0 (0,0)	5 (17,2)	1 (4,5)	5 (17,2)
Mundulzera	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	8 (25,0)	22 (75,9)	8 (25,0)	25 (86,2)	9 (40,9)	25 (86,2)
Reaktion an der Injektionsstelle	0 (0,0)	7 (24,1)	0 (0,0)	7 (24,1)	2 (9,1)	7 (24,1)
Erythem an der Injektionsstelle	0 (0,0)	6 (20,7)	0 (0,0)	9 (31,0)	5 (22,7)	9 (31,0)
Pyrexie	6 (18,8)	15 (51,7)	6 (18,8)	16 (55,2)	2 (9,1)	16 (55,2)
Juckreiz an der Injektionsstelle	0 (0,0)	3 (10,3)	0 (0,0)	3 (10,3)	2 (9,1)	3 (10,3)
Schwellung an der Injektionsstelle	0 (0,0)	2 (6,9)	0 (0,0)	3 (10,3)	1 (4,5)	3 (10,3)
Schmerz	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)	0 (0,0)	3 (10,3)
Fatigue	2 (6,3)	1 (3,4)	2 (6,3)	1 (3,4)	0 (0,0)	2 (6,9)
Erosion an der Injektionsstelle	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)
Hautausschlag an der Injektionsstelle	0 (0,0)	3 (10,3)	0 (0,0)	3 (10,3)	0 (0,0)	3 (10,3)
Urtikaria an der Injektionsstelle	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	9 (28,1)	19 (65,5)	9 (28,1)	21 (72,4)	3 (13,6)	23 (79,3)
Husten	6 (18,8)	12 (41,4)	6 (18,8)	15 (51,7)	1 (4,5)	15 (51,7)
Oropharyngealer Schmerz	1 (3,1)	4 (13,8)	1 (3,1)	5 (17,2)	1 (4,5)	5 (17,2)
Nasenverstopfung	1 (3,1)	3 (10,3)	1 (3,1)	5 (17,2)	0 (0)	6 (20,7)
Rhinorrhoe	2 (6,3)	5 (17,2)	2 (6,3)	7 (24,1)	1 (4,5)	7 (24,1)
Asthma	1 (3,1)	4 (13,8)	1 (3,1)	4 (13,8)	0 (0,0)	4 (13,8)

Studie UX023-CL301	Baseline zu Woche 40		Baseline zu Woche 64		Extensionsphase	
	Aktive Kontrolle ^a	Burosumab	Aktive Kontrolle ^a	Burosumab	Akt. Kontrolle ^a → Burosumab	Burosumab → Burosumab
	(N=32) n (%)	(N=29) n (%)	(N=32) n (%)	(N=29) n (%)	(N=22) n (%)	(N=29) n (%)
Erkrankungen des Nervensystems	8 (25,0)	10 (34,5)	9 (28,1)	12 (41,4)	2 (9,1)	12 (41,4)
Kopfschmerzen	5 (15,6)	8 (27,6)	6 (18,8)	10 (34,5)	2 (9,1)	10 (34,5)
Schwindel	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)
Migräne	2 (6,3)	1 (3,4)	2 (6,3)	1 (3,4)	0 (0,0)	1 (3,4)
Skelettmuskulatur und Bindegewebserkrankungen	13 (40,6)	15 (51,7)	15 (46,9)	17 (58,6)	3 (13,6)	17 (58,6)
Schmerzen in den Extremitäten	9 (28,1)	10 (34,5)	10 (31,3)	11 (37,9)	2 (9,1)	11 (37,9)
Gelenkschmerzen (Arthralgie)	7 (21,9)	11 (37,9)	10 (31,3)	13 (44,8)	0 (0,0)	13 (44,8)
Fisteln	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)
Rückenschmerzen	1 (3,1)	1 (3,4)	3 (9,4)	2 (6,9)	0 (0,0)	2 (6,9)
Nackenschmerzen	1 (3,1)	0 (0,0)	1 (3,1)	1 (3,4)	0 (0,0)	1 (3,4)
Enthesopathie	1 (3,1)	0 (0,0)	1 (3,1)	0 (0,0)	0 (0,0)	0 (0,0)
Unfälle, Vergiftungs- und Handhabungskomplikationen	2 (6,3)	8 (27,6)	2 (6,3)	10 (34,5)	0 (0,0)	10 (34,5)
Prellung	0 (0,0)	4 (13,8)	0 (0,0)	4 (13,8)	0 (0,0)	4 (13,8)
Erkrankungen der Haut und Unterhautzellgewebe	2 (6,3)	10 (34,5)	4 (12,5)	11 (37,9)	0 (0,0)	13 (44,8)
Hautauschlag	2 (6,3)	13 (10,3)	2 (6,3)	3 (10,3)	0 (0,0)	4 (13,8)
Trockene Haut	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)
Erythem	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)	0 (0,0)	2 (6,9)
Untersuchungen	4 (12,5)	8 (27,6)	4 (12,5)	8 (27,6)	0 (0,0)	8 (27,6)
Verringertes Vitamin D	1 (3,1)	6 (20,7)	1 (3,1)	6 (20,7)	0 (0,0)	6 (20,7)
Verringertes 25-Hydroxycholecalciferol im Blut	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)	0 (0,0)	1 (3,4)
Erkrankungen des Immunsystems	5 (15,6)	5 (17,2)	5 (15,6)	5 (17,2)	0 (0,0)	5 (17,2)
Saisonale Allergie	3 (9,4)	4 (13,8)	2 (6,3)	4 (13,8)	0 (0,0)	4 (13,8)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie UX023-CL301	Baseline zu Woche 40		Baseline zu Woche 64		Extensionsphase	
	Aktive Kontrolle ^a	Burosumab	Aktive Kontrolle ^a	Burosumab	Akt. Kontrolle ^a → Burosumab	Burosumab → Burosumab
	(N=32) n (%)	(N=29) n (%)	(N=32) n (%)	(N=29) n (%)	(N=22) n (%)	(N=29) n (%)
Hypersensitivität	1 (3,1)	1 (3,4)	2 (6,3)	1 (3,4)	0 (0,0)	1 (3,4)
Erkrankungen des Ohrs und Labyrinths	2 (6,3)	5 (17,2)	3 (9,4)	6 (20,7)	2 (9,1)	6 (20,7)
Ohrschmerz	1 (3,1)	3 (10,3)	1 (3,1)	4 (13,8)	2 (9,1)	4 (13,8)
Metabolische und ernährungsbedingte Störungen	3 (9,4)	5 (17,2)	3 (9,4)	6 (20,7)	0 (0,0)	6 (20,7)
Vitamin-D-Mangel	1 (3,1)	4 (13,8)	1 (3,1)	5 (17,2)	0 (0,0)	5 (17,2)
Reduzierter Appetit	2 (6,3)	1 (3,4)	2 (6,3)	1 (3,4)	0 (0,0)	1 (3,4)
N: Anzahl der Studienpopulation; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N						
^a : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						

Die häufigsten UEs waren entweder typisch bei einer Manifestation der XLH (Arthralgie, Schmerzen in den Extremitäten, Zahnabszesse und Vitamin D-Mangel), typisch für eine pädiatrische Population (Pyrexie, Husten, Erbrechen, Nasopharyngitis, Kopfschmerzen und Durchfall) oder zu erwartende Ereignisse (Subkutane Injektion: Reaktionen und Erytheme an der Injektionsstelle; frühere klinische Studien: Kopfschmerzen, Schwindel, Zahnschmerzen, Hautausschlag, Myalgie und Schmerzen in den Extremitäten).

Zu den am häufigsten berichteten UEs gehörten typische Kindheitssymptome: Pyrexie (Burosumab: 55 % vs. aktive Kontrolle: 19 %), Husten (52 % vs. 19 %), Erbrechen (41 % vs. 25 %), Kopfschmerzen (35 % vs. 19 %) und Durchfall (24 % vs. 6 %). Die Inzidenz typischer Kinderinfektionen war zwischen den Behandlungsgruppen ähnlich wie z. B. Nasopharyngitis (38 % vs. 44 %) und Grippe (14 % vs. 19 %), ebenso wie die Inzidenz von UEs bei SOC-Infektionen und Infektionen (79 % vs. 75 %). Keines der Ereignisse von Pyrexie, Husten, Erbrechen oder Durchfall wurden vom Prüfarzt als schwerwiegendes Ereignis und auch nicht als mit der Burosumab-Behandlung zusammenhängend bewertet. Pyrexie ging in der Burosumab-Gruppe (16 Kinder) zeitgleich mit Erbrechen (5 Kinder), Durchfall (3 Kinder) und Husten (3 Kinder) einher. 4 Kinder mit Pyrexie hatten auch gleichzeitig Infektionen: Mittelohrentzündung (2 Kinder), Gastroenteritis (1 Kind) und Ohrinfektion (1 Kind). Ereignisse von Pyrexie, Husten, Erbrechen oder Durchfall waren zeitlich nicht mit der Verabreichung von Burosumab verbunden, mit Ausnahme von 4 Patienten. Diese berichteten von einer Pyrexie am selben Tag der Burosumab-Injektion. Auch Kopfschmerzen wurden in der Burosumab-Gruppe oft gleichzeitig mit anderen Krankheiten wie Pyrexie, Erbrechen und Zahnschmerzen berichtet.

Die Gründe für die unterschiedlichen Häufigkeiten dieser typischen Kindheitssymptome (Pyrexie, Husten, Erbrechen, Kopfschmerzen und Durchfall) zwischen den Behandlungsgruppen sind nicht vollständig verstanden. In den meisten Fällen standen die Ereignisse jedoch nicht mit der Verabreichung von Burosumab in einem zeitlichen Kausalzusammenhang. Ein weiterer wichtiger Faktor ist das Studiendesign dieser Open-Label-Studie: Die Patienten der Burosumab-Gruppe haben von einer bekannten, oral verabreichten Therapie zu einem neuartigen Prüfpräparat gewechselt, welches außerdem subkutan verabreicht wurde.

Die Studienteilnehmer mussten vor der Aufnahme orales Phosphat und aktives Vitamin D (≥ 12 Monate für Kinder ≥ 3 Jahre; ≥ 6 Monate für Kinder < 3 Jahre) erhalten haben. Die durchschnittliche Dauer der Substitutionstherapie betrug 3,8 Jahre vor Beginn der Studie. Die Studienpopulation ist daher an eine mehrmals tägliche Dauertherapie mit oralem Phosphat und aktivem Vitamin D gewöhnt, was die Wahrnehmung und Meldung von UEs im Sinne eines potenziellen *Underreporting* beeinflussen kann. Eine weitere Überlegung ist, dass erhöhte FGF23-Konzentrationen die Rekrutierung von Leukozyten begrenzen und eine Immunabwehr beeinträchtigen, indem sie die Aktivierung und Rekrutierung von Leukozytenintegrin stören [57]. Da Burosumab FGF23 hemmt, wird angenommen, dass Patienten, die Burosumab erhalten, eine robustere Immunantwort erfahren könnten. Dies führt möglicherweise zu einer häufigeren Ausprägung der typischen Kindheitssymptome und weiterer Erkrankungen wie

Zahnabszesse und Überempfindlichkeitsereignisse (z. B. Hautausschlag), die in der Burosumab-Gruppe beobachtet wurden. Häufige Manifestationen von XLH und/oder bekannten Nebenwirkungen von Burosumab, wie sie aus früheren klinischen Studien identifiziert wurden (Zahnabszesse, Kopfschmerzen, Schwindel, Zahnschmerzen, Hautausschlag, Myalgie, Schmerzen in den Extremitäten, Reaktion an der Injektionsstelle, Vitamin-D-Mangel), gehörten zu den am häufigsten berichteten UEs: Arthralgie (Burosumab: 45 % vs. aktive Kontrolle: 31 %), Schmerzen in den Extremitäten (38 % vs. 21 %), Zahnabszesse (28 % vs. 9 %) und Vitamin-D-Mangel (21 % vs. 3 %). Arthralgie kann auf die Rachitissymptomatik zurückzuführen sein oder mit dem Wachstum und der körperlichen Aktivität zusammenhängen. Kinder berichteten von Beinschmerzen, die als "Wachstumsschmerzen" beschrieben werden [58]. Außerdem nahm die Körpergröße der Kinder in der Burosumab-Gruppe gegenüber Baseline signifikant stärker zu als in der aktiven Kontrollgruppe. Eine größere Schmerzzinzidenz könnte auch auf eine erhöhte körperliche Aktivität in der Burosumab-Gruppe zurückzuführen sein, die durch die 6MWT-Daten unterstützt wird. Die Verbesserung der PROMIS-Schmerzdomäne deutet ebenfalls darauf hin, dass die Kinder aufgrund des verringerten Schmerzempfindens aktiver sein könnten.

Die UEs waren leicht oder moderat, mit Ausnahme der SUEs, die bei 7 Patienten auftraten (4 in der Burosumab-Gruppe, 3 in der aktiven Kontrollgruppe). Ausschließlich das oben beschriebene UE Papillenödem (Burosumab-Gruppe während der Extensionsphase) war schwerwiegend.

Keines der SUEs wurde vom Prüfarzt als mit dem Studienmedikament assoziiert angesehen, mit Ausnahme eines Ereignisses mit Arthralgie Grad 3 in der Burosumab-Gruppe während der Behandlungsphase (Woche 0-64) und des oben beschriebenen UEs des Papillenödems (Burosumab-Gruppe) während der Extensionsphase. Daraus resultierten keine Änderungen der Verabreichung der Studienmedikation.

Zahnmedizinische Ereignisse

Mit der XLH geht regelhaft eine schlechte Zahngesundheit einher [10]. Über Zahnabszesse, übermäßige Karies, Wurzelkanal oder Zahnextraktion (aufgrund von Karies oder Abszess) wurden bei 10/29 Patienten (34 %) in der Burosumab-Gruppe und 15/32 (47 %) in der aktiven Kontrollgruppe vor Einschluss in die Studie berichtet. Bei jeder Visite wurden die Patienten mündlich gefragt, ob seit der letzten Visite eins der folgenden Ereignisse aufgetreten ist: Zahnabszess, Zahnextraktion (durch Abszess oder Karies), Wurzelbehandlung und/oder Gingivitis.

Die Ergebnisse dieser mündlichen Prüfung wurden getrennt von den UEs tabellarisch dargestellt. Der Prüfarzt beurteilte, ob es sich um Zahnbeschwerden oder ein UE handelt. Zu Woche 64 wurden insgesamt bei 12/29 Patienten (41 %) in der Burosumab-Gruppe und bei

5/32 Patienten (16 %) in der aktiven Kontrollgruppe Zahnbeschwerden gemeldet. Die Zahnbeschwerden traten in beiden Gruppen meist bei Milchzähnen und nicht bei den bleibenden Zähnen auf. Die Burosumab-Gruppe war zudem im Mittel ca. 6 Monate jünger als die Kontrollgruppe. Von den Patienten mit Zahnbeschwerden (Post-Baseline) hatten 7/12 (58 %) in der Burosumab-Gruppe und 3/5 (60 %) in der aktiven Kontrollgruppe eine Vorgeschichte von zahnärztlichen Abszessen, exzessive Karies, Wurzelbehandlungen und/oder Zahnextraktionen (siehe Tabelle 4-43).

Tabelle 4-43: Zusammenfassung der zahnmedizinischen Untersuchungen

Studie UX023-CL301	Aktive Kontrolle ^a	Burosumab
	(N=32) n (%)	(N=29) n (%)
Mündliche zahnmedizinische Befragung		
Jegliche Zahnbeschwerden zu Woche 64	5 (15,6)	12 (41,4)
Zahnabszess	2 (6,3)	10 (34,5)
Zahnextraktion (durch Abszess oder Karies)	3 (9,4)	5 (17,2)
Wurzelbehandlung	0 (0,0)	3 (10,3)
Karies	1 (3,1)	6 (20,7)
Gingivitis	0 (0,0)	0 (0,0)
N: Anzahl der Studienpopulation; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N		
^a : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D		

Dental-assoziierte-UEs wie z. B. Zahnabszesse, Zahnschmerzen und Parodontitis wurden bei 15/29 Kindern (52 %) in der Burosumab-Gruppe und bei 10/32 Kindern (31 %) in der aktiven Kontrollgruppe berichtet (siehe Tabelle 4-44). Dentale UEs wurden bei allen Kindern dokumentiert, die nach Aussage der mündlichen zahnmedizinischen Befragung an Zahnbeschwerden litten ($n_{\text{Burosumab}} = 12$; $n_{\text{aktive Kontrolle}} = 5$). Ebenfalls wurden dentale UEs für Kinder notiert, die in der mündlichen zahnmedizinischen Befragung keine Zahnbeschwerden angegeben hatten ($n_{\text{Burosumab}} = 4$; $n_{\text{aktive Kontrolle}} = 5$). Nicht berücksichtigt wurde hierbei das dentale UE „Zahnen“, welches bei einem Kind in der Burosumab-Gruppe aufgetreten war. Von diesen Kindern mit UEs hatten 7/15 (48 %) in der Burosumab-Gruppe und 6/10 (60 %) in der aktiven Kontrollgruppe eine Vorgeschichte von Zahnabszessen, übermäßigen Karies, Wurzelbehandlung und/oder Zahnextraktionen.

Tabelle 4-44: Ergebnisse der dentalen unerwünschten Ereignisse

Studie UX023-CL301	Aktive Kontrolle ^a	Burosumab
	(N=32) n (%)	(N=29) n (%)
Zähne betreffende unerwünschte Ereignisse		
Gesamtrate dentale UEs	10 (31,3)	15 (51,7)
Zahnabszesse	3 (9,4)	8 (27,6)
Karies	2 (6,3)	9 (31,0)
Zahnschmerz	1 (3,1)	4 (13,8)
Periodontitis	1 (3,1)	1 (3,4)
Zahnschmelzhypoplasie	0 (0,0)	2 (6,9)
Loser Zahn	0 (0,0)	2 (6,9)
Zahnverlust	0 (0,0)	1 (3,1)
Malokklusion	1 (3,1)	0 (0,0)
Zahnfehlstellung	1 (3,1)	0 (0,0)
Zahnverfärbung	1 (3,1)	0 (0,0)
Mündliche zahnmedizinische Befragung (jegliche Zahnbeschwerden zu Woche 64)	5 (15,6)	12 (41,4)
N: Anzahl der Studienpopulation; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N ^a : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D		

Zahnärztliche Komplikationen traten in der Burosumab-Gruppe häufiger auf. Bei der Zahngesundheit werden verschiedene anatomische Merkmale berücksichtigt, darunter Zahnschmelzhypoplasie, geringe Mineralisierung des Dentins, vergrößerte Pulpakammern und parodontales Gewebe. Diese abnormalen anatomischen Merkmale ermöglichen eine stärkere bakterielle Infiltration, die zu Zahnabszessen und Karies führen kann. Das Fehlen von UEs bei der Substitutionstherapie sowie die höhere Inzidenz der am häufigsten gemeldeten UEs in der Burosumab-Gruppe stehen wahrscheinlich im Zusammenhang mit dem Design dieser offenen Studie im Hinblick auf die Fortsetzung oder den Übergang von der vorherigen Therapie (von einer oralen Substitutionstherapie zu einer subkutanen neuartigen Burosumab-Therapie).

Das Sicherheitsprofil von Burosumab war in dieser Studie mit den früheren Studien (UX023-CL201 und UX023-CL205) bei Kindern mit XLH vergleichbar.

Unerwünschtes Ereignis von besonderem Interesse

Reaktionen an der Injektionsstelle und damit Ereignisse von besonderem Interesse, die nur bei Patienten in der Burosumab-Gruppe untersucht wurden, wurden bei 15/29 Kindern (52 %) während der Behandlungsphase berichtet. Während der Extensionsphase stieg die Anzahl der Fälle in der Burosumab-Gruppe nicht weiter an. Nachdem die aktive Kontrollgruppe während

der Extensionsphase Burosumab erhielt, wurde bei 7/22 Kindern (32 %) eine Reaktion an der Injektionsstelle berichtet.

Hypersensitivität war ebenfalls ein UE von besonderem Interesse, welches bei 11/29 Patienten (38 %) in der Burosumab-Gruppe und 6/32 Patienten (19 %) in der aktiven Kontrollgruppe beobachtet wurde. Während der Extensionsphase wurde bei zwei weiteren Patienten in der Burosumab-Burosumab-Gruppe eine Hypersensitivität beobachtet. Die Substitutions-Burosumab-Gruppe blieb unverändert.

Bis zum Datenschnitt erlebte kein Kind der beiden Gruppen ein UE von besonderem Interesse wie eine Hyperphosphatämie, eine ektopische Mineralisierung oder ein Restless-Legs-Syndrom.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 9: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es liegt lediglich eine einzige randomisierte und kontrollierte Studie mit der Patientenpopulation vor, daher kann keine Metaanalyse durchgeführt werden.

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Die Darstellung der Subgruppen erfolgt im vorliegenden Dossier auf Grundlage der offenen, randomisierten und aktiv kontrollierten Phase-III-Studie. Die folgenden Subgruppen wurden für alle patientenrelevanten Endpunkte untersucht:

- Krankheitsschwere (RSS Total Score $\geq 2,0$ vs. RSS Total Score $< 2,0$)
- Alter (< 5 Jahre vs. ≥ 5 Jahre)
- Geschlecht (weiblich vs. männlich)
- Region (Japan vs. Rest der Welt)
- *PHEX*-Mutation (positiv vs. nicht-positiv)

Die Prüfung des Vorliegens einer möglichen Effektmodifikation erfolgte, wie in Abschnitt 4.2.5.5 beschrieben, innerhalb eines strukturierten, mehrstufigen Prozesses, dessen zentraler Ausgangspunkt die Durchführung eines geeigneten Interaktions- oder Homogenitätstests ist.

Bei der Interpretation der Ergebnisse der Interaktionstests sind folgende Probleme unbedingt zu berücksichtigen:

- Das Problem der sehr kleinen Patientenpopulation pro Behandlungsarm. Dies liegt bei den Subgruppenanalysen *PHEX*-Mutation und Region vor.

- Das Problem der Multiplizität, dass bei der Testung einer größeren Anzahl statistischer Hypothesen – wie hier durch die Durchführung der Interaktionstests – eine gewisse Anzahl der Tests bereits aufgrund der angenommenen Irrtumswahrscheinlichkeit von 5 % zu einem falsch positiven Ergebnis führt.

Im Folgenden werden im Abschnitt 4.3.1.3.2.1 zunächst die Ergebnisse der Interaktionstests für alle untersuchten Subgruppen dargestellt. Entsprechend dem strukturierten, mehrstufigen Vorgehen zur Bewertung einer möglichen Effektmodifikation erfolgt anschließend im Abschnitt 4.3.1.3.2.2 eine ausführliche Ergebnisdarstellung für diejenigen Subgruppen, bei denen ein signifikanter Interaktionstest ($p < 0,05$) vorliegt. Die Ergebnisse der Subgruppen mit nicht signifikantem Interaktionstest ($p \geq 0,05$) wurden der Vollständigkeit halber aber dennoch dem Nutzendossier beigelegt.

4.3.1.3.2.1 Überblick über Ergebnisse der Interaktionstests aus Subgruppenanalysen

Die folgende Tabelle 4-45 zeigt in einer Übersicht die Ergebnisse sämtlicher Interaktionstests, die für Subgruppenanalysen der UX023-CL301 durchgeführt wurden.

Insgesamt wurden 10 positive Interaktionstests beobachtet:

- Subgruppe Geschlecht:

Positive Interaktionstests für FPS-R zu Woche 40 ($p = 0,0043$) und zu Woche 64 ($p = 0,0004$); für Serumphosphat – Prozentsatz erreichter Normwert zu Woche 40 ($p = 0,0175$) und zu Woche 64 ($p = 0,0329$) [59, 60]

- Subgruppe Region:

Positive Interaktionstests für RGI-C zu Woche 64 ($p = 0,0136$) [61]; für Körpergröße zu Woche 64 ($p = 0,0168$) [62]; für Serumphosphat zu Woche 40 ($p = 0,0279$) und Woche 64 ($p = 0,0414$) [63]; für Serumphosphat – Prozentsatz erreichter Normwert zu Woche 40 ($p = 0,0037$) und zu Woche 64 ($p = 0,0040$) [64]

Wie bereits im Abschnitt 4.2.5.5 beschrieben, erfolgt aufgrund der zu geringen Patientenzahl pro Behandlungsarm der Subgruppenanalyse Region keine Interpretation. Trotz des signifikanten Interaktionstest ist hier von einem homogenen Effekt auszugehen, weshalb ausschließlich das Ergebnis der Gesamtstudienpopulation relevant ist, welche unter den Abschnitten 4.3.1.3.1.1, 4.3.1.3.1.2 und 4.3.1.3.1.4 dargestellt ist.

Tabelle 4-45: Subgruppenanalysen für die Endpunkte der Studie UX023-CL301

<i>OUTCOMES</i>	p-Werte der Interaktion									
	Subgruppe	Rachitis-Schwere		Alter		Geschlecht		Region		PHEX Mutation
Endpunkt										
Datenschnitt	W40	W64	W40	W64	W40	W64	W40	W64	W40	W64
Morbidität										
RGI-C Global Score	0,9828	0,7551	0,8392	0,8777	0,3834	0,7074	0,1414	0,0136	0,7140	0,5295
RGI-C Global Score =+ 2,0	0,9784	0,9744	0,7035	0,8945	0,3925	0,9685	0,9953	0,9915	0,9822	0,9807
RSS	0,9910	0,9528	0,2974	0,1872	0,9794	0,8650	0,2010	0,0930	0,9040	0,5885
Körpergröße (Standhöhe/Lieglänge)	0,7451	0,9458	0,1254	0,1816	0,6319	0,8412	0,0855	0,0168	0,6107	0,6179
6MWT [m]	0,7670	0,6173	-	-	0,6273	0,7328	0,9779	0,9696	0,1280	0,4032
6MWT [vorhergesagter Prozentsatz]	0,9053	0,7956	-	-	0,7356	0,8572	0,9650	0,9837	0,0978	0,3320
6MWT [m] ^a	0,7985	0,9762	-	-	0,2726	0,3538	0,9586	0,9843	0,1597	0,5905
6MWT [vorhergesagter Prozentsatz] ^a	0,8130	0,9110	-	-	0,3016	0,3733	0,9404	0,9908	0,1518	0,5485
Serumphosphat	0,1432	0,1612	0,2802	0,2669	0,0754	0,1412	0,0279	0,0414	0,3299	0,2945
Serumphosphat – Prozentsatz erreichter Normwert	0,4077	0,2517	0,8321	0,9439	0,0175	0,0329	0,0037	0,0040	0,4788	0,7242
PROMIS – Pain Domain	0,4626	0,7860	-	-	0,7002	0,3785	0,5652	0,8922	0,6038	0,6084
PROMIS – Physical Function and Morbidity Domain	0,5295	0,5334	-	-	0,8162	0,8180	0,2342	0,2365	0,5386	0,5449
PROMIS – Fatigue Domain	0,1683	0,3849	-	-	0,7523	0,7071	0,6717	0,6026	0,2281	0,2721
FPS-R	0,2424	0,6453	-	-	0,0043	0,0004	0,3390	0,1923	0,6124	0,6687
Lebensqualität										
SF-10 – Physical Health Summary Score	0,2775	0,5106	-	-	0,8104	0,7502	0,7652	0,9565	0,2130	0,2298
SF-10 – Psychosocial Summary Score	0,7624	0,7028	-	-	0,2667	0,2360	0,4844	0,8277	0,3205	0,5409

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

<i>OUTCOMES</i>		p-Werte der Interaktion									
Subgruppe	Endpunkt	Rachitis-Schwere		Alter		Geschlecht		Region		PHEX Mutation	
		W40	W64	W40	W64	W40	W64	W40	W64	W40	W64
Sicherheit – Allgemeine Sicherheitsendpunkte											
	Unerwünschte Ereignisse (UEs)	0,8048	0,8184	0,0688	0,0518	0,8708	0,9971	0,2190	0,1765	0,1383	0,5058
	Schwerwiegende unerwünschte Ereignisse (SUEs)	0,9954	0,8150	0,9965	0,9951	0,9956	0,9000	0,9999	1,0000	0,9999	0,9999
	UEs Grad 3, 4 und 5	0,9958	0,9950	0,9961	0,9957	0,2061	0,2703	1,0000	1,0000	0,9997	0,9999
	Therapie-assoziierte UEs	0,9787	0,9991	0,9760	0,9998	0,9786	0,9985	0,9753	0,9985	0,9759	0,9987
	Therapie-assoziierte SUEs	0,9768	0,9781	0,9762	0,9734	0,9706	0,7535	0,9986	0,9999	0,9989	0,9999
	Therapie-assoziierte UEs Grad 3, 4 und 5	0,9741	0,9729	0,9788	0,9778	0,2491	0,3704	0,9996	0,9999	0,9986	0,9992
	Dentale Ereignisse: Karies	0,9968	0,9953	0,9968	0,9963	0,9966	0,9959	0,9999	0,9998	0,9967	0,9951
	Dentale Ereignisse: Zahnextraktion	0,4467	0,7036	0,9999	0,9955	0,9970	0,9950	1,0000	1,0000	0,9999	0,9998
	Dentale Ereignisse: Wurzelbehandlung	1,0000	1,0000	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
	Dentale Ereignisse: Zahnabszess	0,9947	0,9946	0,9926	0,9931	0,9946	0,9939	0,9940	0,9935	0,8753	0,9922
	Dentale Ereignisse: Gingivitis	_b	_b	_b	_b	_b	_b	_b	_b	_b	_b
<p>RGI-C: <i>Radiographic Global Impression of Change</i>; RSS: <i>Rickets Severity Score</i>; 6MWT: 6-Minuten-Gehdistanz Test; PROMIS: <i>Patient Reported Outcomes Measurement Information System</i>; FPS-R: <i>Face Pain Scale – Revised</i>; SF-10: <i>Short Form (SF)-10 Health Survey</i>; PHEX: Phosphat regulierendes Gen mit Homologie zu Endopeptidasen auf dem X-Chromosom</p> <p>^a: Die korrigierte Analyse des 6MWT wurde dargestellt, da zwei Patienten einen fehlerhaften Datensatz aufwiesen (zu Baseline und zu Woche 40). Eine detaillierte Beschreibung erfolgt im Abschnitt 4.3.1.3.1.3.</p> <p>^b: Keine Analyse möglich, da keine Ereignisse aufgetreten sind</p>											

4.3.1.3.2.2 Ergebnisse für Subgruppenanalysen mit positivem Interaktionstest (p<0,05)

Die folgenden Abschnitte enthalten die ausführliche Ergebnisdarstellung von Subgruppenanalysen, für die ein positiver Interaktionstest ($p < 0,05$) vorliegt. Entsprechend dem in Abschnitt 4.2.5.5 beschriebenen strukturierten, mehrstufigen Vorgehen wird im nächsten Schritt anhand der Lage des Effektschätzers und des KI geprüft, ob eine abweichende Interpretation des Zusatznutzens und damit eine bewertungsrelevante Effektmodifikation grundsätzlich vorliegen könnte. Ist dies der Fall, wird abschließend anhand der in Abschnitt 4.2.5.5 beschriebenen Detailkriterien untersucht, ob es sich um eine tatsächliche Effektmodifikation oder etwa um ein falsch positives Ergebnis infolge des multiplen Testens und der geringen Patientenzahl pro Behandlungsarm handelt.

Endpunkt: FPS-R

Tabelle 4-46: Subgruppenanalysen – Ergebnisse für FPS-R aus RCT mit dem zu bewertenden Arzneimittel

Intervention (N)	Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder ≥ 5 Jahre)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n MW (SD)	n MWD (SE) LsMean ^a (SE)	n MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Subgruppe: Geschlecht							
Weibliche Patienten							
Aktive Kontrolle ^b (N=18)	11 0,7 (1,35)	11 0,0 (0,0)	11 0,0 (0,0)	11 -0,7 (0,41) -0,49 (0,032)	11 -0,7 (0,41) -0,50 (0,024)	1,16 [-0,16; 2,47] 0,0845	1,19 [0,43; 1,95] 0,0022 H'g = 1,330 [0,177; 2,484]
Burosumab (N=16)	5 0,0 (0,0)	5 1,2 (0,80)	5 1,2 (1,10)	5 1,2 (0,80) 0,67 (0,683)	5 1,2 (0,49) 0,69 (0,400)		
Männliche Patienten							
Aktive Kontrolle ^b (N=14)	9 0,7 (1,00)	9 1,3 (2,24)	8 1,3 (1,49)	9 0,7 (0,47) 0,72 (0,592)	8 0,8 (0,37) 0,68 (0,397)	-1,09 [-2,21; 0,04] 0,0581	-1,05 [-1,99; -0,11] 0,0284 H'g = -0,755 [-1,717; 0,207]
Burosumab (N=13)	10 0,6 (1,35)	10 0,2 (0,63)	10 0,2 (0,63)	10 -0,4 (0,27) -0,37 (0,128)	10 -0,4 (0,50) -0,37 (0,318)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention	Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder ≥ 5 Jahre)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
(N)	n	n	n	n	n	LsMean ^a [95% KI]	LsMean ^a [95% KI]
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	p-Wert falls p < 0,05: Hedges'g [95% KI]	p-Wert falls p < 0,05: Hedges'g [95% KI]
<p>MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i>; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)</p> <p>^a: Berechnung mit GEE-Modell zu Woche 40 und Woche 64.</p> <p>^b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D</p> <p>GEE: Veränderung der FPS-R-Baseline als abhängige Variable; Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-Alter-Stratifikationsfaktor als Faktoren und Baseline RSS als kontinuierliche Kovariate</p>							

Für den Endpunkt FPS-R lassen sich potenzielle Effektmodifikationen in der Subgruppe Geschlecht zu Woche 40 und 64 beobachten.

Bei einem p-Wert im Interaktionstest von $p_{40} = 0,0043$ bzw. $p_{64} = 0,0004$ ergibt sich in der Subgruppe Geschlecht eine gegensätzliche Veränderung des akuten Schmerzempfindens der Patientengruppen. Während die aktive Kontrollgruppe in der Subgruppe der weiblichen Patienten ein verstärktes akutes Schmerzempfinden berichtet, empfindet die männliche Gruppe weniger Akutschmerz. Die weiblichen Patienten der Burosumab-Gruppe zeigen eine Verbesserung des Schmerzempfindens, während die männlichen Patienten der Burosumab-Gruppe über mehr Akutschmerz berichten. Für die Subgruppe der weiblichen als auch der männlichen Patienten ist der Unterschied zu Woche 64 zwischen den Behandlungsgruppen statistisch signifikant, jedoch nicht klinisch relevant (Woche 64: Differenz_{weiblich}: 1,19; $p = 0,0022$ und Differenz_{männlich} -1,05; $p = 0,0284$).

Es besteht keine unmittelbar erkennbare medizinische, biologische oder physiologische Rationale, die eine solche Effektmodifikation erklären könnte. Innerhalb der Studie UX023-CL301 ergab sich ansonsten nur noch für den Endpunkt Serumphosphat (Prozentsatz erreichter Normwert) eine potenzielle Effektmodifikation durch den Faktor Geschlecht.

Aufgrund der unklaren Datenlage zu Geschlecht als Effektmodifikator und im Hinblick auf die Problematik des multiplen Testens und der hier sehr kleinen weiblichen Patientenzahl in der Burosumab-Gruppe ($n < 10$) wird insgesamt nur von einer scheinbaren Effektmodifikation des Endpunktes FPS-R durch den Faktor Geschlecht als Folge eines falsch positiven Interaktionstests ausgegangen.

Endpunkt: Serumphosphat – Prozentsatz erreichter Normwert

Tabelle 4-47: Subgruppenanalysen – Ergebnisse für Serumphosphat – Prozentsatz erreichter Normwert aus RCT mit dem zu bewertenden Arzneimittel

Intervention		Serumphosphat – Prozentsatz erreichter Normwert [3,2 – 6,1 mg/dl]				
(N)	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n LsMean ^a (SE)	n LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Subgruppe: Geschlecht						
Weibliche Patienten						
Aktive Kontrolle ^b (N=18)	18 7,78 (12,733)	18 8,55 (14,313)	18 7,32 (3,740)	18 8,29 (4,027)	76,22 [65,13; 87, 31] < 0,0001 H'g = 4,708 [3,402; 6,015]	73,27 [61,33; 85,21] < 0,0001 H'g = 4,262 [3,046; 5,479]
Burosumab (N=16)	16 83,71 (17,472)	16 81,25 (18,016)	16 83,54 (3,906)	16 81,55 (4,206)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention		Serumphosphat – Prozentsatz erreichter Normwert [3,2 – 6,1 mg/dl]				
(N)	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n LsMean ^a (SE)	n LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]	LsMean ^a [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]
Männliche Patienten						
Aktive Kontrolle ^b (N=14)	14 4,08 (11,792)	14 3,97 (12,020)	14 5,37 (8,038)	14 5,49 (7,938)	48,34 [24,77, 71,90] 0,0003 H'g = 1,554 [0,693; 2,415]	48,31 [25,03; 71,58] 0,0003 H'g = 1,559 [0,698; 2,421]
Burosumab (N=13)	13 51,65 (39,356)	13 51,28 (38,899)	13 53,70 (8,792)	13 53,80 (8,683)		
MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H'g: Hedges'g (wurden post-hoc berechnet)						
^a : Berechnung mit ANCOVA-Modell zu Woche 40 zu Woche 64.						
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						
ANCOVA: Behandlungsgruppe, Baseline-Alter-Stratifizierungsfaktor als unabhängige Variable und Baseline RSS als kontinuierliche Kovariate						

Für den Endpunkt Serumphosphat – Prozentsatz erreichter Normwert lassen sich potenzielle Effektmodifikationen in der Subgruppe Geschlecht zu Woche 40 und 64 beobachten.

Bei einem p-Wert im Interaktionstest von $p = 0,0175$ bzw. $p = 0,0329$ ergibt sich in der Subgruppe Geschlecht eine Erhöhung der relativen Serumphosphat-Konzentration, die bei weiblichen Patienten deutlicher ist (Woche 64: Differenz_{weiblich}: 73,27; $p < 0,0001$; Differenz_{männlich} 48,31; $p = 0,0003$). Für die Subgruppe der weiblichen und männlichen Patienten ist der Unterschied zwischen den Behandlungsgruppen statistisch signifikant und klinisch relevant.

Es besteht keine unmittelbar erkennbare medizinische, biologische oder physiologische Rationale, die eine solche Effektmodifikation erklären könnte. Innerhalb der Studie UX023-CL301 ergab sich ansonsten nur noch für den Endpunkt FPS-R eine potenzielle Effektmodifikation durch den Faktor Geschlecht [59].

4.3.1.3.2.3 Ergebnisse für präspezifizierten Subgruppenanalysen

Nachfolgend sind die präspezifizierten Subgruppenanalysen für die Studie UX023-CL301 dargestellt. Für die hier aufgeführten Subgruppenanalysen finden sich keine Hinweise oder Belege auf Effektmodifikationen anhand der Interaktionstests (alle $p > 0,05$). Im Fall eines negativen Interaktionstests ist ausschließlich das Ergebnis der Gesamtstudienpopulation relevant. Auf eine Darstellung der Ergebnisse der Subgruppenanalyse im Dossier kann somit verzichtet werden. Zur formalen Vollständigkeit und um eine vollständige Nutzenbewertung bei abweichender Entscheidung des G-BA zu ermöglichen, wurden die Ergebnisse dieser Subgruppenanalysen dennoch in dem Dossier dargestellt.

Endpunkt: Rachitissymptomatik

Tabelle 4-48: Subgruppenanalyse des Endpunktes RGI-C Total Score

Intervention (N)	Radiographic Global Impression of Change (RGI-C) Total Score					
	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	LsMean ^a [95% KI]	LsMean ^a [95% KI]
	MWD (SE)	LsMean ^a (SE)	MWD (SE)	LsMean ^a (SE)		
Baseline RSS Total Score > 2,5 (High RSS)						
Aktive Kontrolle ^b (N=20)	20 0,82 (0,172)	20 0,82 (0,167)	20 1,17 (0,180)	20 1,17 (0,181)	1,14 [0,77; 1,51]	0,93 [0,52; 1,33]
Burosumab (N=19)	19 1,96 (0,092)	19 1,96 (0,088)	19 2,11 (0,092)	19 2,10 (0,091)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)						
Aktive Kontrolle ^b (N=12)	12 0,75 (0,179)	12 0,75 (0,200)	12 0,83 (0,190)	12 0,83 (0,175)	1,04 [0,51; 1,57]	1,09 [0,63; 1,55]
Burosumab (N=10)	10 1,83 (0,159)	10 1,79 (0,159)	10 1,97 (0,126)	10 1,93 (0,141)		
Kinder < 5 Jahren						
Aktive Kontrolle ^b (N=12)	12 0,67 (0,246)	12 0,67 (0,229)	12 1,14 (0,274)	12 1,14 (0,274)	1,18 [0,66; 1,69]	0,90 [0,34; 1,46]
Burosumab (N=14)	14 1,86 (0,134)	14 1,85 (0,126)	14 2,05 (0,069)	14 2,04 (0,063)		
Kinder ≥ 5 Jahren						
Aktive Kontrolle ^b (N=20)	20 0,87 (0,138)	20 0,87 (0,130)	20 0,98 (0,144)	20 0,99 (0,142)	1,10 [0,80; 1,40]	1,07 [0,71; 1,43]
Burosumab (N=15)	15 1,98 (0,095)	15 1,97 (0,082)	15 2,07 (0,131)	15 2,06 (0,115)		

Intervention (N)	Radiographic Global Impression of Change (RGI-C) Total Score					
	Änderung von Baseline zu Woche 40 n MWD (SE)	Änderung von Baseline zu Woche 40 n LsMean ^a (SE)	Änderung von Baseline zu Woche 64 n MWD (SE)	Änderung von Baseline zu Woche 64 n LsMean ^a (SE)	Behandlungs- differenz von Baseline zu Woche 40 LsMean ^a [95% KI]	Behandlungs- differenz von Baseline zu Woche 64 LsMean ^a [95% KI]
Weibliche Patienten						
Aktive Kontrolle ^b (N=18)	18 0,78 (0,185)	18 0,78 (0,192)	18 1,20 (0,153)	18 1,20 (0,156)	1,22 [0,78; 1,67]	0,94 [0,59; 1,30]
Burosumab (N=16)	16 2,00 (0,086)	16 2,00 (0,089)	16 2,15 (0,043)	16 2,15 (0,060)		
Männliche Patienten						
Aktive Kontrolle ^b (N=14)	14 0,81 (0,167)	14 0,78 (0,153)	14 0,83 (0,231)	14 0,80 (0,216)	0,96 [0,52; 1,40]	1,06 [0,55; 1,58]
Burosumab (N=13)	13 1,82 (0,144)	13 1,74 (0,159)	13 1,95 (0,155)	13 1,86 (0,164)		
MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SE: <i>standard error</i> (Standardfehler)						
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.						
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-Alter-Stratifikationsfaktor als Faktoren und Baseline RSS als kontinuierliche Kovariate.						

Für die Subgruppenanalyse nach Rachitis-Schwere, Alter und Geschlecht des Endpunktes „RGI-C Total Score“ sind die Effektschätzer der Subgruppen konsistent zum Ergebnis für die Gesamtpopulation. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

Tabelle 4-49: Subgruppenanalyse des Endpunktes RSS Total Score

Intervention		Rickets Severity Score (RSS) Total Score					
(N)	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs-differenz von Baseline zu Woche 40	Behandlungs-differenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n MW (SD)	n MWD (SE) LsMean ^a (SE)	n MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=20)	20 3,78 (1,057)	20 2,65 (1,204)	20 2,33 (1,104)	20 -1,13 (0,217) -1,11 (0,201)	20 -1,45 (0,248) -1,44 (0,209)	-1,36 [-1,85; -0,87]	-1,18 [-1,72; -0,64]
Burosumab (N=19)	19 3,61 (0,937)	18 1,22 (0,844)	19 1,03 (0,874)	18 -2,39 (0,159) -2,47 (0,145)	19 -2,58 (0,203) -2,62 (0,175)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=12)	12 2,21 (0,257)	12 2,17 (0,835)	12 1,92 (0,557)	12 -0,04 (0,217) -0,05 (0,226)	12 -0,29 (0,156) -0,30 (0,148)	-1,32 [-1,84; -0,80]	-1,22 [-1,59; -0,85]
Burosumab (N=10)	10 2,35 (0,242)	10 0,95 (0,369)	10 0,80 (0,258)	10 -1,40 (0,125) -1,37 (0,118)	10 -1,55 (0,117) -1,52 (0,114)		
Kinder < 5 Jahren							
Aktive Kontrolle ^b (N=12)	12 3,50 (1,477)	12 2,54 (1,339)	12 2,04 (1,137)	12 -0,96 (0,356) -0,90 (0,290)	12 -1,46 (0,387) -1,40 (0,276)	-1,11 [-1,75; -0,47]	-0,86 [-1,45; -0,27]

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	Rickets Severity Score (RSS) Total Score						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Burosumab (N=14)	14 3,29 (1,139)	14 1,32 (0,868)	14 1,07 (0,781)	14 -1,96 (0,185) -2,01 (0,143)	14 -2,21 (0,179) -2,26 (0,122)		
Kinder ≥ 5 Jahren							
Aktive Kontrolle ^b (N=20)	20 3,00 (0,874)	20 2,43 (0,950)	20 2,25 (0,835)	20 -0,58 (0,200) -0,59 (0,182)	20 -0,75 (0,187) -0,76 (0,158)	-1,57 [-2,00; -1,13]	-1,44 [-1,91; -0,96]
Burosumab (N=15)	15 3,07 (0,821)	14 0,93 (0,475)	15 0,83 (0,673)	14 -2,11 (0,223) -2,16 (0,129)	15 -2,23 (0,280) -2,20 (0,184)		
Weibliche Patienten							
Aktive Kontrolle ^b (N=18)	18 3,14 (1,082)	18 2,36 (1,293)	18 2,11 (1,079)	18 -0,78 (0,263) -0,73 (0,256)	18 -1,03 (0,244) -0,98 (0,216)	-1,35 [-1,97; -0,74]	-1,25 [-1,82; -0,68]
Burosumab (N=16)	16 2,94 (0,727)	15 0,93 (0,372)	16 0,75 (0,408)	15 -1,97 (0,165) -2,09 (0,140)	16 -2,19 (0,198) -2,23 (0,159)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	Rickets Severity Score (RSS) Total Score						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Männliche Patienten							
Aktive Kontrolle ^b (N=14)	14 3,25 (1,252)	14 2,61 (0,789)	14 2,25 (0,778)	14 -0,64 (0,254) -0,68 (0,151)	14 -1,00 (0,319) -1,04 (0,207)	-1,32 [-1,80; -0,85]	-1,12 [-1,68; -0,56]
Burosumab (N=13)	13 3,46 (1,181)	13 1,35 (0,944)	13 1,19 (0,947)	13 -2,12 (0,248) -2,01 (0,200)	13 -2,27 (0,287) -2,16 (0,213)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler)							
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-Alter-Stratifikationsfaktor als Faktoren und Baseline RSS als kontinuierliche Kovariate							

Insgesamt zeigen die Subgruppenanalysen (Rachitis-Schwere, Alter und Geschlecht) des Endpunktes RSS ein konsistentes Ergebnis zu dem der Gesamtpopulation mit gleichgerichteten Effektschätzern. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

Tabelle 4-50: Subgruppenanalyse des Endpunktes Responderanalyse des RGI-C

Radiographic Global Impression of Change (RGI-C) – Responderanalyse				
Population (N)	Woche 40		Woche 64	
	Patienten n (%)	OR^a [95% KI]	Patienten n (%)	OR^a [95% KI]
Baseline RSS Total Score > 2,5 (High RSS)				
Aktive Kontrolle ^b (20)	2 (10,0)	35,589 [4,572; 276,995]	6 (30,0)	19,929 [3,065; 129,605]
Burosumab (19)	15 (78,9)		17 (89,5)	
Baseline RSS Total Score ≤ 2,5 (Low RSS)				
Aktive Kontrolle ^b (12)	0 (0,0)	n.b.	0 (0,0)	n.b.
Burosumab (10)	6 (60,0)		8 (80,0)	
Kinder < 5 Jahren				
Aktive Kontrolle ^b (12)	1 (8,3)	29,447 [1,925; 450,359]	4 (33,3)	27,153 [1,591; 463,277]
Burosumab (14)	10 (71,4)		13 (92,9)	
Kinder ≥ 5 Jahren				
Aktive Kontrolle ^b (20)	1 (5,0)	126,131 [2,895; >999,999]	2 (10,0)	84,563 [5,239; >999,999]
Burosumab (15)	11 (73,3)		12 (80,0)	
N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Responder; %: Anteil von N; n.b.: nicht berechenbar; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; OR: <i>Odds Ratio</i> ; ^a : Das <i>generalized linear mixed model</i> (GLMM) beinhaltet die Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-Alter-Stratifikationsfaktor als Faktoren und Baseline RSS als kontinuierliche Kovariate. Die geschätzte Responder-Rate, das 95 %-KI und Odds Ratio aus dem GLMM. ^b : Aktive Kontrolle: orales Phosphat und aktives Vitamin D				

Für die Subgruppenanalysen nach Rachitis-Schwere und Alter des Endpunktes Responderanalyse des RGI-C ergeben sich konsistente Ergebnisse zu dem der Gesamtpopulation mit gleichgerichteten Effektschätzern. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

Endpunkt: Körpergröße (Stehhöhe/Liegelänge)

Tabelle 4-51: Subgruppenanalyse des Endpunktes Körpergröße

Intervention (N)	Körpergröße (Stehhöhe/Liegelänge) [Z-Score]						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=20)	20 -1,96 (0,986)	20 -1,97 (0,990)	20 -1,99 (0,971)	20 -0,01 (0,047) 0,00 (0,044)	20 -0,03 (0,052) -0,01 (0,050)	0,14 [-0,02; 0,30]	0,17 [-0,02; 0,35]
Burosumab (N=19)	18 -2,35 (1,351)	19 -2,12 (1,396)	19 -2,12 (1,225)	18 0,16 (0,068) 0,14 (0,068)	18 0,17 (0,093) 0,16 (0,080)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=12)	12 -2,18 (0,638)	12 -2,10 (0,573)	12 -2,09 (0,550)	12 0,09 (0,051) 0,08 (0,046)	12 0,09 (0,055) 0,08 (0,049)	0,09 [-0,09; 0,26]	0,10 [-0,13; 0,34]
Burosumab (N=10)	10 -2,26 (0,793)	10 -2,11 (0,867)	10 -2,08 (0,916)	10 0,15 (0,076) 0,16 (0,074)	10 0,17 (0,112) 0,18 (0,108)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	Körpergröße (Stehhöhe/Liegelänge) [Z-Score]						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Kinder < 5 Jahren							
Aktive Kontrolle ^b (N=12)	12 -2,27 (1,011)	12 -2,35 (0,999)	12 -2,33 (0,288)	12 -0,07 (0,064) -0,07 (0,060)	12 -0,06 (0,067) -0,05 (0,068)	0,22 [0,00; 0,45]	0,20 [-0,07; 0,47]
Burosumab (N=14)	13 -2,27 (1,149)	14 -2,03 (1,268)	14 -2,04 (0,290)	13 0,15 (0,102) 0,15 (0,101)	13 0,15 (0,140) 0,15 (0,123)		
Kinder ≥ 5 Jahren							
Aktive Kontrolle ^b (N=20)	20 -1,91 (0,764)	20 -1,82 (0,701)	20 -1,85 (0,675)	20 0,09 (0,037) 0,10 (0,032)	20 0,06 (0,047) 0,08 (0,041)	0,04 [-0,06; 0,14]	0,09 [-0,03; 0,22]
Burosumab (N=15)	15 -2,36 (1,221)	15 -2,20 (1,216)	15 -2,17 (1,169)	15 0,16 (0,039) 0,14 (0,040)	15 0,19 (0,058) 0,17 (0,050)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	Körpergröße (Stehhöhe/Liege­länge) [Z-Score]						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Weibliche Patienten							
Aktive Kontrolle ^b (N=18)	18 -2,29 (0,808)	18 -2,21 (0,861)	18 -2,23 (0,860)	18 0,07 (0,051) 0,07 (0,044)	18 0,06 (0,053) 0,06 (0,049)	0,10 [-0,06; 0,25]	0,14 [-0,06; 0,33]
Burosumab (N=16)	15 -2,15 (1,078)	16 -1,92 (1,169)	16 -1,91 (0,958)	15 0,16 (0,067) 0,17 (0,067)	15 0,18 (0,100) 0,19 (0,091)		
Männliche Patienten							
Aktive Kontrolle ^b (N=14)	14 -1,73 (0,868)	14 -1,77 (0,794)	14 -1,76 (0,735)	14 -0,04 (0,046) 0,00 (0,035)	14 -0,03 (0,058) 0,00 (0,044)	0,09 [-0,06; 0,25]	0,10 [-0,09; 0,29]
Burosumab (N=13)	13 -2,51 (1,278)	13 -2,36 (1,290)	13 -2,35 (1,272)	13 0,15 (0,080) 0,09 (0,070)	13 0,16 (0,104) 0,10 (0,082)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler)							
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Alter, Baseline-Liege­länge/Stehhöhe (Z-Scores) als kontinuierliche Kovariate							

Unterschiede des Z-Scores zwischen den Behandlungsgruppen wurden in den Subgruppen Rachitis-Schwere ($\leq 2,5$ Baseline RSS Total Score vs. $> 2,5$ Baseline RSS Total Score), Alter (< 5 Jahre vs. ≥ 5 Jahre) und Geschlecht (weiblich vs. männlich) analysiert. Die Körpergröße war in den Subgruppen und mit der Studienpopulation (FAS) ähnlich. Alle Subgruppenergebnisse besitzen erwartungsgemäß gleichgerichtete Effektschätzer analog zum Ergebnis der Gesamtpopulation. Die Burosumab-Gruppe zeigt unabhängig vom Alter ein besseres Wachstum im Vergleich zur aktiven Kontrollgruppe. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

Endpunkt: Gehfähigkeit

Tabelle 4-52: Subgruppenanalyse des Endpunktes Gehfähigkeit

Intervention (N)	6 Minuten Gehdistanz (6MWT) (Baseline Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n	n	n	n	n	LsMean ^a [95% KI] p-Wert	LsMean ^a [95% KI] p-Wert
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)		
Messung in Meter							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 448,17 (118,826)	12 466,67 (96,557)	12 492,25 (117,332)	12 18,50 (16,249) 29,97 (15,760)	12 44,08 (28,524) 55,55 (25,464)	30,77 [-18,14; 79,68] 0,2176	50,88 [-2,28; 104,05] 0,0607
Burosumab (N=10)	10 362,00 (139,681)	10 436,50 (84,293)	10 482,20 (83,777)	10 74,50 (34,744) 60,74 (20,321)	10 120,20 (22,675) 106,44 (9,559)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 454,00 (92,387)	8 443,50 (98,858)	8 464,88 (112,011)	8 -10,50 (20,288) -8,45 (19,200)	8 10,88 (14,571) 12,92 (15,199)	62,58 [-2,65; 127,80] 0,0600	37,00 [-21,74; 95,74] 0,2170
Burosumab (N=5)	5 373,80 (69,190)	5 431,20 (77,606)	5 427,00 (72,080)	5 57,40 (31,500) 54,12 (26,962)	5 53,20 (29,410) 49,92 (25,258)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	6 Minuten Gehdistanz (6MWT) (Baseline Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n MW (SD)	n MWD (SE) LsMean ^a (SE)	n MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI] p-Wert	LsMean ^a [95% KI] p-Wert
Messung Prozentsatz							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 75,76 (15,807)	12 76,78 (13,512)	12 79,98 (18,014)	12 1,02 (2,617) 4,05 (2,703)	12 4,22 (4,621) 7,25 (4,266)	2,93 [-4,39; 10,26] 0,4323	6,15 [-2,87; 15,18] 0,1814
Burosumab (N=10)	10 59,64 (21,475)	10 70,28 (10,583)	10 76,70 (11,374)	10 10,63 (6,000) 6,99 (2,820)	10 17,05 (4,004) 13,41 (1,541)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 76,85 (14,287)	8 73,03 (15,831)	8 75,37 (17,586)	8 -3,82 (3,348) -3,62 (3,141)	8 -1,48 (2,314) -1,29 (2,371)	11,19 [0,09; 22,29] 0,0481	6,77 [-2,28; 15,82] 0,1427
Burosumab (N=5)	5 67,11 (11,428)	5 74,99 (15,088)	5 72,90 (14,294)	5 7,88 (5,563) 7,57 (4,774)	5 5,80 (4,730) 5,48 (3,991)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention	6 Minuten Gehdistanz (6MWT) (Baseline Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungsdifferenz von Baseline zu Woche 40	Behandlungsdifferenz von Baseline zu Woche 64
(N)	n	n	n	n	n	LsMean ^a [95% KI] p-Wert	LsMean ^a [95% KI] p-Wert
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)		
Korrigierte Analyse^c							
Messung in Meter							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 448,17 (118,826)	12 466,67 (96,557)	12 492,25 (117,332)	12 18,50 (16,249) 20,10 (16,115)	12 44,08 (28,524) 45,68 (27,041)	23,27 [-22,20; 68,94] 0,3148	51,04 [-7,28; 109,36] 0,0863
Burosumab (N=10)	8 392,00 (99,217)	8 437,88 (95,402)	8 491,13 (81,676)	8 45,88 (14,505) 43,47 (13,668)	8 99,13 (10,199) 96,72 (9,334)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 454,00 (92,387)	8 443,50 (98,858)	8 464,88 (112,011)	8 -10,50 (20,288) -8,45 (19,200)	8 10,88 (14,571) 12,92 (15,199)	62,58 [-2,65; 127,80] 0,0600	37,00 [-21,74; 95,74] 0,2170
Burosumab (N=5)	5 373,80 (69,190)	5 431,20 (77,606)	5 427,00 (72,080)	5 57,40 (31,500) 54,12 (26,962)	5 53,20 (29,410) 49,92 (25,258)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	6 Minuten Gehdistanz (6MWT) (Baseline Alter der Kinder ≥ 5 Jahren)						
	Baseline n MW (SD)	Woche 40 n MW (SD)	Woche 64 n MW (SD)	Änderung von Baseline zu Woche 40 n MWD (SE) LsMean ^a (SE)	Änderung von Baseline zu Woche 64 n MWD (SE) LsMean ^a (SE)	Behandlungsdifferenz von Baseline zu Woche 40 LsMean ^a [95% KI] p-Wert	Behandlungsdifferenz von Baseline zu Woche 64 LsMean ^a [95% KI] p-Wert
Messung Prozentsatz							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 75,76 (15,807)	12 76,78 (13,512)	12 79,98 (18,014)	12 1,02 (2,617) 1,65 (2,696)	12 4,22 (4,621) 4,85 (4,477)	3,07 [-4,80; 10,94] 0,4440	7,57 [-2,55; 17,69] 0,1428
Burosumab (N=10)	8 63,80 (13,129)	8 69,47 (11,781)	8 77,17 (10,257)	8 5,68 (2,329) 4,72 (2,223)	8 13,37 (1,753) 12,42 (1,696)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 76,85 (14,287)	8 73,03 (15,831)	8 75,37 (17,586)	8 -3,82 (3,348) -3,62 (3,141)	8 -1,48 (2,314) -1,29 (2,371)	11,19 [0,09; 22,29] 0,0481	6,77 [-2,28; 15,82] 0,1427
Burosumab (N=5)	5 67,11 (11,428)	5 74,99 (15,088)	5 72,90 (14,294)	5 7,88 (5,563) 7,57 (4,774)	5 5,80 (4,730) 5,48 (3,991)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Ricketts Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler)							
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
^c : Bei dieser Analyse wurden 2 Patienten nicht berücksichtigt, da diese einen fehlerhaften Datensatz 1-mal zu Baseline und 1-mal zu Woche 40 aufweisen.							
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline RSS-Stratifikation als Faktoren und Baseline-6MWT als Kovariate							

Für die Subgruppenanalysen Rachitis-Schwere des Endpunktes 6MWT finden sich keine Hinweise oder Belege auf Effektmodifikation. Die Ergebnisse der Subgruppen sind konsistent zum Ergebnis der Gesamtanalyse. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

Endpunkt: Serumphosphat

Tabelle 4-53: Subgruppenanalyse des Endpunktes Serumphosphat

Intervention (N)	Serumphosphat [mg/dl]						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=20)	20 2,29 (0,295)	20 2,59 (0,328)	20 2,55 (0,452)	20 0,29 (0,094) 0,26 (0,077)	20 0,25 (0,094) 0,22 (0,092)	0,51 [0,29; 0,73]	0,63 [0,37; 0,90]
Burosumab (N=19)	19 2,40 (0,229)	19 3,15 (0,384)	19 3,24 (0,434)	19 0,75 (0,075) 0,77 (0,076)	19 0,84 (0,105) 0,85 (0,095)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=12)	12 2,32 (0,190)	12 2,45 (0,355)	12 2,54 (0,278)	12 0,13 (0,098) 0,10 (0,101)	12 0,23 (0,074) 0,19 (0,078)	1,09 [0,74; 1,43]	0,81 [0,50; 1,11]
Burosumab (N=10)	10 2,46 (0,280)	10 3,59 (0,360)	10 3,40 (0,389)	10 1,13 (0,149) 1,19 (0,129)	10 0,94 (0,116) 1,00 (0,116)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	Serumphosphat [mg/dl]						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Kinder < 5 Jahren							
Aktive Kontrolle ^b (N=12)	12 2,39 (0,219)	12 2,54 (0,332)	12 2,61 (0,478)	12 0,15 (0,083) 0,18 (0,088)	12 0,22 (0,126) 0,25 (0,125)	0,62 [0,38; 0,86]	0,53 [0,24; 0,81]
Burosumab (N=14)	14 2,54 (0,224)	14 3,32 (0,468)	14 3,29 (0,358)	14 0,79 (0,092) 0,81 (0,088)	14 0,76 (0,066) 0,78 (0,058)		
Kinder ≥ 5 Jahren							
Aktive Kontrolle ^b (N=20)	20 2,25 (0,269)	20 2,53 (0,353)	20 2,51 (0,336)	20 0,28 (0,100) 0,25 (0,079)	20 0,26 (0,072) 0,23 (0,070)	0,74 [0,49; 1,00]	0,78 [0,50; 1,05]
Burosumab (N=15)	15 2,31 (0,217)	15 3,29 (0,398)	15 3,29 (0,482)	15 0,97 (0,120) 1,00 (0,110)	15 0,98 (0,137) 1,00 (0,124)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Rickets Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler); H ^g : Hedges'g ^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt. ^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D GEE: Behandlungsgruppe, Baseline-Alter, Baseline RSS-Stratifizierung als Faktoren und Baseline-Serumphosphat als Kovariate							

Insgesamt zeigen die Subgruppenanalysen ein konsistentes Ergebnis zu dem der Gesamtpopulation. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

Endpunkt: Schmerzrehebung

Tabelle 4-54: Subgruppenanalyse des Endpunktes PROMIS

Intervention (N)	Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
PROMIS Pediatric Pain Interference Domain Score							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 50,0 (11,62)	12 52,10 (9,51)	12 50,0 (10,48)	12 2,1 (1,70) 2,18 (1,333)	12 -0,1 (1,61) 0,04 (1,423)	-7,76 [-12,45; -3,08]	-0,87 [-6,03; 4,29]
Burosumab (N=10)	10 49,3 (7,39)	10 43,9 (7,54)	10 48,6 (6,67)	10 -5,5 (2,37) -5,58 (1,992)	10 -0,7 (2,76) -0,83 (2,221)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 49,6 (13,48)	8 47,8 (9,52)	8 48,5 (8,49)	8 -1,9 (4,39) -3,95 (2,737)	8 -1,1 (3,33) -3,19 (1,895)	1,74 [-4,10; 7,57]	-3,51 [-10,72; 3,71]
Burosumab (N=5)	5 60,7 (13,78)	5 55,1 (10,26)	5 50,7 (11,15)	5 -5,50 (2,80) -2,22 (1,823)	5 -10,0 (4,66) -6,70 (3,256)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention	Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
(N)	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
PROMIS Pediatric Physical Function Mobility Domain Score							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 44,2 (11,37)	12 45,0 (11,54)	12 45,9 (11,15)	12 0,8 (1,27) 0,63 (1,243)	12 1,8 (1,13) 1,60 (1,083)	1,21 [-2,48; 4,89]	0,15 [-4,12; 4,42]
Burosumab (N=10)	10 46,6 (8,44)	10 48,2 (7,48)	10 48,2 (8,55)	10 1,6 (1,77) 1,83 (1,457)	10 1,6 (2,24) 1,75 (1,937)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 47,4 (7,33)	8 46,3 (6,74)	8 46,9 (7,47)	8 -1,1 (1,68) -0,68 (1,411)	8 -0,5 (1,77) -0,05 (1,598)	4,89 [-0,34; 10,11]	4,58 [-1,37; 10,54]
Burosumab (N=5)	5 42,3 (10,52)	5 47,2 (10,74)	5 47,5 (11,57)	5 4,9 (2,70) 4,21 (2,304)	5 5,3 (3,02) 4,53 (2,624)		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Intervention (N)	Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n	n	n	n	n		
	MW (SD)	MW (SD)	MW (SD)	MWD (SE) LsMean ^a (SE)	MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
PROMIS Pediatric Fatigue Domain Score							
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 50,0 (15,28)	12 50,1 (9,95)	12 47,2 (13,13)	12 0,1 (2,46) 0,76 (1,437)	12 -2,8 (2,53) -2,15 (2,052)	-5,38 [-9,99; -0,76]	-0,86 [-7,37; 5,64]
Burosumab (N=10)	10 46,3 (6,89)	10 42,5 (7,96)	10 44,10 (9,08)	10 -3,9 (2,10) -4,62 (1,891)	10 -2,3 (3,05) -3,01 (2,654)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N=8)	8 42,6 (10,24)	8 41,2 (10,12)	8 41,6 (6,79)	8 -1,4 (4,34) -2,85 (4,114)	8 -1,0 (2,89) -2,49 (2,697)	0,67 [-10,67;12,01]	-1,35 [-11,32; 8,62]
Burosumab (N=5)	5 53,6 (13,08)	5 49,1 (14,33)	5 47,4 (14,32)	5 -4,5 (1,90) -2,18 (2,844)	5 -6,2 (3,43) -3,84 (3,539)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: Rachitis-Schweregrad-Skala; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: <i>standard error</i> (Standardfehler)							
^a : Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt.							
^b : Aktive Kontrolle: Orales Phosphat und aktives Vitamin D							
GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-RSS-Stratifikation als Faktoren und Baseline-FPS-R als Kovariate							

Zusammenfassend zeigen die Subgruppenanalysen nach Rachitis-Schwere der Domänen des PROMIS (*Pediatric Pain Interference, Physical Function Mobility und Fatigue*) gleichgerichtete Effektmodifikatoren zu den Ergebnissen der Gesamtanalyse.

Tabelle 4-55: Subgruppenanalyse des Endpunktes FPS-R

Intervention (N)	Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder ≥ 5 Jahre)						
	Baseline	Woche 40	Woche 64	Änderung von Baseline zu Woche 40	Änderung von Baseline zu Woche 64	Behandlungs- differenz von Baseline zu Woche 40	Behandlungs- differenz von Baseline zu Woche 64
	n MW (SD)	n MW (SD)	n MW (SD)	n MWD (SE) LsMean ^a (SE)	n MWD (SE) LsMean ^a (SE)	LsMean ^a [95% KI]	LsMean ^a [95% KI]
Baseline RSS Total Score > 2,5 (High RSS)							
Aktive Kontrolle ^b (N=12)	12 0,7 (1,30)	12 0,5 (1,24)	11 0,2 (0,60)	12 -0,2 (0,39) 0,03 (0,281)	11 -0,4 (0,36) -0,26 (0,102)	-0,22 [-0,84; 0,40]	0,47 [-0,14; 1,07]
Burosumab (N=10)	10 0,2 (0,63)	10 0,2 (0,63)	10 0,6 (0,97)	10 0,0 (0,30) -0,19 (0,204)	10 0,4 (0,40) 0,21 (0,305)		
Baseline RSS Total Score ≤ 2,5 (Low RSS)							
Aktive Kontrolle ^b (N= 8)	8 0,8 (1,04)	8 0,8 (2,12)	8 1,0 (1,51)	8 0,00 (0,65) -0,02 (0,688)	8 0,3 (0,59) 0,23 (0,497)	0,44 [-1,50; 2,39]	-0,61 [-1,85; 0,64]
Burosumab (N=5)	5 0,8 (1,79)	5 1,2 (1,79)	5 0,4 (0,89)	5 0,4 (0,98) 0,43 (0,705)	5 -0,4 (0,98) -0,37 (0,386)		
MW: Mittelwert; MWD: Mittelwertdifferenz; N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind; RSS: <i>Ricketts Severity Score</i> ; [95% KI]: 95%-iges Konfidenzintervall; SD: <i>standard deviation</i> (Standardabweichung); SE: standard error (Standardfehler) a: Berechnung mit <i>generalized estimation equation</i> (GEE)-Modell zu Woche 40 und 64. In beiden Auswertungen werden die Daten zu Woche 64 durch das GEE-Modell berücksichtigt. b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D GEE: Behandlungsgruppe, Visite, Interaktion Behandlungsgruppe bei Visite und Baseline-RSS-Stratifikation als Faktoren und Baseline-FPS-R als Kovariate							

Insgesamt zeigt die Subgruppenanalyse des Endpunktes FPS-R ein konsistentes Ergebnis zu der PROMIS-Schmerzdomäne (*Pediatric Pain Interference*). Ebenso sind die jeweiligen Effektmodifikatoren der Subgruppen gleichgerichtet zur Gesamtanalyse. Die Ableitung des Zusatznutzens erfolgt daher ausschließlich aus den Ergebnissen für die Gesamtpopulation.

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Mortalität

Mortalität ist ein patientenrelevanter Endpunkt, der jedoch nicht für die XLH-Therapie geeignet ist, da es sich hierbei um eine schwere, chronische Erkrankung des Knochenskeletts und des Stoffwechsels handelt, die allerdings erwartungsgemäß im Verlauf einer Studiendauer von 40/64 Wochen nicht zum Tode führt. Darüber hinaus ist bei Burosumab kein direkter Einfluss auf die Mortalität bekannt. Die Mortalität wurde durch die Datenerhebung der schwerwiegenden unerwünschten Ereignisse in der Studie UX023-CL301 berücksichtigt. In der Studie trat kein Todesfall während der Studiendauer ein.

Morbidität

Es wurden unterschiedliche patientenrelevante Endpunkte, die der Morbidität zuzuordnen sind, im Rahmen der UX023-CL301 erhoben:

Rachitissymptomatik

Burosumab erzielte eine Verbesserung der Rachitissymptomatik und reduzierte die Schwere der Rachitis stärker als die aktive Kontrolle nach 40 Wochen. Dabei wurden zwei Methoden (RGI-C und RSS) gewählt, um eine quantitative und qualitative Aussage treffen zu können. Bis Woche 64 gab es eine kontinuierliche Verbesserung der Rachitissymptomatik für beide Behandlungsgruppen, wobei die Burosumab-Gruppe eine signifikant höhere Verbesserung als die aktive Kontrollgruppe erzielen konnte. Während der Studie waren die unabhängigen Radiologen der beiden Messmethoden verblindet. Bei der RGI-C-Methode deuten positive Werte auf eine Verbesserung („Heilung“ der Rachitissymptomatik) hin. Bei der RSS-Methode spiegeln kleinere Zahlen einen geringeren Schweregrad der Rachitissymptomatik wider.

Der primäre Wirksamkeitsendpunkt, der RGI-C Global Score, zeigte statistisch signifikante Ergebnisse, die auf eine bessere Heilung in der Burosumab-Gruppe im Vergleich zur aktiven Kontrollgruppe zu Woche 40 hindeuten. In der Primäranalyse des primären Wirksamkeitsendpunktes betrug die adjustierte mittlere Änderung +1,92 in der Burosumab-Gruppe und +0,77 in der aktiven Kontrollgruppe, was eine Behandlungsdifferenz von +1,14 (95%-KI: [+0,83; +1,45]; $p < 0,0001$, ANCOVA Modell) bedeutet. In Woche 64 zeigten beide

Behandlungsgruppen eine kontinuierliche Verbesserung durch die Behandlung (Burosumab: +2,06; Aktive Kontrolle: +1,03). Zu Woche 64 konnte die Burosumab-Gruppe verglichen zur aktiven Kontrollgruppe jedoch eine deutlich größere Veränderung mit einer Behandlungsdifferenz von +1,02 (95%-KI: [+0,72; +1,33]; $p < 0,0001$, GEE-Modell) aufweisen. Somit konnte diese Studie die Wirksamkeit von Burosumab hinsichtlich einer Verbesserung der Rachitissymptomatik im Vergleich zur Substitutionstherapie belegen. Der Unterschied beider Behandlungsgruppen war zu Woche 40 als auch zu Woche 64 klinisch relevant.

Eine substantielle Heilung der Rachitissymptomatik (Responderanalyse des RGI-C Global Score $\geq +2,0$) wurde zu Woche 40 bei 72,4 % bzw. 6,3 % der Kinder in der Burosumab- und aktiven Kontrollgruppe beobachtet (Odds Ratio [95%-KI]: 39,139 [7,238; 211,656]; $p < 0,0001$, logistisches Regressionsmodell). Zu Woche 64 konnte bei 86,2 % bzw. 18,8 % der Kinder in der Burosumab- und aktiven Kontrollgruppe eine Heilung mit einem RGI-C Global Score $\geq +2,0$ erzielt werden (Odds Ratio [95%-KI]: 34,075 [5,627; 206,340]; $p = 0,0002$).

Der RSS Total Score sank in der Burosumab-Gruppe von 3,17 zu Baseline über 1,13 zu Woche 40 auf 0,95 zu Woche 64, was eine Gesamtreduktion von 70 % bedeutet. In der aktiven Kontrollgruppe verringerte sich der RSS Total Score von 3,19 zu Baseline über 2,47 zu Woche 40 auf 2,17 zu Woche 64, was einer Reduktion von 32 % entspricht. In der Burosumab-Gruppe war die Reduktion somit mindestens doppelt so hoch wie in der aktiven Kontrollgruppe. Die Behandlungsdifferenz zwischen den Gruppen von -1,34 war bereits zu Woche 40 statistisch signifikant ($p < 0,0001$; ANCOVA-Modell). Dieser signifikante Unterschied ist mit einer Behandlungsdifferenz von -1,21 auch zu Woche 64 zu sehen ($p < 0,0001$; GEE-Modell). Der Unterschied beider Behandlungsgruppen war zu Woche 40 als auch zu Woche 64 klinisch relevant.

Körpergröße

Die Analyse der Veränderung der Körpergröße ermittelt aus der Stehhöhe/Liegelänge [Z-Score] zeigte in beiden Gruppen ein Wachstum. Zu Woche 64 zeigte die Veränderung der Stehhöhe/Liegelänge ein signifikant stärkeres Wachstum in der Burosumab-Gruppe als in der aktiven Kontrollgruppe.

Die Stehhöhe/Liegelänge [Z-Score] in der Burosumab-Gruppe betrug -2,32 zu Baseline, -2,12 zu Woche 40 und -2,11 zu Woche 64. Dies spiegelt eine mittlere adjustierte Veränderung von +0,16 bzw. +0,17 wider. In der aktiven Kontrollgruppe betrug die Stehhöhe/Liegelänge [Z-Score] -2,05 zu Baseline, -2,02 zu Woche 40 und -2,03 zu Woche 64.

Die primäre Analyse des patientenrelevanten Endpunktes Körpergröße wurde zu Woche 64 durchgeführt. In der Burosumab-Gruppe war die Veränderung der Stehhöhe/Liegelänge [Z-Score] mit einer Behandlungsdifferenz von +0,14 signifikant größer als in der aktiven Kontrollgruppe ([95%-KI]: [0,00; 0,29]; $p = 0,0490$, GEE-Modell) Der Vorteil unter

Burosumab-Behandlung konnte bereits zu Woche 40 gezeigt werden (Differenz [95%-KI]: +0,12 [0,01; 0,24]; $p = 0,0408$, GEE-Modell).

Gehfähigkeit

Der 6MWT ist ein Messinstrument, das die körperliche Leistungsfähigkeit durch Ermittlung einer Laufstrecke misst. Da der 6MWT nicht für alle Kinder der verschiedenen Altersklassen der Studie UX023-CL301 geeignet ist, wurde der Endpunkt nur bei Kindern ermittelt, die bei der Screeningvisite ≥ 5 Jahre alt waren. Die folgende Zusammenfassung des 6MWT bezieht sich auf die Analyse, die zwei Kinder in der Burosumab-Gruppe ausschließt, da diese nicht in der Lage waren, den 6MWT zu Baseline bzw. zu Woche 40 abzuschließen.

In der Burosumab-Gruppe stieg die Mobilität im 6MWT (gemessen als absoluter Wert in Gesamtmeter und relativer Wert als Prozentsatz der vorhergesagten normalen Laufstrecke) von 385 m zu Baseline über 435 m zu Woche 40 auf 466 m zu Woche 64. Die aktive Kontrollgruppe zeigte einen geringeren Anstieg mit 451 m zu Baseline, 457 m zu Woche 40 und 481 m zu Woche 64.

Nach der Kontrolle der Differenz der Ausgangswerte zwischen den Behandlungsgruppen zeigte das GEE-Modell, dass die Behandlung mit Burosumab zu einer signifikant höheren Verbesserung der Mobilität als die aktive Kontrolle führte. Zu Woche 64 konnte eine Behandlungsdifferenz bei der absoluten 6MWT von 46,55 m ($p = 0,0399$) und der relativen 6MWT von 7,27 % ($p = 0,0496$) erzielt werden.

Serumphosphat

Die Messung des Serumphosphats wurde bei den Patienten im nüchternen Zustand durchgeführt. Bereits nach der ersten Dosis von Burosumab stiegen die mittleren Serumphosphat-Konzentrationen rasch zum unteren Bereich der Norm (LLN: 3,2 mg/dl) und konnten bis Woche 64 in diesem Bereich gehalten werden (mit einer Ausnahme). Bei der Burosumab-Gruppe erhöhte sich die Serumphosphat-Konzentration von 2,42 mg/dl zu Baseline auf 3,38 mg/dl zu Woche 40 und auf 3,36 mg/dl zu Woche 64. Die aktive Kontrollgruppe zeigte ebenfalls einen Anstieg der Serumphosphat-Konzentration. Dieser war verglichen zur Burosumab-Gruppe jedoch sehr langsam und deutlich unter dem LLN. Die Werte lagen zu Baseline bei 2,30 mg/dl, zu Woche 40 bei 2,55 mg/dl und zu Woche 64 bei 2,56 mg/dl. Die Veränderungen der Serumphosphat-Konzentration von Baseline zu Post-Baseline bis Woche 64 waren in der Burosumab-Gruppe größer als in der aktiven Kontrollgruppe (Differenz₆₄ [95%-KI]: 0,74 mg/dl [0,58; 0,91]; $p < 0,0001$, ANCOVA-Modell). Alle Kinder der Burosumab-Gruppe erreichten eine klinisch signifikante Erhöhung der Serumphosphat-Konzentration. Dabei erreichten 28/29 Kinder (97 %) mindestens einmalig während der Studiendauer einen Wert im Normbereich [3,2-6,1 mg/dl]. In der aktiven Kontrollgruppe war dies nur bei 8/32 Kindern (25 %) der Fall.

Patient-Reported Outcomes Measurement Information System (PROMIS): Pediatric Pain Interference, Physical Function Mobility and Fatigue domains

Der PROMIS-Fragebogen wurde für Kinder verwendet und ausgefüllt, die zur Screeningvisite ≥ 5 Jahre alt waren ($N_{\text{(Burosumab)}} = 15$; $N_{\text{(aktive Kontrolle)}} = 20$). Zu Baseline lagen die mittleren T-Scores für die PROMIS-Domänen Schmerz, physische Funktionsmobilität und Fatigue innerhalb von 1 Standardabweichung (SD) des standardisierten Mittelwerts von 50, was auf keine signifikante Beeinträchtigung in beiden Behandlungsgruppen hinweist. Höhere T-Scores stellen eine größere Auswirkung auf das zu messende Konstrukt dar; positive Auswirkungen werden durch erhöhte Mobilitätswerte und verringerte Werte für Schmerzen und Fatigue ausgedrückt.

Der Schmerz T-Score zeigte einen allgemeinen Rückgang in der Burosumab-Gruppe, mit 53,1 zu Baseline, 47,6 zu Woche 40 (LsMean: -5,23) und 49,3 zu Woche 64 (LsMean: -3,55). Die aktive Kontrollgruppe erreichte kaum eine Änderung bei 49,9 zu Baseline, 50,4 zu Woche 40 (LsMean: -0,31) und 49,4 zu Woche 64 (LsMean: -1,29). Der Behandlungsunterschied zwischen den Gruppen in der Veränderung der T-Scores der Schmerzdomäne des PROMIS zeigte konträre Ergebnisse zu den zwei Auswertungszeitpunkten. Zu Woche 40 zeigte sich mit -4,91 ein statistisch signifikanter Behandlungsunterschied ($p = 0,0245$; GEE-Modell). Zu Woche 64 dagegen war der Behandlungsunterschied mit -2,26 nicht mehr statistisch signifikant ($p = 0,3091$, GEE-Modell).

Der *Physical Function Mobility* T-Score stieg von Baseline bis Woche 40 in der Burosumab-Gruppe um 2,76 und zeigte in der aktiven Kontrollgruppe nur eine minimale Veränderung von +0,09. Zu Woche 64 konnte in der Burosumab-Gruppe im Vergleich zur aktiven Kontrollgruppe ebenfalls eine größere Veränderung beobachtet werden (LsMean_(Burosumab): +2,8; LsMean_(aktive Kontrolle): +0,92). Der Behandlungsunterschied zwischen den Behandlungsgruppen in der Veränderung des T-Scores der *Physical Function Mobility*-Domäne des PROMIS war in beiden Wochen nicht statistisch signifikant.

Der Fatigue T-Score sank in der Burosumab-Gruppe von Baseline bis Woche 40 um 4,17. Die aktive Kontrollgruppe zeigte nur eine Reduktion von 0,97. Ähnlich verhielten sich die Werte zu Woche 64, mit einer stärkeren Abnahme der Müdigkeit in der Burosumab-Gruppe (LsMean: -3,65) als in der aktiven Kontrollgruppe (LsMean: -2,57). Der Unterschied zwischen den Behandlungsgruppen in der Veränderung der Fatigue-Domäne des PROMIS war weder in Woche 40 noch in Woche 64 statistisch signifikant.

Face-Pain-Scale-Revised (FPS-R)

Der Schmerzintensitätsfragebogen FPS-R zeigte im Zeitverlauf keine nennenswerten Veränderungen innerhalb der Behandlungsgruppen oder Unterschiede zwischen den Behandlungsgruppen. Die meisten Kinder in beiden Behandlungsgruppen berichteten weder zu Baseline noch zu Post-Baseline von einem akuten Schmerz. Vermutlich kann durch die

dauerhafte Belastung der Kinder durch Schmerz keine genaue Einstufung in „Schmerzfrei“ ermittelt werden.

Lebensqualität

Short Form-10 (SF-10)

Die funktionelle Komponente (PHS-10) des Gesundheitsfragebogens SF-10 zeigte zu Baseline einen Mittelwert, der etwa 1 Standardabweichung unter dem Bevölkerungsmittelwert beider Behandlungsgruppen lag ($MW_{\text{(Burosumab)}} = 40,03$; $MW_{\text{(aktive Kontrolle)}} = 40,74$). Dies deutet auf eine leichte Beeinträchtigung der berichteten körperlichen Gesundheit im Vergleich zur US-Allgemeinbevölkerung hin. Die Burosumab-Gruppe zeigte zu Woche 40 gegenüber Baseline eine größere Verbesserung der mittleren adjustierten Veränderung des PHS-10-Scores verglichen zur aktiven Kontrollgruppe (6,03 vs. 1,69). Diese Ergebnisse bestätigen sich ebenfalls zu Woche 64 gegenüber Baseline (5,99 vs. 0,48). Der Unterschied der Behandlungsdifferenz war nicht signifikant.

Diese Ergebnisse deuten auf eine größere Verbesserung der berichteten körperlichen Gesundheit der Kinder in der Burosumab-Gruppe hin. Dieser Trend wird durch die Ergebnisse des PROMIS *Pain Interference*- sowie *Physical Function Mobility*- und des 6MWT gestützt.

Die psychosoziale Komponente (PSS-10) des Gesundheitsfragebogens SF-10 zeigte zu Baseline einen Mittelwert, der in der Nähe des standardisierten Mittelwerts von 50 für beide Behandlungsgruppen lag ($MW_{\text{(Burosumab)}} = 50,76$; $MW_{\text{(aktive Kontrolle)}} = 52,79$). Dies deutet auf keine Beeinträchtigung des psychosozialen Bereichs im Vergleich mit der US-Bevölkerung hin.

Die mittlere adjustierte Veränderung der PSS-10-Werte von Baseline zu Woche 40 betrug in der Burosumab-Gruppe 1,67 und in der aktiven Kontrollgruppe -0,58. Von Baseline zu Woche 64 betrug diese 1,07 bzw. 1,52. Auch hier lag kein statistisch signifikanter Unterschied der Behandlungsdifferenz vor.

Sicherheit

Die Auswertung der unerwünschten Ereignisse (UEs) ergab keine signifikanten Sicherheitsbedenken (Ausnahme: Therapie-assoziierte UEs). Während der Studie wurde kein Kind aufgrund von UEs oder aus einem anderen Grund aus der Studie ausgeschlossen. Kein Studienteilnehmer verstarb. Die meisten Studienteilnehmer hatten mindestens ein UE (Burosumab: 100 %; Aktive Kontrolle: 84 %). Bei 7 Patienten ($n_{\text{(Burosumab)}} = 4$; $n_{\text{(aktive Kontrolle)}} = 3$) traten UEs Grad 3 auf. Die restlichen UEs waren mild oder moderat (Grad 1 oder 2) und es wurden keine lebensbedrohlichen UEs Grad 4 gemeldet.

Unerwünschte Ereignisse, die durch den Prüfarzt als möglicherweise mit der Studienmedikation in Zusammenhang gebracht werden konnten, wurden bei 59 % der Kinder in der Burosumab-Gruppe und bei 22 % in der aktiven Kontrollgruppe festgestellt. Der Großteil der therapie-assoziierten UEs waren Reaktionen an der Injektionsstelle.

Zu den am häufigsten berichteten UEs gehörten typische Kindheitssymptome: Pyrexie (Burosumab: 55 % vs. aktive Kontrolle: 19 %), Husten (52 % vs. 19 %), Erbrechen (41 % vs. 25 %), Kopfschmerzen (35 % vs. 19 %) und Durchfall (24 % vs. 6 %). Die Inzidenz typischer Kinderinfektionen war zwischen den Behandlungsgruppen ähnlich wie z. B. Nasopharyngitis (38 % vs. 44 %) oder Grippe (14 % vs. 19 %), ebenso wie die Inzidenz von UEs bei SOC-Infektionen und Infektionen (79 % vs. 75 %). Keines der Ereignisse von Pyrexie, Husten, Erbrechen oder Durchfall wurde vom Prüfarzt als schwerwiegendes Ereignis bewertet und auch nicht mit der Burosumab-Behandlung in Zusammenhang gebracht.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.**

Nicht zutreffend.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-56: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

Nicht zutreffend

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-57: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs-therapie 1>	<Vergleichs-therapie 2>	<Vergleichs-therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2> <Studie 3>	• •		• •	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-58: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-59: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-60: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse:* Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.
- *Ergebnisse zu den Effekten:* Stellen Sie die gepoolten Ergebnisse dar.

- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht zutreffend.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-61: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-62: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-63: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie***

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen**4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen**

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-64: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Bei Burosumab (Crysvita[®]) handelt es sich um ein Arzneimittel zur Behandlung eines seltenen Leidens gemäß § 35a Abs. 1 Satz 10 SGB V. Der Zusatznutzen gilt demnach bereits durch die Zulassung als belegt. Nachweise gemäß § 35a Abs. 1 Satz 3 Nummer 2 und SGB V über den medizinischen Nutzen und Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie müssen nicht erbracht werden. Lediglich das Ausmaß des Zusatznutzens ist für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, nachzuweisen (5. Kapitel § 12 Nr. 1 Satz 2 VerfO des G-BA).

Die Zulassung für Crysvita[®] als Medikament zur Behandlung seltener Krankheiten (Orphan-Arzneimittel) in der europäischen Union wurde am 19. Februar 2018 erteilt [65]. Aufgrund des hohen medizinischen Bedarfs zur wirksamen Behandlung der XLH, was auch die Teilnahme am Arzneimittel-Härtefallprogramm bestätigte, wurde eine bedingte Zulassung (*conditional marketing authorisation*) durch die Europäische Zulassungsbehörde (EMA) gewährt. Diese Entscheidung beruhte auf zwei einarmigen, offenen Phase-II-Studien (UX023-CL201 und UX023-CL205) sowie einer retrospektiven, nicht-interventionellen Kontrollstudie (UX023-CL002). Diese Studien wurden im erstmaligen Nutzenbewertungsverfahren vom 15. April 2018 zur Beurteilung des Zusatznutzens von Burosumab herangezogen (Vorgangsnummer 2018-04-15-D-349 [66]). Außerdem erfolgte in der schriftlichen Stellungnahme zum offenen Verfahren eine Nachreichung des 40-Wochen-Datenschnitts der multizentrischen, randomisierten, offenen und aktiv kontrollierten Phase-III-Zulassungsstudie UX023-CL301. Aufgrund der bedingten Zulassung und der noch laufenden RCT-Studie konnten so zum Zeitpunkt des 15. April 2018 lediglich Unterlagen zur Bewertung mit der Evidenzstufe IV (Fallserien und andere nicht vergleichende Studien) eingereicht werden. Bereits diese Studiendaten zeigten, dass der Anti-FGF23-Antikörper die Folgeerkrankungen der Dysfunktion im Phosphatstoffwechsel sowie die Rachitissymptomatik bei Kindern mit XLH deutlich verbessert. Zu dieser Symptomatik gehören eine verringerte Wachstumsgeschwindigkeit, stark eingeschränkte Gehfähigkeit, erhebliche Funktions- und Mobilitätsstörungen, einhergehende starke Schmerzen und eine daraus resultierende verminderte Lebensqualität.

In diesem Dossier werden nun die Studiendaten zur abgeschlossenen Behandlungsphase (Woche 40 und Woche 64) der Phase-III-Zulassungsstudie UX023-CL301 mit einer Evidenzstufe Ib (RCT) für die erneute Bewertung von Burosumab herangezogen. Die in Modul 4 dargestellten Daten der UX023-CL301 belegen eindrucksvoll die bereits beachtlichen Ergebnisse zur Wirkung von Burosumab bei pädiatrischen XLH-Patienten im Alter von 1-12 Jahren aus den pivotalen Studien des ersten Nutzenbewertungsverfahrens zu Burosumab (Vorgangsnummer 2018-04-15-D-349 [66]).

Die erhobenen patientenrelevanten Endpunkte stimmen überwiegend mit den Endpunkten der pivotalen Studien aus dem ersten Verfahren überein. Abweichend wurde auf Empfehlung der FDA für die Schmerzerhebung bei der Studie UX023-CL301 der PROMIS-Fragebogen statt des POSNA-PODCI angewandt, sowie ergänzend der FPS-R erhoben. Zudem kam es bei der Erhebung der Endpunkte wie Gehfähigkeit, Fragebögen der funktionellen Behinderung, Schmerz und Lebensqualität (POSNA-PODCI, PROMIS, FPS-R, 6MWT oder SF-10) zu altersbedingten Abweichungen der Auswertung, da diese für Patienten im Alter von 1-4 Jahren nur bedingt oder gar nicht geeignet waren.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Burosumab (Crysvita®) ist zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischem Nachweis einer Knochenerkrankung zugelassen [1]. Bisher gibt es keine kausale und ausreichend wirksame Therapie für diese

Patienten. Burosumab ist daher das erste Arzneimittel, das speziell für Patienten mit XLH entwickelt wurde und direkt an der Ursache, dem genetisch bedingten Phosphatverlust, ansetzt.

Im Rahmen des ersten Dossiers vom 15. April 2018 wurden zur Beurteilung der Behandlung bei Kindern mit XLH die Daten von zwei einarmigen, offenen Studien (UX023-CL201, UX023-CL205), von einer retrospektiven nicht-interventionellen Beobachtungsstudie (UX023-CL002), sowie der Datenschnittpunkt zu Woche 40 von der offenen und aktiv kontrollierten Phase-III-Studie UX023-CL301 herangezogen [1]. Da die RCT-Studie noch laufend war, erfolgte am 04. Oktober 2018 ein befristeter Beschluss des G-BA, um Einblicke in die 64-Wochen-Daten der Studie zu gewährleisten. Im Beschluss wurde ein nicht quantifizierbarer Zusatznutzen vergeben, den die Patientenvertretung „nur schweren Herzens“ aufgrund der Befristung zustimmte, da sie in den Endpunkten Serumphosphat und Rachitissymptomatik einen beträchtlichen Zusatznutzen sehe.

Die Ergebnisse aus der abgeschlossenen Behandlungsphase (Woche 0 bis 64) stehen nun zur Verfügung und werden in diesem Dossier dargestellt und zur Beurteilung des Zusatznutzens von Burosumab zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischem Nachweis einer Knochenerkrankung herangezogen.

Die folgende Tabelle 4-65 zeigt in einer Übersicht die zugrunde liegenden Ergebnisse der einzelnen Endpunktkategorien bezüglich des Vergleichs der Behandlungsgruppen mit denen das Ausmaß des Zusatznutzens der Burosumab-Behandlung abgeleitet werden konnte.

Tabelle 4-65: Zusammenfassung der Ergebnisse zur Bestimmung des Ausmaßes des Zusatznutzens von Burosumab für Patienten mit XLH

Patientenrelevante Endpunkte	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 40	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 64
Rachitissymptomatik		
<i>Radiographic Global Impression of Change (RGI-C) Score</i>		
LsMean^a	1,14	1,02
[95% KI]	[0,83; 1,45]	[0,72; 1,33]
p-Wert	< 0,0001	< 0,0001
Hedges'g^b [95% KI]	H'g = 1,846 [1,246; 2,446]	H'g = 1,600 [1,023; 2,178]
<i>Rickets Severity Score (RSS)</i>		
LsMean^a	-1,34	-1,21
[95% KI]	[-1,74; -0,94]	[-1,59; -0,83]
p-Wert	< 0,0001	< 0,0001
Hedges'g^b [95% KI]	H'g = -1,399 [-1,964; -0,833]	H'g = -1,119 [-1,710; -0,529]
Körpergröße		
<i>Stehhöhe/Liegelänge [Z-Score]</i>		
LsMean^a	0,12	0,14
[95% KI]	[0,01; 0,24]	[0,00; 0,29]
p-Wert	0,0408	0,0490
Hedges'g^b [95% KI]	H'g = 0,537 [0,021; 1,053]	H'g = 0,490 [-0,025; 1,005]
Gehfähigkeit^c		
<i>6-Minuten-Gehdistanz^c (6MWT) [absolut - in Meter]</i>		
LsMean^a	38,30	45,55
[95% KI]	[-2,28; 78,98]	[2,09; 89,02]
p-Wert	0,0650	0,0399
Hedges'g^b [95% KI]	-	H'g = 0,679 [-0,038; 1,396]
<i>6-Minuten-Gehdistanz^c (6MWT) [relativ - in Prozent]</i>		
LsMean^a	6,19	7,27
[95% KI]	[-0,68; 13,06]	[0,01; 14,52]
p-Wert	0,0775	0,0496
Hedges'g^b [95% KI]	-	H'g = 0,703 [-0,015; 1,422]
Biochemischer Parameter		
<i>Serumphosphat-Konzentration [mg/dl]</i>		
LsMean^a	0,77	0,74
[95% KI]	[0,60; 0,94]	[0,58; 0,91]
p-Wert	< 0,0001	< 0,0001
Hedges'g^b [95% KI]	H'g = 2,047 [1,427; 2,667]	H'g = 1,981 [1,368; 2,595]

Patientenrelevante Endpunkte	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 40	UX023-CL301 Behandlungsdifferenz (B-P) Kontrolle vs. Burosumab Woche 64
Schmerzerhebung und funktionelle Behinderung		
<i>Patient Reported Outcomes Measurement Information System (PROMIS) Pediatric Pain Interference Domain Score</i>		
LsMean^a	-4,91	-2,26
[95% KI]	[-9,19; -0,63]	[-6,61; 2,09]
p-Wert	0,0245	0,3091
Hedges'g^b [95% KI]	H ^g = -0,697 [-1,390; -0,008]	-
<i>Patient Reported Outcomes Measurement Information System (PROMIS) Physical Function Mobility Domain Score</i>		
LsMean^a	2,68	1,90
[95% KI]	[-0,53; 5,88]	[-1,80; 5,59]
p-Wert	0,1014	0,3145
Hedges'g^b [95% KI]	-	-
<i>Patient Reported Outcomes Measurement Information System (PROMIS) Fatigue Domain Score</i>		
LsMean^a	-3,20	-1,08
[95% KI]	[-7,79; 1,39]	[-6,21; 4,06]
p-Wert	0,1724	0,6810
Hedges'g^b [95% KI]	-	-
<i>Face Pain Scale – Revised (FPS-R)</i>		
LsMean^a	-0,00	0,05
[95% KI]	[-0,80; 0,79]	[-0,58; 0,68]
p-Wert	0,9905	0,8786
Hedges'g^b [95% KI]	-	-
Lebensqualität		
<i>Short Form (SF)-10 Health Survey Physical Health Summary Score (PHS-10)</i>		
LsMean^a	4,35	5,51
[95% KI]	[-1,15; 9,85]	[-0,17; 11,18]
p-Wert	0,1213	0,0571
Hedges'g^b [95% KI]	-	-
<i>Short Form (SF)-10 Health Survey Psychosocial Summary Score (PSS-10)</i>		
LsMean^a	2,24	-0,44
[95% KI]	[-1,86; 6,35]	[-4,59; 3,70]
p-Wert	0,2841	0,8337
Hedges'g^b [95% KI]	-	-
(B-P): Behandlungsdifferenz betrachtet von Baseline zu Post-Baseline (Woche 40 bzw. 64) der Burosumab-Gruppe mit der aktiven Kontrollgruppe (orales Phosphat und aktives Vitamin D); [95 %-KI]: 95 %-Konfidenzintervall		
^a : LsMean: Statistische Methode für den jeweiligen Endpunkt siehe Abschnitt 4.3.1.3.1		
^b : Hedges'g-Berechnung erfolgt nur, falls p < 0,05		
^c : Die korrigierte Analyse des 6MWT wurde dargestellt, da zwei Patienten einen fehlerhaften Datensatz aufwiesen (zu Baseline und zu Woche 40). Eine detaillierte Beschreibung erfolgt im Abschnitt 4.3.1.3.1.3.		

Interpretation der Ergebnisse zu Morbidität

Rachitissymptomatik

Die Veränderung der Rachitissymptomatik von Baseline zu Woche 40 sowie Woche 64 wurde sowohl anhand des RGI-C-Scores als auch des RSS-Scores beurteilt. Bei der RGI-C-Methode wurde die „Heilung“ der Rachitissymptomatik in einem direkten Vergleich zwischen zwei Zeitpunkten untersucht, während die RSS-Methode die Rachitis-Schwere aufzeigte und somit die Abnahme der Rachitis-Schwere von Baseline zu Woche 40 bzw. Woche 64 untersucht werden konnte.

Die Verbesserung der Rachitissymptomatik durch Burosumab-Behandlung wurde bereits im ersten Nutzenbewertungsverfahren anhand der sowohl statistisch signifikanten als auch klinisch relevanten Ergebnisse des RSS- und RGI-C-Scores der zwei einarmigen Phase-II-Studien (UX023-CL201 und UX023-CL205) bestätigt. Auch der nicht-adjustierte Vergleich der Studie UX023-CL201 mit der Substitutionstherapie der retrospektiven Beobachtungsstudie UX023-CL002 zeigte hierbei, dass die Behandlung mit Burosumab zu einer signifikanten und noch nie zuvor erreichten Verbesserung der Rachitissymptomatik (RGI-C- und RSS-Methode) führte.

Der erstmalige kontrollierte Vergleich von Burosumab mit der aktiven Kontrolle im Hinblick auf die „Heilung“ der Rachitissymptomatik in XLH-Patienten in der hier betrachteten prospektiven Studie UX023-CL301 beweist abermals die beachtliche Wirkung von Burosumab bei XLH-Patienten. Sowohl beim RGI-C- als auch RSS-Score sind statistisch hochsignifikante und klinisch relevante Unterschiede in der Verbesserung der Rachitissymptomatik im Vergleich der beiden Behandlungsgruppen an beiden Zeitpunkten zu erkennen ($p < 0,0001$). So zeigt die Burosumab-Behandlung bei beiden Messmethoden zu Woche 40 sowie Woche 64 eine deutliche Überlegenheit für die „Heilung“ der Rachitissymptomatik gegenüber der aktiven Kontrollgruppe (orale Gabe von Phosphat und aktivem Vitamin D). So bestätigt die randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie UX023-CL301 die Ergebnisse des Zulassung-begründenden Vergleichs mit der historischen Kontrolle (UX023-CL201 und UX023-CL002).

Körpergröße

Die ermittelten Körpergrößen durch Messung der Stehhöhe bzw. Liegelänge wurden als Z-Scores angegeben und zeigten bereits bei der pivotalen Studie UX023-CL201 unter der Burosumab-Behandlung ein statistisch signifikant schnelleres Wachstum, welches verglichen zur retrospektiven Beobachtungsstudie UX023-CL002 mit der Substitutionstherapie ebenfalls schneller war. Diese Erkenntnisse konnten in der randomisierten, offenen aktiv kontrollierten Phase-III-Studie UX023-CL301 bestätigt werden. Die adjustierte Behandlungsdifferenz des Z-Scores der Körpergröße zeigte sowohl zu Woche 40 als auch Woche 64 einen signifikanten Unterschied zwischen den beiden Behandlungsgruppen zugunsten der Burosumab-Behandlung in Form eines verbesserten Wachstums. Hier ist allerdings anzumerken, dass die

Wachstumsanalyse durch die ungleichen Baseline-Werte verzerrt wurde, wobei die Burosumab-Gruppe einen geringeren Z-Score der Körpergröße zu Baseline zeigte.

Gehfähigkeit

Gangschwierigkeiten sind ein typisches klinisches Merkmal in pädiatrischen XLH-Patienten. Aus diesem Grund wurde die Gehfähigkeit sowohl in den pivotalen Phase-II-Studien als auch in der randomisierten, offenen und aktiv kontrollierten Studie mithilfe des 6MWT, der ab einem Alter von 5 Jahren geeignet ist, betrachtet. Bereits im ersten Nutzenbewertungsverfahren von Burosumab konnte in der Studie UX023-CL201 eine statistisch signifikante Verbesserung der Gehfähigkeit nach 40 Wochen unter der Burosumab-Behandlung gezeigt werden. Die Daten der UX023-CL301 bestätigen die beobachteten Effekte aus der pivotalen Studie. Die ursprünglich geplanten und dargestellten statistischen Analysen zum 6MWT umfassten eine Burosumab-Gruppe von 15 Patienten. Zu zwei Patienten liegt kein vollständiger Datensatz vor, da diese den 6MWT zu jeweils einer Messung nicht abgeschlossen hatten. So fehlen bei einem Patienten die Daten zu Baseline beim anderen die Daten zu Woche 40. Aufgrund dessen ist in Tabelle 4-65 die statistische Analyse abzüglich der zwei Patienten für die Patientenpopulation $n=13$ mit vollständigem Datensatz dargestellt. Die Ergebnisse des absoluten 6MWT demonstrieren sowohl zu Woche 40 als auch Woche 64 eine deutliche Steigerung der absoluten Gehdistanz bei der Burosumab-Gruppe und eine geringe Steigerung der absoluten Gehdistanz bei der aktiven Kontrollgruppe. Der relative 6MWT basierend auf normativen Werten zeigt bei der Burosumab-Gruppe, jedoch nicht in der aktiven Kontrollgruppe, einen Anstieg von Baseline über Woche 40 bis Woche 64. Der Unterschied zwischen den Behandlungsgruppen weist nach 64 Wochen einen statistisch signifikanten Unterschied für die Überlegenheit der Burosumab-Behandlung bei der Verbesserung der absoluten und relativen Gehdistanz verglichen zur aktiven Kontrolle auf ($p_{\text{absolut}}=0,0399$; $p_{\text{relativ}}=0,0496$) auf. Die Analyse zu Woche 40 war statistisch nicht signifikant. Die Analyse wurde durch die abweichenden Baseline-Werte der Behandlungsgruppen zugunsten der aktiven Kontrollgruppe verzerrt, da die Burosumab-Gruppe bereits zu Baseline eine schlechtere Gehfähigkeit im Vergleich zur Kontrollgruppe zeigte.

Serumphosphat

Mit den zwei einarmigen Studien (UX023-CL201 und UX023-CL205) konnte bereits im ersten Nutzenbewertungsverfahren gezeigt werden, dass sich durch die Burosumab-Behandlung die Serumphosphat-Konzentration, die 1,25-Dihydroxyvitamin-D-Konzentration, die Alkalische Phosphatase und die renale Phosphat-Rückresorption (TmP/GFR) regulieren und sich im altersspezifischen Normbereich einstellen. Mit den Ergebnissen der abgeschlossenen Behandlungsphasen der aktiv kontrollierten Phase-III-Studie UX023-CL301 werden die Erkenntnisse aus den pivotalen Studien zu allen biochemischen Parametern bestätigt. Ferner ist es nun möglich den beträchtlichen Zusatznutzen im Vergleich zur Substitutionstherapie mit

oralem Phosphat und aktivem Vitamin D zu veranschaulichen. In der UX023-CL301 konnte innerhalb der Burosumab-Gruppe ein schneller Anstieg der Serumphosphat-Konzentration in den altersspezifischen Normbereich erreicht und dieser konstant bis Woche 64 gehalten werden. Bei der aktiven Kontrollgruppe dagegen wurde lediglich ein kleiner Anstieg der Serumphosphat-Konzentration beobachtet, welcher jedoch ausschließlich aus der oralen Einnahme des Phosphats resultiert. Dieser statistisch hochsignifikante und klinisch relevante Unterschied ($p < 0,0001$; ANCOVA-Modell) zugunsten von Burosumab im Vergleich zur Substitutionstherapie der aktiven Kontrolle zeigt eindrucksvoll den beträchtlichen Zusatznutzen. Burosumab ist damit die erste kausale Therapie zur Behandlung der XLH, die die zugrunde liegende Pathophysiologie der FGF23-induzierten Hypophosphatämie verbessert.

Schmerzerhebung und funktionelle Behinderung

Die Erhebung des Schmerzes wurde in den Studien UX023-CL201 und UX023-CL301 mit unterschiedlichen Messinstrumenten erhoben. Der POSNA-PODCI-Fragebogen diente in der UX023-CL201-Studie zur Messung der funktionellen Behinderung und des Schmerzes. Bereits im ersten Nutzenbewertungsverfahren wurden die signifikanten Verbesserungen unter Burosumab-Behandlung auf der Skala „Sport und körperliche Funktion“ und „Schmerz und Wohlbefinden“ aus der Studie UX023-CL201 gezeigt. Auf Empfehlung der FDA wurde in der randomisierten, offenen und aktiv kontrollierten Phase-III-Studie UX023-CL301 der PROMIS-Fragebogen zur Ermittlung des Schmerzempfindens, der Körperfunktion und des Ermüdungsempfindens bei Patienten ab einem Alter von 5 Jahren verwendet. Dabei zeigte sich zu Baseline, dass die Studienteilnehmer im Gegensatz zu gesunden Kindern (standardisierter Mittelwert von 50) eine deutlich beeinträchtigte Körperfunktion aufwiesen. Obwohl in der Burosumab-Gruppe zu Baseline über mehr Schmerzen als in der aktiven Kontrollgruppe berichtet wurde, wurden in Woche 40 verglichen zur Kontrollgruppe weniger Schmerzen dokumentiert. Zu Woche 64 berichteten beide Behandlungsgruppen in etwa von ähnlichen Schmerzen. Zu Woche 40 zeigt sich demnach eine statistisch signifikant ($p = 0,0245$) bessere Wirkung auf die Schmerzempfindung von Burosumab verglichen zur aktiven Kontrolle, welche zu Woche 64 jedoch nicht mehr statistisch signifikant ist. Bei den Domänen Körperfunktion sowie Fatigue konnte kein signifikanter Behandlungsunterschied nachgewiesen werden. Doch ist eine geringe Verbesserung der Körperfunktion im Studienverlauf unter Burosumab-Behandlung zu erwähnen. Der FPS-R Schmerzfragebogen, der ebenfalls zur Schmerzbeurteilung erhoben wurde, weist keine merklichen Veränderungen innerhalb der Behandlungsgruppen auf. Die meisten der Patienten gaben zu allen Erhebungszeitpunkten keinen akuten Schmerz (Score = 0) an. Dies lässt sich möglicherweise mit einer Gewöhnung der Patienten an die häufig nachgewiesenen Schmerzen bei der XLH erklären. Ferner geht aus der retrospektiven Studie UX023-CL002 hervor, dass Patienten, die mit einer Substitutionstherapie behandelt wurden, unter solchen Schmerzen leiden, dass eine Schmerzmedikation erforderlich ist.

Zusammenfassung Morbidität

Burosumab war in der randomisierten, offenen und aktiv kontrollierten Studie UX023-CL301 gegenüber der aktiven Kontrolle mit oralem Phosphat und aktivem Vitamin D bei der „Heilung“ der Rachitissymptomatik zu beiden Zeitpunkten deutlich überlegen. Diese Ergebnisse stützen die bereits vorhandenen Erkenntnisse aus dem Vergleich der einarmigen Phase-II-Studien (UX023-CL201 und UX023-CL205) mit der retrospektiven Langzeitstudie UX023-CL002 mit einer Substitutionstherapie, die für das erste Nutzenbewertungsverfahren herangezogen wurden [67]. Hier konnte ebenfalls eine Wirksamkeit der aktiven Kontrolle beobachtet werden. Somit wurde die Wirksamkeit von Burosumab im Vergleich zu einer wirksamen Behandlung (Substitutionstherapie) in der aktiv kontrollierten Studie UX023-CL301 gezeigt.

Burosumab ist die erste kausale Therapie zur Behandlung der XLH. Durch die Studie UX023-CL301 konnte bestätigt werden, dass im Gegensatz zur Burosumab-Behandlung eine Substitutionstherapie ausschließlich die Dysregulation des Phosphatstoffwechsels ausgleicht, jedoch den pathologischen renalen Phosphatverlust weder verringert noch positiv beeinflusst. Somit handelt es sich bei der Substitutionstherapie lediglich um ein Entgegenwirken der Symptomatik und nicht um einen kausalen Eingriff an die pathophysiologischen Ursachen der XLH. In den dargestellten Studien konnte dagegen durchweg gezeigt werden, dass Burosumab den pathophysiologischen Defekt der XLH aufhebt, die Serumphosphat-Konzentration normalisiert und dadurch die Rachitissymptomatik verbessert und folglich das Wachstum, die Gehfähigkeit sowie die Mobilität normalisiert und Schmerzen reduziert. Diese Effekte konnten durch die UX023-CL301-Studie bei Kindern mit XLH reproduziert werden, die die meiste Zeit ihres Lebens mit einer Substitutionstherapie behandelt wurden.

Interpretation der Ergebnisse zur Lebensqualität

Die Erhebung der Lebensqualität erfolgte durch den pädiatrischen SF-10-Fragebogen, da dieser zur Beurteilung körperlicher und psychosozialer Funktionen bei Kindern ab einem Alter von 5 Jahren geeignet ist. Bei der Studie UX023-C301 zeigten die Patienten der aktiven Kontrollgruppe weder eine Verbesserung der Körperfunktion noch der psychosozialen Domäne. Unter Burosumab-Behandlung dagegen ist eine Verbesserung der Körperfunktion sowohl zu Woche 40 als auch Woche 64 gegeben. Die psychosoziale Domäne weist dagegen nur auf eine geringfügige Verbesserung bei der Burosumab-Gruppe hin. Trotz fehlender signifikanter Unterschiede in den Behandlungsgruppen ist eine deutliche Tendenz zugunsten der Burosumab-Behandlung sichtbar, was die Ergebnisse der einarmigen Studie UX023-CL201 bestätigt, welche eine signifikante Verbesserung der körperlichen Funktion unter Burosumab-Behandlung aufweist. Weiterhin decken sich diese Ergebnisse mit den Morbiditätsendpunkten (UX023-CL201: POSNA-PODCI Sport und körperliche Funktion, Schmerz und Wohlbefinden und 6MWT; UX023-CL301: PROMIS *Physical Function Mobility* und 6MWT) und deuten darauf hin, dass die Burosumab-Behandlung einen höheren Grad an Verbesserungen der Morbidität und der somit einhergehenden Verbesserung der Lebensqualität erzielt.

Interpretation der Ergebnisse zur Sicherheit

Die folgende Tabelle 4-66 zeigt in einer Übersicht die häufigsten unerwünschten Ereignisse nach PT ($\geq 20\%$), welche aufgrund dessen besondere Relevanz bei der Beurteilung des Ausmaßes des Zusatznutzens haben.

Tabelle 4-66: Übersicht der unerwünschten Ereignisse nach PT ($\geq 20\%$) in beiden Behandlungsarmen der Studie UX023-CL301 zu Woche 40 und Woche 64

Preferred Term	UX023-CL301 Woche 40		UX023-CL301 Woche 64	
	Aktive Kontrolle ^a (N = 32) n (%)	Burosumab (N = 29) n (%)	Aktive Kontrolle ^a (N = 32) n (%)	Burosumab (N = 29) n (%)
Pyrexie	6 (18,8)	15 (51,7)	6 (18,8)	16 (55,2)
Husten	6 (18,8)	12 (41,4)	6 (18,8)	15 (51,7)
Gelenkschmerzen (Arthralgie)	7 (21,9)	11 (37,9)	10 (31,3)	13 (44,8)
Erbrechen	8 (25,0)	11 (37,9)	8 (25,0)	12 (41,4)
Schmerzen in den Extremitäten	9 (28,1)	10 (34,5)	10 (31,3)	11 (37,9)
Nasopharyngitis	13 (40,6)	10 (34,5)	14 (43,8)	11 (37,9)
Kopfschmerzen	5 (15,6)	8 (27,6)	6 (18,8)	10 (34,5)
Karies	2 (6,3)	5 (17,2)	2 (6,3)	9 (31,0)
Erythem an der Injektionsstelle	0 (0,0)	6 (20,7)	0 (0,0)	9 (31,0)
Zahnabszesse	3 (9,4)	7 (24,1)	3 (9,4)	8 (27,6)
Reaktionen an der Injektionsstelle	0 (0,0)	7 (24,1)	0 (0,0)	7 (24,1)
Rhinorrhoe	2 (6,3)	5 (17,2)	2 (6,3)	7 (24,1)
Durchfall	2 (6,3)	7 (24,1)	2 (6,3)	7 (24,1)
Verringertes Vitamin D	1 (3,1)	6 (20,7)	1 (3,1)	6 (20,7)

N: Anzahl der Patienten; n: Anzahl der Patienten, die analysiert worden sind
^a: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

Sowohl in den beiden Phase-II-Studien (UX023-CL201 und UX023-CL205) als auch in der aktiv kontrollierten Phase-III-Studie UX023-CL301 ergeben sich aus der Bewertung von unerwünschten Ereignissen (UEs) keine signifikanten Sicherheitsbedenken. In allen drei Studien wurde weder ein Therapie-/Studienabbruch noch Todesfall aufgrund von UEs berichtet. Demnach erweist sich Burosumab insgesamt als wirksames, gut verträgliches und sichereres Arzneimittel zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischen Nachweis einer Knochenerkrankung.

In der Behandlungsphase (Woche 0 bis 64) berichteten die meisten Patienten sowohl der Burosumab-Gruppe (100 %) als auch der aktiven Kontrollgruppe (84 %) mindestens ein UE, wobei die Prüferärzte bei 59 % bzw. 22 % (Burosumab vs. aktive Kontrolle) einen Kausalzusammenhang zur Behandlung als möglich erachteten. Der Großteil der berichteten UEs wurde als mild oder moderat eingestuft. Bei 3 Patienten der aktiven Kontrollgruppe sowie

4 Patienten der Burosumab-Gruppe traten jedoch UEs vom Grad 3 auf. Diese Vorkommnisse wurden von den Prüfern als nicht therapie-assoziiert angesehen, mit Ausnahme eines Ereignisses Arthralgie. In der Extensionsphase trat in der Burosumab-Gruppe ein Papillenödem von Grad 3 auf. Im Fall des Papillenödems handelte es sich um einen Patienten mit vorheriger Kraniosynostose, welche häufig mit erhöhtem kranialem Druck und daraus resultierenden ophthalmologischen Problemen assoziiert ist. Es wurden keine Änderungen bei der Verabreichung der Burosumab-Therapie auf diese zwei Ereignisse durchgeführt.

Die häufigsten unerwünschten Ereignisse ($\geq 20\%$) sind in Tabelle 4-66 dargestellt. Hierzu gehören unter anderem Pyrexie, Husten, Gelenkschmerzen, Erbrechen, Karies aber auch Reaktionen an der Injektionsstelle. Die aufgeführten UEs wie Husten, Nasopharyngitis, Erbrechen, Kopfschmerzen, Durchfall und Pyrexie werden als typisch für pädiatrische Patienten angesehen. Schmerzen in Extremitäten, Arthralgien, Zahnabszesse sowie verringertes Vitamin D wiederum sind eng mit dem Krankheitsbild der XLH assoziiert. So lassen sich die UEs mit dem Erkrankungsbild der XLH und der pädiatrischen Patientenpopulation erklären. In der aktiv kontrollierten Studie UX023-CL301 traten in beiden Behandlungsgruppen die häufig berichteten UEs auf, die in den einarmigen Studien beobachtet wurden. Mit Ausnahme der Reaktion an der Injektionsstelle waren diese UEs typisch für eine pädiatrische Population oder ein häufiges Symptom der XLH-Erkrankung.

Die häufigsten Therapie-assoziierten UEs in allen drei zulassungsrelevanten Studien waren Reaktionen an der Injektionsstelle, welche nur leicht ausgeprägt und gut zu behandeln waren. Eine Reaktion an der Injektionsstelle ist eine bekannte und zu erwartende Folge einer subkutanen Verabreichung von Proteintherapeutika (Antikörpertherapeutika). Die aufgetretenen Reaktionen wurden vom Prüfer als allgemeine lokale Reizung und nicht als schwere Überempfindlichkeitsreaktion angesehen. Die Dauer der meisten Reaktionen an der Injektionsstelle betrug etwa ein bis zwei Tage. Bei der Studie UX023-CL301 waren alle genannten UEs aus ehemaligen, pivotalen klinischen Studien im Zusammenhang mit der XLH und Burosumab-Behandlung bekannt.

Insbesondere bei den für eine pädiatrische Patientenpopulation nicht ungewöhnlichen UEs wie beispielsweise Pyrexie, Husten oder Erbrechen zeigte sich eine verstärkte Häufigkeit bei der Burosumab-Gruppe gegenüber der aktiven Kontrollgruppe. Obwohl dieses Vorkommnis nicht plausibel erscheint, wurden die meisten dieser aufgetretenen UEs vom Prüfer nicht mit der Burosumab-Behandlung assoziiert. So waren Kopfschmerz, Pyrexie oder Erbrechen häufig als Symptome einer gleichzeitigen Erkrankung des Kindes aufgetreten.

Im Vergleich zur historischen Kontrolle UX023-CL002 mit einer Substitutionstherapie (orales Phosphat und aktives Vitamin D) wurde in den Burosumab-Studien kein Ereignis berichtet, bei dem die Serumphosphat-Konzentration über der Norm lag [3,2-6,1 mg/dl]. Nierenultraschalluntersuchungen waren in den meisten Fällen in beiden Behandlungsgruppen der Studie UX023-CL301 über die 64 Wochen unauffällig (Burosumab: 90 %; Aktive Kontrolle: 78 %). Ebenfalls wurden keine klinisch bedeutsamen Erhöhungen von Nephrokalzinose-Werten beobachtet. Hervorzuheben ist, dass unter der Therapie mit

Burosumab die Folgeerkrankungen wie eine Nephrokalzinose und ein sekundärer Hyperparathyreoidismus, die häufig durch eine Substitutionstherapie auftreten, ausbleibt.

Übertragbarkeit auf den deutschen Versorgungskontext

Da die Mehrheit der Patienten kaukasischer Herkunft war und eine genetische Erkrankung vorliegt, welche mit großer Wahrscheinlichkeit weitestgehend unabhängig von äußeren Einflüssen ist, gibt es keine Anzeichen, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext schließen lassen.

Abschließende Bewertung

XLH ist eine chronische und genetisch bedingte Erkrankung des Skelettstoffwechsels. Durch den Überschuss an FGF23 kommt es zu einem renalen Phosphatverlust, welcher zu einer Hypophosphatämie führt und verhindert, dass die Epiphysen (Wachstumsplatten) bei pädiatrischen Patienten mineralisieren und sich normal entwickeln. Die Folgen sind Kleinwuchs, schwere Knochendeformationen, Mobilitätseinschränkungen und Schmerzen. Dadurch, dass die Knochendeformationen auf Lebzeiten in der Pubertät fixiert werden, ist eine frühzeitige wirksame Behandlung der XLH dringend notwendig, um die Anhäufung der Knochendeformationen und starken körperlichen Beeinträchtigungen zu verzögern bzw. zu verhindern und eine maximal mögliche Lebensqualität der Patienten zu gewährleisten.

Die Ergebnisse legen nahe, dass Burosumab durch seinen Wirkmechanismus die Phosphathomöostase normalisiert und die damit verbundenen weiteren Stoffwechselkreisläufe wie die des Calciums und Vitamin D positiv beeinflusst und normalisiert. Ebenfalls scheint die Niere durch das Entgegenwirken des pathologischen Phosphatverlusts entlastet und Folgeerkrankungen, wie Nierenschäden und Nephrokalzinosen, die häufig bei einer Substitutionstherapie beobachtet werden, bleiben aus. Daher hat die Burosumab-Behandlung ein besseres Sicherheitsprofil als die Behandlung mit einer Substitutionstherapie. Die Ergebnisse unterstützen den Zusatznutzen von Burosumab hinsichtlich der besseren Wirksamkeit unter ähnlichem Nebenwirkungsprofil bei der Behandlung von Kindern mit XLH.

Medizinischer Zusatznutzen von Burosumab

Der medizinische Zusatznutzen von Burosumab wurde bereits im ersten Dossier vom 15. April 2018 durch die Ergebnisse der Studien UX023-CL201 und UX023-CL205 dargelegt. Eindrucksvoll gestützt werden diese nun im aktuellen Dossier mit den hier dargestellten Ergebnissen der randomisierten, offenen und aktiv kontrollierten Studie UX023-CL301 mit der Evidenzstufe 1b. In der Gesamtschau der Endpunkte Morbidität, Lebensqualität und Sicherheit wird für Burosumab zur Behandlung von Kindern ab 1 Jahr und Jugendlichen in der

Skelettwachstumsphase mit XLH mit röntgenologischem Nachweis einer Knochenerkrankung insgesamt ein beträchtlicher Zusatznutzen abgeleitet. Es handelt sich bei der Burosumab-Behandlung um eine bisher noch nicht erreichte deutliche Verbesserung der Erkrankung basierend insbesondere auf einer Abschwächung schwerwiegender Symptomatik (Kleinwuchs, stark eingeschränkte Gehfähigkeit, erhebliche Mobilitäts- und Funktionsstörungen, einhergehende starke Schmerzen, geringere Lebensqualität). Die Überprüfung der Sicherheit zeigt keine auffälligen, neuen unerwünschten Ereignisse. Ferner weist die Burosumab-Behandlung ein deutlich besseres Nebenwirkungsprofil im Vergleich zur Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D, welche schwerwiegende Nebenwirkungen wie Nierenschäden, Nephrokalzinose und einen sekundären Hyperparathyreoidismus zur Folge haben kann, auf. Burosumab stellt die erste und einzige kausale Therapie zur Behandlung der XLH dar und behandelt somit die zugrunde liegende Pathophysiologie der FGF23-induzierten Hypophosphatämie. Aufgrund der raschen Normalisierung der Mineral- und Vitamin-Dysregulation und der damit verbundenen normalisierten Knochenmineralisierung kommt es zur Rückbildung skelettaler Veränderungen und Reduktion der Knochen- und Gelenkschmerzen. Die für die Patienten spürbare Linderung der Erkrankung durch die Verringerung der Schmerzen und Mobilitätseinschränkungen führt zu einem deutlichen Anstieg ihrer Lebensqualität. Die Initiierung eines Arzneimittel-Härtefallprogramms beim Paul-Ehrlich-Institut (PEI) unterstreicht die fehlenden therapeutischen Alternativen. Dies gilt als Voraussetzung für die Teilnahme am Arzneimittel-Härtefallprogramm. Es zeigt einerseits, dass unabhängige Fachexperten wie das PEI ebenfalls von Burosumab überzeugt sind und andererseits, dass Burosumab eine erhebliche Versorgungslücke geschlossen hat. Außerdem ist in diesem Zusammenhang auch die Ansicht der Patientenvertretung während des ersten Nutzenbewertungsverfahrens zu erwähnen, welche sich ebenfalls für einen beträchtlichen Zusatznutzen von Burosumab bei der Behandlung der XLH ausgesprochen hatte. Denn bereits im ersten Nutzenbewertungsverfahren zeigten die Ergebnisse der pivotalen Studien (UX023-CL201 und UX023-CL205) eindeutige Hinweise auf einen beträchtlichen Zusatznutzen der Burosumab-Behandlung. Durch die weiteren Ergebnisse von Woche 40 und Woche 64 der randomisierten, offenen und aktiv kontrollierten Studie UX023-CL301 können die vorherigen Annahmen mehr als nur eindrucksvoll bestätigt werden. Es lässt sich somit konstatieren, dass die Ergebnisse der pivotalen Studien (UX023-CL201 und UX023-CL205) und die in diesem Dossier dargestellten Ergebnisse der RCT UX023-CL301 den geforderten **beträchtlichen Zusatznutzen** für den Wirkstoff Burosumab zur Behandlung der XLH bei Kindern vollumfänglich rechtfertigen.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des

Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-67: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Kinder ab 1 Jahr und Jugendlichen in der Skelettwachstumsphase mit XLH und röntgenologischen Nachweis einer Knochenerkrankung	Beträchtlicher Zusatznutzen

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹³, Molenberghs 2010¹⁴). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁵) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006¹⁶) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

¹³ Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁴ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: *A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials*. *Stat Methods Med Res* 2010; 19(3): 205-236.

¹⁵ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

¹⁶ Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht zutreffend.

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Tabelle 4-68: Liste der zur Bewertung des Zusatznutzens eingeschlossenen Studien

Studie	Zitate			Titel
	Studien- unterlagen	Registereinträge	Publika- tionen	
- Randomisierte kontrollierte Studien				
UX023-CL301	[19-21]	- Clinicaltrials.gov [51] - EU Clinical Trials Registry [52] - WHO ICTRP [53] - Pharmnet.bund [54] - EMA <i>clinical data</i> ^a	[50]	Efficacy and Safety of Burosumab (KRN23) Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X Linked Hypophosphatemia (XLH)
Studien für indirekte Vergleiche				
Keine				
Nicht randomisierte vergleichende Studien				
Keine				
Weitere Untersuchungen				
Keine				
^a Die Suche unter EMA <i>clinical data</i> am 16.07.2019 ergab keine Treffer.				

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Kyowa Kirin Holdings B.V. (2018): CRYSVITA 10/20/30 mg Injektionslösung; Fachinformation. Stand: Januar 2019 [Zugriff: 22.07.2019]. URL: <http://www.fachinfo.de>.
2. Mohnike K, Klingbiel K-H (2004): Familiäre hypophosphatämische Rachitis - Diagnostik, Betreuung und Langzeitkonsequenzen. Berlin: ABW Wissenschaftsverlag GmbH.
3. Haffner D, Waldegger S (2016): Chapture 35: Disorders of Phosphorus Metabolism. In: Geary D.F. , Schaefer F: Pediatric Kidney Disease. Berlin Heidelberg: Springer-Verlag; 953-72.
4. Ultragenyx pharmaceutical inc. (2016): SF-10_Health Survey for Children_ Questionnaire.
5. Ultragenyx pharmaceutical inc. (2016): UX023-CL301 PROMIS v2 Parent report (5-7 years) Questionnaire.
6. Ultragenyx pharmaceutical inc. (2016): UX023-CL301 PROMIS v2 Pediatric Self-Report (8-12 years) Questionnaire.
7. Ultragenyx pharmaceutical inc. (2016): Clinical Evaluator Training Certificate V1.0 for protocol UX023-CL301.
8. Ultragenyx pharmaceutical inc. (2016): Biomedical Systems: Independent Review Manual X-ray RSS.
9. Ultragenyx pharmaceutical inc. (2017): Biomedical Systems: Independent Review Manual X-ray RGI-C.
10. Carpenter T, Imel E, Holm I, de Beur S, Insogna K (2011): A Clinician's Guide to X-Linked Hypophosphatemia. Journal of for Bone and Mineral Research; 26(7):1381–8.
11. Linglart A, Biosse-Duplan M, Briot K, Chaussain C, Esterle L, Guillaume-Czitrom S, et al. (2014): Therapeutic management of hypophosphatemic rickets from infancy to adulthood. Endocrine connections; 3(1):R13-30.
12. Gemeinsamer Bundesausschuss (G-BA) (2018): Mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses hier: Wirkstoff Burosumab vom 21. August 2018. [Zugriff: 05.08.2019]. URL: https://www.g-ba.de/downloads/91-1031-359/2018-08-21_Wortprotokoll_Burosumab_D-349.pdf.
13. Deutsche Gesellschaft für Kinderendokrinologie und -diabetologie (DGKED) e. V. (2016): S1-Leitlinie 174-008: Hereditäre hypophosphatämische Rachitiden; Stand: 03.2016. [Zugriff: 22.07.2019]. URL: https://www.awmf.org/uploads/tx_szleitlinien/174-008l_S1_Hereditaere_hypophosphataemische_Rachitiden_2016-05.pdf.
14. Gemeinsamer Bundesausschuss (G-BA) (2017): Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 AM-NutzenV Beratungsanforderung 2017-B-176.

15. Gemeinsamer Bundesausschuss (G-BA) (2015): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Secukinumab. [Zugriff: 12.07.2019]. URL: https://www.g-ba.de/downloads/40-268-3450/2015-11-27_AM-RL-XII_Secukinumab_2015-06-01-D167_TrG.pdf.
16. International Conference on Harmonization (ICH) (1995): Guideline for industry - The extent of population exposure to assess clinical safety: for drugs intended for long-term treatment of non-life-threatening conditions E1. [Zugriff: 16.08.2019]. URL: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E1/Step4/E1_Guideline.pdf.
17. Ultragenyx pharmaceutical (2017): Validation of the Rickets Severity Scale (RSS) and Radiographic Global Impression of Change Scale (RGI-C) for the Assessment of Rickets in Children with X-Linked Hypophosphatemia.
18. Ultragenyx pharmaceutical (2017): Ultragenyx RSS Variability Assessment for Study UX023-CL201 Version 1.0 (11 July 2017).
19. Ultragenyx pharmaceutical (2017): Study Protocol UX023-CL301: A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH).
20. Ultragenyx pharmaceutical (2019): Clinical Study Report UX023-CL301: A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH).
21. Ultragenyx pharmaceutical Inc. (2018): Statistical Analysis Plan: Study Number UX023-CL301: A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH).
22. Thacher TD, Pettifor JM, Tebben PJ, Creo AL, Skrinar A, Mao M, et al. (2019): Rickets severity predicts clinical outcomes in children with X-linked hypophosphatemia: Utility of the radiographic Rickets Severity Score. *Bone*; 122:76-81.
23. Ultragenyx pharmaceutical (2016): WhitePaper: Rickets in X-linked Hypophosphatemia Biology, Clinical Features, and Radiographic Findings.
24. Gemeinsamer Bundesausschuss (G-BA) (2016): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Asfotase alfa. [Zugriff: 08.07.2019]. URL: https://www.g-ba.de/downloads/40-268-3662/2016-03-17_AM-RL-XII_Asfotase-alfa_D-188_TrG.pdf.
25. Ultragenyx pharmaceutical inc. (2016): Clinical Evaluator Reference Manual Version 1.2 (15 July 2016), Protocol: UX023-CL301: A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH).
26. Robert-Koch-Institut (RKI) (2013): Referenzperzentile für anthropometrische Maßzahlen und Blutdruck aus der Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland (KiGGS) - 2. erweiterte Auflage. [Zugriff: 24.05.2019]. URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/KiGGS_Referenzperzentile.pdf?blob=publicationFile.

27. Centers for Disease Control and Prevention (CDC) (2009): National Center for Health Statistics: Data Tables: selected percentiles and LMS parameters, selected Z-score values, 1977 NCHS Growth Chart Equations. [Zugriff: 19.03.2018]. URL: https://www.cdc.gov/growthcharts/data_tables.htm.
28. Gemeinsamer Bundesausschuss (G-BA) (2016): Nutzenbewertung Dossierbewertung für Orphan Drugs Asfotase alfa (Strensiq®) von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V; Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO: Wirkstoff: Asfotase alfa. [Zugriff: 12.07.2019]. URL: https://www.g-ba.de/downloads/92-975-1192/2015-12-23_Nutzenbewertung-G-BA_Asfotase%20alfa.pdf.
29. Rosario AS, Schienkiewitz A, Neuhauser H (2011): German height references for children aged 0 to under 18 years compared to WHO and CDC growth charts. *Ann Hum Biol*; 38(2):121-30.
30. Carpenter TO, Whyte MP, Imel EA, Boot AM, Hogler W, Linglart A, et al. (2018): Burosumab Therapy in Children with X-Linked Hypophosphatemia. *The New England journal of medicine*; 378(21):1987-98.
31. American Thoracic Society (2002): American Thoracic Society (ATS) Statement: Guidelines for the Six-Minute Walk Test *Am J Respir Crit Care Med*; 166:111-7.
32. Ultragenyx pharmaceutical (2017): Patient-Reported Outcome Measures: Clinical Outcome Assessment Dossier: POSNA-PODCI, PROMIS®, FPS-R and 6MWT for Pediatric patients with X-linked hypophosphatemia (Version: 5.0, Last Updated: 6 July 2017).
33. Ultragenyx pharmaceutical (2016): For Protocol UX023-CL201: Clinical Evaluator Training Certificate V1.0 20 August 2016.
34. Zivicnjak M, Schnabel D, Billing H, Staude H, Filler G, Querfeld U, et al. (2011): Age-related stature and linear body segments in children with X-linked hypophosphatemic rickets. *Pediatr Nephrol*; 26(2):223-31.
35. Schnabel D (2013): Genetische Störungen der Knochenmineralisation. In: Burckhardt P: *Osteologie des Kindes- und Jugendalters*. Schattauer.
36. Broderick JE, DeWitt EM, Rothrock N, Crane PK, Forrest CB (2013): Advances in Patient-Reported Outcomes: The NIH PROMIS((R)) Measures. *EGEMS (Wash DC)*; 1(1):1015.
37. Health Measures (2019): List of Pediatric Measures: Available PROMIS® Measures for Pediatric Self-Report (ages 8-17) and Parent Proxy Report (ages 5-17) [Zugriff: 09.07.2019]. URL: <http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-pediatric-measures>.
38. Lai JS, Stucky BD, Thissen D, Varni JW, DeWitt EM, Irwin DE, et al. (2013): Development and psychometric properties of the PROMIS((R)) pediatric fatigue item banks. *Qual Life Res*; 22(9):2417-27.
39. Varni JW, Stucky BD, Thissen D, Dewitt EM, Irwin DE, Lai JS, et al. (2010): PROMIS Pediatric Pain Interference Scale: an item response theory analysis of the pediatric pain item bank. *J Pain*; 11(11):1109-19.
40. DeWitt EM, Stucky BD, Thissen D, Irwin DE, Langer M, Varni JW, et al. (2011): Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: built using item response theory. *J Clin Epidemiol*; 64(7):794-804.
41. Adelphi Values (2017): UX7575A | Version 2_0: Qualitative research in pediatric XLH: Combined concept elicitation and cognitive debriefing interviews to understand the

- patient and caregiver experience of XLH, and assess the face and content validity of the PROMIS items, POSNA-PODCI and FPS-R.
42. Irwin DE, Varni JW, Yeatts K, DeWalt DA (2009): Cognitive interviewing methodology in the development of a pediatric item bank: a patient reported outcomes measurement information system (PROMIS) study. *Health Qual Life Outcomes*; 7:3.
 43. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B (2001): The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain*; 93(2):173-83.
 44. Orpha net (2012): Hypophosphatämie, X-chromosomale. [Zugriff: 17.06.2019]. URL: http://www.orpha.net/consor/cgi-bin/OC_Exp.php?Expert=89936.
 45. European Medicinal Agency (EMA) (2008): CHMP Assessment report for sovrima, International Nonproprietary Name: idebenone, Procedure No. EMEA/H/C/000908 [Zugriff: 25.07.2018]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/000908/WC500070576.pdf.
 46. European Medicinal Agency (EMA) (2014): CHMP assessment report Tracleer, International non-proprietary name: Bosentan, Procedure No. EMEA/H/C/000401/II/0066. [Zugriff: 25.07.2018]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000401/WC500184279.pdf.
 47. Saris-Baglama R, DeRosa M, Raczek A (2006): Development, Validation, and Norming of the SF-10 for Children of Health Survey. *Quality of Life Research*; 15(S1):A-145.
 48. Saris-Baglama R, DeRosa M, Raczek A, Bjorner J, Ware J (2007): The SF-10TM Health Survey for Children: A User's Guide. *Quality Metric Incorporated* 235-74.
 49. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2017): IQWiG-Methodenpapier Version 5.0. [Zugriff: 19.07.2018]. URL: <https://www.iqwig.de/de/methoden/methodenpapier.3020.html>.
 50. Imel EA, Glorieux FH, Whyte MP, Munns CF, Ward LM, Nilsson O, et al. (2019): Burosumab versus conventional therapy in children with X-linked hypophosphataemia: a randomised, active-controlled, open-label, phase 3 trial. *Lancet*; 393(10189):2416-27.
 51. Ultragenyx Pharmaceutical Inc|Kyowa Kirin Co, Ltd (2019): UX023-CL301 - Efficacy and Safety of Burosumab (KRN23) Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients With X Linked Hypophosphatemia (XLH) - CTgov (NCT02915705). URL: <https://ClinicalTrials.gov/show/NCT02915705>
 52. Ultragenyx Pharmaceuticals Inc (2016): UX023-CL301 - A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH) - EU (2016-000600-29). URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000600-29
 53. Ultragenyx Pharmaceuticals Inc (2016): UX023-CL301;EMA/902676 - A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH) (EUCTR2016-000600-29-DK). Stand des Eintrags: 3 July 2017. URL: <https://www.clinicaltrialsregister.eu/ctr-search/trial/2016-000600-29/DK>
 54. Ultragenyx Pharmaceuticals Inc (2016): A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients with X-linked Hypophosphatemia (XLH) -

- PharmNet.Bund (2016-000600-29). [Zugriff: 09.07.2019]. URL: <https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html>
55. Geiger R, Strasak A, Treml B, Gasser K, Kleinsasser A, Fischer V, et al. (2007): Six-minute walk test in children and adolescents. *The Journal of pediatrics*; 150(4):395-9, 9 e1-2.
 56. Thissen D, Liu Y, Magnus B, Quinn H, Gipson DS, Dampier C, et al. (2016): Estimating minimally important difference (MID) in PROMIS pediatric measures using the scale-judgment method. *Qual Life Res*; 25(1):13-23.
 57. Rossaint J, Oehmichen J, Van Aken H, Reuter S, Pavenstadt HJ, Meersch M, et al. (2016): FGF23 signaling impairs neutrophil recruitment and host defense during CKD. *J Clin Invest*; 126(3):962-74.
 58. Evans AM (2008): Growing pains: contemporary knowledge and recommended practice. *J Foot Ankle Res*; 1(1):4.
 59. Kyowa Kirin Pharmaceutical Development Inc. (2019): Interaction test and Subgroup Analysis (Gender) on Percentage of time in normal range (3.2 – 6.1 mg/dL) of serum phosphorus - week 40 and 64.
 60. Kyowa Kirin Pharmaceutical Development Inc. (2019): Interaction test and Subgroup Analysis (Gender) on Change in Faces Pain Scale – Revised (FPS-R) from Baseline - week 40 and 64.
 61. Kyowa Kirin Pharmaceutical Development Inc. (2019): Interaction test and Subgroup Analysis (Region) on change in rickets assessed by the RGI-C global score from Baseline - week 64.
 62. Kyowa Kirin Pharmaceutical Development Inc. (2019): Interaction test and Subgroup Analysis (Region) on change in Standing Height/Recumbent Length Z Score from Baseline - week 64.
 63. Kyowa Kirin Pharmaceutical Development Inc. (2019): Interaction test and Subgroup Analysis (Region) on Change in serum phosphorus [mg/dL] to mean postbaseline values from Baseline - 40 and 64.
 64. Kyowa Kirin Pharmaceutical Development Inc. (2019): Interaction test and Subgroup Analysis (Region) on Proportion of subjects reaching the normal range (3.2 - 6.1 mg/dL) of serum phosphorus - week 40 and 64.
 65. Europäische Kommission (2018): Durchführungsbeschluss der Kommission vom 19.2.2018 über die Erteilung einer bedingten Zulassung für das Humanarzneimittel für seltene Leiden "CRYSVITA - Burosumab" gemäß Verordnung (EG) Nr. 726/2004 des Europäischen Parlaments und des Rates.
 66. Gemeinsamer Bundesausschuss (G-BA) (2018): Nutzenbewertung nach § 35a SGB V - Nutzenbewertungsverfahren zum Wirkstoff Burosumab. [Zugriff: 22.07.2019]. URL: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/359/>.
 67. Kyowa Kirin GmbH (2018): Dossier zur Nutzenbewertung gemäß § 35a SGB V; Burosumab (Crysvita®); X-chromosomal gebundene Hypophosphatämie (XLH); Modul 4A vom 27.03.2018. [Zugriff: 25.07.2018]. URL: https://www.g-ba.de/downloads/92-975-2391/2018-03-27_Modul4_Burosumab.pdf.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMWI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁷ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-69: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren in EMBASE

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	11.07.2019	
Zeitsegment	1974 to 2019 July 10	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁸] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	krn23.mp.	45
2	exp krn23/	85
3	krn 23.mp.	16
4	exp krn 23/	85
5	burosumab.mp.	89
6	exp burosumab/	85
7	anti-FGF 23.mp.	9
8	anti-FGF23.mp.	87
9	crysvita.mp.	10
10	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	159
11	random*.tw. or placebo*.mp. or double-blind*.tw.	1677739
12	10 and 11	39

Die Suche in EMBASE ergibt insgesamt 39 Treffer. Einer der Treffer erfüllt die Einschlusskriterien aus Tabelle 4-4.

¹⁸ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Tabelle 4-70: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren in MEDLINE

Datenbankname	MEDLINE	
Suchoberfläche	Ovid	
Datum der Suche	11.07.2019	
Zeitsegment	Ovid MEDLINE(R) 1946 to July Week 1 2019, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations 1946 to July 10, 2019 Ovid MEDLINE(R) Daily Update July 10, 2019	
Suchfilter	Modifizierter* Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁹] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	krn23.mp.	14
2	krn 23.mp.	1
3	burosumab.mp.	24
4	anti-FGF 23.mp.	6
5	anti-FGF23.mp.	40
6	crysvita.mp.	5
7	1 or 2 or 3 or 4 or 5 or 6	67
8	randomi#ed controlled trial.pt. or randomi#ed.mp. or placebo.mp.	870602
9	7 and 8	6
* Modifikation berücksichtigt sowohl britische als auch amerikanische Schreibweise (siehe # 8)		

Die Suche in MEDLINE ergibt insgesamt 6 Treffer. Einer der Treffer erfüllt die Einschlusskriterien aus Tabelle 4-4.

¹⁹ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Tabelle 4-71: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren in Cochrane

Datenbankname	Cochrane	
Suchoberfläche	Cochrane Library	
Datum der Suche	11.07.2019	
Zeitsegment	Keine Einschränkung	
Suchfilter	Keine Suchfilter	
#	Suchbegriffe	Ergebnis
1	krn23	18
2	krn 23	6
3	burosumab	14
4	anti-FGF 23	1
5	anti-FGF23	18
6	crysvita	0
7	#1 or #2 or #3 or #4 or #5 or #6	31

Die Suche in Cochrane ergibt insgesamt 31 Treffer, wovon 28 Treffer unter Trials eingeordnet sind.

Die Suche ergab eine relevante Studie, die in der Suchoberfläche EMBASE und MEDLINE gefunden wurde. Diese Studie ist die in diesem Dossier dargestellte offene, randomisierte und aktiv kontrollierte Phase-III-Studie UX023-CL301 [50].

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Entfällt. Es werden keine indirekten Vergleiche auf Basis von RCTs durchgeführt.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Entfällt. Es werden keine weiteren nicht randomisierten vergleichenden Studien im vorliegenden Dossier dargestellt.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Entfällt. Es werden keine weiteren Untersuchungen im vorliegenden Dossier durchgeführt.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. *clinicaltrials.gov*), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. *http://www.clinicaltrials.gov*), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-72: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in *clinicaltrials.gov*

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	09.07.2019
Suchstrategie	[OTHER TERMS] KRN23 OR KRN 23 OR burosumab OR anti-FGF 23 OR anti-FGF23 OR crysvita
Treffer	19

Die Suche in *clinicaltrials.gov* ergibt insgesamt 19 Treffer. Keiner der Treffer erfüllt die Einschlusskriterien aus der Tabelle 4-4.

Tabelle 4-73: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EU Clinical Trials Register (EU-CTR)

Studienregister	EU Clinical Trials Register (EU-CTR)
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	09.07.2019
Suchstrategie	KRN23 OR "KRN 23" OR burosumab OR "anti-FGF 23" OR anti-FGF23 OR crysvita
Treffer	5

Die Suche in EU-CTR ergibt insgesamt 5 Treffer. Keiner der Treffer erfüllt die Einschlusskriterien aus der Tabelle 4-4.

Tabelle 4-74: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in WHO International Clinical Trials Registry Platform (WHO ICTRP)

Studienregister	WHO International Clinical Trials Registry Platform (WHO ICTRP)
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	10.07.2019
Suchstrategie	KRN23 OR KRN 23 OR burosumab OR anti-FGF 23 OR anti-FGF23 OR crysvita
Treffer	36 Treffer 21 Studien

Die Suche in WHO ICTRP ergibt insgesamt 36 Treffer zu 21 Studien. Keiner der Treffer erfüllt die Einschlusskriterien aus der Tabelle 4-4.

Tabelle 4-75: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in Pharmnet.Bund

Studienregister	Pharmnet.Bund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html
Datum der Suche	09.07.2019
Suchstrategie	KRN23 [Textfelder] ODER KRN 23 [Textfelder] ODER burosumab [Textfelder] ODER anti-FGF 23 [Textfelder] ODER anti-FGF23 [Textfelder] ODER crysvita
Treffer	1

Die Suche in Pharmnet.Bund ergibt insgesamt 1 Treffer. Keiner der Treffer erfüllt die Einschlusskriterien aus der Tabelle 4-4.

Tabelle 4-76: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EMA *clinical data*

Studienregister	EMA <i>clinical data</i>	
Internetadresse	https://clinicaldata.ema.europa.eu/web/cdp/home	
Datum der Suche	16.07.2019	
#	Suchbegriff	Treffer
1	krn23	0
2	krn 23	0
3	burosumab	0
4	anti-FGF 23	0
5	anti-FGF23	0
6	crysvita	0

Die Suche in EMA *clinical data* ergab keine relevanten Treffer, welche die Einschlusskriterien aus der Tabelle 4-4 erfüllen.

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Entfällt. Es wird kein indirekter Vergleich auf Basis randomisierter kontrollierter Studien durchgeführt.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Entfällt. Es liegen keine nicht randomisierten vergleichenden Studien in dem hier betrachteten Anwendungsgebiet von Burosumab vor.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Entfällt. Es wurden keine weiteren Untersuchungen in dem hier betrachteten Anwendungsgebiet von Burosumab durchgeführt.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Die genaue Formulierung der Ausschlusskriterien ist der Tabelle 4-5 zu entnehmen.

Es wurde im Rahmen der bibliographischen Literaturrecherche keine Dokumente im Volltext gesichtet. Alle Dokumente wurden bereits im Titel- oder Abstract-Screening ausgeschlossen.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Entfällt. Es werden keine indirekten Vergleiche auf Basis von RCTs durchgeführt.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Entfällt. Es werden keine weiteren nicht randomisierten vergleichenden Studien im vorliegenden Dossier dargestellt.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Entfällt. Es werden keine weiteren Untersuchungen im vorliegenden Dossier dargestellt.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Die genaue Formulierung der Ausschlusskriterien ist der Tabelle 4-5 zu entnehmen.

In den folgenden Tabellen finden sich die in Studienregistern identifizierten und ausgeschlossenen Studien. Wurde die gleiche Studie in unterschiedlichen Registern identifiziert ist diese als Duplikat ausgeschlossen aber dennoch in jedem Register aufgelistet.

Tabelle 4-77: Ausgeschlossene Studien_ClinicalTrials.gov

Nr.	Titel	Register-nummer	Ausschlussgrund
1	Study of the Anti-FGF23 Antibody, Burosumab, in Adults With XLH	NCT03920072	A1
2	Burosumab for CSHS	NCT03993821	A1
3	Open Label Trial Assessing Safety and Efficacy of Burosumab (KRN23), in a Patient With ENS and Hypophosphatemic Rickets	NCT03581591	A5
4	A Study of KRN23 in Pediatric Patients With X-linked Hypophosphatemic Rickets/Osteomalacia	NCT03233126	A5
5	Study of the Safety, Pharmacodynamics (PD) and Efficacy of KRN23 in Children From 1 to 4 Years Old With X-linked Hypophosphatemia (XLH)	NCT02750618	A5
6	A Study of KRN23 in Subjects With Tumor-Induced Osteomalacia or Epidermal Nevus Syndrome	NCT02722798	A1
7	Open Label Study of KRN23 on Osteomalacia in Adults With X-linked Hypophosphatemia (XLH)	NCT02537431	A1
8	Study of KRN23 in Adults With X-linked Hypophosphatemia (XLH)	NCT02526160	A1
9	Long-Term Extension Study of KRN23 in Adult Subjects With X-Linked Hypophosphatemia (XLH)	NCT02312687	A5

Nr.	Titel	Register- nummer	Ausschluss- grund
10	Study of KRN23 in Adult Subjects With Tumor-Induced Osteomalacia (TIO) or Epidermal Nevus Syndrome (ENS)	NCT02304367	A1
11	Study of KRN23 (Burosumab), a Recombinant Fully Human Monoclonal Antibody Against Fibroblast Growth Factor 23 (FGF23), in Pediatric Subjects With X-linked Hypophosphatemia (XLH)	NCT02163577	A3
12	A Study of KRN23 in Subjects With X-linked Hypophosphatemic Rickets/Osteomalacia	NCT02181764	A1
13	An Extension Study of KRN23 in Adults With X-Linked Hypophosphatemia	NCT01571596	A1
14	A Repeated Study of KRN23 in Adults With X-Linked Hypophosphatemia	NCT01340482	A1
15	Individual Patient Compassionate Use of Burosumab	NCT03775187	A5
16	A Study of KRN23 in X-linked Hypophosphatemia	NCT00830674	A1
17	X-linked Hypophosphatemia Disease Monitoring Program	NCT03651505	A5
18	Phase 2 Study of KHK7580	NCT02216656	A1

Tabelle 4-78: Ausgeschlossene Studien_ EU Clinical Trials Register (EU-CTR)

Nr.	Titel	Register- nummer	Ausschluss- grund
19	A Phase 3b Open-label Study of the Anti-FGF23 Antibody, Burosumab (KRN23) in Adult Patients with X-linked Hypophosphatemia (XLH)	2018-000202-37	A1
20	A Randomized, open Label, Dose Finding, Phase 2 Study to Assess the Pharmacodynamics and Safety of the anti-FGF23 antibody, KRN23, in Pediatric Patients with X-linked Hypophosphatemia (XLH)	2014-000406-35	A5
21	A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study with Open-Label Extension to Assess the Efficacy and Safety of KRN23 in Adults with X-linked Hypophosphatemia (XLH)	2014-005529-11	A1
22	An Open-Label, Single-Arm, Phase 3 Study to Evaluate the Effects of KRN23 on Osteomalacia in Adults with X-linked Hypophosphatemia (XLH)	2015-001775-41	A5

Tabelle 4-79: Ausgeschlossene Studien_ WHO International Clinical Trials Registry Platform (WHO ICTRP)

Nr.	Titel	Register-nummer	Ausschluss-grund
23	Burosumab for CSHS	NCT03993821	A1
24	Individual Patient Compassionate Use of Burosumab	NCT03775187	A5
25	Study of the Anti-FGF23 Antibody, Burosumab, in Adults With XLH	NCT03920072	A1
26	Open Label Trial Assessing Safety and Efficacy of Burosumab (KRN23), in a Patient With ENS and Hypophosphatemic Rickets	NCT03581591	A5
27	A Study of KRN23 in Pediatric Patients With X-linked Hypophosphatemic Rickets/Osteomalacia	NCT03233126	A5
28	A Study of KRN23 in Pediatric Patients with X-linked Hypophosphatemic Rickets/Osteomalacia	JPRN-JapicCTI-173614	A5
29	A study to examine the efficacy of KRN23 on bone quality	EUCTR2015-001775-41-FR	A1
30	A Phase 2 Open-Label Trial of KRN23	JPRN-JapicCTI-163191	A1
31	A Study of KRN23 in Subjects With Tumor-Induced Osteomalacia or Epidermal Nevus Syndrome	NCT02722798	A1
32	Study of the Safety, Pharmacodynamics (PD) and Efficacy of KRN23 in Children From 1 to 4 Years Old With X-linked Hypophosphatemia (XLH)	NCT02750618	A5
33	Open Label Study of KRN23 on Osteomalacia in Adults With X-linked Hypophosphatemia (XLH)	NCT02537431	A1
34	Study of KRN23 in Adults With X-linked Hypophosphatemia (XLH)	NCT02526160	A1
35	Long-Term Extension Study of KRN23 in Adult Subjects With X-Linked Hypophosphatemia (XLH)	NCT02312687	A5
36	Study of KRN23 in Adult Subjects With Tumor-Induced Osteomalacia (TIO) or Epidermal Nevus Syndrome (ENS)	NCT02304367	A1
37	A Phase 1 study of KRN23	JPRN-JapicCTI-142601	A5

Nr.	Titel	Register- nummer	Ausschluss- grund
38	A Study of KRN23 in Subjects With X-linked Hypophosphatemic Rickets/Osteomalacia	NCT02181764	A5
39	A clinical study (phase 2) in hereditary Rickets (XLH) in children to assess the safety and working of the antibody anti-FGF23	EUCTR2014-000406-35-GB	A5
40	An Extension Study of KRN23 in Adults With X-Linked Hypophosphatemia	NCT01571596	A1
41	A Repeated Study of KRN23 in Adults With X-Linked Hypophosphatemia	NCT01340482	A1
42	A Study of KRN23 in X-linked Hypophosphatemia	NCT00830674	A1

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Entfällt. Es werden keine indirekten Vergleiche durchgeführt.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Entfällt. Es werden keine weiteren nicht randomisierten vergleichenden Studien im vorliegenden Dossier dargestellt.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Entfällt. Es werden keine weiteren Untersuchungen im vorliegenden Dossier dargestellt.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-80 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-80 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-80 (Anhang): Studiendesign und -methodik für Studie UX023-CL301

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Ziel dieser Studie war die Bewertung von Burosumab im Vergleich zu der aktiven Kontrolle (orales Phosphat/aktives Vitamin D) in Bezug auf die Verbesserung der Rachitis-symptomatik bei Kindern im Alter von 1 – 12 Jahren mit XLH.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Es handelt sich um eine multizentrische, randomisierte, offene, aktiv kontrollierte (Phase 3) Studie mit einem Zuteilungsverhältnis von 1:1 auf subkutane Behandlung mit Burosumab oder auf per oralen Einnahme von oralem Phosphat/aktivem Vitamin D (aktive Kontrolle). Die Randomisierung erfolgte mithilfe eines von einem unabhängigen Drittanbieter entwickelten Randomisierungsplans über ein Interaktives Web-Response-System (IWRS). Die Randomisierung wurde nach Rachitis-Schwere, Alter und Region stratifiziert.</p> <p>Die Studie besteht aus einer Screeningvisite inkl. Wash-out-Phase, in der die Kinder die Behandlung mit oralem Phosphat/aktivem Vitamin D für 7 Tage abbrechen, einer 64-wöchigen (Woche 0 - 64) Behandlungsphase und einer 76-wöchigen Extensionsphase (Woche 140).</p> <p><u>Behandlungsphase Burosumab (Woche 0 – 64):</u> Die Initialdosis von Burosumab war 0,8 mg/kg KG alle 2 Wochen s.c. Die Dosis konnte auf 1,2 mg/kg KG basierend auf den Nüchtern-Serumphosphatwert erhöht werden.</p> <p><u>Behandlungsphase aktive Kontrolle (Woche 0 – 64):</u></p>

Item ^a	Charakteristikum	Studieninformation
		<p>Mehrere tägliche Dosen von oralem Phosphat/aktivem Vitamin D. Aufgrund der Variabilität der Dosen und Einnahmezeiten erfolgte eine patientenindividuelle Gabe nach Ermessen des Arztes.</p> <p>Zu Verfügung gestellte Leitlinien: EU-Leitlinie: Orales Phosphat 45 – 70 mg/KG/Tag in 3–4 geteilten täglichen Dosen und Alfacalcidol 1 – 2 µg/Tag 1 x täglich US-Leitlinie: Orales Phosphat 20 – 40 mg/KG/Tag in 2–5 geteilten täglichen Dosen und Calcitriol 20 – 30 ng/KG/Tag in 2–3 geteilten täglichen Dosen</p> <p>Die Behandlung mit Alfacalcidol oder Calcitriol erfolgt basierend auf klinischen Parametern.</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Das Original-Studienprotokoll ist auf den 04. April 2016 datiert.</p> <p><u>Amendment 1 (Global) datiert auf den 03. November 2017:</u></p> <ul style="list-style-type: none"> • Die Behandlungs-/Extensionsphase wurde für Studienteilnehmer in Europa, USA, Kanada und Australien erweitert. Für Studienteilnehmer aus Japan und Korea ist das Studienende nach Woche 64 (finale Dosis Burosumab zu Woche 62); Patienten in diesen Ländern werden in eine separate klinische Studie mit Burosumab aufgenommen oder erhalten Burosumab auf andere Weise. Für Patienten in Europa, den USA, Kanada und Australien besteht die Studie aus der Behandlungsphase (64 Wochen) und der Extensionsphase (bis zu 76 Wochen) für eine Gesamtbehandlungsdauer von bis zu 140 Wochen. Für Patienten in Europa und Australien endet die Extensionsphase im September 2018; Die Patienten werden in separate klinische Studien mit Burosumab aufgenommen oder erhalten Burosumab auf andere Weise. Für Patienten in den USA und Kanada endet die Extensionsphase im September 2018 bzw. Juni 2019, wenn erwartet wird, dass kommerziell verfügbares Burosumab zur Verfügung steht; Die Patienten erhalten entweder kommerzielles Burosumab oder Burosumab auf andere Weise. Die Gesamtdauer der Behandlung in dieser Studie für Patienten in Europa, den USA, Kanada und Australien variiert je nach dem Anfangsdatum der Studieneinschreibung; jedoch wird Woche 140 nicht überschritten. • Behandlung im Zeitraum der Extensionsphase (nur in Europa, den USA,

Item ^a	Charakteristikum	Studieninformation
		<p>Kanada und Australien): Nach Abschluss der 64-wöchigen Behandlungsphase werden die Patienten, die zu Burosumab randomisiert wurden, die Behandlung mit Burosumab fortsetzen; Patienten, die nach Randomisierung einer aktiven Kontrolle unterzogen wurden, konnten zu der Burosumab-Behandlungsgruppe mit der Anfangsdosis/Behandlungsschema wechseln. Patienten in der aktiven Kontrollgruppe brechen die Behandlung nach der 64. Woche ab. Nach der Wash-out-Phase (von oralem Phosphats/aktivem Vitamin D) konnte die erste Dosis von Burosumab zu Woche 66 appliziert werden.</p> <ul style="list-style-type: none"> • Das maximale Geschlechterverhältnis wurde von 60% auf 70% weiblicher Patienten erhöht, um das reale Geschlechterverhältnis der XLH-Patientenpopulation in der Studie besser widerzuspiegeln. • Es wurde klargestellt, dass „mindestens“ und nicht „in etwa“ 20 Studienteilnehmer im Alter von 1 bis <5 Jahre sein müssen. <p>Auswahl der Studienpopulation:</p> <ul style="list-style-type: none"> • Das Einschlusskriterium # 4 wurde auf "Serumkreatinin unter dem altersadjustierten oberen Normalgrenzwert" revidiert. • Das Einschlusskriterium # 6 wurde überarbeitet, um anzuzeigen, dass die konventionelle Therapie "... 7 Tage vor der Randomisierungsvisite" und nicht "... vor der Screeningvisite" abgebrochen werden sollte. • Das Einschlusskriterium # 10 wurde aktualisiert, um sexuell aktive männliche Studienteilnehmer mit weiblichen Partnern im gebärfähigen Alter für die Dauer der Studie plus 12 Wochen nach Beendigung der Studie zu zwingen, ein Kondom mit Spermizid oder eine hochwirksame Empfängnisverhütungsmethode (statt 2 Methoden) zu verwenden. • Ausschlusskriterium # 1 wurde überarbeitet, um zu spezifizieren, dass das Tanner-Stadium 4 oder höher mithilfe von körperlicher Untersuchung in einem der folgenden Fälle beurteilt wird: Genitalien, Brust- oder Schamhaare. • Das Ausschlusskriterium # 7 wurde in den ersten 40 Wochen der Studie auf "Geplante orthopädische Chirurgie einschließlich Osteotomie, Implantation oder Entfernung von Klammernahtgeräten (staples), 8-Platten oder andere orthopädische Hardware"

Item ^a	Charakteristikum	Studieninformation
		<p>korrigiert.</p> <ul style="list-style-type: none"> • Einführung von zusätzlichen Urin-Schwangerschaftstests bei Frauen, die Burosumab erhielten, die die Menarche erreicht hatten (Visiten in Wochen 52, 76, 88, 100, 112, 124 und 140) • Es wurde klargestellt, dass für alle Studienteilnehmer eine zulässige Abweichung von ± 3 Tagen für jede Visite (mit Ausnahme der Screening- und Sicherheits-Follow-up-Visite) erlaubt ist. Weiterhin wurde klargestellt, dass die Burosumab-Applikation frühestens 8 Tage nach der letzten verabreichten Dosis erfolgen sollte und dass die Sicherheits-Follow-up-Visite eine zulässige Abweichung von ± 7 Tagen aufweist. <p>Sicherheitsanalysen</p> <ul style="list-style-type: none"> • Es wurde die Definition einer hochwirksamen Empfängnisverhütungsmethode hinzugefügt. • "Zahnereignisse" wie "Karies, Zahnextraktion, Wurzelkanal, Zahnabszess und Zahnfleischentzündung" wurden erklärt, um diese Ereignisse für die Beurteilung der Zahngesundheit zu spezifizieren. • Die UE-Berichterstattung für die aktive Kontrollgruppe wurde präzisiert und Anlage 1 hinzugefügt, die als Referenzsicherheitsinformation (RSI) für die oralen Phosphat/aktiven Vitamin-D-Präparate in der aktiven Kontrollgruppe dieser Studie dient. <p>Verbotene Medikation</p> <ul style="list-style-type: none"> • Es wurde klargestellt, dass in der Burosumab-Gruppe aktives Vitamin D, nicht die Vitamin D Supplementation, verboten ist. <p>Studienmedikation</p> <ul style="list-style-type: none"> • Es wurde spezifiziert, dass die Patienten, die aufgrund einer erhöhten Serumphosphat-Konzentration über dem ULN Burosumab absetzen mussten, mit der Hälfte der letzten verabreichten Dosis (d.h. der halben Dosis von entweder 0,8 oder 1,2 mg/kg) und einer Maximaldosis von 40 mg, Burosumab wiederaufnehmen dürfen. Die maximal zulässige Dosis von Burosumab wurde als 90 mg pro Applikation definiert. • Das Gesäß wird als potenzielle Injektionsstelle für die Applikation von Burosumab eingeschlossen.
4	Probanden / Patienten	Es wurden 32 Kinder in den aktiven Kontrollarm

Item ^a	Charakteristikum	Studieninformation
		und 29 Kinder in den Burosumab-Arm im Alter zwischen 1 – 12 Jahren mit X-chromosomaler Hypophosphatämie randomisiert.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Kinder, die in die Studie aufgenommen werden sollten, mussten alle Kriterien erfüllen:</p> <ul style="list-style-type: none"> - Männliche und weibliche Patienten im Alter von 1-12 Jahren mit radiologischem Hinweis auf eine Rachitis mit einem RSS Total Score von mind. 2,0 - <i>PHEX</i>-Mutation oder direktes Familienmitglied mit entsprechender X-chromosomaler Vererbung - Biochemische Befunde zu Baseline und/oder Screeningvisite, die mit XLH assoziiert sind (basierend auf Nüchternwerte [mindestens 4 Stunden]): Serumphosphat < 3,0 mg/dL (0,97 mmol/l) - Serum-Kreatinin zu Baseline und/oder Screeningvisite unter dem Alter entsprechenden oberen Normwert (ULN) (basierend auf Nüchternwerte [mindestens 4 Stunden]) - Serum 25(OH)D über dem unteren Normwert LLN; ≥ 16 ng/ml) zur Screeningvisite. (Falls 25(OH)D unter dem Normwert liegt, wird eine 25(OH)D-Substitution vorgeschrieben. Dies gilt unter der Annahme, dass der Patient sonst alle anderen Voraussetzungen erfüllt und erneut nach mindestens 7 Behandlungstagen gescreent wird - Der Patient hat für mindestens 12 aufeinanderfolgende Monate (bei Kindern ≥ 3 Jahre) oder für mind. 6 aufeinanderfolgende Monate (bei Kindern < 3 Jahre) 7 Tage vor der Randomisierungs-Visite sowohl orales Phosphat- als auch aktives Vitamin-D erhalten - Bereitschaft den Zugang zu früheren medizinischen Aufzeichnungen für die Erfassung von historischen Wachstums- und radiographischen Daten und dem Krankheitsverlauf zu gewähren - Schriftliche oder mündliche Zustimmung und schriftliche Einverständniserklärung eines rechtlich Bevollmächtigten nach Erläuterung der Studie und aller Messverfahren - Der Patient muss nach Ansicht des Prüfarztes bereit und in der Lage sein, alle Aspekte der Studie einzuhalten, den Studien-Visiten-Plan einzuhalten und die Anweisungen zu befolgen. - Frauen, die die Menarche erreicht haben, müssen während des Screenings einen negativen Schwangerschaftstest haben und sich während der Studie zusätzlichen Schwangerschaftstests unterziehen. Frauen im gebärfähigen Alter müssen bereit sein eine hochwirksame

Item ^a	Charakteristikum	Studieninformation
		<p>Empfängnisverhütungsmethode für die Dauer der Studie plus 12 Wochen nach Absetzung der Studienmedikation anzuwenden. Sexuell aktive männliche Patienten mit weiblichen Partnern im gebärfähigen Alter müssen zustimmen, ein Kondom mit Spermizid oder eine hochwirksame Verhütungsmethode für die Dauer der Studie plus 12 Wochen nach Absetzen der Studienmedikation zu verwenden.</p> <p>Ausschlusskriterien: Kinder, die eines der folgenden Ausschlusskriterien erfüllten, waren nicht zur Teilnahme an der Studie berechtigt:</p> <ul style="list-style-type: none"> - Tanner Stufe 4 oder höher in einem der folgenden Merkmale: Genitalien, Brust oder Schamhaare, basierend auf einer körperlichen Untersuchung - Größen-Perzentil > 50 basierend auf länderspezifischen Normen - Verwendung von Aluminiumhydroxid-Antazida (z. B. Maalox[®] und Mylanta[®]), systemischen Kortikosteroiden, Acetazolamid und Thiaziden innerhalb von 7 Tagen vor der Screeningvisite - Aktuelle oder vorherige Verwendung von Leuprorelin (z. B. Lupron[®], Viadur[®], Eligard[®]), Triptorelin (Trelstar[®]), Goserelin (Zoladex[®]) oder anderen Arzneimitteln, von denen bekannt ist, dass sie die Pubertät verzögern - Verwendung von Wachstumshormontherapie innerhalb von 12 Monaten vor der Screeningvisite - Vorhandensein einer Nephrokalzinose mit Nierenultraschall Graduierung von 4 basierend auf folgender Skala: <ul style="list-style-type: none"> 0 = normal 1 = schwacher hyperechogener Rand um die medulläre Pyramide 2 = intensiverer echogener Rand mit Echos, mit schwacher Füllung der Pyramide 3 = gleichmäßig intensive Echos in der ganzen Pyramide 4 = Steinbildung: alleiniger Fokus der Echos an der Spitze der Pyramide - Geplante oder empfohlene orthopädische Chirurgie; einschließlich Osteotomie, Implantation oder Entfernung von Klammernahtgeräten (staples), 8-Platten oder andere orthopädische Hardware innerhalb der ersten 40 Wochen der klinischen Studie - Hypo-/ oder Hyperkalzämie, definiert als Serumcalcium-Wert innerhalb der

Item ^a	Charakteristikum	Studieninformation
		<p>altersabhängigen Normalwerte (basierend auf Nüchternwerten [Mindestens 4 Stunden] der Screeningvisite und/oder Baseline</p> <ul style="list-style-type: none"> - Nachweis eines Hyperparathyreoidismus (iPTH-Konzentration 2,5X ULN) - Einnahme/ Verwendung von Medikamenten zur Unterdrückung von PTH (z. B. Cinacalcet, Kalzिमimetika) innerhalb von 2 Monaten vor der Screeningvisite - Vorliegen oder damaliges Vorliegen einer Bedingung des XLH-Patienten, die der Prüfarzt als ein hohes Risiko für eine schlechte Compliance oder einem möglichen Studienabbruch einstuft - Bestehen einer Begleiterkrankung oder eines Gesundheitszustandes, die die Teilnahme an der Studie beeinträchtigen oder die Patientensicherheit beeinflussen würde - Damalige rezidivierende oder Prädisposition für Infektionen oder bekannte Immundefizienz - Verwendung eines therapeutischen monoklonalen Antikörpers innerhalb von 90 Tagen vor der Screeningvisite oder vorherige allergische oder anaphylaktische Reaktionen auf einen monoklonalen Antikörper - Vorherige oder Vorliegen einer Hypersensitivität von KRN23 Hilfsstoffen, die nach Urteil des Prüfarztes für die Studie mit einem erhöhten Risiko für unerwünschte Effekte einhergehen - Verwendung eines Prüfprodukts oder Prüfmedikaments innerhalb von 30 Tagen vor dem Screening oder Voraussetzung eines Prüfpräparats vor Abschluss aller geplanten Studienbewertungen <p>oder <u>in Japan</u> die Verwendung eines Prüfpräparats oder eines Prüfmedikaments innerhalb von 4 Monaten vor der Screeningvisite oder die Notwendigkeit des Prüfpräparates vor Abschluss aller Studienbewertungen.</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>Sponsor war Ultragenyx Pharmaceutical Inc. Novato, CA, USA.</p> <p>In Studie UX023-CL301 wurden insgesamt 61 Kinder im Alter von 1-12 Jahren in 16 klinischen Prüfzentren in weltweit 7 Ländern aufgenommen (4 in USA, 3 in Kanada, 3 in Japan, 2 in Großbritannien, 1 in Australien, 1 in Korea und 1 in Schweden)</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur	Burosumab (KNR23) ist ein rekombinanter, vollständig humaner, monoklonaler Antikörper (IgG1), der die überschüssige biologische

Item ^a	Charakteristikum	Studieninformation
	Administration etc.	<p>Aktivität des Fibroblasten-Wachstumsfaktors 23 (FGF23) bindet und hemmt.</p> <p>Je 1 ml Burosumab wurde als sterile, klare, farblose und konservierungsmittelfreie Lösung in 5 ml-Einweg-Durchstechflaschen mit einer Konzentration von 10 mg/ml oder 30 mg/ml verwendet. Die Dosierung von Burosumab wird auf das Körpergewicht des Kindes berechnet. Die Injektion von Burosumab wurde abwechselnd subkutan (von geschultem Personal der Studie) in Bauch, Oberarm, Oberschenkel oder Gesäß verabreicht.</p> <p><u>Burosumab-Gruppe:</u> Die Initialdosis von Burosumab war 0,8 mg/kg KG, alle 2 Wochen s.c. Die Dosis konnte auf 1,2 mg/kg KG basierend auf den Nüchtern-Serumphosphatwert erhöht werden.</p> <p><u>Aktive Kontrollgruppe:</u> Diese erhielt mehrere tägliche Dosen von oralem Phosphat und aktivem Vitamin D. Aufgrund der Variabilität der Dosen und Einnahmezeiten erfolgte eine patientenindividuelle Gabe nach Ermessen des Arztes.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p><u>Wirksamkeit:</u></p> <p>Primärer Wirksamkeitsendpunkt: Evaluation der Wirksamkeit von Burosumab auf die Rachitis-Schwere bei Kindern mit XLH anhand des RGI-C Total Scores zu Woche 40 verglichen mit der aktiven Kontrolle</p> <ul style="list-style-type: none"> ○ Erhebungszeitpunkte: Screeningvisite, w40, w64 ○ Schulung und Verblindung der 3 Auswerter sowie Mehrfachbeurteilung <p>Sekundärer Wirksamkeitsendpunkt: Evaluation des Effekts von Burosumab verglichen mit der aktiven Kontrolle:</p> <p><u>Wesentliche sekundäre Endpunkte:</u></p> <ul style="list-style-type: none"> ● Wachstum, Wachstumsgeschwindigkeit und Deformation der unteren Extremitäten (gemessen durch RGI-C-Score des gestreckten Beins, sowie Z-Scores der Liegelänge/Stehhöhe) <ul style="list-style-type: none"> ○ Erhebungszeitpunkte: Screeningvisite, w40, w64 (RGI-C) bzw. Baseline, w24, w40, w64 (Stehhöhe/Liegelänge) ○ Schulung und Verblindung der Auswerter sowie

		<p>Mehrfachbeurteilung von 3 Auswertern (RGI-C)</p> <ul style="list-style-type: none"> • Änderung des RSS Total Score <ul style="list-style-type: none"> ○ Erhebungszeitpunkte: Screeningvisite, w40, w64 ○ Schulung und Verblindung des einzelnen Auswertern • Änderung der Blut-Parameter Serum-Phosphat und Alkalische Phosphatase (ALP) zur Bewertung der Knochenmineralisierung und daraus folgend der Rachitis-Schwere <p><u>Weitere sekundäre Endpunkte:</u></p> <ul style="list-style-type: none"> • Beurteilung der Rachitissymptomatik anhand von RGI-C- und RSS-Handgelenk- sowie Kniescores <ul style="list-style-type: none"> ○ Erhebungszeitpunkte: Screeningvisite, w40, w64 ○ Schulung und Verblindung des einzelnen Auswertern • Wachstum anhand der Wachstumsgeschwindigkeit (Z-Scores) • Pharmakodynamische Marker, die den Zustand der Phosphathomöostase über den Studienverlauf reflektieren (Phosphat im Serum und Urin, Serum-1,25(OH)₂-Vitamin D₃, TmP/GFR und TRP) • Gehfähigkeit (6MWT), Patienten-/Elternberichte zu Schmerz, Fatigue und körperliche Funktion/Mobilität (PROMIS; FPS-R) <ul style="list-style-type: none"> ○ Erhebungszeitpunkte: Baseline, w24, w40, w64 ○ Schulungen des klinischen Personals auf Basis der ATS Guidelines (6MWT) ○ Validierung des PROMIS-Fragebogens durch Interviewstudie ○ Schulungen des klinischen Personals (FPS-R) <p>Explorative Endpunkte:</p> <ul style="list-style-type: none"> • Gesundheitsbezogene Lebensqualität mittels Short Form 10 (SF-10) for Children Health Survey (SF-10: für Kinder ab 5 Jahren zur Screeningvisite; Domänen: <i>Physical Health Summary Score</i> (PHS-10) und <i>Psychosocial Summary Score</i> (PSS-10)) <ul style="list-style-type: none"> ○ Erhebungszeitpunkte: Baseline, w24, w40, w64 ○ Schulungen des klinischen Personals • Dentale Untersuchungen: Anzahl der Zahnereignisse von Karies, Zahnextraktion, Wurzelkanal, Zahnabszess und Gingivitis, die zu Baseline und nach Post-Baseline beurteilt wurden
--	--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Item ^a	Charakteristikum	Studieninformation
		<p>Sicherheitsendpunkte:</p> <ul style="list-style-type: none"> • Frequenz und Schwere der unerwünschten Ereignisse (UEs) und schwerwiegenden unerwünschten Ereignissen (SUEs) • Frühere und begleitende Medikation • Vitalzeichen und Gewicht • Körperliche Untersuchungen • Geschätzte, glomeruläre Filtrationsrate (eGFR) • Nierenultraschall mit Nephrokalzinosen Score • Echokardiogramm (ECHO) • Chemische, hämatologisch und urinalytische Parameter einschließlich zusätzliche Burosumab/XLH spezifische Parameter von Interesse (Serum-25-hydroxyvitamin-D, Lipase, Amylase, Kreatinin, Calcium und iPTH im Serum und Urin) • Anti-Burosumab Antikörperstatus • Elektrokardiogramm (EKG) <p>Pharmakokinetik Sicherheitsendpunkte:</p> <ul style="list-style-type: none"> • Serum Burosumab
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Diese Phase-III-Studie wurde durchgeführt, um die Wirkung von Burosumab auf die Verbesserung der Rachitissymptomatik mithilfe des RGI-C Global Scores zu Woche 40 im Vergleich zur aktiven Kontrolle (orales Phosphat/aktives Vitamin D) zu untersuchen. Dabei basiert die Annahme auf einen mittleren RGI-C-Global Score von 1,80 in der Burosumab-Gruppe, 1,40 in der aktiven Kontrollgruppe und eine Standardabweichung von 0,50.</p> <p>Unter dieser Annahme wurde eine Gesamtstichprobengröße von 60 pädiatrischen Patienten (30 pro Behandlungsgruppe) festgelegt, um einen solchen Unterschied im mittleren RGI-C-Global Score zu Woche 40 zwischen den Behandlungsgruppen unter Verwendung eines t-Tests (zweiseitiger α-Wert von 0,05) mit einer ungefähren Power von 80 % nachzuweisen. Eine Drop-out-Rate von 10 % wurde in die Berechnung der Stichprobengröße einbezogen.</p> <p>Das Alter der teilnahmeberechtigten Personen wird auf Altersverteilung überwacht. Etwa 20</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Patienten im Alter von 1 bis < 5 Jahren werden eingeschlossen (etwa 10 in jeder Behandlungsgruppe), und nicht mehr als 10 Patienten im Alter von 1 bis < 3 Jahren werden aufgenommen. Nicht mehr als 70 % weibliche Patienten werden aufgenommen.</p> <p>Final wurden 61 Patienten in die Studie aufgenommen und 1:1 auf die zwei Behandlungsgruppen (Burosumab (n=29) und aktive Kontrolle: orales Phosphat/aktives Vitamin D (n=32)) randomisiert.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Die Primäranalyse erfolgte zu Woche 40.</p> <p>Eine zusätzliche Analyse wird zu Woche 64 durchgeführt, um Langzeitbehandlungseffekte, zusätzliche Wirksamkeitsergebnisse und langfristige Sicherheit zu bewerten.</p> <p>Es kam zu keinen vorzeitigen Studienabbrüchen.</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	Siehe Antwort in Feld 8a und 8b
8a	Methode zur Generierung der zufälligen Zuteilung	Die Patienten wurden 1:1 auf die zwei Behandlungs-Gruppen Burosumab und aktive Kontrollgruppe (orales Phosphat/aktives Vitamin D) über ein Interaktives Web Response System (IWRS) aufgeteilt. Der entwickelte Randomisierungsplan wurde von einem unabhängigen Drittanbieter erstellt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung wurde nach Baseline-Rachitis-Schwere (RSS Total Score $\leq 2,5$ vs. $> 2,5$), Alter (< 5 Jahre vs. ≥ 5 Jahren) und Region (Japan vs. Rest der Welt) stratifiziert.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Da die randomisierte und aktiv kontrollierte Studie UX023-CL301 offen war, wusste sowohl Studienpersonal, Prüfarzt als auch der Patient selbst in welcher Behandlungsgruppe dieser zugeteilt wurde.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierungsliste mit zugewiesenen Behandlungskohorten wurde über ein IWRS, welches von einem Drittanbieter bereitgestellt wurde, erstellt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>Zu a) und b)</p> <p>Es wurde keine Verblindung vorgenommen, da eine Verblindung einer solchen Studie aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen Behandlung und häufigen Dosisanpassungen, die bei einer konventionellen</p>

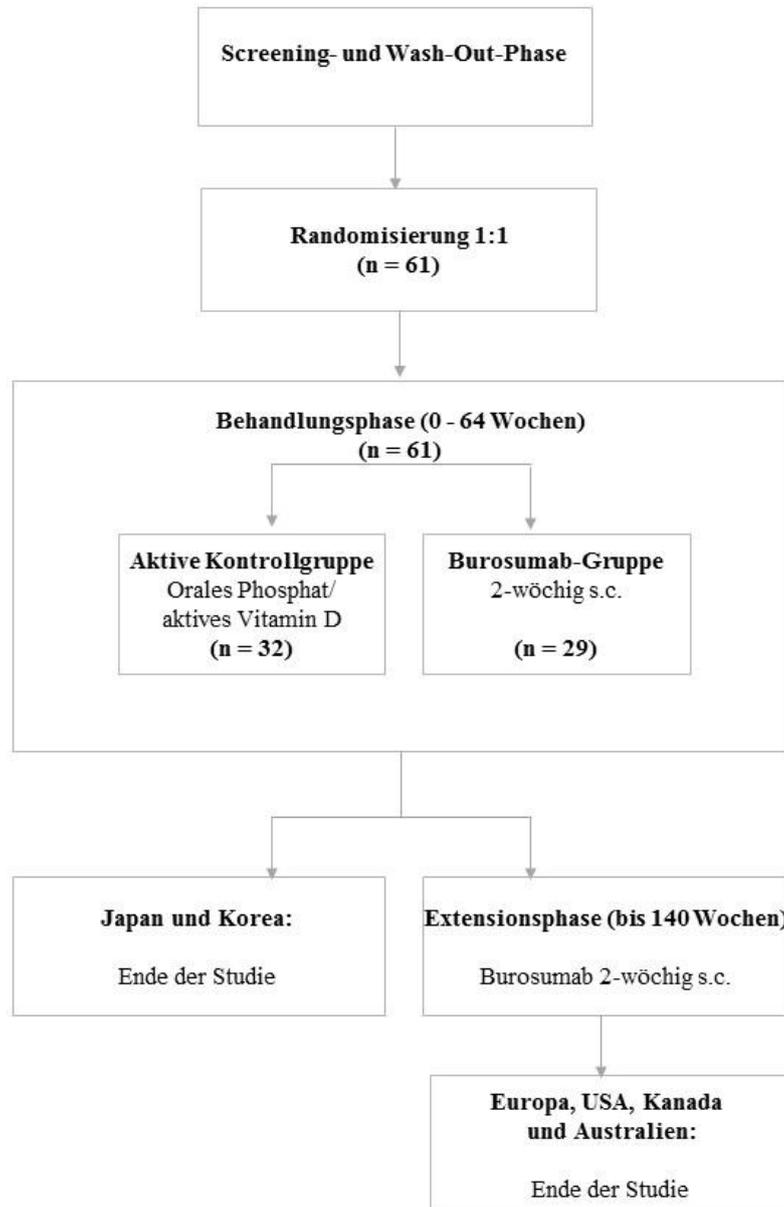
Item ^a	Charakteristikum	Studieninformation
		<p>Therapie (oralen Phosphat/aktives Vitamin D) erforderlich sind, problematisch wäre.</p> <p>Aus diesem Grund war dies eine offene Studie, weshalb sowohl das Studienpersonal als auch die teilnehmenden Patienten wussten zu welcher Behandlungsgruppe (Burosumab vs. aktive Kontrolle) sie zugeteilt wurden.</p> <p>Zu c)</p> <p>Eine Ausnahme bei der Verblindung stellen die Auswerter des RSS und RGI-C-Scores dar, welche bezüglich Dosierung, Dosierungsregime, Befolgen des Studienprotokolls, Behandlungsdauer und Vorbehandlung mit einer Substitutionstherapie verblindet waren.</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Wirksamkeits-Analyse:</p> <p>Bei wiederholten Messungen der klinischen Endpunkte wurde das <i>Generalized Estimating Equations</i> (GEE)-Modell zur Bewertung der Veränderung im Laufe der Zeit verwendet. Das GEE-Modell beinhaltet die Behandlungsgruppe, den Studienbesuch und die Interaktion zwischen Behandlungsgruppe und Studienbesuch als kategorische Variablen. Die Stratifizierungsfaktoren Baseline RSS Total Score und das Baseline-Alter werden ebenfalls einbezogen, sofern nicht anders angegeben. Modellbasierte Schätzungen der Veränderungen von Baseline, Standardfehler (SE) und entsprechende 95% -Konfidenzintervallen (KIs) werden ebenfalls zusammen mit p-Werten zur Bewertung der statistischen Signifikanz berechnet. Weitere unabhängige Variablen werden im Analyseabschnitt für die einzelnen Endpunkte angegeben.</p> <p>Analyse des primären Wirksamkeitsendpunktes:</p> <p>Der primäre Wirksamkeitsendpunkt wurde getestet, um den mittleren RGI-C-Global-Score zwischen der Burosumab-Gruppe und der aktiven Kontrollgruppe zu vergleichen.</p> <p>Bei der Analyse der Woche 40 wurde eine Kovarianzanalyse (ANCOVA-Modell) durchgeführt und beinhaltet den RGI-C-Global-Score der Behandlungsgruppe und Baseline-Altersstratifikationsfaktor als unabhängige Variablen, Baseline RSS Total Score als kontinuierliche Kovariate. Der Baseline RSS Total Score wurde nicht in das Modell aufgenommen. Zu Woche 64 wurde das <i>repeat</i></p>

Item ^a	Charakteristikum	Studieninformation
		<p><i>measurement model</i>, welches die Faktoren Behandlungsgruppe, Studienbesuch, Interaktion zwischen Behandlungsgruppe und Studienbesuch angepasst an RSS (Baseline) und Alter beinhaltet, durchgeführt.</p> <p>Der RSS-Gesamtscore wurde analog zum RGI-C analysiert.</p> <p>Analyse der sekundären Wirksamkeitsendpunkte:</p> <p>Analysen zu den weiteren Endpunkten wurden zu Woche 40 und Woche 64 durchgeführt. Diese werden deskriptiv für jede Behandlungsgruppe zusammengefasst. Der Vergleich der Behandlungs- und Kontrollgruppe erfolgt mittels <i>repeat measurement model</i>, welches die Faktoren Behandlungsgruppe, Studienbesuch, Interaktion zwischen Behandlungsgruppe und Studienbesuch angepasst an RSS (Baseline) und Alter beinhaltet. Weitere Baseline-Kovariaten werden gegebenenfalls zur Korrektur berücksichtigt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Responderanalysen:</p> <p>Die RGI-C Global Score Responderanalyse wurde für Woche 40 und wird ebenfalls für Woche 64 durchgeführt.</p> <p>RGI-C-Responder sind Patienten mit einem mittleren RGI-C Global Score $\geq + 2,0$.</p> <p>Die Woche 40 Analyse wurde mithilfe eines logistischen Regressionsmodells mit der Behandlungsgruppe und dem Baseline-Altersstratifikationsfaktor als unabhängige Variablen und der Baseline RSS Total Score als kontinuierliche Kovariante durchgeführt. Der Baseline RSS Total Score wurde nicht in das Modell aufgenommen.</p> <p>In der Woche 64 Analyse wurde ein ähnliches logistisches Regressionsmodell angewendet, das sowohl die Daten der Woche 40 als auch der Woche 64 als abhängige Variable, Behandlungsgruppe, Studienbesuch, Interaktion zwischen Behandlungsgruppe und Studienbesuch und Baseline-Altersstratifikationsfaktor als unabhängige Variablen enthält. Der Baseline RSS Total Score diente als kontinuierliche Kovariate. Der Baseline RSS Total Score wurde ebenfalls nicht in das Modell aufgenommen. Das Odds Ratio, bei der die Burosumab-Behandlung mit der konventionellen Therapie verglichen wurde, wurde zusammen mit 95% Konfidenzintervall (KI) und p-Wert angegeben. Die erwartete Responder-Rate mit 95% KI der individuellen Behandlungsgruppen wurde ebenfalls angegeben.</p> <p>Subgruppenanalysen:</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Für die Wirksamkeitsdaten wurden für folgende Endpunkte Subgruppen definiert:</p> <ul style="list-style-type: none"> - Rickets Severity Score RSS: Subgruppen definiert nach RSS Total Score zu Baseline. <ul style="list-style-type: none"> ○ Baseline RSS Total Score $\geq 2,5$ ○ Baseline RSS Total Score $< 2,5$ - Alter: Subgruppen definiert nach Alter zu Baseline. <ul style="list-style-type: none"> ○ Alter zu Baseline ≥ 5 ○ Alter zu Baseline < 5 - Herkunft: Subgruppen definiert nach Herkunft <ul style="list-style-type: none"> ○ Japan ○ Rest der Welt - Geschlecht: Subgruppen definiert nach Geschlecht. <ul style="list-style-type: none"> ○ männlich ○ weiblich <p>Sensitivitätsanalysen</p> <p>Zur Beurteilung der Robustheit des primären Analyseergebnisses der RSS- und RGI-C-Endpunkte für fehlende Werte wurden Sensitivitätsanalysen durchgeführt. Thesen können beinhalten, sind aber nicht beschränkt auf:</p> <ul style="list-style-type: none"> • Last Observation Carry Forward (LOCF): Die letzte nicht fehlende Beobachtung (Post-Baseline) wurde auf den entsprechenden Endpunkt zur Bewertung übertragen. Wenn keine Post-Baseline-Beobachtung verfügbar ist, wurde die Baseline-Beobachtung für die Auswertung verwendet; Wenn keine Baseline verfügbar ist, wurde der LOCF nicht durchgeführt. • Multiple Imputation (MI): Das PROC MI-Modell in SAS wurde verwendet um fehlende Werte unter Verwendung eines Markov-Chain-Monte-Carlo-Modells (MCMC) 100 Mal zu imitieren. Der kalkulatorische Datensatz wurde mit dem ANCOVA-Modell in der Woche 40 Analyse oder dem GEE-Modell in Woche 64 analysiert. Die Ergebnisse wurden mit dem PROC MIANALYZE-Modell in SAS kombiniert. Dabei wurden der Baseline RSS Total Score, Behandlungsgruppe und Baseline-Altersstratifikationsfaktor als Einflussvariablen verwendet. Andere Variablen wie auch die Änderung von ALP können ebenfalls als Einflussvariable betrachtet werden.
Resultate		

Item ^a	Charakteristikum	Studieninformation
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	61 Kinder im Alter von 1-12 Jahren wurden in die Studie aufgenommen und 1:1 zu den Behandlungsgruppen (Burosumab und aktive Kontrolle) zugeteilt.
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) Alle 61 Patienten wurden auf die zwei Behandlungsgruppen (Burosumab und aktive Kontrolle) aufgeteilt. b) 29 Patienten erhielten Burosumab subkutan; 32 Patienten wurden mit oralem Phosphat und aktivem Vitamin D behandelt. c) Alle 61 Patienten wurden für das primäre Zielkriterium herangezogen.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Es ging keiner der 61 Teilnehmer nach Randomisierung verloren oder wurde ausgeschlossen.
14	Aufnahme / Rekrutierung	s. nachstehendes Flussdiagramm
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	s. nachstehendes Flussdiagramm
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie läuft derzeit noch. Der Studienbericht mit den 64-Wochen-Daten liegt vor. Der dazugehörige Datenschnittpunkt beruht auf dem 30. Juli 2018. Derzeit befinden sich die Patienten in der Extensionsphase.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.



Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Tabelle 4-81 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UX023-CL301

Studie: UX023-CL301

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
UX023-CL301 Protocol (week 64) + Amendment	
UX023-CL301 Statistical Analysis Plan	UX023-CL201 SAP
UX023-CL301 Clinical Study Report	UX023-CL201 CSR

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie bestehend aus zwei Behandlungsgruppen. Eine Behandlungsgruppe erhielt Burosumab subkutan alle zwei Wochen (Q2W), während die aktive Kontrollgruppe eine patientenindividuelle tägliche orale Gabe von Phosphat und aktivierten Vitamin D bekam. Die Randomisierung der beiden Behandlungsgruppen erfolgte elektronisch.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Randomisierung der Patienten wurde 1:1 mittels *Interactive Web Response System (IWRS)* durchgeführt. Die Randomisierung wurde nach Rachitis-Schwere, Alter und Region stratifiziert.

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Nicht zutreffend.

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Randomisierung für die Studie UX023-CL301 erfolgte elektronisch über das *Interactive Web Response System* (IWRS).

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Nicht zutreffend.

3. Verblindung von Patienten und behandelnden Personen

Patient:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung der Patienten oder der behandelnden Personen. Die Patienten der Kontrollgruppe mit der oralen Gabe von Phosphat/aktivierten Vitamin D sowie die Teilnehmer der Burosumab-Gruppe waren sich ihrer Medikation bewusst. Eine Verblindung der Patienten stellt sich aufgrund der unterschiedlichen Verabreichungsmethoden (subkutane Injektion vs. mehrfach tägliche orale Gabe) sowie der häufigen, individuellen Dosisanpassungen der oralen Phosphat- und aktivierten Vitamin D-Therapie als schwer durchführbar dar.

behandelnde bzw. weiterbehandelnde Personen:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Aufteilung der Patientengruppen war dem behandelnden Personal bekannt. Eine Ausnahme stellen die Auswerter der Rachitissymptomatik dar, welche sowohl beim RSS als auch bei RGI-C verblindet waren. Eine allgemeine Verblindung des behandelnden Personals stellt sich aufgrund der unterschiedlichen Verabreichungsmethoden (subkutane Injektion vs. mehrfach tägliche orale Gabe) sowie der häufigen, individuellen Dosisanpassungen der oralen Phosphat- und aktiviertem Vitamin D-Therapie als schwer durchführbar dar.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Studie UX023-CL301 erreichte die in dem Studienprotokoll und im Statistical Analysis Plan (SAP) spezifizierten Aufnahmeziele. Die Ergebnisse aller im Studienprotokoll und SAP beschriebenen Tests sind im Studienbericht (CSR) dargestellt. Die ergebnisunabhängige Berichterstattung ist durch das teilweise eigenständige Ausfüllen der Fragebögen durch die Patienten gewährleistet. Außerdem besteht durch die Erhebung biochemischer, numerischer Parameter sowie der verblindeten Auswertung der Röntgenbilder eine objektive Beurteilung.

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Bei der Studie UX023-CL301 handelt es sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie, bestehend aus einer 64-wöchigen Behandlungs- und einer bis Woche 140 laufenden Extensionsphase. Eine Behandlungsgruppe erhielt Burosumab subkutan alle zwei Wochen (Q2W), während die aktive Kontrollgruppe eine patientenindividuelle tägliche orale Gabe von Phosphat und aktivierten Vitamin D als Vergleichstherapie bekam. Die Studie umfasst eine heterogene Population (n=61) bezogen auf Alter (1 – 12 Jahren) und Wachstum. Dabei erfolgte eine Stratifizierung nach dem Alter, sowie der Region und der Rachitis-Schwere. Die geringe Populationsgröße lässt sich mit dem Krankheitsbild der XLH als seltenes Leiden erklären. Die Randomisierung der Patienten in die Burosumab-Gruppe (n=29) und die Kontrollgruppe (n=32) wurde 1:1 elektronisch mittels *Interactive Web Response System* (IWRS) durchgeführt. Eine ergebnisunabhängige Berichterstattung ist bei jedem Endpunkt gewährleistet. Es erfolgte weder eine Verblindung der Patienten noch des behandelnden Personals (mit Ausnahme der Auswerter beim RSS und RGI-C), da dies aufgrund der unterschiedlichen Applikationsformen und der erforderlichen individuellen und häufigen Dosisanpassungen bei der oralen Phosphat/aktivierten Vitamin D-Gabe insbesondere bei Kindern problematisch gewesen wäre. Trotz dessen liegt bei den Endpunkten der Studie UX023-CL301 aufgrund der objektiv messbaren, numerischen Endpunkte und der verblindeten Auswertung der Rachitissymptomatik ein niedriges Verzerrungspotenzial vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**Endpunkt: Rachitissymptomatik (RGI-C, RSS)****1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Bewertung des RSS wurde mit der Messmethodik nach Thacher durchgeführt. Die Auswertung erfolgte ausschließlich durch den unabhängigen, geschulten und verblindeten Auswerter Thomas Thacher. Die bilateralen Röntgenaufnahmen wurden einzeln in einer randomisierten Reihenfolge präsentiert. Der Auswerter war bezüglich des Zeitpunktes der Röntgenbildaufnahme und des Behandlungsstatus inklusive Informationen zu Patientenalter/-geschlecht, Dosierungsregime oder Patientenhistorie verblindet.

Die Auswertung des RGI-C-Scores wurde zentral von drei unabhängigen, geschulten und verblindeten pädiatrischen Radiologen durchgeführt. Die Auswerter waren bezüglich Dosierung, Dosierungsregime, Befolgen des Studienprotokolls, Behandlungsdauer und Vorbehandlung mit einer Substitutionstherapie verblindet.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Alle Patienten der FAS-Population wurden für die Analyse herangezogen. Diese ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugewiesene Medikation erhalten sowie mindestens eine Bewertung Post-Baseline bekommen haben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Bei der RSS-Auswertung wurden die bilateralen Röntgenaufnahmen einzeln in einer randomisierten Reihenfolge präsentiert, sodass kein Vergleich zu einer weiteren Röntgenbildaufnahme möglich war. Sowohl beim RSS als auch bei RGI-C wurde jede Bewertung direkt in ein elektronisches Datenerfassungssystem (*EDC system – electronic data capture system*) eingetragen und elektronisch an eine zentrale Bildgebungseinrichtung übermittelt. Diese Daten konnten nachträglich nicht mehr geändert werden. So war eine ergebnisabhängige nachträgliche Korrektur der Punktzahlen nicht möglich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Der RSS-Gesamtscore summiert sich aus dem Handgelenks- und Kniescore und dient einer quantitativen Beurteilung der Rachitissymptomatik. Bei der Interpretation des Scores sind die häufig geringeren Symptomausprägungen der XLH-bedingten Rachitis im Vergleich zur ernährungsbedingten Rachitis zu beachten. Beim RGI-C-Score wurde ein krankheitsspezifisches, RGI-C-Punktesysteme zur qualitativen Bewertung der Rachitissymptomatik zu verschiedenen Zeitpunkten (Baseline und Post-Baseline) verwendet. Die Erhebungszeitpunkte lagen zu Baseline, Woche 40 und 64. Die Auswertung erfolgte ausschließlich durch den unabhängiger, zentralen Auswerter Thomas Thacher (RSS) oder zentral durch drei unabhängige, verblindete pädiatrische Radiologen (RGI-C). Bei allen vier Auswertern der RSS- und RGI-C-Methode erfolgte eine Verblindung bezüglich Dosierung, Dosierungsregime, Befolgen des Studienprotokolls, Behandlungsdauer und Vorbehandlung mit einer Substitutionstherapie. Zudem erhielten alle eine Schulung für den Ablauf der Bewertung nach einer vordefinierten Methodik (*Biomedical Systems, Independent Review Manual*). Demnach wurden die Scores nach SOP (*standard operation procedure*) erhoben. Des Weiteren wurde ein Mittelwert aus den drei RGI-C-Einzelscores bestimmt. Eine ergebnisunabhängige Berichterstattung ist durch Verblindung, die randomisierte Präsentation der Röntgenbilder und den direkten Upload der Ergebnisse in das elektronische Datenerfassungssystem (*EDC system – electronic data capture system*) gegeben. Demnach weist der Endpunkt **Rachitissymptomatik** ein niedriges Verzerrungspotenzial auf.

Endpunkt: Körpergröße (Stehhöhe/Liegelänge)

1. Verblindung der Endpunkterheber

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung bei der Endpunkterhebung.

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Alle Patienten der FAS-Population wurden für die Analyse herangezogen. Diese ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugewiesene Medikation erhalten sowie mindestens eine Bewertung Post-Baseline bekommen haben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist gewährleistet.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Sowohl das ITT-Prinzip als auch eine ergebnisunabhängige Berichterstattung wurden umgesetzt. Es sind keine sonstigen (endpunktspezifischen) Aspekte, die zur Verzerrung der Endpunktergebnisse führen können, gegeben. Aufgrund der Verwendung des gleichen Radiometers zur Messung der Stehhöhe zu Baseline, Woche 24, 40 und 64 in der Behandlungsphase und der Dreifachmessung sowie Mittelwert-Berechnung um das Risiko von Messfehlern zu minimieren, weist der objektive, numerische Endpunkt **Körpergröße**, trotz des offenen Studiendesigns, ein niedriges Verzerrungspotenzial auf.

Endpunkt: 6-Minuten-Gehdistanz Test**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung bei der Endpunkterhebung.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde eine Modifikation der ITT-Population, welche für die Analyse herangezogen wurde, vorgenommen. Diese umfasst nur Patienten der FAS-Population, die zum Zeitpunkt der

Screeningvisite ≥ 5 Jahre alt waren. Die Erhebung des Endpunktes ist altersabhängig und für Kinder unter 5 Jahre nicht geeignet.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist gewährleistet.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Sowohl das ITT-Prinzip als auch eine ergebnisunabhängige Berichterstattung wurden umgesetzt. Es sind keine sonstigen (endpunktspezifischen) Aspekte, die zur Verzerrung der Endpunktergebnisse führen können, gegeben. Über die Rundenanzahl und durch eine objektive Messung der Gehstrecke mittels eines Maßbandes wurde dieser numerische Endpunkt in Metern bestimmt. Es wurden die Kriterien der *ATS Guidelines* als qualitätssichernde Maßnahme berücksichtigt. Diese wurden dem klinischen Personal in persönlichen Schulungen an allen teilnehmenden Studienzentren durch Ultragenyx erläutert. Nach Erhalt des Schulungszertifikats wurde der Test von dem klinischen Personal mithilfe des *Clinical Evaluator Manuals* nach SOP ausgeführt. Hierzu gehörten beispielsweise die standardisierten Formulierungen für die Kommunikation mit dem Patienten. Ferner betrug die Anzahl des klinischen Personals in der Regel lediglich zwei Personen pro Standort, um die Messvariabilität während der gesamten Studie klein zu halten. Folglich ist trotz des offenen Studiendesigns von einem niedrigen Verzerrungspotenzial für den Endpunkt **6-Minuten-Gehdistanz Test** auszugehen.

Endpunkt: Biochemischer Laborparameter: Serumphosphat

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung bei der Endpunkterhebung.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Alle Patienten der FAS-Population wurden für die Analyse herangezogen. Diese ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugewiesene Medikation erhalten sowie mindestens eine Bewertung Post-Baseline bekommen haben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist gewährleistet.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Sowohl das ITT-Prinzip als auch eine ergebnisunabhängige Berichterstattung wurden umgesetzt. Es sind keine sonstigen (endpunktspezifischen) Aspekte, die zur Verzerrung der Endpunktergebnisse führen können, gegeben. Beim Serumphosphat handelt es sich um einen objektiven numerischen Laborparameter, mit einem exakt definierten Normalbereich, bei dem das Risiko einer Beeinträchtigung durch das Studienpersonal, die Erziehungsberechtigten oder den Patienten minimal ist. Ferner wurde das Blutserum durch ausgebildetes klinisches Personal entnommen und mit validierten Analysegeräten untersucht. Dementsprechend ist trotz des offenen Studiendesigns von einem niedrigen Verzerrungspotenzial beim Endpunkt **Serumphosphat** auszugehen.

Endpunkt: Patient-Reported Outcomes Measurement Information System (PROMIS): Pediatric Pain Interference, Physical Function Mobility and Fatigue Domain Score**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung bei dem Ausfüllen der Fragebögen, da dieses sowohl von den Patienten selbstständig (Selbstbericht-Version bei Kindern ≥ 8 Jahre), von den Erziehungsberechtigten (Eltern-Version bei Kindern ≤ 7 Jahre) oder in Form eines Interviews durch geschultes klinisches Personal (Ausnahmefälle: Selbstbericht-Version bei Kindern ≥ 8 Jahre) gemacht wurde.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde eine Modifikation der ITT-Population, welche für die Analyse herangezogen wurde, vorgenommen. Diese umfasst nur Patienten der FAS-Population, die zum Zeitpunkt der Screeningvisite ≥ 5 Jahre alt waren. Die Erhebung des Endpunktes ist altersabhängig und für Kinder unter 5 Jahre nicht geeignet.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist unklar.
Die Beantwortung der Fragebögen erfolgte altersabhängig durch die Patienten selbstständig (Selbstbericht-Version bei Kindern ≥ 8 Jahre), die Erziehungsberechtigten (Eltern-Version bei Kindern ≤ 7 Jahre) oder in Form eines Interviews durch geschultes klinisches Personal (Ausnahmefälle: Selbstbericht-Version bei Kindern ≥ 8 Jahre). Bei der Eltern-Version liegt keine objektive Einschätzung der Kindersymptomatik vor, da diese auf der subjektiven Wahrnehmung der Eltern beruht. Die Entscheidung über Ausnahmefälle, wie mit Kindern zu verfahren ist, wenn sie die Selbstbericht-Version nicht selbstständig ausfüllen konnten, oblag dem behandelnden Arzt. Neben dem Einsatz des klinischen Personals war eine andere Option, dass die Erziehungsberechtigten die Eltern-Version für ihr Kind ausfüllten.
Nach Beantwortung der Fragebögen in Papierform wurden die Ergebnisse direkt in das PROMIS *Online Scoring System* hochgeladen, sodass eine nachträgliche Änderung nicht möglich war.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Das ITT-Prinzip wurde eingehalten. Es sind keine sonstigen (endpunktspezifischen) Aspekte, die zur Verzerrung der Endpunktergebnisse führen können, gegeben. Der Einsatz des PROMIS-Fragebogens erfolgte auf Empfehlung der FDA, welche die Ansicht vertritt, dass der PROMIS-Fragebogen den Gesundheitszustand der XLH-Patienten ab einem Alter von 5 Jahren eindeutiger erfasse als der POSNA-PODCI aus der pivotalen Studie UX023-CL201. Eine Verblindung war aufgrund des selbstständig auszufüllenden Fragebogens nicht gegeben. Aufgrund der unterschiedlichen Personengruppen, welche die Fragebögen auszufüllen hatten, ist eine ergebnisunabhängige Berichterstattung unklar. Durch den direkten Upload der Antworten nach Ausfüllen der Papierform in das PROMIS *Online Scoring System*, wird eine nachträgliche ergebnisabhängige Änderung verhindert. Des Weiteren musste die Entscheidung über das Vorgehen bei den Ausnahmefällen vom Prüfarzt ausführlich begründet und dokumentiert werden. Zudem wurde das klinische Personal in der Durchführung der Interviews geschult. Ferner erfolgte eine Validierung des PROMIS-Fragebogens durch die Interviewstudie UX7575A, welche mit einer kognitiven Nachbesprechung das Verständnis der Fragen und Antwortmöglichkeiten sowie die Relevanz der Fragen in Bezug auf die XLH bestätigen konnte. Aufgrund der genannten Aspekte ist trotz des offenen Studiendesigns und der unklaren ergebnisunabhängigen Berichterstattung von einem niedrigen Verzerrungspotenzial für den Endpunkt **Schmerz und funktionelle Behinderung (PROMIS)** auszugehen.

Endpunkt: Faces Pain Scale – Revised (FPS-R)

1. Verblindung der Endpunkterheber

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung bei der Endpunkterhebung.

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde eine Modifikation der ITT-Population, welche für die Analyse herangezogen wurde, vorgenommen. Diese umfasst nur Patienten der FAS-Population, die zum Zeitpunkt der Screeningvisite ≥ 5 Jahre alt waren, da der FPS-R für die Verwendung bei Kindern im Alter von 5 bis 16 Jahren validiert wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist gewährleistet.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Sowohl das ITT-Prinzip als auch eine ergebnisunabhängige Berichterstattung wurden umgesetzt. Es sind keine sonstigen (endpunktspezifischen) Aspekte, die zur Verzerrung der Endpunktergebnisse führen können, gegeben. Beim FPS-R handelt es sich um ein selbstberichtendes Instrument in Form von einem Fragebogen, welcher speziell für Kinder entwickelt und für die Verwendung bei Kindern im Alter von 5 bis 16 Jahren validiert wurde. Das klinische Personal wurde im Umgang mit den Kindern geschult, sodass eine standardisierte Kommunikation/Anleitung gegeben war. Trotz des offenen Studiendesigns ist aufgrund dessen von einem niedrigen Verzerrungspotenzial beim Endpunkt **FPS-R** auszugehen.

Endpunkt: Lebensqualität (SF-10)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung beim Vervollständigen des Fragebogens, da dieser in Form einer Fremdeinschätzung von den Erziehungsberechtigten der Patienten ausgefüllt wurde.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde eine Modifikation der ITT-Population, welche für die Analyse herangezogen wurde, vorgenommen. Diese umfasst nur Patienten der FAS-Population, die zum Zeitpunkt der

Screeningvisite ≥ 5 Jahre alt waren, da der SF-10 für die Verwendung bei Kindern im Alter von 5 bis 18 Jahren entwickelt wurde. .

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist unklar.
Beim SF-10 handelt es sich um eine Fremdeinschätzung der Eltern bezüglich der Lebensqualität des Kindes. Demnach liegt keine objektive Einschätzung der Lebensqualität der Patienten vor, da diese auf der subjektiven Wahrnehmung der Eltern beruht.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es erfolgte keine explizite Validierung des Instruments an Patienten mit dem Krankheitsbild der XLH.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Das ITT-Prinzip wurde umgesetzt. In Kommunikation mit dem G-BA, welcher einen Fragebogen zur Einschätzung der Lebensqualität wie den SF-36 forderte, wurde hierzu der SF-10 verwendet. Beim SF-10 handelt es sich im Gegensatz zum SF-36 um einen Fragebogen zur Fremdeinschätzung der Lebensqualität bei Kindern im Alter von 5 bis 18 Jahren. Dieser ist in pädiatrischen klinischen Forschungseinrichtungen weit verbreitet und wurde von der EMA in verschiedenen Zulassungsverfahren akzeptiert (siehe Abschnitt 4.2.5.2). Durch die subjektive Fremdeinschätzung des Gesundheitszustandes des Kindes besteht eine unklare Berichterstattung. Außerdem handelt es sich um ein offenes Studiendesign und eine spezifische Validierung auf die XLH fehlt. In der Gesamtschau ist das Verzerrungspotenzial beim Endpunkt **SF-10** als hoch einzuschätzen.

Endpunkt: Sicherheit

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es erfolgte keine Verblindung bei der Endpunkterhebung.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Alle Patienten der FAS-Population wurden für die Analyse herangezogen. Diese ist definiert als alle Patienten, die randomisiert wurden und mindestens eine zugewiesene Medikation erhalten sowie mindestens eine Bewertung Post-Baseline bekommen haben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine ergebnisunabhängige Berichterstattung ist gewährleistet.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Keine sonstigen (endpunktübergreifenden) Aspekte vorhanden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Es handelt sich um eine randomisierte, offene und aktiv kontrollierte Phase-III-Zulassungsstudie. Sowohl das ITT-Prinzip als auch eine ergebnisunabhängige Berichterstattung wurden umgesetzt. Es sind keine sonstigen (endpunktspezifischen) Aspekte, die zur Verzerrung der Endpunktergebnisse führen können, gegeben. Entsprechend den Vorgaben des G-BA werden die Nebenwirkungen über die UEs und SUEs sowohl als Gesamtraten als auch aufgeteilt nach SOC und PT dargestellt. Es erfolgte eine Einstufung des kausalen Zusammenhangs der UEs mit der Prüfmedikation sowie des Schweregrads der UEs gemäß der *National Cancer Institute Common Terminology Criteria for Adverse Events (NST CTCAE)* durch den verantwortlichen Prüfarzt. Trotz des offenen Studiendesigns kann von einem niedrigen Verzerrungspotenzial für den Endpunkt **Sicherheit** ausgegangen werden.

Endpunkt: Zahnärztliche Untersuchungen

Anmerkung: In der Studie UX023-CI301 wird der Endpunkt **zahnärztliche Untersuchungen** unter dem Aspekt **Sicherheit** berücksichtigt.