Dokumentvorlage, Version vom 21.02.2019

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Burosumab (Crysvita®)

Kyowa Kirin GmbH

Modul 4 A

Zur Behandlung von Kindern und Jugendlichen (1 – 17 Jahre) mit X-chromosomaler Hypophosphatämie (XLH)

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Stand: 31.01.2022

Inhaltsverzeichnis

Seite

Tabellenverzeichnis **4**
Abbildungsverzeichnis **9**
Abkürzungsverzeichnis **10**
4 Modul 4 – allgemeine Informationen **14**
4.1 Zusammenfassung der Inhalte von Modul 4 15
4.2 Methodik 28
4.2.1 Fragestellung 29
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung 30
4.2.3 Informationsbeschaffung 34
4.2.3.1 Studien des pharmazeutischen Unternehmers 35
4.2.3.2 Bibliografische Literaturrecherche 35
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken 36
4.2.3.4 Suche auf der Internetseite des G-BA 38
4.2.3.5 Selektion relevanter Studien 39
4.2.4 Bewertung der Aussagekraft der Nachweise 39
4.2.5 Informationssynthese und -analyse 41
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien 41
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien 41
4.2.5.3 Meta-Analysen 68
4.2.5.4 Sensitivitätsanalysen 69
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren 72
4.2.5.6 Indirekte Vergleiche 76
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen 79
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel 79
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel 79
4.3.1.1.1 Studien des pharmazeutischen Unternehmers 79
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche 81
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken 83
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA 84
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel 85
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 86
4.3.1.2.1 Studiendesign und Studienpopulationen 86
4.3.1.2.2 Verzerrungspotenzial auf Studienebene 95
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien 96
4.3.1.3.1 Ergebnisse patientenrelevanter Endpunkte – RCT 98
4.3.1.3.1.1 Rachitissymptomatik – RCT 101

4.3.1.3.1.2 Körpergröße (Stehhöhe/Liegelänge und Sitzhöhe) – RCT 112
4.3.1.3.1.3 Gehfähigkeit – RCT 118
4.3.1.3.1.4 Serumphosphat – RCT 125
4.3.1.3.1.5 Schmerzerhebung und funktionelle Behinderung – RCT 129
4.3.1.3.1.6 Lebensqualität – RCT 149
4.3.1.3.1.7 Sicherheit – RCT 155
4.3.1.3.2 Subgruppenanalysen – RCT 185
4.3.1.3.2.1 Ergebnisse für Subgruppenanalysen mit positivem Interaktionstest (p < 0,05) 192
4.3.1.4 Liste der eingeschlossenen Studien - RCT 197
4.3.2 Weitere Unterlagen 197
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 197
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 197
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 198
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 198
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT 198
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT 201
4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT 201
4.3.2.2 Nicht randomisierte vergleichende Studien 201
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 201
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien 202
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 203
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien 203
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien 204
4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien 204
4.3.2.3 Weitere Untersuchungen 205
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 205
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 205
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen 206
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen 206
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen 207
4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen 207
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens 208
4.4.1 Beurteilung der Aussagekraft der Nachweise 208
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß 209
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht 224
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte 224
4.5.1 Begründung für die Vorlage indirekter Vergleiche 224
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen 225
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen 225

4.5.4 Verwendung von Surrogatendpunkten 225
4.6 Referenzliste 226
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche 234
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken 238
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche) 240
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken) 241
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT 247
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 263
Anhang 4-G : Ergänzende Analysen 283

Tabellenverzeichnis

Seite

Tabelle 4-1: Einschlusskriterien für RCT mit Burosumab als Intervention 17
Tabelle 4-2: Ausschlusskriterien für RCT mit Burosumab als Intervention 18
Tabelle 4-3: Zusammenfassung der Ergebnisse der Kategorie *Morbidität* zum medizinischen Zusatznutzen von Burosumab gegenüber der zweckmäßigen Vergleichstherapie (zVT) 19
Tabelle 4-4: Zusammenfassung der Ergebnisse der Kategorie *Lebensqualität* zum medizinischen Zusatznutzen von Burosumab gegenüber der zVT 24
Tabelle 4-5: Zusammenfassung der Ergebnisse der Kategorie *Sicherheit* zum medizinischen Zusatznutzen von Burosumab gegenüber der zVT 25
Tabelle 4-6: Einschlusskriterien für RCT mit Burosumab als Intervention 32
Tabelle 4-7: Ausschlusskriterien für RCT mit Burosumab als Intervention 33
Tabelle 4-8: Patientenrelevante Endpunkte für die Bewertung von Burosumab in der Therapie der XLH bei Kindern ab einem Jahr 43
Tabelle 4-9: Kriterien für die Berechnung der Interaktionstests in Anlehnung an das Methodenpapier 6.0 des IQWiG 73
Tabelle 4-10: Übersicht der für die Endpunkte der Studie UX023-CL301 durchgeführten Subgruppenanalysen 75
Tabelle 4-11: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 80
Tabelle 4-12: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel 81
Tabelle 4-13: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel 84
Tabelle 4-14: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel 85
Tabelle 4-15: Studienpool – RCT mit dem zu bewertenden Arzneimittel 86
Tabelle 4-16: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 87
Tabelle 4-17: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel 89
Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel 90
Tabelle 4-19: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel 95
Tabelle 4-20: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel 97

Tabelle 4-21: Operationalisierung von *Rachitissymptomatik* 101

Tabelle 4-22: Bewertung des Verzerrungspotenzials für *Rachitissymptomatik* in RCT mit dem zu bewertenden Arzneimittel..... 103

Tabelle 4-23: Ergebnisse für *Radiographic Global Impression of Change (RGI-C) Total Score* aus RCT mit dem zu bewertenden Arzneimittel 104

Tabelle 4-24: Ergebnisse für *Radiographic Global Impression of Change (RGI-C)* - Responderanalyse aus RCT mit dem zu bewertenden Arzneimittel 106

Tabelle 4-25: Ergebnisse für *Rickets Severity Score (RSS) – Total Score* aus RCT mit dem zu bewertenden Arzneimittel 108

Tabelle 4-26: Ergebnisse der LOCF-Analyse für *Rickets Severity Score (RSS)* – Total Score aus RCT mit dem zu bewertenden Arzneimittel 110

Tabelle 4-27: Operationalisierung von *Körpergröße (Stehhöhe/Liegelänge und Sitzhöhe)* .. 112

Tabelle 4-28: Bewertung des Verzerrungspotenzials für *Körpergröße* in RCT mit dem zu bewertenden Arzneimittel 113

Tabelle 4-29: Ergebnisse für *Körpergröße (Stehhöhe/Liegelange)* aus RCT mit dem zu bewertenden Arzneimittel 114

Tabelle 4-30: Ergebnisse für *Sitzhöhe* aus RCT mit dem zu bewertenden Arzneimittel..... 116

Tabelle 4-31: Operationalisierung von *Gehfähigkeit* 118

Tabelle 4-32: Bewertung des Verzerrungspotenzials für *Gehfähigkeit* in RCT mit dem zu bewertenden Arzneimittel 119

Tabelle 4-33: Ergebnisse für *Gehfähigkeit* aus RCT mit dem zu bewertenden Arzneimittel 120

Tabelle 4-34: Operationalisierung von *Serumphosphat* 125

Tabelle 4-35: Bewertung des Verzerrungspotenzials für *Serumphosphat* in RCT mit dem zu bewertenden Arzneimittel 125

Tabelle 4-36: Ergebnisse für *Serumphosphat* aus RCT mit dem zu bewertenden Arzneimittel..... 127

Tabelle 4-37: Operationalisierung von *Schmerzerhebung* und *funktionelle Behinderung* 129

Tabelle 4-38: Bewertung des Verzerrungspotenzials für *Schmerzerhebung* und funktionelle Behinderung in RCT mit dem zu bewertenden Arzneimittel 131

Tabelle 4-39: Ergebnisse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel (GEE-Analyse) 132

Tabelle 4-40: Sensitivitätsanalyse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel unterteilt nach Alter (5 – 7 Jahre, Fremdeinschätzung) 137

Tabelle 4-41: Sensitivitätsanalyse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel unterteilt nach Alter (8 – 12 Jahre, Selbsteinschätzung) 139

Tabelle 4-42: Ergebnisse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel (MMRM-Analyse)..... 143

Tabelle 4-43: Ergebnisse für *Schmerzerhebung* mittels FPS-R aus RCT mit dem zu bewertenden Arzneimittel 146

Tabelle 4-44: Operationalisierung von *Lebensqualität* 149

Tabelle 4-45: Bewertung des Verzerrungspotenzials für *Lebensqualität* in RCT mit dem zu bewertenden Arzneimittel 150

Tabelle 4-46: Ergebnisse für *Lebensqualität (SF-10)* aus RCT mit dem zu bewertenden Arzneimittel 151

Tabelle 4-47: Operationalisierung von *Sicherheit* 155

Tabelle 4-48: Bewertung des Verzerrungspotenzials für *Sicherheit* in RCT mit dem zu bewertenden Arzneimittel 158

Tabelle 4-49: Ergebnisse für *unerwünschte Ereignisse nach Beginn der Behandlung* (*TEAE*) aus RCT mit dem zu bewertenden Arzneimittel (*Safety*-Population) 159

Tabelle 4-50: Deskriptive Darstellung der *unerwünschten Ereignisse nach Beginn der Behandlung (TEAE)* aus RCT mit dem zu bewertenden Arzneimittel zu Woche 40, 64 und in der Extensionsphase 161

Tabelle 4-51: Ergebnisse für jegliche *unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)* nach *System Organ Class* (SOC) und *Preferred Term* (PT) (bei ≥ 10 % der Patienten in mindestens einer Behandlungsgruppe) aus RCT mit dem zu bewertenden Arzneimittel (Baseline – W 40, *Safety*-Population) 163

Tabelle 4-52: Ergebnisse für jegliche *unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)* nach *System Organ Class* (SOC) und *Preferred Term* (PT) (bei ≥ 10 % der Patienten in mindestens einer Behandlungsgruppe) aus RCT mit dem zu bewertenden Arzneimittel (Baseline – Woche 64, *Safety*-Population) 168

Tabelle 4-53: Zusammenfassung der *zahnmedizinischen Untersuchungen* nach Beginn der Behandlung 177

Tabelle 4-54: Ergebnisse der *dentalen unerwünschten Ereignisse nach Beginn der Behandlung (TEAE)* 179

Tabelle 4-55: Ergebnisse zu *unerwünschten Ereignissen von besonderem Interesse nach Beginn der Behandlung* 184

Tabelle 4-56: Matrix der durchgeführten Subgruppenanalysen 187

Tabelle 4-57: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für UX023-CL301 189

Tabelle 4-58: Subgruppenanalysen – Ergebnisse für *FPS-R* aus RCT mit dem zu bewertenden Arzneimittel 193

Tabelle 4-59: Subgruppenanalysen – Ergebnisse für *PT Fieber* aus RCT mit dem zu bewertenden Arzneimittel 196

Tabelle 4-60: Liste der zur Bewertung des Zusatznutzens eingeschlossenen Studien - RCT 197

Tabelle 4-61: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche 198

Tabelle 4-62: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 199

Tabelle 4-63: Operationalisierung von <Endpunkt xxx> ... 199
Tabelle 4-64: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche ... 200
Tabelle 4-65: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche ... 200
Tabelle 4-66: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien ... 202
Tabelle 4-67: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien ... 203
Tabelle 4-68: Operationalisierung von <Endpunkt xxx> ... 203
Tabelle 4-69: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 204
Tabelle 4-70: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen ... 206
Tabelle 4-71: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen ... 206
Tabelle 4-72: Zusammenfassung der Ergebnisse der Kategorie *Morbidität* zum medizinischen Zusatznutzen von Burosumab ... 211
Tabelle 4-73: Zusammenfassung der Ergebnisse der Kategorie *Lebensqualität* zum medizinischen Zusatznutzen von Burosumab ... 217
Tabelle 4-74: Zusammenfassung der Ergebnisse für *unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)* aus RCT mit dem zu bewertenden Arzneimittel ... 218
Tabelle 4-75: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens ... 224
Tabelle 4-76: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EMBASE ... 235
Tabelle 4-77: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in MEDLINE ... 236
Tabelle 4-78: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in Cochrane ... 236
Tabelle 4-79: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in clinicaltrials.gov ... 238
Tabelle 4-80: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EU Clinical Trials Register (EU-CTR) ... 239
Tabelle 4-81: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in *WHO International Clinical Trials Registry Platform* (WHO ICTRP) ... 239
Tabelle 4-82: Übersicht der Treffer in den Studienregistern ... 241
Tabelle 4-83: Ausgeschlossene Studien ClincialTrials.gov ... 241
Tabelle 4-84: Ausgeschlossene Studien EU-CTR ... 243
Tabelle 4-85: Ausgeschlossene Studien WHO ICTRP ... 244
Tabelle 4-86: Studiendesign und - methodik für Studien UX023-CL301 ... 247

Tabelle 4-87: (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UX023-CL301 264

Abbildungsverzeichnis

Seite

Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 82

Abbildung 2: Studiendesign der UX023-CL301-Studie 92

Abbildung 3: Veränderung zur Baseline der *Pediatric Pain Interference Domain* des PROMIS-Fragebogens 134

Abbildung 4: Veränderung zur Baseline der *Pediatric Physical Function Mobility Domain* des PROMIS-Fragebogens 134

Abbildung 5: Veränderung zur Baseline der *Pediatric Fatigue Domain Score* des PROMIS-Fragebogens 135

Abbildung 6: Veränderung zur Baseline des *FPS-R (Schmerzerhebung)*..... 147

Abbildung 7: Veränderung zu Baseline für den Endpunkt *Physical Health Summary Score* über den Verlauf der 64-wöchigen aktiv-kontrollierten Phase der Studie UX023-CL301.... 153

Abbildung 8: Veränderung zu Baseline für den Endpunkt *Psychosocial Summary Score* über den Verlauf der 64-wöchigen aktiv-kontrollierten Phase der Studie UX023-CL301.... 153

Abbildung 9: Subgruppenanalyse *FPS-R*..... 195

Abbildung 10: Patientenfluss der Studie UX023-CL301 gemäß CONSORT 262

Abkürzungsverzeichnis

Abkürzung	Bedeutung
1,25(OH)$_2$-Vitamin D$_3$	Aktives Vitamin D
6MWT	6-Minuten-Gehdistanz Test (*engl.:* 6-*Minute Walk Test*)
ALP	Alkalische Phosphatase
AMNOG	Arzneimittelmarktneuordnungsgesetz
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse (*engl.: analysis of covariance*)
ARD	Absolute Risikodifferenz (*engl.: absolute risk difference*)
ARR	Adjustiertes Relatives Risiko (*engl. adjusted relative risk*)
ATS	*American Thoracic Society*
B – P	Baseline zu Post-Baseline
BMI	Body-Mass-Index (*engl.: body mass index*)
CDC	*Center of disease control and prevention*
CHQ	*Child Health Questionnaire*
cm	Zentimeter
COMP	*Committee of Orphan Medicinal Products*
CONSORT	*Consolidated Standards of Reporting Trials*
CRF	*Case Report Form*
CSR	Studienbericht *(engl. Clinical study report)*
CTCAE	*Common Terminology Criteria for Adverse Events*
DIMDI	Deutsches Institut für Medizinische Dokumentation
ECHO	Echokardiogramm
EDC system	*Electronic data capture system*
EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm
EMA	Europäische Arzneimittelagentur (*engl.: European Medicines Agency*)
EU	Europäische Union
EU-CTR	*EU Clinical Trials Registry*
FAS	*Full Analysis Set*
FDA	*Food Drug Administration*

Abkürzung	**Bedeutung**
FGF23	Fibroblasten-Wachstumsfaktor 23 (*engl.: Fibroblast growth factor* 23)
FPS-R	*Face Pain Scale-Revised*
G-BA	Gemeinsamer Bundesausschuss
GEE	*Generalized estimation equation*
GFR	Glomeruläre Filtrationsrate
GLMM	*Generalized linear mixed models*
H‘g	Hedges‘g
IgG1	Immunglobulin G1
iPTH	Intakt-Parathormon
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	*Intention to treat*
IWRS	*Interactive Web Response System*
KG	Körpergewicht
KI	Konfidenzintervall
KRN23	Synonym für Burosumab
LOCF	*Last Observation Carried Forward*
LS Mean	*Least Squares Means*
LSMD	*Least squares mean difference*
MAR	Behandlung fehlender Werte (*engl.: missing at random*)
MCMC	Markov Chain Monte Carlo
MedDRA	*Medical Dictionary for Regulatory Activities*
MeSH	*Medical Subject Headings*
MI	*Multiple Imputation*
MID	Minimal wichtiger Unterschied (*engl. minimal important difference*)
MMRM	*Mixed model repeated measures*
MTC	*Mixed Treatment Comparison*
MW	Mittelwert
MWD	Mittelwertdifferenz
NaPi-2	Natrium-Phosphat-Kotransporter
NCI CTCAE	*National Cancer Institute Common Terminology Criteria for Adverse Events*

Abkürzung	**Bedeutung**
NIH	*National Institutes of Health*
OHIP-14	*Oral Health Impact Profile-14*
OR	*Odds ratio*
PEI	Paul-Ehrlich-Institut
PHEX	*Phosphat regulierendes Gen mit Homologie zu Endopeptidasen auf dem X-Chromosom (engl.: Phosphate-regulating neutral endopeptidase)*
PHS	*Physical Health Summary Score*
PODCI	*Pediatric Outcomes Data Collection Instrument*
POSNA	*Pediatric Orthopedic Society of North America*
PROMIS	*Patient Reported Outcomes Measurement Information System*
PSS	*Psychosocial Summary Score*
PT	*Preferred Terms* nach MedDRA
PTH	Parathormon
RCT	*Randomized Controlled Trial*
RGI-C	*Radiographic Global Impression of Change*
ROW	Rest der Welt *(engl. Rest of the World)*
RR	Relatives Risiko (*engl.: relative risk, risk ratio*)
RSS	*Rickets Severity Score*
s.c.	Subkutan
SAP	Statistischer Analyseplan (*engl.: Statistical Analysis Plan*)
SD	Standardabweichung (*engl. standard deviation*)
SE	Standardfehler (*engl. standard error*)
SF-10	*Short form 10*
SF-36	*Short form 36*
SGB	Sozialgesetzbuch
SMQs	*Standardised MedDRA Queries*
SOC	*System Organ Class* nach MedDRA
SOP	*Standard Operating Procedure*
SRM	*Standardized Response Mean*
STE	*Surrogate Threshold Effects*
STROBE	*Strengthening the Reporting of Observational Studies in Epidemiology*

Abkürzung	**Bedeutung**
SUE	Schwerwiegendes UE
TEAE	Unerwünschtes Ereignis nach Beginn der Behandlung (*engl.: treatment emergent adverse event*)
TmP	Maximale tubuläre Rückresorption von Phosphat
TREND	*Transparent Reporting of Evaluations with Non-Randomized Design*
UE	Unerwünschtes Ereignis
ULN	Untere Grenze des Normbereichs (*engl.: upper limit normal*)
USA oder US	Vereinigte Staaten von Amerika (*engl.: The United States of America*)
VerfO	Verfahrensordnung
WHO	*World Health Organization*
WHO ICTRP	*WHO International Clinical Trials Registry Platform*
XLH	X-chromosomale Hypophosphatämie (*engl.: X-linked hypophosphatemia*)
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 4.1)

– 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 4.4)

– 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.

Hintergrund

Bei Burosumab handelt es sich um ein Arzneimittel zur Behandlung eines seltenen Leidens. Für Arzneimittel zur Behandlung eines seltenen Leidens gilt der Zusatznutzen durch die Zulassung als belegt. Nachweise gemäß § 35a Abs. 1 Satz 3 Nummer 2 und 3 SGB V über den medizinischen Nutzen und Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie (zVT) 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).

Burosumab wurde bereits zweimal vom Gemeinsamen Bundesausschuss (G-BA) bewertet. Im ersten Beschluss vom 04. Oktober 2018 bewertete der G-BA die beiden pivotalen Phase-II-Studien UX023-CL201 und UX023-CL205, sowie einen historischen Vergleich (UX023-CL002). Ebenfalls konnte der G-BA während dieses Verfahrens Einblicke in die damals neu verfügbaren 40-Wochen-Daten der randomisierten und aktiv kontrollierten Phase-III-Studie UX023-CL301 erhalten. Da die Studie UX023-CL301 noch laufend war, wurde der Beschluss des G-BA vom 04. Oktober 2018 befristet und ein nicht quantifizierbarer Zusatznutzen vergeben, dem die Patientenvertretung „nur schweren Herzens“ zustimmte, da sie in den Endpunkten Serumphosphat und Rachitissymptomatik einen beträchtlichen Zusatznutzen sehe (Verfahrensnummer 2018-04-15-D-349). Im zweiten Beschluss vom 02. April 2020 bewertete der G-BA die Studiendaten der abgeschlossenen Studie UX023-CL301 (Verfahrensnummer 2019-10-01-D-492). Die Ergebnisse bestätigten eindrucksvoll die bereits beachtlichen Ergebnisse zur Wirkung von Burosumab bei pädiatrischen XLH-Patienten aus dem ersten Nutzenbewertungsverfahren und belegten den beträchtlichen Zusatznutzen von Burosumab gegenüber einer weniger wirksamen Therapie, einer Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D.

Übersteigt der Umsatz des Arzneimittels für seltene Leiden mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten zwölf Kalendermonaten einen Betrag von 50 Millionen Euro, hat der pharmazeutische Unternehmer gemäß § 35a Absatz 1 Satz 12 SGB V innerhalb von drei Monaten nach Aufforderung durch den G-BA Nachweise gemäß 5. Kapitel § 5 Absatz 1 bis 6 VerfO des G-BA zu übermitteln.

Der G-BA hat mit Schreiben vom 21.10.2021 die Kyowa Kirin GmbH darüber informiert, dass der Umsatz mit Burosumab (Crysvita®) zu Lasten der GKV einen Betrag von 50 Millionen Euro überschritten hat. Maßgeblicher Zeitpunkt für die Einreichung der Nachweise gemäß § 35a Abs. 1 Satz 12 SGB V ist damit der 01.02.2022. Im vorliegenden Dossier werden nun die erforderlichen Nachweise des Zusatznutzens von Burosumab gegenüber der zVT dargestellt.

Fragestellung

Wie groß ist das Ausmaß des Zusatznutzens von Burosumab (Crysvita®) bei der Behandlung der XLH von Kindern und Jugendlichen im Alter von 1 – 17 Jahren mit 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 einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung. Bei der XLH handelt es sich um eine seltene Erkrankung, die schwerwiegende Behinderungen zur Folge hat.

Laut Anwendungsgebiet ist Burosumab zur Behandlung der XLH bei Kindern und Jugendlichen im Alter von 1 – 17 Jahren mit röntgenologischem Nachweis einer Knochenerkrankung, sowie bei Erwachsenen angezeigt [1]. In diesem Dossier wird ausschließlich das Anwendungsgebiet der Kinder dargestellt.

Die XLH stellt die häufigste erbliche Form der Rachitis dar. Sie ist gekennzeichnet durch eine Störung der renalen Phosphat-Rückresorption, was zu 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 biologische Aktivität hemmt. Durch die Hemmung von FGF23 stellt Burosumab die tubuläre Rückresorption von Phosphat aus der Niere wieder her. Zudem wird die Synthese von $1,25(OH)_2$-Vitamin D_3 gesteigert und die Darmabsorption von Calcium und Phosphat erhöht. Gemeinsam tragen diese Maßnahmen zur Normalisierung der Serumphosphat-Konzentration bei und stimulieren die Knochenmineralisierung. Burosumab ist die einzige pathophysiologisch kausale Therapiemöglichkeit für Kinder mit XLH und deckt den hohen medizinischen Bedarf.

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, untersucht. 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 Europa, den 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 KG im Abstand von zwei Wochen subkutan verabreicht wurde und sofern die Studien mindestens einen patientenrelevanten Endpunkt beinhalteten (Tabelle 4-1).

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 (Tabelle 4-2).

Tabelle 4-1: Einschlusskriterien für RCT mit Burosumab als Intervention

Einschlusskriterien		
E1	**Patientenpopulation**	• Patienten mit X-chromosomaler Hypophosphatämie (XLH) • Kinder und Jugendliche im Alter von 1 – 17 Jahren
E2	**Intervention**	• Burosumab
E3	**Komparator**	• Phosphatsubstitution[a] (in Kombination mit aktivem Vitamin D)
E4	**Endpunkte**	• 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**	• Randomisierte kontrollierte Studien (RCT) • Volltextpublikation oder Studienbericht • Veröffentlichung in deutscher oder englischer Sprache
E6	**Studiendauer**	• Ohne Einschränkung
a: Der G-BA hat als zVT eine Phosphatsubstitution festgelegt. Um die Folgeerkrankungen der Phosphatsubstitution wie z. B. einen sekundären Hyperparathyreoidismus zu vermeiden, geht der G-BA davon aus, dass zusätzlich eine Substitution mit aktivem Vitamin D (Calcitriol oder Alfacalcidol) erfolgt. Daraus ergibt sich eine Kombinationsbehandlung einer mehrmals täglichen oralen Gabe von Phosphat und aktivem Vitamin D		

Tabelle 4-2: Ausschlusskriterien für RCT mit Burosumab als Intervention

Ausschlusskriterien		
A1	**Patientenpopulation nicht entsprechend E1**	• Andere Formen der Rachitis (z. B. Vitamin-D-Mangel, Mangelernährung) • Erwachsene (≥ 18 Jahre) • Säuglinge/Kleinkinder (< 1 Jahr)
A2	**Intervention nicht entsprechend E2**	• Ausschließlich andere Wirkstoffe als Burosumab • Dosierung, Applikation nicht gemäß Zulassung von Burosumab
A3	**Komparator**	• Placebo
A4	**Endpunkte nicht entsprechend E4**	• Studien mit ausschließlich Surrogat-Parametern
A5	**Studien- und Publikationstyp nicht entsprechend E5**	• 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

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 Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Darüber hinaus weist die Studie 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 (Verzerrungspotenzial 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 ausführlich auf der Basis von patientenrelevanten Endpunkten 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 Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen beruhen auf der Studie UX023-CL301, einer multizentrischen, randomisierten und aktiv kontrollierten Studie der Phase III im Parallelgruppendesign. Die aktive Kontrolle erfolgte mit einer Substitutionstherapie bestehend aus oralem Phosphat und aktivem Vitamin D. Um eine hohe Aussagekraft und Ergebnissicherheit zu gewährleisten, wurde für die Herleitung des Zusatznutzens ausschließlich die aktiv kontrollierte Phase (64 Wochen) der Studie berücksichtigt.

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 Zulassungsstudie UX023-CL301 durch die Datenerhebung der schweren unerwünschten Ereignisse berücksichtigt. Es trat kein Todesfall während der Studiendauer auf.

Morbidität

Die Tabelle 4-3 zeigt in einer Übersicht die Ergebnisse der Endpunktkategorie *Morbidität* der Zulassungsstudie UX023-CL301.

Tabelle 4-3: Zusammenfassung der Ergebnisse der Kategorie *Morbidität* zum medizinischen Zusatznutzen von Burosumab gegenüber der zweckmäßigen Vergleichstherapie (zVT)

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64**
Rachitissymptomatik		
Radiographic Global Impression of Change (RGI-C) Score		
LsMean[a] [95% KI] p-Wert	1,14 [0,83; 1,45] < 0,0001	1,02 [0,72; 1,33] < 0,0001
Hedges'g[b] [95% KI]	H'g = 1,846 [1,246; 2,446]	H'g = 1,600 [1,023; 2,178]

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64**
Rickets Severity Score (RSS)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-1,34 [-1,74; -0,94] < 0,0001 H'g = -1,399 [-1,964; -0,833]	-1,21 [-1,59; -0,83] < 0,0001 H'g = -1,119 [-1,710; -0,529]
Körpergröße		
Stehhöhe/Liegelänge [Z-Score]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	0,12 [0,01; 0,24] 0,0408 H'g = 0,537 [0,021; 1,053]	0,14 [0,00; 0,29] 0,0490 H'g = 0,490 [-0,025; 1,005]
Sitzhöhe (cm)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-0,79 [-1,98; 0,39] 0,1901 -	0,08 [-1,58; 1,75] 0,9213 -
Gehfähigkeit[c]		
6-Minuten-Gehdistanz Test[c] (6MWT) [in Meter]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	38,30 [-2,38; 78,98] 0,0650 -	45,55 [2,09; 89,02] 0,0399 H'g = 0,679 [-0,038; 1,396]
6-Minuten-Gehdistanz Test[c] (6MWT) Gesamtrate [Prozentsatz]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	6,19 [-0,68; 13,06] 0,0775 -	7,27 [0,01; 14,52] 0,0496 H'g = 0,703 [-0,015; 1,422]
Serumphosphat		
Serumphosphat-Konzentration [mg/dl]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	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]
Schmerzerhebung		
Patient Reported Outcomes Measurement Information System (PROMIS) Pediatric Pain Interference Domain Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-4,91 [-9,19; -0,63] 0,0245 H'g = -0,697 [-1,390; -0,008]	-2,26 [-6,61; 2,09] 0,3091 -

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64**
Patient Reported Outcomes Measurement Information System (PROMIS) Physical Function Mobility Domain Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	2,68 [-0,53; 5,88] 0,1014 -	1,90 [-1,80; 5,59] 0,3145 -
Patient Reported Outcomes Measurement Information System (PROMIS) Fatigue Domain Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-3,20 [-7,79; 1,39] 0,1724 -	-1,08 [-6,21; 4,06] 0,6810 -
Face Pain Scale – Revised (FPS-R)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-0,00 [-0,80; 0,79] 0,9905 -	0,05 [-0,58; 0,68] 0,8786 -

(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 4.2.5.2 und 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 unvollständigen Datensatz aufwiesen (zu Baseline und zu Woche 40). Eine detaillierte Beschreibung erfolgt in den Abschnitten 4.2.5.4 und 4.3.1.3.1.3.

Verbesserung der Rachitissymptomatik

Erstmalig konnte in der Indikation X-chromosomale Hypophosphatämie (XLH) durch die RCT UX023-CL301 anhand einer aktiven Kontrolle die „Heilung" der Rachitissymptomatik durch eine Burosumab-Behandlung gezeigt werden. Außerdem konnten die ebenfalls eindrucksvollen Ergebnisse der beiden pivotalen Phase -II-Studien UX023-CL201 und UX023-CL205 mit der RCT UX023-CL301 bestätigt werden. Diese zeigen, basierend auf einem historischen Vergleich, dass Burosumab einer schwach wirksamen Substitution aus oralem Phosphat und aktivem Vitamin D überlegen ist. Die Burosumab-Behandlung führte bereits nach 40 Wochen zu einer signifikanten Erhöhung der RGI-C-Werte und zur Reduktion der RSS-Werte. 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 ($LsMean_{W64}$: 1,02; $p < 0,0001$; GEE-Modell) sowie klinisch relevant (H'g = 1,600 [1,023; 2,178]) (Tabelle 4-3).

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 in der Burosumab-Gruppe (MW_{BL}: 3,17; MW_{W40}:1,13), in

der aktiven Kontrollgruppe dagegen lediglich um 23 % (MW_{BL}:3,19; MW_{W40}: 2,47) (Tabelle 4-25). Die Behandlungsdifferenz der beiden Gruppen zu Woche 64 war mit -1,21 statistisch signifikant (p < 0,0001; GEE-Modell) und klinisch relevant (H'g = -1,119 [-1,710; -0,529]) (Tabelle 4-3 und Tabelle 4-25).

Verbesserung der Körpergröße (Stehhöhe und Sitzhöhe)

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 Stehhöhe zeigen eine Veränderung des Wachstums in Richtung Normalisierung durch die Burosumab-Behandlung im Vergleich zur aktiven Kontrolle. Der Unterschied in den Behandlungsgruppen bei der Veränderung des Z-Scores ($LsMean_{W40}$: 0,12; $LsMean_{W64}$: 0,14) war zu Woche 40 und zu Woche 64 statistisch signifikant (p_{40} = 0,0408; p_{64} = 0,0490; GEE-Modell) (Tabelle 4-3).

Anhand der Sitzhöhe in cm (absolute Werte) zeigte sich dagegen kein statistisch signifikanter Unterschied zwischen der Burosumab-Gruppe und der aktiven Kontrolle ($LsMean_{W64}$: 0,08; p = 0,9213) (Tabelle 4-3). Bei der Interpretation dieses Ergebnis ist jedoch zu beachten, dass sich die durch den chronischen Phosphatverlust bedingten Skelettanomalien hauptsächlich anhand von Verformungen der unteren Extremitäten zeigen und durch die Messung der Sitzhöhe nur unzureichend erfasst werden. Darüber hinaus fand keine Transformation der absoluten Werte in z-Scores statt, sodass die Werte nicht für Alter und Geschlecht standardisiert sind. Der Endpunkt Sitzhöhe ist somit mit Unsicherheiten behaftet.

Verbesserung der Gehfähigkeit

Zwei Studienteilnehmer der Burosumab-Gruppe waren nicht in der Lage, zu Baseline bzw. zu Woche 40 die 6 Minuten vollständig zu gehen. Aufgrund dieses unvollständigen Datensatzes wurden die beiden Studienteilnehmer in der vorliegenden Auswertung (korrigierte Analyse) nicht berücksichtigt. Die Ergebnisse des 6MWT (Gesamtrate in Meter und Prozentsatz der vorhergesagten Gehdistanz basierend auf normativen Werten) zeigen eine deutliche Verbesserung der Gehfähigkeit in der Burosumab-Gruppe, allerdings nicht in der aktiven Kontrollgruppe. Der Unterschied der Behandlungsdifferenz war zu Woche 64 signifikant (absoluter 6MWT: $LsMean_{W64}$: 45,55 m; p = 0,0399; Prozentsatz erreichter Normwert: $LsMean_{W64}$: 7,27 Prozentpunkte; p = 0,0496) (Tabelle 4-3). Zusätzlich wurde eine weitere Analyse durchgeführt, bei der die beiden Studienteilnehmer mit dem unvollständigen Datensatz

berücksichtigt wurden (FAS-Analyse). Die Ergebnisse hierzu sind ausführlich in Abschnitt 4.3.1.3.1.3 beschrieben.

Veränderung der Serumphosphat-Konzentration

Burosumab ist die erste, kausale Therapie zur Behandlung der XLH, die die zugrunde liegende Pathophysiologie der FGF23-induzierten Hypophosphatämie korrigiert. Bereits innerhalb kurzer Zeit kann eine Normalisierung der Serumphosphat-Konzentration erzielt werden und diese auch langfristig aufrechterhalten werden. Die Behandlungsdifferenz zu Woche 64 belegt sowohl den statistisch signifikanten als auch klinisch relevanten Vorteil der Burosumab-Behandlung gegenüber der aktiven Kontrolle ($LsMean_{W64}$: 0,74 (mg/dl); $p < 0,0001$; H'g = 1,981 [1,368; 2,595]) (Tabelle 4-3). 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 Kliniker der mündlichen Anhörung 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 [10-13]. Mit der Studie UX023-CL301 konnte nun eindeutig belegt werden, dass die Burosumab-Behandlung zielgerichtet in den Phosphatstoffwechsel eingreift, eine Normalisierung des Serumphosphats induziert und im Gegensatz zur reinen Substitution mit Phosphat 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 aufwiesen. 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 ($LsMean_{W40}$: -4,91; $p = 0,0245$), jedoch nicht mehr zu Woche 64 ($LsMean_{W64}$: -2,26; $p = 0,3091$) (Tabelle 4-3). Für die Domänen *Physical Function Mobility Domain Score* und *Fatigue Domain Score* zeigt sich ein Trend zu einer stärkeren Abnahme der Fatigue ($LsMeanW_{64}$: -1,08; $p = 0,6810$) und Steigerung der Mobilität ($LsMean_{W64}$: 1,90; $p = 0,3145$) zugunsten von Burosumab (Tabelle 4-3). Die akute Schmerzintensität, welche über den FPS-R abgefragt wurde, zeigte keine merklichen Veränderungen innerhalb der Behandlungsgruppen ($LsMean_{W64}$: 0,05; $p = 0,8786$) (Tabelle 4-3). Jedoch wurde bereits zu Baseline kein akuter Schmerz in den Behandlungsgruppen berichtet.

Lebensqualität

Die Tabelle 4-4 zeigt in einer Übersicht die Ergebnisse der Endpunktkategorie *Lebensqualität* der Zulassungsstudie UX023-CL301.

Tabelle 4-4: Zusammenfassung der Ergebnisse der Kategorie *Lebensqualität* zum medizinischen Zusatznutzen von Burosumab gegenüber der zVT

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64**
Lebensqualität		
Short Form (SF)-10 Health Survey Physical Heath Score (PHS-10)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	4,35 [-1,15; 9,85] 0,1213 -	5,51 [-0,17; 11,18] 0,0571 -
Short Form (SF)-10 Health Survey Psychosocial Summary Score (PSS-10)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	2,24 [-1,86; 6,35] 0,2841 -	-0,44 [-4,59; 3,70] 0,8337 -

(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 4.2.5.2 und 4.3.1.3.1

b: Hedges'g-Berechnung erfolgt nur, falls $p < 0,05$

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 nahezu unverändert ($LsMean_{W64}$: -0,44; p = 0,8337) (Tabelle 4-4). Die zweite Komponente des Fragebogens „körperliche Funktion" (PHS) verbesserte sich dagegen in der Burosumab-Gruppe im Vergleich zur aktiven Kontrollgruppe deutlich ($LsMean_{W64}$: 5,51; p = 0,0571) (Tabelle 4-4).

Zusammenfassend lässt sich somit feststellen, dass mit **Burosumab** erstmalig eine **zielgerichtete Therapie** zur Verfügung steht, durch die bereits innerhalb **kurzer Zeit** die **Serumphosphat-Konzentration** ausreichend erhöht und auch **langfristig aufrechterhalten** werden kann. Infolgedessen wird die **Knochenmineralisierung normalisiert** und die „**Heilung"** der **Rachitis** induziert. Dieser positive Effekt wirkt sich direkt auf das **Wachstum** und die **Gehfähigkeit** aus, führt zu einer **Abnahme** von **Schmerzen** und **Steigerung** der **Lebensqualität** von Kindern mit XLH.

Sicherheit

Die Tabelle 4-5 enthält eine Übersicht der Ergebnisse der patientenrelevanten Endpunkte der Kategorie *Sicherheit* in der Studie UX023-CL303.

Tabelle 4-5: Zusammenfassung der Ergebnisse der Kategorie *Sicherheit* zum medizinischen Zusatznutzen von Burosumab gegenüber der zVT

UX023-CL301 *Safety*-Population	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 40**	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 64**
Gesamtrate		
Jegliche TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	28 (96,6) vs. 26 (81,3) 1,188 [0,922; 1,423] 0,1064	29 (100,0) vs. 27 (84,4) 1,185 [1,021; 1,376] 0,0538
Schwerwiegende TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	3 (10,3) vs. 1 (3,1) 3,310 [0,364; 30,076] 0,3385	3 (10,3) vs. 3 (9,4) 1,115 [0,207; 6,018] 1,0000
Schwere TEAE Grad 3 – 4		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	4 (13,8) vs. 2 (6,3) 2,207 [0,436; 11,166] 0,4107	4 (13,8) vs. 3 (9,4) 1,547 [0,316; 7,582] 0,6988
Therapieabbruch durch TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	0 (0) vs. 0 (0) -	0 (0) vs. 0 (0) -
Tod durch TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	0 (0) vs. 0 (0) -	0 (0) vs. 0 (0) -

UX023-CL301 *Safety*-Population	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 40**	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 64**
TEAE von besonderem Interesse		
Reaktionen an der Injektionsstelle		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	12 (41,4) vs. 0 (0) 27,5 [1,70; 444,66] < 0,0001	15 (51,7) vs. 0 (0) 34,1 [2,132; 545,51] < 0,0001
Hypersensitivität		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%)) RR[a,c] [95% KI] p-Wert	9 (31,0) vs. 2 (6,3) 4,966 [1,168; 21,114] 0,0183	11 (37,9) vs. 6 (18,8) 2,023 [0,857; 4,773] 0,1522

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; *TEAE*: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; RR: *Risk Ratio* (relatives Risiko)

Ein RR < 1 zeigt einen Vorteil für Burosumab an und ein RR > 1 einen Vorteil für die aktive Kontrolle.

a: RR: Eine detaillierte Beschreibung des statistischen Modells ist in den Abschnitten 4.2.5.2 und 4.3.1.3.1.7 zu finden.

b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

c: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.

Insgesamt erweist sich Burosumab als wirksames, gut verträgliches und sicheres Arzneimittel zur Behandlung von XLH bei Kindern ab einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung. Es wurde weder ein Therapie- /oder Studienabbruch noch ein Todesfall aufgrund von *TEAE* berichtet.

Die meisten Studienteilnehmer hatten mindestens ein *TEAE* (Burosumab-Gruppe: 100 %; Aktive Kontrollgruppe: 84 %). Bei 7 Patienten (4 in der Burosumab-Gruppe; 3 in der aktiven Kontrollgruppe) traten *TEAE* Grad 3 auf. Keines der *TEAE* Grad 3 war schweren Ausmaßes, außer das schwerwiegende *TEAE* Kraniosynostose in der aktiven Kontrollgruppe. Keines der *TEAE* Grad 3 wurde vom Prüfarzt als therapieassoziiert angesehen, außer das *TEAE* Grad 3 Arthralgie in der Burosumab-Gruppe. Als Reaktion auf diese therapieassoziierten *TEAE* des Grad 3 wurden keine Änderungen in der Verabreichung von Studienmedikamenten vorgenommen. Die restlichen *TEAE* waren mild oder moderat (Grad 1 oder 2) und es wurden keine lebensbedrohlichen *TEAE* Grad 4 gemeldet (Tabelle 4-5).

Die häufigsten *unerwünschten Ereignisse von besonderem Interesse* in der Studie UX023-CL301 waren Reaktionen an der Injektionsstelle, welche nur leicht ausgeprägt und gut zu behandeln waren (Tabelle 4-5). Eine Reaktion an der Injektionsstelle ist eine bekannte und zu erwartende Folge einer subkutanen Verabreichung von Proteintherapeutika (Antikörpertherapeutika). Die Ausprägung dieser *TEAE* an der Injektionsstelle waren in den meisten Fällen mild oder moderat und wurden von den Prüfärzten 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 wurde keine klinisch bedeutsame Erhöhung 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, ausbleiben. Der Wechsel von der Phosphatsubstitution auf die Burosumab-Therapie führt sogar nachweislich zu einer Abnahme der renalen Calcium-Ausscheidung und hat dadurch das Potenzial, langfristig die Entstehung von Nephrokalzinosen zu verhindern [14].

Somit ist die **Sicherheit** und **Verträglichkeit** von Burosumab zur **Langzeitbehandlung** der XLH **bestätigt**.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Burosumab ist das erste Arzneimittel, dass 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 [15]. In der Gesamtschau der Endpunkte Morbidität, Lebensqualität und Sicherheit wird für Burosumab zur Behandlung von XLH bei Kindern ab einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung 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, damit einhergehende starke Schmerzen, geringere Lebensqualität). Schwerwiegende Nebenwirkungen wie eine Nephrokalzinose und einen sekundären oder sogar tertiären Hyperparathyreoidismus, die häufig unter der Phosphatsubstitution zu beobachten sind, treten unter Burosumab nicht auf [16].

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 Wiederherstellung der Knochenmineralisierung kommt es zur Rückbildung skelettaler Veränderungen der Knochen und zur Abnahme von Gelenkschmerzen. Auch der Zahnstatus sollte sich so im Laufe der Zeit verbessern. Die Markteinführung und das zuvor durchgeführte Arzneimittel-Härtefallprogramm beim Paul-Ehrlich-Institut unterstreichen den hohen therapeutischen Bedarf der Kinder mit XLH, der erst durch die Zulassung von Burosumab im Jahr 2018 gedeckt

werden konnte. In der Gesamtschau ergibt sich somit für Burosumab zur Behandlung der XLH bei Kindern ab einem Jahr ein **beträchtlicher Zusatznutzen**.

Fazit

Die für die vorliegende Nutzenbewertung herangezogene Studie UX023-CL301 weist auf Studien- und Endpunktebene ein niedriges Verzerrungspotenzial auf und hat somit eine hohe Aussagekraft. Die Ergebnisse der Studie sind zudem vollständig auf den deutschen Versorgungskontext übertragbar (vgl. Abschnitt 4.3.1.2.1, 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.

Die Ergebnisse der Studie UX023-CL301 belegen eine ausgeprägte Wirksamkeit von Burosumab zur pathologisch kausalen Therapie der X-chromosomalen Hypophosphatämie bei Kindern ab einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung und bestätigen eindrucksvoll die Ergebnisse der Studien UX023-CL201 und UX023-CL205 aus dem ersten Nutzenbewertungsverfahren von Burosumab [17]. Insbesondere führt die Anwendung von Burosumab bei den Endpunkten Gehdistanz, Wachstum, Serumphosphat und Rachitissymptomatik zu einer beträchtlichen Verbesserung bei den XLH-Patienten. Damit ermöglicht Burosumab den XLH-Patienten eine langfristige Freiheit von Knochen- und Gelenkschmerzen, die wiederum durch Deformationen der Extremitäten hervorgerufen werden. Ebenso werden schwerwiegende Symptome, die sowohl durch die XLH selbst aber auch durch eine Substitutionstherapie auftreten können, verbessert bzw. vermieden. Hier ist vor allem die Gefahr einer Nephrokalzinose und eines sekundären Hyperparathyreoidismus zu nennen, die unter einer Burosumab-Therapie nicht zu beobachten sind [14].

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

Unter Berücksichtigung des sehr guten Wirksamkeits- und Sicherheitsprofils, das sich anhand der drei Studien UX023-CL301, UX023-CL201 und UX023-CL205 zeigt, liegt für Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren mit röntgenologischem Nachweis einer Knochenerkrankung nach § 5 Abs. 7 der AM-NutzenV somit ein **Beleg** für einen **beträchtlichen Zusatznutzen** vor.

4.2 Methodik

Abschnitt 4.1 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).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

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 XLH bei Kindern ab einem Jahr mit röntgenologischem Nachweis angezeigt ist. Es handelt sich hierbei um ein Arzneimittel zur Behandlung eines seltenen Leidens gemäß § 35a Abs. 1 Satz 11 SGB V. Der Zusatznutzen von Burosumab (Crysvita®) ist somit bereits durch die Zulassung belegt. Nachweise über den medizinischen Nutzen und Zusatznutzen im Verhältnis zu einer zuvor bestimmten zweckmäßigen Vergleichstherapie (zVT) müssen demnach nicht erbracht werden. Lediglich das Ausmaß des Zusatznutzen ist für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, zu beschreiben.

Übersteigt der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten zwölf Kalendermonaten einen Betrag von 50 Millionen Euro, hat der pharmazeutische Unternehmer gemäß § 35a Absatz 1 Satz 12 SGB V innerhalb von drei Monaten nach Aufforderung durch den G-BA Nachweise gemäß 5. Kapitel § 5 Absatz 1 bis 6 VerfO des G-BA zu übermitteln.

Der G-BA hat mit Schreiben vom 21.10.2021 die Kyowa Kirin GmbH darüber informiert, dass der Umsatz mit Burosumab (Crysvita®) zu Lasten der GKV einen Betrag von 50 Millionen Euro überschritten hat. Maßgeblicher Zeitpunkt für die Einreichung der Nachweise gemäß § 35a Abs. 1 Satz 12 SGB V ist damit der 01.02.2022. Im vorliegenden Dossier werden nun die erforderlichen Nachweise des Zusatznutzens von Burosumab gegenüber der zVT dargestellt.

Die im vorliegenden Dossier zu beantwortende Frage lautet:

Wie groß ist das Ausmaß des Zusatznutzens von Burosumab (Crysvita®) bei der Behandlung der XLH von Kindern und Jugendlichen im Alter von 1 – 17 Jahren mit röntgenologischem Nachweis einer Knochenerkrankung hinsichtlich patientenrelevanter Endpunkte im Sinne des § 35a SGB V (Sozialgesetzbuch)?

Die **Patientenpopulation** umfasst laut Anwendungsgebiet pädiatrische XLH-Patienten ab einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung.

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].

In den beiden Beratungsgesprächen vom 08. November 2017 und 11. März 2020 hat der G-BA eine 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 (RCT) (**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.

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/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt 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. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Um die vorhandene Evidenz von Burosumab in dem vorliegenden Anwendungsgebiet vollständig abzubilden, wird zusätzlich eine systematische Literaturrecherche in Literaturdatenbanken und Studienregistern durchgeführt.

Die Ein- und Ausschlusskriterien der Studien, die für die Bewertung des Ausmaßes des Zusatznutzens von Burosumab herangezogen werden, orientieren sich an der zuvor definierten Fragestellung gemäß Abschnitt 4.2.1. Im vorliegenden Dossier soll sowohl der medizinische Nutzen als auch der Zusatznutzen primär anhand direkter Evidenz nachgewiesen werden. Bei den Ein- und Ausschlusskriterien müssen folgende Parameter (gemäß PICOS = *Population*, *Intervention*, *Comparison*, *Outcome*, *Study design*) berücksichtigt werden die anschließend detailliert beschrieben werden:

- Patientenpopulation
- Intervention
- Vergleichstherapie (Komparator)
- Endpunkte
- Studientyp/-dauer

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

Im Folgenden sind die Ein- und Ausschlusskriterien für randomisierte, kontrollierte Studien mit Burosumab als Intervention dargestellt (Tabelle 4-6 und Tabelle 4-7).

Tabelle 4-6: Einschlusskriterien für RCT mit Burosumab als Intervention

Einschlusskriterien		
E1	**Patientenpopulation**	• Patienten mit X-chromosomaler Hypophosphatämie (XLH) • Kinder und Jugendliche im Alter von 1 – 17 Jahren
E2	**Intervention**	• Burosumab
E3	**Komparator**	• Phosphatsubstitution[a] (in Kombination mit aktivem Vitamin D)
E4	**Endpunkte**	• 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**	• Randomisierte kontrollierte Studien (RCT) • Volltextpublikation oder Studienbericht • Veröffentlichung in deutscher oder englischer Sprache
E6	**Studiendauer**	• Ohne Einschränkung
a: Der G-BA hat als zVT eine Phosphatsubstitution festgelegt. Um die Folgeerkrankungen der Phosphatsubstitution wie z. B. einen sekundären Hyperparathyreoidismus zu vermeiden, geht der G-BA davon aus, dass zusätzlich eine Substitution mit aktivem Vitamin D (Calcitriol oder Alfacalcidol) erfolgt. Daraus ergibt sich eine Kombinationsbehandlung einer mehrmals täglichen oralen Gabe von Phosphat und aktivem Vitamin D		

Tabelle 4-7: Ausschlusskriterien für RCT mit Burosumab als Intervention

Ausschlusskriterien		
A1	**Patientenpopulation nicht entsprechend E1**	• Andere Formen der Rachitis (z. B. Vitamin-D-Mangel, Mangelernährung) • Erwachsene (≥ 18 Jahre) • Säuglinge/Kleinkinder (< 1 Jahr)
A2	**Intervention nicht entsprechend E2**	• Ausschließlich andere Wirkstoffe als Burosumab • Dosierung, Applikation nicht gemäß Zulassung von Burosumab
A3	**Komparator**	• Placebo
A4	**Endpunkte nicht entsprechend E4**	• Studien mit ausschließlich Surrogat-Parametern
A5	**Studien- und Publikationstyp nicht entsprechend E5**	• 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

Erläuterungen zu den Ein- und Ausschlusskriterien

Patientenpopulation (Ein- und Ausschlusskriterium 1)

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:

Pädiatrische XLH-Patienten ab einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung.

Intervention (Ein- und Ausschlusskriterium 2)

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

Die empfohlene Anfangsdosis von Burosumab beträgt 0,8 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].

Komparator (Ein- und Ausschlusskriterium 3)

Die seit über 50 Jahren historisch gewachsene Behandlung von Kindern mit XLH vor Burosumab ist nach Angaben der Leitlinie „Hereditäre hypophosphatämische Rachitiden“ der Deutschen Gesellschaft für Kinderendokrinologie und -diabetologie (DGKED) e.V. eine

Phosphatsubstitution [18]. Um die durch die Phosphatsubstitution bedingte Folgeerkrankung eines sekundären oder tertiären Hyperparathyreoidismus zu vermeiden, empfiehlt die Leitlinie zusätzlich eine Substitution mit aktivem Vitamin D [18, 19]. 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 vorliegenden Anwendungsgebiet von Burosumab ist eine Phosphatsubstitution. Im Beratungsgespräch verdeutlichte der G-BA zudem, dass eine zusätzliche Therapie mit aktivem Vitamin D erfolgen solle, um die Nebenwirkung der Substitutionstherapie zu vermeiden [19, 20].

Endpunkte (Ein- und Ausschlusskriterium 4)

Zur Beurteilung des Ausmaßes des Zusatznutzens wurden die patientenrelevanten Endpunkte im Anwendungsgebiet von XLH betrachtet. Diese umfassen sowohl die Morbidität als auch die Sicherheit. Da es sich bei XLH um eine schwerwiegende chronische Erkrankung handelt, die aber potenziell nicht tödlich verläuft, wird in der Studie UX023-CL301 die Mortalität durch die Datenerhebung der schweren unerwünschten Ereignisse erhoben und deshalb nicht gesondert dargestellt.

Studien- und Publikationstyp (Ein-/Ausschlusskriterium 5)

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

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.

Studiendauer (Einschlusskriterium 6)

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 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.“ [21]). Unterstützt wird dies durch die ICH-E1-Leitlinie der Europäischen Arzneimittelagentur (EMA), die eine Untersuchung von mindestens 6 Monaten bei chronischen Erkrankungen empfiehlt [22].

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 (inklusive „in-process & other non-indexed citations) 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 der Cochrane-Datenbank durchgeführt. Für MEDLINE (inklusive *in-process, in-data-review & 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 2018 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. Der erste 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/ Studienergebnisdatenbanken

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

Eine Suche in Studienregistern/ Studienergebnisdatenbanken 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 sowie Studienergebnisdatenbanken 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/ Studienergebnisdatenbanken clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (https://clinicaldata.ema.europa.eu) sowie dem Arzneimittel-

Informationssystem (AMIS, https://www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken 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, *WHO International Clinical Trials Registry Plattform* (WHO ICTRP), *EU Clinical Trials Registry* (EU-CTR), *EMA Clinical Data und* dem Arzneimittel-Informationssystem AMIce (ehemals AMIS) durchgeführt. Die detaillierte Suchstrategie ist für jedes Studienregister im Anhang 4-B getrennt dargestellt.

Die Suche wurde am 02.11.2021 durchgeführt; es wurden keine generellen Einschränkungen bei den genannten Recherchen vorgenommen.

4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen[1]. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA 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 auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

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

Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Des Weiteren erfolgte eine Recherche auf der Internetseite des G-BA, um weitere Studien, die den Kriterien in Abschnitt 4.2.2 entsprechen, zu identifizieren. Dafür wurde über die Suchmaske der Website des G-BA mit dem Schlagwort „Burosumab“ gesucht. Zusätzlich wurde nach dem Schlagwort „UX023-CL301“ gesucht, um eventuell weitere Einträge zu der zuvor identifizierten Studie UX023-CL301 zu identifizieren. Der Wirkstoff Burosumab wurde bereits vom G-BA bewertet. Die Vorgangsnummern lauten: 2018-04-15-D-349 (Erstbewertung) und 2019-10-01-D-492 (Neubewertung nach Fristablauf). Die Suche wurde am 05.11.2021 durchgeführt.

[1] Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staeck K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. BMJ 2015;350:h796

4.2.3.5 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 beschriebenen Rechercheschritte. 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 wurden die Titel und Abstracts der gefundenen Treffer gesichtet und anhand der vordefinierten in Abschnitt 4.2.2 dargelegten Kriterien selektiert. Im Weiteren wurden 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 sind im Abschnitt 4.3.1.1 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):

<u>A: Verzerrungsaspekte der Ergebnisse auf Studienebene</u>

– 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

<u>B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene</u>

– 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 eine Verblindung der Endpunkterheber, die Umsetzung des *Intension-to-treat* (ITT)-Prinzips, eine 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)[2]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[3] bzw. STROBE-Statements[4] 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.

Das Design und die Methodik der eingeschlossenen Studie werden in den Abschnitten 4.3.1.2.1 und 4.3.1.2.2 sowie den zugehörigen Anhängen dokumentiert. Die Bewertung folgt den *Concolidated 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 (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des

[2] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[3] 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.
[4] 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.

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.

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

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 ≥ 5 Jahre)
- Ethnizität und Ethnie
- Gewicht, *body mass index* (BMI) und Größe

Krankheitsspezifische Parameter

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

Stratifizierungscharakteristika der Studie

- Baseline RSS Total Score: ≤ 2,5 (Low RSS) vs. > 2,5 (High RSS)
- Alter: < 5 Jahre und ≥ 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 XLH bei Kindern ab einem Jahr mit röntgenologischem Nachweis einer Knochenerkrankung wurden Endpunkte bezüglich der Morbidität, Lebensqualität und Sicherheit untersucht. Die Mortalität wurde im Rahmen der Sicherheit erhoben.

Tabelle 4-8: Patientenrelevante Endpunkte für die Bewertung von Burosumab in der Therapie der XLH bei Kindern ab einem Jahr

Endpunktkategorie und Ausprägung	Beispiele für Erhebung/Messinstrument
Mortalität	
Im Rahmen der Sicherheit erhoben	-
Morbidität	
Rachitissymptomatik	• *Radiographic Global Impression of Change (RGI-C) Score* • *Rickets Severity Score (RSS)*
Körpergröße	• Stehhöhe/Liegelänge [Z-Score] • Sitzhöhe (cm)

Gehfähigkeit	• *Six-Minute-Walk-Test* (6-Minuten-Gehdistanz Test, 6MWT)
	  ○ Absolute Gehdistanz (Gesamtgehdistanz in Metern)
	  ○ Relative Gehdistanz (zurückgelegte Strecke ausgedrückt als Prozentsatz der vorhergesagten Werte auf Grundlage normativer Daten)

Serumphosphat	• Serumphosphat [mg/dl]

Schmerzerhebung und funktionelle Behinderung	• *Patient-Reported Outcomes Measurement Information System*
	  ○ *Pediatric Pain Interference domain scores*
	  ○ *Physical Function Mobility domain scores*
	  ○ *Fatigue domain scores*
	• *Face Pain Scale – Revised (FPS-R)*

Lebensqualität	
Spezifischer Lebensqualitätsfragebogen für Kinder	• *Short Form (SF)-10 Health Survey*
Sicherheit	
Unerwünschte Ereignisse nach Beginn der Behandlung	Gesamtrate von Patienten mit Ereignis
Schwerwiegende TEAE	Gesamtrate von Patienten mit Ereignis
UE mit CTCAE-Grad 3 oder höher: Grad 3: Unterbrechung von täglichen Aktivitäten des Patienten	Gesamtrate von Patienten mit Ereignis

Endpunktkategorie und Ausprägung	Beispiele für Erhebung/Messinstrument
(Therapie/Behandlung erforderlich) Grad 4: Sterberisiko Grad 5: Tod	
UE, die zum Studienabbruch führten	Gesamtrate von Patienten mit Ereignis
Zahnärztliche Untersuchungen	Gesamtrate von Patienten mit Ereignis
UE von besonderem Interesse	• Reaktion an der Injektionsstelle • Hypersensibilität • Ektopische Mineralisierung • Hyperphosphatämie • *Restless-Legs-Syndrom*
CTCAE: *Common Terminology Criteria for Adverse Events;* TEAE: *treatment emergent adverse events*	

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 schweren 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 [23-27].

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 gemäß § 2 Satz 3 AM-NutzenV als patientenrelevant einzustufen [23, 28, 29].

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 [23, 30]. 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 [23].

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 bei der 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 [23, 28].

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 [23, 26, 28].

Validierung des RSS-Scores

Auf Grundlage der Daten aus der Phase-II-Studie UX023-CL201 (siehe Verfahrensnummer 2018-04-15-D-349) erfolgte eine Validierung des RSS-Scores, um die Reliabilität und die Korrelation des RSS-Scores mit dem Schweregrad XLH-bedingter Anomalien zu untersuchen [28]. Als Grundlage dienten Röntgenbildaufnahmen der 52 Studienteilnehmer (5 – 12 Jahre), die zu Baseline, zu Woche 40 und Woche 64 durchgeführt wurden, und sowohl Aufnahmen von den Handgelenken als auch von den Knien beinhalteten. Die Validierung erfolgte in einem mehrstufigen Prozess, der im Folgenden detailliert beschrieben wird.

Interrater-Reliabilität

Im ersten Schritt wurde die Interrater-Reliabilität, d. h. die Übereinstimmung der Einschätzung des RSS-Scores durch unabhängige Experten, überprüft. Hierzu wurden 104 Röntgenaufnahmen (52 Aufnahmen zu Baseline und 52 Aufnahmen zu Woche 40) der Studienteilnehmer (5 – 12 Jahre) der Studie UX023-CL201 durch vier Experten evaluiert. Um eine hohe Aussagesicherheit zu erzielen, waren die Auswerter gegenüber der Behandlung verblindet. Das gewichtete Cohens Kappa von 0,45 bis 0,65 und der Spearman-Korrelationskoeffizient von 0,69 bis 0,82 zeigen eine hohe Übereinstimmung zwischen den Auswertungen durch die unabhängigen Experten [28].

Intrarater-Reliabilität

Zusätzlich zur Interrater-Reliabilität wurde die Intrarater-Reliabilität bestimmt, um das Ausmaß der Übereinstimmung von sich wiederholenden Messungen zu evaluieren. Hierbei wertete einer der vier Experten, der bereits zuvor an der Auswertung der Röntgenaufnahmen in der Studie UX023-CL201 beteiligt war, die Röntgenaufnahmen erneut aus, um den RSS-Score zu bestimmen. Die daraus resultierenden RSS-Scores wurden anschließend mit der Erstauswertung aus der Studie UX023-CL201 verglichen. Das gewichtete Cohens Kappa von 0,66 und der Pearson- (0,91) sowie Spearman- (0,9) Korrelationskoeffizient belegen ein hohes Maß an interner Konsistenz bei der erneuten Auswertung von Röntgenaufnahmen durch den gleichen Betrachter [28].

Korrelation des RSS-Scores mit XLH-Parametern und patientenrelevanten Endpunkten

Anschließend wurde die Korrelation des RSS-Scores mit XLH-spezifischen Parametern und patientenrelevanten Endpunkten untersucht. Anhand der Alkalischen Phosphatase (ALP) zeigte sich, dass höhere Serumkonzentrationen der ALP zu Baseline signifikant mit höheren RSS-Scores korrelierten ($r = 0,47$; $p = 0,0005$). Des Weiteren wurde festgestellt, dass höhere RSS-Scores mit einer stärker ausgeprägten Symptomatik, einem größeren Ausmaß an Schmerzen und einer Einschränkung der physischen Funktion einhergehen. Studienteilnehmer mit einem RSS-Score ≥ 1,5 zu Baseline wiesen eine geringere Stehhöhe (Z-Score: -2,1) auf als die Gruppe mit einem RSS-Score < 1,5 (Z-Score: -1,4; $p = 0,0183$). Des Weiteren war eine stärkere Einschränkung der Gehfähigkeit in der Gruppe mit einem RSS ≥ 1,5 feststellbar. Während die Studienteilnehmer dieser Gruppe durchschnittlich lediglich 77 % (< 80% bedeutet eine deutliche körperliche Einschränkung) der vorausgesagten Gehdistanz in 6 Minuten zurücklegen konnten, erreichten die Studienteilnehmer mit einem RSS von < 1,5 zu eine Distanz, die 86 % der vorausgesagten Gehdistanz entsprach ($p = 0,0349$). Übereinstimmend

hiermit gaben die Studienteilnehmer mit einem RSS ≥ 1,5 ein höheres Maß an Einschränkungen der physischen Funktion und dem Schmerzlevel (POSNA-PODCI) an (physische Funktion: p = 0,0275; Schmerzen: p = 0,0006). Das Ausmaß der Rachitis kann somit direkt von den Patienten wahrgenommen werden und äußert sich anhand von Schmerzen und Beeinträchtigungen der physischen Funktion, des Wachstums sowie der Gehfähigkeit.

Änderungssensitivität

Schlussendlich wurde die Änderungssensitivität anhand der ALP und des RGI-C zu Woche 64 bestimmt. Eine Verbesserung der Rachitis und damit eine Abnahme des RSS-Scores korrelierte hierbei signifikant mit einer Abnahme der ALP-Aktivität (r = 0,47; p = 0,0008) und zeigt, dass eine Normalisierung der ALP mit einer deutlichen Verbesserung der Rachitis einhergeht. Übereinstimmend hiermit war auch anhand des RGI-C eine hohe Korrelation mit dem RSS-Score messbar. Eine Abnahme des RSS-Scores zu Woche 64 korrelierte signifikant mit höheren RGI-C-Werten (r = -0,65; p < 0,0001). Diese Analysen unterstreichen, dass durch den Endpunkt RSS relevante Änderungen des Gesundheitszustandes von XLH-Patienten gemessen werden können.

Zusammenfassend lässt sich somit feststellen, dass der RSS-Score ein valides Messinstrument darstellt, um das Ausmaß der Rachitis bei XLH-Patienten zu evaluieren. Die Analysen zeigen, dass der RSS-Score mit der Aktivität der ALP korreliert und ein höherer RSS-Score mit stärkeren Wachstumseinschränkungen, Schmerzen und Beeinträchtigungen der physischen Funktion einhergeht. Eine Verbesserung des RSS-Scores kann somit direkt vom Patienten wahrgenommen werden und ist daher unmittelbar als patientenrelevant einzustufen.

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 [23, 30]. 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 [23, 26].

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 Rachitis-symptomatik

+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 Auswerters bezüglich der Veränderungen der Röntgenbildaufnahmen [26].

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 [26].

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 [23, 26, 30].

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 [26].

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 [23, 26, 28, 30]. Bildgebende Verfahren wie der RGI-C und der RSS sind daher geeignete und valide Messinstrumente, um bei einer metabolischen Knochenerkrankung das Ansprechen auf die Therapie zu evaluieren. Ein Rückgang der Knochendeformation und der Anomalien äußert sich u. a. in einer verbesserten Mobilität und der körperlichen Funktion (siehe 6MWT) sowie Abnahme von Gelenkschmerzen und Vermeidung von lebenslangen schwerwiegenden Knochen- und Gelenkkomplikationen. Des Weiteren bedeutet ein Rückgang der Knochendeformationen eine Normalisierung der Physiognomie der betroffenen Patienten, die durch ihre offensichtlichen Fehlbildungen und Gangauffälligkeiten immer wieder einer Stigmatisierung im öffentlichen Alltag ausgesetzt sind. Gerade für Kinder können solche Merkmale immer wieder Anlass für soziale Ausgrenzungen sein und unter Umständen schwere psychische Erkrankungen nach sich ziehen.

Validierung des RGI-Scores

Analog zum RSS-Score erfolgte ebenfalls eine Validierung des RGI-Scores bei Kindern mit XLH. Als Datengrundlage dienten die Röntgenaufnahmen der Phase-II-Studie UX023-CL201 (siehe Verfahrensnummer 2018-04-15-D-349) und der Phase-III-Studie UX023-CL301 [29]. Neben der Reliabilität und der Validität untersuchte diese umfassende Analyse auch die Änderungssensitivität. Dieser mehrstufige Prozess wird im Folgenden detailliert dargestellt.

Intrarater-Reliabilität

Zunächst wurde die Intrarater-Reliabilität bestimmt, um das Ausmaß der Übereinstimmung des RGI-C-Scores bei sich wiederholenden Messungen zu evaluieren. Hierfür wurden Röntgenaufnahmen des Handgelenks und des Knies von Baseline (n = 52) und Woche 40 (n = 52) durch drei unabhängige Experten ausgewertet, die diese Aufnahmen bereits zuvor in der Studie UX023-CL201 analysiert hatten. Die Reevaluierung fand frühestens 6 Wochen nach der Erstauswertung statt. Den Experten war weder das Alter, noch das Geschlecht oder die Behandlung (Phosphat oder Burosumab) der Patienten bekannt. In 90 % der Fälle wurde bei

der erneuten Auswertung der Röntgenaufnahmen nahezu übereinstimmende Ergebnisse (Abweichung ±1 Punkt) erzielt. Das gewichtete Cohens Kappa von > 0,5 unterstreicht hierbei die mäßige bis hohe Übereinstimmung der Auswertungen durch die gleichen Betrachter [29].

Interrater-Reliabilität

Zusätzlich wurde die Interrater-Reliabilität analysiert. Bei über 90 % der Röntgenaufnahmen konnte eine Übereinstimmung zwischen den unabhängigen Experten mit einer maximalen Abweichung von einem Punkt (±1) festgestellt werden. Das gewichtete Cohens Kappa lag im Bereich von 0,49 bis 0,52 und indiziert eine moderate Übereinstimmung.

Konvergenzvalidität

Im nächsten Schritt wurde die Konvergenzvalidität bestimmt, um zu überprüfen, ob der RGI-C-Score geeignet ist, Veränderungen im Ausmaß der Rachitis zu messen. Als Grundlage dienten konzeptionell verknüpfte klinische Parameter sowie patientenberichtete Endpunkte. Hierbei konnte ein signifikanter Zusammenhang zwischen dem RGI-C-Score und dem Serumphosphat (r = 0,397), der ALP (r = -0,611), dem RSS-Gesamtscore (r = -0,672) und der Stehhöhe (r = 0,268) detektiert werden. Ein weiterer signifikanter Zusammenhang wurde auch für die Endpunkte SF-10 (PHS) (r = 0,293) und die zwei Domänen des POSNA-PODCI-Fragebogens (r = 0,409) festgestellt. Im Gegensatz dazu lag kein signifikanter Zusammenhang zwischen dem RGI-C-Gesamtscore und der Gehfähigkeit (6MWT) sowie dem PROMIS-Fragebogen vor.

Known-Groups-Validität

Des Weiteren konnte gezeigt werden, dass sich typische klinische Merkmale der XLH-Erkrankung signifikant zwischen Patienten mit einem RGI-Score von ≥ 2 und < 2 unterscheiden. Dies traf auf den RSS-Gesamtscore (p < 0,001), die Serumphosphat-Konzentration (p < 0,001), die ALP (p < 0,001), die Wachstumsgeschwindigkeit (p = 0,017) und die Stehhöhe (p = 0,002) zu. Patienten mit einem RGI-C-Score zu Woche 64 von ≥ 2 wiesen eine höhere Serumphosphat-Konzentration sowie eine größere Wachstumsgeschwindigkeit und stärkere Zunahme der Stehhöhe auf. Zudem zeichneten sich Studienteilnehmer mit einem RGI-C-Score von ≥ 2 durch eine stärkere Abnahme des RSS-Gesamtscores und der ALP-Aktivität aus. Ein signifikanter Unterschied war auch für die Domäne PHS des SF-10 Fragebogens (p = 0,029) und für die Domäne Schmerz/Komfort des POSNA-PODCI-Fragebogens (p = 0,025) feststellbar.

Änderungssensitivität

Schlussendlich wurde die Änderungssensitivität bestimmt, um zu überprüfen, ob anhand des RGI-C relevante Änderungen im Gesundheitszustand der XLH-Patienten gemessen werden können. Hierfür wurden, basierend auf den Daten der Studie UX023-CL201 und UX023-CL301 zu Woche 64, die Studienteilnehmer in vier Gruppen eingeteilt:

- Gruppe 1: RSS-Gesamtscore verschlechtert (≥ +1,0)
- Gruppe 2: RSS-Gesamtscore unverändert (> -1,0 bis < +1,0)

- Gruppe 3: RSS-Score verbessert (≤ -1 bis > -2)
- Gruppe 4: RSS-Score deutlich verbessert (≤ -2)

Als Maß für die Änderungssensitivität wurde anschließend die *standardized response mean* (SRM) bestimmt. Dies war jedoch nur für die Gruppen 2 – 4 möglich, da die Gruppe 1 lediglich zwei Studienteilnehmer umfasste. Die SRM lag zwischen 1,13 und 7,54 und war bei der Gruppe, die eine deutliche Verbesserung des RSS-Scores aufwies, am größten.

Zusammenfassend lässt sich feststellen, dass der RGI-C eine valide und sensitive Messmethode darstellt, um den Einfluss der Burosumab-Behandlung auf die Rachitis-Schwere zu evaluieren. Analog zum RSS korreliert eine Veränderung des RGI-C direkt mit der klinischen Symptomatik wie z. B. einer verzögerten Wachstumsgeschwindigkeit, Schmerzen und Einschränkungen in der körperlichen Funktion. Eine Veränderung des RGI-Scores kann somit direkt vom Patienten wahrgenommen werden und ist daher als patientenrelevant einzustufen.

Körpergröße (Stehhöhe/Liegelänge und Sitzhöhe)

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 [18, 26]. In der Vergangenheit wurden bereits durch den G-BA anthropometrische Parameter, insbesondere bei Kindern mit charakteristischen, krankheitsbedingten Wachstumsstörungen, als patientenrelevante Morbiditätsparameter eingeschätzt [31]. Aufgrund dessen gilt die Verbesserung des Wachstums gemäß § 2 Satz 3 AM-NutzenV als ein patientenrelevanter Endpunkt.

Die Körpergröße in Zentimeter wurde bei Patienten ≥ 2 Jahre durch Messung der Stehhöhe mit einem Stadiometer ermittelt. Bei Patienten < 2 Jahre bzw. bei Kindern, welche nicht fähig oder nicht bereit waren, für die Messung zu stehen, wurde die Liegelänge bestimmt [32]. 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. 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 [33]. 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*) [26, 27, 34].

Zusätzlich wurde die absolute Sitzhöhe in Zentimeter bei Patienten ab zwei Jahren mittels eines Radiometers erhoben. Um eine ausreichend hohe Validität der Messergebisse zu erhalten, erfolgten Dreifachmessungen, aus denen der Mittelwert gebildet wurde. Die Sitzhöhe wurde

analog zur Stehhöhe zu Woche 24, 40 und 64 bestimmt. In der Extensionsphase wurden die Messungen zu Woche 76, 88, 100, 112, 124 und Woche 140 durchgeführt.

Bei der Körpergröße, der Sitzhöhe und dem 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 [33]. 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 [26]. Ferner wurden in dem früheren Nutzenbewertungsverfahren des Wirkstoffs Asfotase alfa die Z-Scores auf Datengrundlage des CDC vom G-BA als plausibel anerkannt [31, 35]**.** Weiterhin bestätigt die vergleichende Studie von *Rosario et al.*, dass die Wachstumscharts der CDC für den Vergleich mit deutschen Wachstumskurven geeignet sind [36]. Die Eignung der Wachstumscharts der CDC wurde zudem auch in der mündlichen Anhörung zum zweiten Nutzenbewertungsverfahren von Burosumab durch die medizinischen Experten bestätigt [13]. 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 Knochendeformationen, 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 [37]. Eine vom Patienten bewusst wahrnehmbare Verbesserung der körperlichen Leistungsfähigkeit unter der 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. Diese normativen Daten beziehen sich auf gesunde Kinder, die den 6MWT mit einem Messrad absolviert haben. Die Altersadjustierung erfolgte in der Einteilung dieser Kinder in Altersgruppen (3 – 5, 6 – 8 und 9 – 11 Jahren). 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 [27, 37-39]. Laut ATS (*American Thoracic Society*) ist die Angabe der absoluten Werte zu empfehlen [38]. Die absoluten Werte der Veränderung der 6MWT zwischen Baseline und Post-Baseline zeigen dem Patienten die veränderte Gehfähigkeit unter der 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 gemäß § 2 Satz 3 AM-NutzenV 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) [26]. 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/min und ein Blutdruck von > 180/100 mmHg, um das Auftreten von Arrhythmien sowie negativen, kardiovaskulären Ereignissen zu vermeiden [32, 38].

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 [32].

Um eine möglichst hohe Validität zu erhalten, werden die Kriterien der *ATS Guidelines* als qualitätssichernde Maßnahme berücksichtigt [32, 38]. 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 [32, 38, 39]. 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, 32]. 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 Gleichgewichtsanpassung 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 [32, 38]. Trotz dieser Modifikationen der *ATS Guidelines* lässt sich konstatieren, dass die Datenvalidität aufgrund des geschulten klinischen Personals gewährleistet ist. Eine mögliche Verzerrung der relativen Analyse liegt durch die gesunde Vergleichsgruppe, welche den 6MWT mit einem Messrad absolviert hat, gegebenenfalls bei beiden Behandlungsgruppen vor.

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 [40]. Durch die *PHEX*-Mutation wird die Aktivität einer Endopeptidase verringert, die normalerweise einen phosphaturischen Faktor (FGF23) inaktiviert [41]. Die dadurch erhöhte FGF23-Konzentration hemmt den für die Phosphat-Rückresorption wichtigen Natrium-Phosphat-Kotransporter (NaPi-2). Somit spiegelt sich die pathophysiologische Ursache des Phosphatverlust und die Manifestation der XLH-Erkrankung in der Serumphosphat-Konzentration wider. Folglich handelt es sich bei der Serumphosphat-Konzentration gemäß § 2 Satz 3 AM-NutzenV um einen patientenrelevanten Endpunkt.

Der Indikator für eine erfolgreiche Therapie mit Burosumab ist die Serumphosphat-Konzentration, die als Zielparameter gilt. Das Erreichen einer normal-physiologischen Serumphosphat-Konzentration bei einer Substitutionstherapie mit oralem Phosphat und aktivem Vitamin D ist nicht möglich [12, 13]. Die Behandlung mit Burosumab stellt gegenüber der Substitutionstherapie mehrere Vorteile dar. Im Vergleich zu einer 2-wöchigen subkutanen Injektion von Burosumab muss bei einer Substitutionstherapie eine 3 – 5 Mal tägliche Gabe von Phosphat und aktivem Vitamin D erfolgen [10, 11]. Die häufige Einnahme führt zu einer geringen Compliance und die erforderlichen hohen Dosen an Phosphat schädigen die Nieren bereits innerhalb kurzer Zeit. Hinzu kommt, dass die Substitution mit Phosphat die Zirkulation von freiem FGF23 fördert und dadurch den renalen Phosphatverlust zusätzlich verstärkt. Die therapieassoziierten Nebenwirkungen wie z. B. die Ausbildung eines sekundären oder tertiären Hyperparathyreoidismus oder die Entstehung von Nephrokalzinosen erfordern zudem ein strenges Monitoren der XLH-Patienten [2, 3, 41].

Die Serumphosphat-Konzentration (mg/dl) wurde sowohl zu Baseline als auch zu Woche 1, 2, 4, 8, 12, 16, 24, 32, 33, 40, 52, 64, 74, 76, 88, 100, 112, 124 und Woche 140 nach einer Fastenzeit von mindestens 4 h über Nacht gemessen.

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 ab dem adoleszenten Alter. 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. Durch die Beinachsenfehlstellung kommt es zu einer Adaptation der nachfolgenden Gelenke, und folglich zu einer präarthrotischen Deformität, die, je älter der Patient ist, erhebliche Schmerzen bereitet. Daher ist die Reduktion von Schmerzen und eingeschränkter Mobilität gemäß § 2 Satz 3 AM-NutzenV ein patientenrelevanter Endpunkt. Dieser Endpunkt wurde in der Studie UX023-CL301 anhand der *Patient Reported Outcomes Measurement Information System* (PROMIS) pädiatrischen Schmerzskala gemessen [13, 39].

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 [39].

Der PROMIS-Fragebogen wurde von dem *National Institutes of Health* (NIH) entwickelt und verwendet spezifische Domänen zur Beurteilung des Wohlbefindens der Patienten [42]. 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 [39]. 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

[26, 27]. 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 [26]. Die PROMIS-Referenzpopulation der Domänen beider Versionen war der allgemeine US-Zensus von 2000. Für die PROMIS Domänen *Pain Interference, Physical Function Mobility* und *Fatigue* wurde die Durchschnittsbevölkerung sowohl für die Selbstbericht- als auch Elternversion verwendet [43].

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]. Der Unterschied zwischen dem Selbstbericht- und Fremdeinschätzungsfragebogen liegt nur in der Wortwahl und nicht in der Bedeutung. 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, die mindestens 8 Jahre alt zur Screeningvisite waren, 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, sodass 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 [25, 26].

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. Die Abfrage der Schmerzintensität erfolgte

allerdings nicht über den PROMIS-Fragebogen, sondern ausschließlich über das Messinstrument *Face Pain Scale-Revised* (FPS-R). Das heißt, dass die Frage 23 der Selbstbericht-Fragebogenversion bei den Kindern nicht abgefragt wurde. Ursprünglich war dies angedacht, erfolgte jedoch nicht, weil eine einheitliche selbstberichtende Abfrage gewünscht war. 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 die letzte Woche. 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 [39].

In der Studie UX023-CL301 wurde ausschließlich die Version 2 des PROMIS-Fragebogens verwendet, obwohl ursprünglich die Verwendung der Version 1 geplant war, da zum Zeitpunkt der Erstellung des Studienprotokolls die Version 2 des PROMIS-Fragebogens noch nicht öffentlich zugänglich war. Somit wurde auf Anraten der PROMIS-Entwickler die aktualisierte Version des Fragebogens (Version 2) eingesetzt. Die Freigabe der Version 2 des PROMIS-Fragebogens erfolgte vor der Einschreibung des ersten Studienteilnehmers und wurde in einer späteren Novellierung des Protokolls dokumentiert. Um jedoch den Studienzentren schnellstmöglich die Änderungen und Aktualisierungen des Studienprotokolls mitzuteilen, wurden Memos bereits vor dem Amendment 1 an die Studienzentren versandt, um eine lückenlose und schnelle Kommunikation zu gewährleisten [44-46]. 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 [47-49]. Die Domäne Mobilität wurde dagegen 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 [47, 50].

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 [26, 39].

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 [26, 39]. 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 [51]. 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 [52]. Demnach ist die Validität der erstellten Fragebögen mithilfe des pädiatrischen PROMIS-Verzeichnisses gegeben.

Thissen et al. untersuchte die *minimally important difference* (MID). Dabei wurden Jugendlichen, Eltern und Klinker (n = 246) zwei ausgefüllte PROMIS-Fragebögen, welche die Domänen Depression, Schmerz, Fatigue und Mobilität abfragten, vorgelegt. Die Teilnehmer sollten nun angeben, ob die Person, die diesen Fragebogen ausfüllte, eine bedeutende Änderung hatte oder nicht. Das Ergebnis dieser Umfrage zeigte eine MID bei 1,9 – 2,4 [53].

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 und für die Verwendung bei Kindern im Alter von 5 bis 16 Jahren validiert ist [54]. Dabei wird das 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 [39]. In der Studie UX023-CL301 wurde das Kind aufgefordert, auf ein Gesicht zu zeigen, dass am besten die Intensität des aktuellen Schmerzes abbildet [39].

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) [25-27]. Es wurden nicht, wie im CAO-Manual [39] 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 der 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 die Verbesserung des Gesundheitszustandes gibt, handelt es sich hierbei gemäß § 2 Satz 3 AM-NutzenV 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.

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 [20]. 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 einem Jahr und Jugendlichen, sodass 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 [55, 56]. Dieser Fragebogen leitet sich vom SF-36 ab und wurde speziell für Kinder entwickelt [57]. Die Items wurden aus dem *Child Health Questionnaire* (CHG) 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:
  Allgemeiner Gesundheitszustand, Einschränkungen durch Gesundheits-probleme und Schmerzen
- Psychosoziale Gesamtpunktzahl (PSS-10): Items 4, 6 bis 9:
  Psychische Probleme, Verhaltensauffälligkeit, Zufriedenheit und Freundschaft

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 sind auf einen Mittelwert von 50 und eine Standardabweichung von 10 in der kombinierten US-Allgemeinbevölkerung aus gesunden, chronisch kranken Kindern und Kindern mit Behinderungen standardisiert. Die Bewertung des SF-10 erfolgt nach der in *Saris-Baglama et al.* (2007) beschriebenen Methode. Höhere globale Punktzahlen sind mit einer besseren Lebensqualität verbunden [58].

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 Studie 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 gemäß § 2 Satz 3 AM-NutzenV 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 UE 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. In die Auswertung wurden jedoch nur *UE* eingeschlossen, die nach der erstmaligen Verabreichung der Prüfsubstanz (Burosumab) oder der aktiven Kontrolle (orale Phosphatsubstitution und aktives Vitamin D) auftraten (*Treatment emergent adverse events* (*TEAE*)), um die Inzidenz, Häufigkeit und Schwere der *UE* sowie einen möglichen Zusammenhang mit der Behandlung zu bewerten. *UE*, die zwischen dem Zeitpunkt der Einwilligung und der Baseline-Visite auftraten, wurden hierbei nicht berücksichtigt, da diese *UE* in keinem Bezug zur Behandlung mit der Prüfsubstanz standen und die objektive Beurteilung des Sicherheitsprofils in beiden Behandlungsarmen beeinträchtigt hätten.

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 vor allem Patienten, die nicht in die Extensionsphase aufgenommen werden konnten (Kinder aus Japan oder Korea).

Neben den Klinik-Visiten wurden auch telefonische Visiten oder Hausbesuche zur Dokumentation von UEs und Begleitmedikationen gesammelt. Die Hausbesuche bei den Kindern in der Burosumab-Gruppe während der Behandlungsperiode und bei allen Kindern während der Extensionsphase dienten außerdem der subkutanen Injektion von Burosumab. Hausbesuche konnten für Studienteilnehmer in Japan oder Korea nicht angeboten werden (Japan: $N_{(Burosumab)} = 2$; $N_{(aktive\ Kontrolle)} = 3$; Korea: $N_{(Burosumab)} = 0$; $N_{(aktive\ Kontrolle)} = 2$). Bei diesen Patienten wurden ausschließlich Klinik-Visiten und keine telefonischen Visiten durchgeführt. Für Kinder in der Kontrollgruppe wurden Informationen zu UEs und Begleitmedikation während der Behandlungsperiode per Telefonanruf abgefragt. Diese Anrufe erfolgten exakt zu dem gleichen Zeitpunkt wie die Hausvisiten der Patienten in der Burosumab-Gruppe und wurden gemäß dem Studienprotokoll durchgeführt (siehe Tabelle 2.2 im Studienprotokoll [17]). Durch die unterschiedliche Art der Visiten in den Behandlungsarmen ist deshalb von einer detaillierteren Erfassung der Sicherheitsinformationen in der Burosumab-Gruppe auszugehen.

Der Prüfarzt beurteilt weiterhin die Schwere, die Kausalität (Zusammenhang mit der Studienmedikation) und die Bedeutung jedes UE.

Schwere der UEs

Der Prüfarzt schätzt die Schwere aller UEs gemäß der *National Cancer Institute Common Terminology Criteria for Adverse Evenst (NCI CTCAE)* 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.

- Schwer (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).

Schwerwiegendes UEs

Ein schwerwiegendes UE 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 ein medizinisches oder chirurgisches Eingreifen benötigen könnte, um eines der oben gelisteten Ereignisse zu verhindern.

Krankenhausaufenthalte, die vor der Studienteilnahme geplant waren, wurden nicht als schwerwiegendes UE betrachtet. Krankenhausaufenthalte aufgrund von Vorerkrankungen, die nach Eintritt in die Studie geplant wurden, galten als schwerwiegende UEs.

UE von besonderem Interesse

Als UE von besonderem Interesse waren folgende Ereignisse vordefiniert:

- Reaktionen an der Injektionsstelle
- Hypersensibilität
- Hyperphosphatämie
- Ektopische Mineralisierung
- *Restless-Legs-Syndrom*

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 der Anwendung der Studienmedikation und dem Auftreten eines UE nicht zwangsläufig mit einem kausalen Zusammenhang zwischen beiden gleichzusetzen ist.

Entsprechend der Vorgaben des G-BA werden im vorliegenden Nutzendossier *TEAE* und *schwerwiegende TEAE* sowohl als Gesamtraten als auch aufgeteilt nach Systemorganklassen (SOC) bzw. *Preferred Terms* (PT) dargestellt. Da durch den Endpunkt *Sicherheit* eine Verringerung von Nebenwirkungen beurteilt werden kann, handelt es sich hierbei um einen validen und patientenrelevanten Endpunkt.

Die UE wurden bei allen Visiten (Klinik-Visiten, Hausbesuchen und telefonische Visiten) und im gesamten Studienverlauf erhoben. Alle UE wurden ab dem Zeitpunkt der Unterzeichnung der Einverständniserklärung bis zu 12 Wochen nach der letzten Gabe aufgezeichnet.

Zahnmedizinische Untersuchung

XLH ist eine genetisch bedingte Erkrankung, die durch eine Dysregulation der Mineral- und Vitamin-Homöostase gekennzeichnet ist. Diese führt neben einer Störung der Mineralisierung von Knochen auch zu Defekten im Dentin. Aufgrund dessen weisen XLH-Patienten sowohl bei den Milch- als auch bei den bleibenden Zähnen häufig Zahnabszesse auf [11], ohne dass vorab Karies oder Traumata festgestellt wurden [59]. Zahnabszesse können sowohl durch eine Infektion, als auch unbehandelten Karies oder Defekten in Schmelz, Dentin und Zement verursacht werden. Die verminderte Mineralisierung fördert die Ausbreitung der eitrigen und schmerzhaften Infektion, die sogar zum Zahnverlust führen kann [10, 11, 59].

75 % der Patienten mit XLH weisen eine Parodontal- oder Zahnfleischerkrankung auf, die möglicherweise mit der Burosumab-Behandlung in Zusammenhang gebracht werden könnte. Derzeit gibt es jedoch keine Möglichkeit das Ausmaß des Schmelz- oder Dentindefekts zu bestimmen [59]. Aus diesem Grund wurde sowohl ein möglicher Zahnverlust sowie das Auftreten von Karies und Zahnabszessen in der Fachinformation von Crysvita® als Nebenwirkungen mit aufgenommen [1].

Die Ursache für das häufige Auftreten von Zahnproblemen ist noch nicht vollständig geklärt. Bekannt ist jedoch, dass das PHEX-Protein in den Knochen und Zähnen exprimiert wird. Durch die gestörte Mineralisierung der Zähne kommt es zu einer Fehlbildung der Zahnhartsubstanz (Dentin) der Milchzähne und der bleibenden Zähne. Weiterhin ist bekannt, dass sich auch ohne jegliche Vorgeschichte von Karies oder Zahnproblemen, spontane periapikale Abszesse und Fisteln bilden können. Da das Dentin den größten Anteil des Zahnes ausmacht und bei XLH-Patienten am stärksten betroffen ist, sind die Auswirkungen einer verzögerten und unvollständigen Mineralisierung des Dentins in Bezug auf die Festigkeit und Funktion der Zähne gravierend. Somit kommt es neben Zahnabszessen häufig auch zu anderen zahnmedizinischen Befunden wie z. B. Zahnschmelzdefekte und -frakturen, Wurzeldysplasien, vergrößerte Pulpa-Kammern mit langen Pulpa-Hörnern, parodontale Erkrankungen, Hypomineralisierung des Alveolarknochens, Verlust der Dura-Lamelle und Malokklusion [10, 11, 59, 60].

Hanisch et al. veröffentlichte 2019 die Ergebnisse einer Studie, in der 43 XLH-Patienten (32 Frauen und 11 Männer) zu ihrer mundgesundheitsbezogenen Lebensqualität befragt wurden. Zur Bewertung der Zahngesundheit wurde ein Fragebogen entwickelt und validiert [61], der aus offenen Fragen und Fragen der standardisierten deutschen Version des *Oral Health Impact Profile-14* (OHIP-14) besteht. Für jede Antwort wurden Punkte vergeben. Im Gesamtergebnis konnten Werte zwischen 0 Punkten (bestmögliche mundgesundheitsbezogene Lebensqualität) und 56 Punkten (schlechteste mundgesundheitsbezogene Lebensqualität) erzielt werden. Die im Rahmen dieser Studie ermittelten Werte wurden mit den bekannten Werten aus der deutschen Normbevölkerung verglichen, hier lag der durchschnittliche Score bei 4,09 Punkten. Die Mehrheit der Studienteilnehmer (77,5 %) berichteten über Zahnprobleme, die in Zusammenhang mit der XLH-Erkrankung stehen. Dabei wurden vor allem die Störung der Zahnmineralisierung (n = 26) sowie Abszess- und Fistelbildung (n = 21) genannt. Studienteilnehmer mit Zahnproblemen zeigten eine erhöhte Tendenz zu einer

schlechteren Zahngesundheit als diejenigen ohne orale Symptomatik. In Deutschland waren die OHIP-14-Scores für diese XLH-Patienten mit 10,30 Punkten deutlich schlechter als die Werte der Allgemeinbevölkerung [62].

Die zahnärztliche Untersuchung war ein explorativer Endpunkt in der Studie UX023-CL301. Hierbei wurde die Anzahl der dentalen Ereignisse wie Karies, Zahnextraktion, Wurzelbehandlung, Zahnabszess und Gingivitis zu Baseline, zu Woche 4, 8, 16, 24, 32, 40, 52, 64 und in der Extensionsphase zu Woche 74 abgefragt. Bei jeder Klinik-Visite wurde eine mündliche Befragung durch einen ausgebildeten Arzt durchgeführt. Dieser Arzt hatte die Fähigkeit, Zahnabszesse oder ausgeprägte Karies zu identifizieren. Eine spezifische Ausbildung/Schulung des Studienpersonals für die zahnmedizinische Beurteilung erfolgte jedoch nicht. Die Kinder oder deren Eltern wurden dabei proaktiv gefragt, ob seit ihrer letzten Visite ein dentales Ereignis wie Karies, Zahnextraktion (aufgrund von Abszess oder Karies), Wurzelbehandlung, Zahnabszess oder Gingivitis aufgetreten war bzw. behandelt wurde. Die Ergebnisse dieser mündlichen Befragung wurden im eCRF (*electronic case report forms*) erfasst und getrennt von den therapieassoziierten unerwünschten Ereignissen (TEAEs) tabellarisch dargestellt. Die mündliche Befragung beurteilte proaktiv den dentalen Zustand, während die TEAEs sich auf Zahnprobleme bezogen, die außerhalb der geplanten Klinik-Visiten auftraten und von den Kindern oder ihren Eltern berichtet und vom Prüfarzt dokumentiert wurden.

Während bei Patienten mit XLH Zahnprobleme wie Abszesse und vergrößerte Pulpakammern auftreten, gibt es nur wenige klinische Daten über den Zeitpunkt, die Häufigkeit und den Verlauf von dentalen Anomalien bei pädiatrischen Patienten. In den verfügbaren Berichten wurden dentale Befunde meist bei einer kleinen pädiatrischen XLH-Population berichtet, die mit einer Substitutionstherapie behandelt wurden. Rückschlüsse auf die zugrundeliegende Zahnerkrankung durch die XLH, ohne eine Behandlung, können dadurch nicht gezogen werden. Aus diesem Grund wurde die zahnmedizinische Untersuchung als explorativer Endpunkt aufgenommen, um neue Erkenntnisse über die Zahnerkrankung der pädiatrischen XLH-Patienten zu gewinnen. Die Auswirkungen von Burosumab auf die Zahngesundheit bei XLH-Patienten wurde daher nicht nach einem systematischen Vorgehen durch formale zahnärztliche Untersuchungen, einschließlich Röntgenbildaufnahmen und Routineüberwachung bewertet.

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 gemäß § 2 Satz 3 AM-NutzenV um einen patientenrelevanten Endpunkt.

Verwendete statistische Methoden

Auswertungspopulationen Full Analysis Set (FAS)

Die FAS-Population war definiert als alle randomisierten Patienten, die mindestens eine Injektion der Studienmedikation erhalten hatten und mindestens einen Post-Baseline Wert haben. 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 (FAS: N(Burosumab) = 29; N(aktive Kontrolle) = 32). Eine Ausnahme stellen die Endpunkte dar, bei denen die Kinder ein Mindestalter von 5 Jahren zur Screeningvisite haben mussten (FAS ≥ 5 Jahre: N(Burosumab) = 15; N(aktive Kontrolle) = 20).

Auswertungspopulation Safety Anaylsis Set

Die *Safety*-Population war definiert als alle randomisierten Patienten mit mindestens einer Dosis der zugeteilten Studienmedikation. Die Auswertung der Patienten erfolgt gemäß der tatsächlichen erhaltenen Behandlung.

Auswertungsmethoden

Analysen mit stetigen Messwerten

Die Ergebnisse der Studie UX023-CL301 werden durch adäquate Effektschätzer mit zugehörigem zweiseitigem 95% KI und zweiseitigen p-Werten geeigneter statistischer Tests zu einem Signifikanzniveau von 5 % dargestellt.

Für stetige Variablen wurden Änderungen zu Baseline anhand von Mittelwerten inklusive SD berechnet und Unterschiede zwischen den Behandlungsarmen mittels ANCOVA (analysis of covariance), GEE (*generalized estimation equations*) oder MMRM (*mixed model repeated measures*) analysiert.

Endpunkte ohne Berücksichtigung von Messwiederholungen

Die Auswertung der Endpunkte ohne Berücksichtigung von Messwiederholungen erfolgte mittels ANCOVA. Bei den ANCOVA-Modellen waren die Behandlungsgruppe, die Baseline-Alters-Stratifizierungsfaktoren (< 5 Jahre vs. ≥ 5 Jahren) als Faktoren und der Baseline-Wert des jeweiligen Endpunkts als kontinuierliche Kovariate vorgesehen (Ausnahme RGI-C: hier diente der RSS Total Score als kontinuierliche Kovariate). Bei dem Endpunkt Serumphosphat-Konzentration wurde statt der Baseline-Alters-Stratifizierungsfaktoren das Alter zu Baseline berücksichtigt.

Endpunkte unter Berücksichtigung von Messwiederholungen

Endpunkte unter Berücksichtigung von Messwiederholungen wurden unter der Verwendung des GEE-Modells ausgewertet, dass die patientenspezifische Heterogenität berücksichtigt und somit ein geeignetes Auswertungsmodell darstellt. Bei den GEE-Modellen wurden die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten als Faktoren, sowie der Baseline-Wert des jeweiligen Endpunkts als kontinuierliche Kovariate berücksichtigt (Ausnahme RGI-C: hier wurde der RSS Total Score verwendet). Die Baseline-RSS-Stratifizierungsfaktoren (≤ 2,5 vs. > 2,5) (Ausnahme: Endpunkt RSS, RGI-C) bzw. die Baseline-Alters-Stratifizierungsfaktoren (Ausnahme: 6MWT, PROMIS, FPS-R und SF-10) wurden ebenfalls in das Modell aufgenommen. Bei dem Endpunkt Körpergröße/Sitzhöhe wurden die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren als Faktoren aufgenommen, das Alter zu Baseline und die Baseline

Körpergröße/Sitzhöhe als kontinuierliche Kovariate. Für die Kovarianz wurde zunächst eine austauschbare (*exchangeable*) Struktur gewählt. Führte diese zur Nichtkonvergenz des Modells, wird eine unabhängige Kovarianzstruktur verwendet. Nur Patienten mit Werten zur Baseline und dem Auswertungszeitpunkt wurden berücksichtigt.

Zusätzlich zu dem GEE-Modell wurde bei Messwiederholungen das MMRM verwendet, um die Vorgaben der Dossiervorlage zu erfüllen. Bei diesen Auswertungen (trifft nur auf den Endpunkt PROMIS zu) wurden die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren als Faktoren in das Modell aufgenommen, sowie der Baseline-Wert des PROMIS als Kovariate. Als Kovarianzstruktur wurde die *compound symmetry* verwendet.

Um die klinische Relevanz des gemessenen Effekts zu bestimmen, wurde entsprechend den Empfehlungen des IQWiG Methodenpapiers die SMD *post-hoc* in Form von Hedges`g berechnet [63]. Als Irrelevanzschwelle wurde je nach Richtung der Effektschätzer +0,2 oder -0,2 gewählt. Lag das 95% KI vollständig oberhalb oder unterhalb dieser Irrelevanzschwelle, wurde von einem klinisch relevanten Effekt ausgegangen. Obwohl sich durch die Vielzahl an *post-hoc* durchgeführten statistischen Tests ein multiples Testproblem ergibt, wurde keine Adjustierung des Signifikanzniveaus vorgenommen.

Analyse mit binären Messwerten

Für dichotome Zielvariablen wurden, zusätzlich zu den absoluten und prozentualen Anteilen, das *Odds Ratio* (OR), das *Relative Risiko* (RR), das adjustierte RR (ARR) und die absolute Risikodifferenz (ARD) inklusive 95% KI und p-Werten bestimmt.

- Das OR wurde mittels der logistischen Regression bestimmt. Die Baseline-Alters-Stratifizierungsfaktoren und die Behandlungsgruppe wurden als Faktoren in das Modell aufgenommen, sowie der RSS Total Score als kontinuierliche Kovariate.

- Das adjustierte RR und die dazugehörigen Konfidenzintervalle wurden *post-hoc* nach der Methode von Zhang und Yu bestimmt (adjustiert für die Baseline-Alters-*Stratifizierungsfaktoren, die Behandlungsgruppe und den RSS Total Score*) [64]. Hierfür wurde das adjustierte OR sowie die Grenzen des entsprechenden Konfidenzintervalls mit folgender Formel transformiert:

$$\mathrm{ARR} = \frac{\mathrm{OR}}{(1 - \mathrm{P}_0) + (\mathrm{P}_0 \times \mathrm{OR})}$$

  OR = *Odds Ratio*
  P_0 = Risiko der Kontrollgruppe

- Die RR und ARD wurden *post-hoc* unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation und der p-Wert mittels Fishers exaktem Test berechnet.

Das RR, ARR und OR wurden als Burosumab gegenüber der aktiven Kontrolle berechnet. Die ARD wurde als Burosumab minus aktive Kontrolle berechnet.

Für die Auswertungen der Gesamtraten der jeglichen *TEAE*, *schweren TEAE* und *schwerwiegenden TEAE* sowie für die Responderanalyse des primären Endpunkts *RGI-C* wurde zu Woche 64 das GLMM (*Generalised linear mixed models*) verwendet, dass die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren berücksichtigte und den Baseline RSS Total Score als kontinuierliche Kovariate. Patientenspezifische Heterogenität wurde mit Hilfe von zufälligen Effekten für die Modellkonstante für die Patienten berücksichtigt.

Umgang mit fehlenden Werten

In der Studie UX023-CL301 wurden im Allgemeinen fehlende Daten – sofern nicht anders angegeben – als fehlend behandelt. Bei den Endpunkten, bei denen die Änderung zu Baseline evaluiert wurde, wurden nur die Studienteilnehmer mit in die Analysen aufgenommen, bei denen sowohl ein Baseline- als auch ein Post-Baseline-Wert vorhanden war. Für wiederholende Messungen werden bei der *generalized estimated equation* (GEE)-Analyse die Modellparameter gleichzeitig unter Verwendung aller beobachteten Daten geschätzt.

Abhängig von der Anzahl an fehlenden Werten wurde eine Sensitivitätsanalyse durchgeführt und die fehlenden Werte unter Verwendung von verschiedenen Methoden ersetzt. Eine ausführliche Beschreibung der durchgeführten Sensitivitätsanalysen ist in Abschnitt 4.2.5.4 zu finden.

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 Effekten nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung[5] erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine

[5] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. Cochrane Database Syst Rev 2015: 25-27.

qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen 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 verwendet[6] 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 geeigneter statistische Maße auf Vorliegen von Heterogenität[7, 5] erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0{,}05$), 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.

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.

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

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 ($N_{(Burosumab)}$ = 29; $N_{(aktive\ Kontrolle)}$ = 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 ($N_{(Burosumab)}$ = 15; $N_{(aktive\ Kontrolle)}$ = 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 [26]:

- *Last Oberservation Carried Forward* (LOCF): Die letzte nicht fehlende Post-Baseline-Beobachtung 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 mit Hilfe eines *Markov Chain Monte Carlo* (MCMC)-Ansatzes 100 vollständige Datensätze mit imputierten Werten zu erstellen. Die imputierten Daten werden mit dem ANCOVA-Modell zu Woche 40 oder dem GEE-Modell zu Woche 64 analysiert. Die jeweiligen Ergebnisse der 100 Modelle werden mit dem PROC MIANALYZE-Verfahren in SAS kombiniert. Der Baseline RSS Total Score, die Behandlungsgruppe und der Baseline-Alter-Stratifizierungsfaktor werden als Einflussvariablen in der Imputation verwendet. Andere Variablen wie beispielsweise die Veränderung der alkalischen Phosphatase (ALP) zu Baseline können ebenfalls als Einflussvariable berücksichtigt werden.

Sensitivitätsanalysen

Um die Robustheit der Endpunkte 6MWT, RGI-C und RSS zu überprüfen, wurde in der Studie UX023-CL301 für diese Endpunkte jeweils eine Sensitivitätsanalyse durchgeführt.

Sensitivitätsanalyse 6MWT

Für den Endpunkt **6MWT** wurden zwei Analysen dargestellt. Die erste Analyse schließt alle Patienten ein, die zur Screeningvisite mindestens 5 Jahre alt waren (FAS ≥ 5 Jahre). Hier sind auch zwei Patienten in der Burosumab-Gruppe enthalten, die den 6MWT nicht vollständig abgeschlossen haben. Ein Kind in der Burosumab-Gruppe absolvierte zu Baseline 54 Meter in 1:17 Minuten. Bereits zu Testbeginn berichtete dieses Kind von anhaltenden Knochenschmerzen und erklärte, dass es an diesem Tag nicht schnell laufen könne. Ein weiteres Kind in der Burosumab-Gruppe lief zu Woche 40 eine Strecke von 440 Metern bevor es nach 5:27 Minuten anfragte den Test zu beenden. Durch die beiden Patienten liegt eine entgegengesetzte Verzerrung in der Burosumab-Gruppe vor. Aus diesem Grund wurde eine zweite Analyse präsentiert, die diese beiden Kinder ausschloss. Eine Unterscheidung zwischen dem Baseline- und 40-Wochen-Wert wurde nicht gemacht, da nicht klar ist, ob das Kind zu Baseline jemals ganze 6 Minuten hätte gehen können (hier erfolgte kein Testlauf zur Screeningvisite). Beide Analysen des 6MWT in Metern zeigen dennoch robuste und konsistente Ergebnisse.

Sensitivitätsanalyse RGI-C und RSS

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 Total Scores 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 Total Score der Woche 40 abzuleiten. Die Ergebnisse dieser LOCF-Analyse waren denen der Primäranalyse sehr ähnlich. Alle Röntgenbildaufnahmen wurden für die RSS-Bewertungen in Woche 64 einbezogen.

Sensitivitätsanalyse PROMIS MMRM-Analyse (mixed model repeated measures)

Für die Auswertungen von mit Skalen erhobenen patientenberichteten Endpunkten war in der Studie UX023-CL301 laut SAP das GEE-Modell vorgesehen. Ähnlich wie das MMRM-Modell, berücksichtigt dieses statistische Modell die über den Studienverlauf ermittelten Informationen vollständig. Um den Anforderungen der mittlerweile gültigen Dossiervorlage gerecht zu werden und um auszuschließen, dass durch die Verwendung des MMRM-Modells ggf. abweichende Ergebnisse erzielt werden, wurde beispielhaft für den patientenberichteten

Endpunkt PROMIS *post-hoc* eine MMRM-Analyse durchgeführt. Diese Ergebnisse waren konsistent zu den Primäranalysen und belegen, dass das GEE-Modell ein adäquates Auswertungsmodell darstellt, das die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt.

Sensitivitätsanalyse PROMIS Selbst- und Fremdeinschätzung

Basierend auf der Nutzenbewertung vom 02.01.2020 wurde für das vorliegende Dossier neben der Hauptanalyse zusätzlich eine Sensitivitätsanalyse durchgeführt, bei der die Auswertung der Fremd- (Kinder im Alter von 5 – 7 Jahren) und Selbsteinschätzung (Kinder im Alter von 8 – 12 Jahre) getrennt voneinander erfolgte [65]. Diese Analyse wurde *post-hoc* durchgeführt und war im SAP nicht präspezifiziert.

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 (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.2).

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 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. 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 6.0 wird der Interaktionstest daher nur dann interpretiert, wenn folgende Kriterien erfüllt sind [63]:

Tabelle 4-9: Kriterien für die Berechnung der Interaktionstests in Anlehnung an das Methodenpapier 6.0 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 tabellarisch 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. Diese Analysen befinden sich daher ausschließlich im Anhang 4-G.

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(Burosumab) = 2; N(aktive Kontrolle) = 3
  - Rest der Welt: N(Burosumab) = 29; N(aktive Kontrolle) = 27
- *PHEX*-Mutation:
  - Positiv: N(Burosumab) = 27; N(aktive Kontrolle) = 28
  - Nicht-Positiv: N(Burosumab) = 2; N(aktive Kontrolle) = 4

Bei diesen Subgruppenanalysen, auch bei signifikantem Ergebnis im Interaktionstest, kann durch die Nicht-Einhaltung der Mindestanforderungen des Methodenpapiers 6.0 keine valide Interpretation erfolgen, weshalb hier nur das Ergebnis der Gesamtstudienpopulation zu berücksichtigen ist [63].

Präspezifizierte Subgruppenanalyse und Rationale

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

- Krankheitsschweregrad: Baseline RSS Total Score (≤ 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)

Die Subgruppenanalysen für die Studie UX023-CL301 wurden im SAP prädefiniert und mit dem Ziel ausgewählt, mögliche Einflussfaktoren auf den Therapieeffekt von Burosumab zu untersuchen. So wurde die Subgruppe RSS Total Score (≤ 2,5 vs. > 2,5) definiert, um zu untersuchen, ob Burosumab bei Patienten mit einer schweren Rachitissymptomatik im Vergleich zu einer milderen Rachitissymptomatik ähnlich wirksam ist. Da jüngere Kinder eine längere Wachstums- und Knochenumbauphase als ältere Kinder haben, wurde die Subgruppe Alter (< 5 vs. ≥ 5 Jahre) definiert. Diese Subgruppe ist insbesondere für die Endpunkte Wachstum und Rachitissymptomatik (RSS und RGI-C) interessant. Die Subgruppe Region (Japan vs. Rest der Welt) wurde aus Gründen von operativen und logistischen Gesichtspunkten gewählt und nur deskriptiv dargestellt.

Zusätzliche Analysen wurden für die Endpunkte PROMIS, FPS-R und 6MWT berechnet. Die Subgruppe RSS Total Score (≤ 2,5 vs. > 2,5) sollte dabei untersuchen, ob die Kinder mit einer ausgeprägteren Rachitissymptomatik mehr Schmerzen, eine verminderte körperliche Funktion

oder eine verringerte Lebensqualität hatten. Um die Unterschiede im zeitlichen Verlauf des Wachstums und der Pubertät zwischen Mädchen und Jungen sowie potentielle Auswirkungen auf das Skelett zu analysieren, wurden Subgruppen nach Geschlecht für die Endpunkte Wachstum und Rachitissymptomatik (RSS und RGI-C) ermittelt. Basierend auf dem *PHEX*-Mutationsstatus wurden ebenfalls Subgruppen zu Woche 40 und 64 für RGI-C, RSS, Wachstum und Wachstumsgeschwindigkeit durchgeführt. Mit dieser Analyse sollte beobachtet werden, ob sich der klinische Phänotyp von Wachstum und Rachitissymptomatik unterscheidet, je nachdem, ob eine *PHEX*-Mutation identifiziert werden konnte oder nicht.

Wie bereits beschrieben, ist die Aussagekraft der Subgruppen Region und *PHEX*-Mutationsstatus aufgrund der Patientenzahlen limitiert.

Da die Subgruppen Region und *PHEX*-Mutation nicht die Mindestanforderungen zur Darstellung von Subgruppen im Dossier, wie im Abschnitt 4.3.1.3.2 aufgeführt, erfüllen, wird auf eine Darstellung im Dossier verzichtet.

Post-hoc definierte Subgruppenanalysen

Post-hoc definierte Subgruppenanalysen wurden für die Endpunkte PROMIS, FPS-R, SF-10 und 6MWT – sofern nicht vollständig im *Clinical Study Report* der UX023-CL301-Studie [26] vorhanden – nachträglich durchgeführt, um die Vorgaben der Dossiervorlage zu erfüllen und eine einheitliche Analyse aller Endpunkte transparent darzustellen. In der nachfolgenden Tabelle 4-10 erfolgt die Darstellung, welche Subgruppen a priori oder *post-hoc* durchgeführt wurden.

Tabelle 4-10: 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
PHEX-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

UX023-CL301
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 Tabelle 4-57 dargestellt. Eine bedeutsame Heterogenität wurde ab einem Interaktions-p-Wert < 0,05 angenommen und als Beleg für eine Effektmodifikation gewertet. Lag ein Beleg für eine Interaktion in einem Endpunkt vor, wurde die Subgruppenanalyse für die entsprechende Population in Abschnitt 4.3.1.3.2.1 dargestellt und diskutiert.

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. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethode darstellen[8]. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache

[8] Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.

adjustierte indirekte Vergleiche[9] oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)[10] und Rücker (2012)[11] vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz 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 ist[12].

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades). Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen[13, 14, 15]

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

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*

[9] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50(6): 683-691.
[10] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.
[11] Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods 2012; 3(4): 312-324.
[12] Schöttker B, Lühmann D, Boulkhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.
[13] 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.
[14] Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study BMJ 2011; 343 :d4909
[15] Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Res Synth Methods 2013; 4(4): 291-323.

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten 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.

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)

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

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, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-11: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungsstudie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer ggf. Datenschnitt**	**Therapiearme**
UX023-CL301[a]	ja	ja	abgeschlossen	64 Wochen[b]	Intervention: Burosumab Aktive Kontrolle: Orales Phosphat und aktives Vitamin D
a: Die primäre Wirksamkeits- und Sicherheits-Analyse nach 40 Wochen ist abgeschlossen und wurde im ersten Verfahren (Verfahrensnummer 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 und wurden bereits durch den G-BA bewertet (Verfahrensnummer D-492). 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.					

Geben Sie an, welchen Stand die Information in Tabelle 4-11 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 02.11.2021.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-11 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-12: 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.

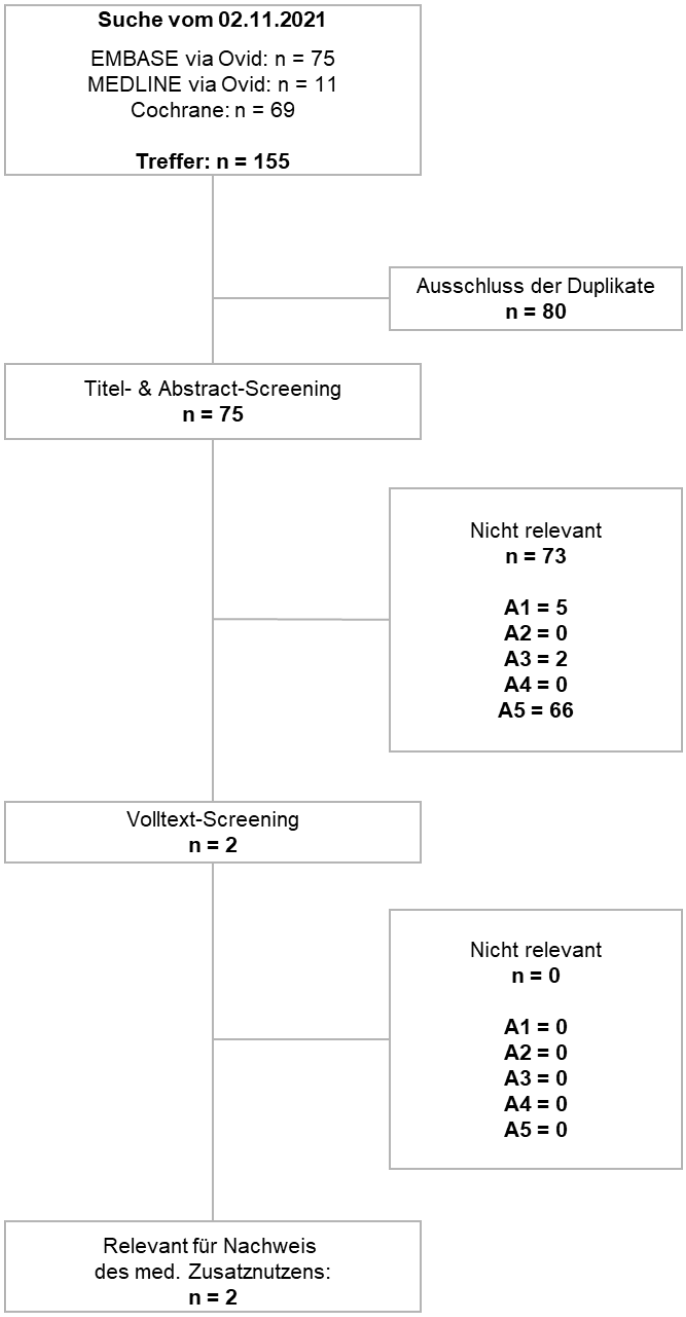

Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die Recherche in bibliographischen Datenbanken wurde am 02.11.2021 durchgeführt. Treffer in Studienregistern wurden dabei nicht berücksichtigt, da diese bereits über die Studienregistersuche abgedeckt sind. Insgesamt ergab diese Recherche 155 Treffer von denen 80 als Duplikate ausgeschlossen wurden. Die verbleibenden 75 Treffer wurden anhand von *Titel* und *Abstract* mit den in Abschnitt 4.2.2 genannten Kriterien bewertet. 73 konnten so ausgeschlossen werden. Zwei Publikationen wurden als relevant für die Neubewertung des medizinischen Zusatznutzens von Burosumab nach Überschreitung der Umsatzgrenze von 50 Millionen Euro erachtet und eingeschlossen. Beide Publikationen beziehen sich auf die pivotale Phase-III-Studie UX023-CL301 mit Datenschnitt zu Woche 64 und dienen somit als

Evidenz für das vorliegende Anwendungsgebiet (*Burosumab versus conventional therapy in children with X-linked hypophosphataemia: a randomised, active-controlled, open-label, phase 3 trial* (Lancet; 393(10189):2416-27), *Patient-Reported Outcomes from a Randomized, Active-Controlled, Open-Label, Phase 3 Trial of Burosumab Versus Conventional Therapy in Children with X-Linked Hypophosphatemia* (Calcified tissue international; 108(5):622-33) [66, 67].

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank 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 4-11) 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 4-13: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	**Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank 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	Clinicaltrials.gov NCT02915705 [68] EU Clinical Trials Registry 2016-000600-29 [69] WHO ICTRP EUCTR2016-000600-29 [70] Arzneimittel-Informationssystem (AMIce) [71] EMA *clinical data* [b]	ja	ja	abgeschlossen
a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: Die Suche unter EMA clinical data am 02.11.2021 ergab keine Treffer.				

Geben Sie an, welchen Stand die Information in Tabelle 4-13 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: 02.11.2021

4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-11) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

[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 4-14: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	**Relevante Quellen**[a]	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Studie durch Suche in Studienregistern / Studienergebnisdatenbanken identifiziert (ja/nein)**
UX023-CL301[b]	Burosumab (Crysvita®) X-chromosomale Hypophosphatämie ≥ 1 bis < 18 Jahre (Vorgangsnummer 2018-04-15-D-349) [17, 72-74]	ja	ja	ja
	Burosumab (Crysvita®) Neubewertung nach Fristablauf: X-chromosomale Hypophosphatämie ≥ 1 bis < 18 Jahre (Vorgangsnummer 2019-10-01-D-492) [65, 75-77]	ja	ja	ja
a: Quellen aus der Suche auf der Internetseite des G-BA b: Burosumab wurde vor Überschreitung der 50 Millionen Euro Umsatzgrenze bereits vom G-BA mit den gleichen Daten Nutzen bewertet.				

Geben Sie an, welchen Stand die Information in Tabelle 4-14 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.

Die Suche erfolgte am 05.11.2021. Bei der identifizierten Studie handelt es sich um die Studie UX023-CL301 die bereits nutzenbewertet wurde und im vorliegenden Dossier für die Wiedereinreichung nach Überschreitung der Umsatzgrenze von 50 Millionen Euro erneut für die Bewertung des Zusatznutzens von Burosumab herangezogen wird.

4.3.1.1.5 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, 4.3.1.1.3 und 4.3.1.1.4) 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 aller verfügbaren Quellen in Abschnitt 4.3.1.4 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 4-15: 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)**	**Studienberichte Studienprotokoll SAP (ja/nein [Zitat])**	**Register-einträge[c] (ja/nein [Zitat])**	**Publikation und sonstige Quellen[d] (ja/nein [Zitat])**
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebo-kontrolliert						
Nicht zutreffend.						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
UX023-CL301	ja	ja	nein	ja [25-27, 78]	ja [68-71]	ja [17, 65-67, 72-77]

a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden.

b: Studie, für die der Unternehmer Sponsor war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.

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. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-16: 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/ Datenschnitte** <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	Intervention: Burosumab subkutan alle 2 Wochen Initialdosis 29 Patienten Kontrolle: Orales Phosphat und aktives Vitamin D Mehrmalig tägliche Gabe 32 Patienten	64 Wochen gefolgt von einer freiwilligen Extensionsphase bis zu Woche 140	16 Studienzentren: USA (5), Kanada (3), Japan (3), Großbritannien (2), Schweden (1), Australien (1) und Korea (1) Studienbeginn: 09/2016 Ende Primär-Analyse: 02/2018 Studienende: 07/2019	**Primärer Endpunkt:** Evaluation der Wirksamkeit von Burosumab auf die Rachitis-Schwere bei Kindern mit XLH anhand des RGI-C Total Scores **Sekundäre Endpunkte:** • Serumphosphat-Konzentration, $1,25(OH)_2$-Vitamin D_3, ALP, TmP/GFR • 6MWT • Körpergröße • Schmerz-erhebung: PROMIS Schmerz, Mobilität und Fatigue und FPS-R **Explorative Endpunkte:**

	• Kinder-spezifischer Gesundheits-fragebogen *Short Form 10* (SF-10) • Zahn-gesundheits-bezogene Befragung **Sicherheitsend-punkte:** • UE • Schwerwiegende UE • Unerwünschte Ereignisse von besonderem Interesse

Tabelle 4-17: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	**Orales Phosphat**	**Burosumab**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
UX023-CL301	**Behandlungsphase (Woche 0 – 64):** Mehrere tägliche Dosen von oralem Phosphat und aktivem Vitamin D. Aufgrund der Variabilität der Dosen und Einnahmezeiten erfolgt eine patienten-individuelle 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	**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.	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. Nicht erlaubt waren folgende Medikamente in beiden Behandlungsarmen: Alumiumhydroxy-Antazida, systemisches Kortison, Acetazolamid, Thiazide, Wachstumshormone, PTH Suppressoren, monoklonale Antikörper-Therapien

Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie	**UX023-CL301**	
Gruppe **N (FAS)**	**Aktive Kontrolle** **(N=32)**	**Burosumab** **(N=29)**
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)
Ethnie I (Asiaten) n (%)	6 (18,8)	2 (6,9)
Ethnie II (Weiße) n (%)	25 (78,1)	25 (86,2)
Ethnie III (Anders) 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 (8,91)	19,59 (8,98)
BMI [kg/m^2][a] MW (SD)	18,22 (2,16)	18,00 (2,45)
Größe [Z-Score][a] MW (SD)	-2,05 (0,87)	-2,32 (1,17)
Krankheitsbezogene Baseline-Charakteristika		
RSS Total Score, MW (SD)	3,19 (1,14)	3,17 (0,98)
Serumphosphat [mg/dl], MW (SD)	2,30 (0,26)	2,42 (0,24)
Serum 1,25(OH)$_2$-Vitamin D$_3$ [pg/ml], MW (SD)	40,18 (14,89)	46,00 (20,06)
Alkalische Phosphatase [U/L], MW (SD)	523,4 (154,42)	510,8 (124,90)
Substitutionstherapie [b] (SOC) erhalten, n (%)	32 (100)	29 (100)
Alter bei Initiierung der SOC [Jahre], MW (SD)	2,07 (2,03)	2,61 (2,56)
Dauer der konventionellen Therapie [Jahre], MW (SD)	4,31 (3,03)	3,32 (3,12)
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)

Studie	**UX023-CL301**	
Gruppe **N (FAS)**	**Aktive Kontrolle** **(N=32)**	**Burosumab** **(N=29)**
Region; n (%)		
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 Baseline-Werte der Burosumab-Gruppe beziehen sich auf N = 28
b: Die Substitutionstherapie besteht aus einer oralen Gabe von Phosphat und aktivem Vitamin D

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. 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.

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, einer Wash-out-Phase, einer 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.

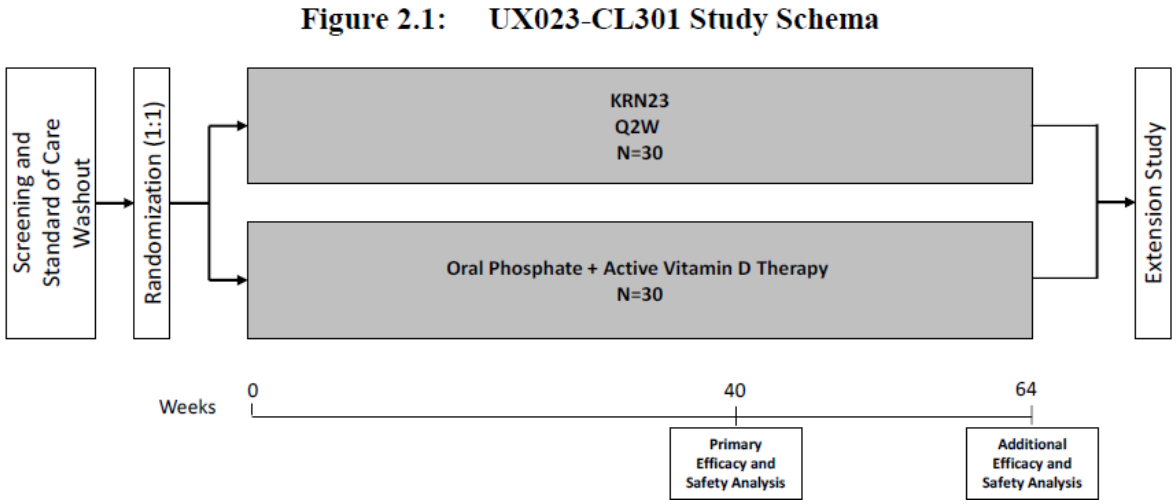

Abbildung 2: Studiendesign der UX023-CL301-Studie

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.

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 [26].

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 – 60 mg/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. 13 der 15 Kinder mit einer Vitamin-D-Dosiserhöhung hatten eine Dosiserhöhung von über 20 % zu Baseline bei einer zweimal täglichen Einnahme [26].

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 % regelkonform 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 Kindern, 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 [26].

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 4-18). Typische biochemische Parameter wie Serumphosphat, Kreatinin und 1,25(OH)2-Vitamin D3 mussten über-/ oder unterhalb der altersspezifischen Normgrenze sein.

Baseline-Charakteristika

Die Patienten wurden nach einem Alter unter und über 5 Jahren stratifiziert, wobei 12 Patienten der aktiven Kontrollgruppe und 14 Patienten 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. Dies 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 als übertragbar auf den deutschen Versorgungskontext 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 4-19: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	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 *RSS Total Scores* und *RGI-C-Scores* 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 (FAS ≥ 5 Jahre). 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 4-20: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Endpunktkategorie	**Endpunkt**	**UX023-CL301**
Mortalität	Im Rahmen der Sicherheit erhoben	- [a]
Morbidität	*Radiographic Global Impression of Change* (RGI-C)	X
	RGI-C-Responder	X
	Rickets Severity Score (RSS)	X
	Körpergröße: Stehhöhe/Liegelänge [Z-Score], Sitzhöhe (cm)	X
	6-Minuten-Gehdistanz Test (6MWT) [m und Prozentsatz]	X
	Serumphosphat [mg/dl]	X
	Patient-Reported Outcomes Measurement Information System (PROMIS): *Pediatric Pain Interference*, *Physical Function Mobility* und *Fatigue Domain Score*	X
	Faces Pain Scale – Revised (FPS-R)	X
Gesundheitsbezogene Lebensqualität	*Short Form 10* (SF-10)	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 inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
– 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.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]).

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,

2. Gesamtrate schwerwiegender UE (SUE),

3. Gesamtrate der Abbrüche wegen UE,

4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,

5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind

- Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind

- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher

Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

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 4-21: 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. Die Erhebung während der aktivkontrollierten Behandlungsphase der Studie erfolgte zu Woche 40 und Woche 64. <u>Analysepopulation:</u> FAS <u>Statistisches Modell:</u> • Hauptanalyse: Auswertungen zur mittleren Änderung des *RGI-C Total Scores* wurden zu Woche 40 mit einer ANCOVA und zu Woche 64 mit einem GEE-Modell bestimmt. Bei ANCOVA wurden die Behandlungsgruppe, die Baseline-Alters-Stratifizierungsfaktoren als Faktoren und der Baseline RSS Total Score als kontinuierliche Kovariate berücksichtigt. Die abhängige Variable war der RGI-C. Bei dem GEE-Modell wurden dagegen die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren aufgenommen, sowie der Baseline RSS Total Score als kontinuierliche Kovariate. Die abhängige Variable war der RGI-C. • Responderanalyse: Der Anteil an Studienteilnehmern mit einem *RGI-C* ≥ +2,0 wurde anhand eines logistischen Regressionsmodells analysiert (Woche 40). Hierfür wurde das *Odds Ratio* (OR) inklusive 95% KI und p-Werten bestimmt. Die Baseline-Alters-Stratifizierungsfaktoren und die Behandlungsgruppe wurden als Faktoren in das Modell aufgenommen, der RSS Total Score dagegen als kontinuierliche Kovariate. Zu Woche 64 erfolgte die Auswertung mittels GLMM. Die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren wurden als Faktoren berücksichtigt und der Baseline RSS Total Score als kontinuierliche Kovariate. Patientenspezifische Effekte für die Modellkonstante wurden als zufällige Effekte in das Modell aufgenommen. Das adjustierte RR (Adjustierung analog zum OR) und die dazugehörigen Konfidenzintervalle wurden *post-hoc* nach der Methode von Zhang und Yu bestimmt (siehe Abschnitt 4.2.5.2) [64]. Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die

	zugehörigen 95% KI wurden mittels Normalverteilungsapproximation und der p-Wert mittels Fishers exaktem Test berechnet. Darstellung im Dossier: • Hauptanalyse: Differenz der mittleren Änderung von Baseline zu Woche 40 und Woche 64 zwischen den Behandlungsgruppen (*RGI-C Total Score*). • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier anhand von Hedges`g geprüft. • Responderanalyse: Anteil an Patienten mit einem *RGI-C* von ≥ +2,0.
Rickets Severity Score (RSS)	
UX023-CL301	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. 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. 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. Die Erhebung während der aktivkontrollierten Behandlungsphase der Studie erfolgte zu Woche 40 und Woche 64. Analysepopulation: FAS Statistisches Modell: • Auswertungen zur mittleren Änderung des *RSS Total Scores* wurden zu Woche 40 mit einer ANCOVA und zu Woche 64 mit einem GEE-Modell bestimmt. Bei ANCOVA wurden die Behandlungsgruppe und die Baseline-Alters-Stratifizierungsfaktoren als unabhängige Variablen und der Baseline-Wert des RSS Total Scores als kontinuierliche Kovariate berücksichtigt. Bei dem GEE-Modell wurden dagegen die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren aufgenommen, der Baseline RSS Total Score dagegen als kontinuierliche Kovariate. Die abhängige Variable war die Änderung zu Baseline im RSS Total Score. Darstellung im Dossier: • Hauptanalyse: Differenz der mittleren Änderung von Baseline zu Woche 40 und Woche 64 zwischen den Behandlungsgruppen (*RSS Total Score*). • Sensitivitätsanalyse LOCF: Von einem Kind wurde der RSS-Kniewert zu Woche 40 nicht bewertet, da die Oberschenkelwachstumsplatten bereits verschmolzen waren und die RSS-Auswertung nicht mehr durchgeführt werden konnte. Der Baseline-Kniewert wurde imputiert, um den Kniewert der Woche 40 zu ersetzen und somit den Gesamtscore der Woche 40 ableiten zu können. • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird

im vorliegenden Dossier anhand des Hedges`g geprüft.

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 4-22: Bewertung des Verzerrungspotenzials für *Rachitissymptomatik* in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunktherheber**	**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.

Tabelle 4-23: Ergebnisse für *Radiographic Global Impression of Change (RGI-C) Total Score* aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Radiographic Global Impression of Change (RGI-C) Total Score**					
(N)	**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 H'g = 1,846 [1,246; 2,446]	1,02 [0,72; 1,33] < 0,0001 H'g = 1,600 [1,023; 2,178]
Burosumab (N=29)	29 1,92 (0,080)	29 1,92 (0,110)	29 2,06 (0,074)	29 2,06 (0,072)		

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

a: Berechnung mit ANCOVA-Modell zu Woche 40 und *generalized estimation equation* (GEE)-Modell zu Woche 64.

b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

ANCOVA: Behandlungsgruppe, Baseline-Alters-Stratifizierungsfaktoren als Faktoren und Baseline RSS Total Score als kontinuierliche Kovariate
GEE: Behandlungsgruppe, Behandlungsvisiten und Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie Baseline-Alters-Stratifizierungsfaktoren als Faktoren, Baseline RSS Total Score 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). Der Trend setzte sich zu Woche 64 weiter fort. Während die Burosumab-Gruppe eine Verbesserung der Rachitissymptomatik um +2,06 (LsMean) aufwies, erreichte die aktive Kontrollgruppe dagegen lediglich eine Verbesserung von +1,03 (LsMean). Die Behandlungsdifferenz von +1,02 zu Woche 64 war statistisch signifikant und klinisch relevant ($p < 0,0001$; H’g = 1,600 [1,023; 2,178]; GEE-Modell) und unterstützt die Überlegenheit von Burosumab, die sich bereits in Woche 40 deutlich abzeichnete (Tabelle 4-23).

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

Tabelle 4-24: 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)	**Patienten n (%)**	**OR[a] [95% KI]**	**ARR[b] [95% KI]**	**p-Wert[a]**	**RR[c] [95% KI]**	**ARD[c] [95% KI]**	**p-Wert[d]**
Anteil der Patienten mit einem RGI-C ≥ +2,0 zu Woche 40							
Aktive Kontrolle[e] (32)	2 (6,3)	39,139 [7,238; 211,656]	11,50 [5,20; 14,83]	< 0,0001	11,586 [2,972; 45,168]	0,662 [0,479; 0,845]	< 0,0001
Burosumab (29)	21 (72,4)						
Anteil der Patienten mit einem RGI-C ≥ +2,0 zu Woche 64							
Aktive Kontrolle[e] (32)	6 (18,8)	34,075 [5,627; 206,340]	4,72 [3,01; 5,21]	0,0002	4,598 [2,203; 9,596]	0,675 [0,490; 0,859]	< 0,0001
Burosumab (29)	25 (86,2)						

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Responder; %: Anteil von N; RSS: *Rickets Severity Score*; [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARR: adjustiertes *Risk Ratio*; ARD: absolute Risikodifferenz

RGI-C-Responder ist definiert als die Patienten mit einem RGI-C ≥ +2,0; ein ARR, RR bzw. OR > 1 bedeutet ein Vorteil für Burosumab und ein ARR, RR bzw. OR < 1 ein Vorteil für Phosphat.

a: Das logistische Regressionsmodell berücksichtigt die Baseline-Alters-Stratifizierungsfaktoren und die Behandlungsgruppe als Faktoren und den RSS Total Score als kontinuierliche Kovariate. Die geschätzte Responder-Rate, das 95 %-KI und der zweiseitiger p-Wert stammen aus dem logistischen Regressionsmodell. Zu Woche 64 wurde ein GLMM verwendet, dass die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren berücksichtigt und den Baseline RSS Total Score als kontinuierliche Kovariate. Zusätzlich wurden patientenspezifische Effekte für die Modellkonstante berücksichtigt.

b: Das adjustierte RR und die dazugehörigen Konfidenzintervalle wurden *post-hoc* nach der Methode von Zhang und Yu bestimmt (adjustiert analog wie das OR) (siehe Abschnitt 4.2.5.2) [64].

c: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet (*post-hoc*). Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet.

d: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

e: Aktive Kontrolle: orales Phosphat und aktives Vitamin D

Eine erhebliche Verbesserung der „Heilung“ der Rachitissymptomatik (≥ +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 % der Patienten in der Burosumab-Gruppe und bei 6,3 % in der aktiven Kontrollgruppe beobachtet (Relatives Risiko (RR) [95% KI]: 11,586 [2,972; 45,168], p < 0,0001; logistisches Regressionsmodell). Nach 64 Wochen konnte bei 86,2 % der Patienten der Burosumab-Gruppe eine „Heilung“ der Rachitissymptomatik festgestellt werden, bei der aktiven Kontrollgruppe betrug der Anteil dagegen nur 18,8 %. Dies entspricht einem Behandlungsunterschied von 67,4 Prozentpunkten. Das RR betrug 4,598 [2,203; 9,596] und war mit p < 0,0001 statistisch signifikant (Tabelle 4-24).

Tabelle 4-25: 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 **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 3,19 (1,141)	32 2,47 (1,092)	32 2,17 (0,947)	32 -0,72 (0,182) -0,71 (0,138)	32 -1,02 (0,192) -1,02 (0,151)	-1,34 [-1,74; -0,94] < 0,0001 H'g = -1,399 [-1,964; -0,833]	-1,21 [-1,59; -0,83] < 0,0001 H'g = -1,119 [-1,710; -0,529]
Burosumab (N=29)	29 3,17 (0,975)	28 1,13 (0,715)	29 0,95 (0,724)	28 -2,04 (0,143) -2,04 (0,145)	29 -2,22 (0,166) -2,23 (0,117)		

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 ANCOVA-Modell zu Woche 40 und *generalized estimation equation* (GEE)-Modell zu Woche 64

b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

ANCOVA: Behandlungsgruppe und Baseline-Alters-Stratifizierungsfaktoren als unabhängige Variablen, Baseline-Wert des RSS Total Scores als kontinuierliche Kovariate
GEE: Behandlungsgruppe, Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren, Baseline RSS Total Score als kontinuierliche Kovariate

Die Abnahme des Schweregrads der Rachitissymptomatik in beiden Behandlungsgruppen wird durch die Veränderung des RSS Total Scores 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 ($LsMean_{40}$: -1,34, $p < 0,0001$, ANCOVA-Modell; $LsMean_{64}$: -1,21 $p < 0,0001$, GEE-Modell) (Tabelle 4-25).

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 hatte nur die Hälfte (16/32) der Kinder in der aktiven Kontrollgruppe eine Verbesserung der Rachitis-Schwere ≥ 1 erlangt.

Tabelle 4-26: Ergebnisse der LOCF-Analyse für *Rickets Severity Score (RSS)* – Total Score aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Rickets Severity Score (RSS) - Total Score**						
	LOCF						
(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 **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 3,19 (1,141)	32 2,47 (1,092)	32 2,17 (0,947)	32 -0,72 (0,182) -0,71 (0,137)	32 -1,02 (0,192) -1,02 (0,151)	-1,35 [-1,74; -0,95] < 0,0001 H'g = -1,399 [-1,964; -0,833]	-1,21 [-1,59; -0,83] < 0,0001 H'g = -1,119 [-1,710; -0,529]
Burosumab (N=29)	29 3,17 (0,975)	29 1,12 (0,703)	29 0,95 (0,724)	29 -2,05 (0,139) -2,05 (0,141)	29 -2,22 (0,166) -2,23 (0,117)		

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 (wurde *post-hoc* berechnet); LOCF*: Last Observation Carried Forward*

a: Berechnung mit ANCOVA-Modell zu Woche 40 und *generalized estimation equation* (GEE)-Modell zu Woche 64

b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

ANCOVA: Behandlungsgruppe, Baseline-Alter-Stratifizierungsfaktoren als Faktoren und Baseline RSS als kontinuierliche Kovariate

GEE: Behandlungsgruppe, Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren, Baseline RSS Total Score als kontinuierliche Kovariate

Sensitivitätsanalyse des Endpunktes RSS (LOCF)

Eine Sensitivitätsanalyse wurde durchgeführt, um die Robustheit des primären Analyseergebnisses gegenüber fehlenden Daten zu bewerten. Von einem Kind fehlten der Knie- und somit der Gesamtscore des RSS zu Woche 40. Der RSS-Kniewert zu Woche 40 wurde nicht bewertet, da die Oberschenkelwachstumsplatten bereits verschmolzen waren und die RSS-Auswertung nicht mehr durchgeführt werden konnte. Der Baseline-Kniewert wurde imputiert, um den Kniewert der Woche 40 zu ersetzen und somit den Gesamtscore der Woche 40 ableiten zu können. Die Ergebnisse dieser LOCF-Analyse waren denen der Primäranalyse sehr ähnlich (Tabelle 4-26).

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.

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 und Sitzhöhe) – RCT

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

Tabelle 4-27: Operationalisierung von *Körpergröße (Stehhöhe/Liegelänge und Sitzhöhe)*

Studie	Operationalisierung
UX023-CL301	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. 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. Das Wachstum wurde unter Verwendung von Z-Scores bewertet, die die Anzahl an Standardabweichungen eines jeden Wertes von den nach Alter und Geschlecht standardisierten normalen Durchschnittswerten widerspiegeln. Als Grundlage diente eine repräsentative Stichprobe gesunder Kinder aus den USA (Wachstumscharts des *Centers for Disease Control*). Zusätzlich wurde bei Patienten ab 2 Jahren die *absolute Sitzhöhe* in Zentimeter gemessen. Die Sitzhöhe wurde während der aktivkontrollierten Behandlungsphase analog zur Stehhöhe zu Baseline, Woche 24, 40 und 64 erhoben. Um robuste Ergebnisse zu erhalten, wurden ebenfalls Dreifachmessungen durchgeführt und daraus der Mittelwert gebildet. Eine Transformation in z-Scores erfolgte nicht. <u>Analysepopulation:</u> FAS <u>Statistisches Modell:</u> • Auswertungen zur mittleren Änderung der *Körpergröße/Sitzhöhe* wurden zu Woche 40 und zu Woche 64 mit einem GEE-Modell bestimmt. Die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren wurden als Faktoren aufgenommen, das Alter und die Baseline Körpergröße/Sitzhöhe als kontinuierliche Kovariate. Die abhängige Variable war die Änderung zu Baseline in der Körpergröße/Sitzhöhe. <u>Darstellung im Dossier:</u> • Differenz der mittleren Änderung von Baseline zu Woche 40 und Woche 64 zwischen den Behandlungsgruppen • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier anhand des Hedges`g geprüft

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 4-28: Bewertung des Verzerrungspotenzials für *Körpergröße* in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunktherheber**	**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.

Tabelle 4-29: Ergebnisse für *Körpergröße (Stehhöhe/Liegelange)* aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Körpergröße (Stehhöhe/Liegelänge) [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 **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 -2,05 (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] 0,0408 H'g = 0,537 [0,021; 1,053]	0,14 [0,00; 0,29] 0,0490 H'g = 0,490 [-0,025; 1,005]
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)		

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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, das Alter und die Baseline Stehhöhe/Liegelänge als kontinuierliche Kovariate

Die XLH ist eine schwere, chronische Erkrankung des Knochenskeletts, die durch Kleinwuchs und Beindeformationen 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/Liegelänge 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 auch nur einen geringen Einfluss auf 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/Liegelänge 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/Liegelänge 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/Liegelänge 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 (Tabelle 4-29).

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 [33]. 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 [26]. Weiterhin bestätigt die vergleichende Studie von *Rosario et al.*, dass die Wachstumscharts der CDC für den Vergleich mit deutschen Wachstumskurven geeignet sind [36]. 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.

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

Intervention	**Sitzhöhe**						
(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 **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)	29 62,96 (8,605)	28 64,08 (9,769)	29 65,26 (9,555)	28 2,52 (0,305) 3,09 (0,428)	29 3,73 (0,315) 4,27 (0,449)	-0,79 [-1,98; 0,39] 0,1901 -	0,08 [-1,58; 1,75] 0,9213 -
Burosumab (N=29)	26 59,33 (10,365)	24 62,43 (9,004)	26 63,56 (9,379)	24 3,13 (0,912) 2,29 (0,412)	26 5,09 (0,999) 4,36 (0,721)		

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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, das Alter und die Baseline Sitzhöhe als kontinuierliche Kovariate

Die absolute Sitzhöhe in Zentimetern unterschied sich zu Baseline zwischen den Behandlungsgruppen (Burosumab: MW: 59,33 cm; aktive Kontrollgruppe: MW: 62,96 cm). Zu Woche 40 betrug die adjustierte mittlere Änderung in der Burosumab-Gruppe 2,29 cm und in der aktiven Kontrollgruppe 3,09 cm. Nach 64 Wochen konnte eine ähnliche Zunahme der Sitzhöhe bei beiden Behandlungsgruppen festgestellt werden. Die Behandlungsdifferenz von 0,08 cm war mit p = 0,9213 nicht statistisch signifikant (Tabelle 4-30).

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.

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 4-31: Operationalisierung von *Gehfähigkeit*

Studie	Operationalisierung
UX023-CL301	Der 6-Minuten-Gehdistanz Test (6MWT) wurde von einem geschulten klinischen Personal in Übereinstimmung mit allgemeinen Prinzipien durchgeführt, die in den Richtlinien der *American Thoracic Society* (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 [39]. 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 sechs 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 aus, die durch den Verzicht auf ein Rad entstanden sein könnte. Die *Gehfähigkeit* wurde während der aktivkontrollierten Behandlungsphase zu Baseline, zu Woche 24, 40 und 64 erhoben. <u>Analysepopulation:</u> FAS (Baseline Alter ≥ 5 Jahre) <u>Statistisches Modell:</u> • Auswertungen zur mittleren Änderung der *Gehfähigkeit* wurden zu Woche 40 und zu Woche 64 mit einem GEE-Modell bestimmt. Bei dem GEE-Modell wurden die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren als Faktoren berücksichtigt, sowie der Baseline 6MWT als Kovariate. Die abhängige Variable war die Änderung zu Baseline im 6MWT. <u>Darstellung im Dossier:</u> • Differenz der mittleren Änderung von Baseline zu Woche 40 und Woche 64 zwischen den Behandlungsgruppen (Gesamtrate in Meter und Prozentsatz) • Hauptanalyse: Die Analyse beinhaltet zwei Patienten der Burosumab-Gruppe, die zu Baseline (1:14 min) bzw. zu Woche 40 (5:27 min) nicht in der Lage waren, die sechs Minuten vollständig zu gehen. • Sensitivitätsanalyse: Bei dieser Analyse wurden die zwei Patienten, die zu Baseline bzw. zu Woche 40 nicht in der Lage waren, die sechs Minuten vollständig zu gehen, nicht berücksichtigt, da diese einen unvollständigen Datensatz aufwiesen. • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier mittels Hedges`g geprüft

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 *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 ≥ 5 Jahre ist definiert als alle Patienten, die randomisiert wurden, ein Mindestalter von 5 Jahre zur Screeningvisite aufwiesen und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten im 6MWT mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS ≥ 5 Jahre alle randomisierten Patienten. Daher wird die FAS ≥ 5 Jahre 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.

Tabelle 4-33: Ergebnisse für *Gehfähigkeit* aus RCT mit dem zu bewertenden Arzneimittel

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**	**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[b]							
Aktive Kontrolle[c] (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[b] (%)							
Aktive Kontrolle[c] (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)		

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**	**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]**
Korrigierte Analyse[d]							
Gesamtrate – Messung in Meter							
Aktive Kontrolle[c] (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] 0,0650 -	45,55 [2,09; 89,02] 0,0399 H'g = 0,679 [-0,038; 1,396]
Burosumab (N=13)	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)		
Gesamtrate – Prozentsatz (%)							
Aktive Kontrolle[c] (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] 0,0775 -	7,27 [0,01; 14,52] 0,0496 H'g = 0,703 [-0,015; 1,422]
Burosumab (N=13)	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)		

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: Bei dieser Analyse wurden alle Patienten berücksichtigt, die mindestens eine zugeteilte Medikation erhalten haben und mindestens 5 Jahre alt waren (FAS). Die Analyse beinhaltet zwei Patienten der Burosumab-Gruppe, die zu Baseline (1:14 min) bzw. zu Woche 40 (5:27 min) nicht in der Lage waren, die 6 Minuten vollständig zu gehen.

c: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

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**	**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]**

d: Bei der korrigierten Analyse wurden die zwei Patienten, die zu Baseline bzw. zu Woche 40 nicht in der Lage waren, die 6 Minuten vollständig zu gehen, nicht berücksichtigt, da diese einen unvollständigen Datensatz aufwiesen.

GEE: Behandlungsgruppe, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, Baseline 6MWT als Kovariate

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 [79]) innerhalb der 6 Minuten ermittelt.

6MWT absolute Gehdistanz (korrigierte Analyse)

Die Burosumab-Gruppe (MW: 385,00 m) wies zu Baseline eine stärkere Einschränkung in der Gehfähigkeit als die aktive Kontrollgruppe (MW: 450 m) auf. Hierbei ist jedoch anzumerken, dass der 6MWT kein Stratifizierungsfaktor war und die Burosumab-Gruppe im Vergleich zur Kontrollgruppe im Durchschnitt 6 Monate jünger war, sodass Unterschiede in der Gehfähigkeit auf den Entwicklungsstatus zurückzuführen sein können. Um die Robustheit der Analysen zu erhöhen und um den Einfluss der unterschiedlichen Baseline-Werte in den beiden Behandlungsgruppen gering zu halten, wurden die Baseline-Werte der Studienteilnehmer als Kovariate mit in das statistische Regressionsmodell einbezogen, um für nicht kontrollierbare Variablen zu adjustieren.

Zwei pädiatrische Patienten der Burosumab-Gruppe (ein Kind zu Baseline und ein Kind zu Woche 40) waren nicht in der Lage, zu jedem Erhebungszeitpunkt die kompletten 6 Minuten zu gehen. (siehe auch Abschnitt 4.2.5.4). Eines der zwei Kinder absolvierte zu Baseline 54 Meter in 1:17 Minuten. Bereits zu Testbeginn berichtete dieses Kind von anhaltenden Knochenschmerzen und erklärte, dass es an diesem Tag nicht schnell laufen könne. Ein Testlauf wurde bei diesem Kind zur Screeningvisite nicht durchgeführt. Das zweite Kind lief zu Woche 40 eine Strecke von 440 Metern bevor es nach 5:27 Minuten anfragte den Test zu beenden. Aufgrund dieses unvollständigen Datensatzes wurden zwei Analysen durchgeführt: eine Analyse, die diese beiden Studienteilnehmer nicht berücksichtigte (korrigiert Analyse) und eine, die auf dem *Full Analysis Set* beruht und die beiden Studienteilnehmer einschloss.

Die Gehdistanz verbesserte sich zu Woche 40 in der Burosumab-Gruppe um durchschnittlich 50,31 m (LsMean: 43,68 m), in der aktiven Kontrollgruppe dagegen lediglich um 6,90 m (LsMean: 5,38 m). Dieser Trend setzte sich zu Woche 64 weiter fort. Während die Studienteilnehmer der Burosumab-Gruppe eine adjustierte mittlere Änderung von 74,83 m aufwiesen, erreichte die aktive Kontrolle eine Steigerung der Gehdistanz von lediglich 29,28 m (LsMean). Die Behandlungsdifferenz von 45,55 m war statistisch signifikant (p = 0,0399) und bestätigt den positiven Effekt von Burosumab auf die Gehfähigkeit von Kindern mit XLH (Tabelle 4-33). Validierte MIDs liegen bis zum heutigen Tag für XLH nicht vor. Daher wurde die klinische Relevanz anhand des Hedges`g bestimmt. Der statistisch signifikante Effekt zugunsten von Burosumab erreichte jedoch zu Woche 64 nicht die klinische Relevanzschwelle (H'g = 0,679 [-0,038; 1,396]) (Tabelle 4-33).

6MWT relative Gehdistanz (korrigierte Analyse)

Ähnliche Ergebnisse wurden in der Auswertung der relativen Gehdistanz erzielt. Nach 40 Wochen zeigte sich in der Burosumab-Gruppe eine Verbesserung der Gehfähigkeit, die sich bis zu Woche 64 weiter steigerte. Während in der Burosumab-Gruppe der relative Wert zu Woche 64 um 9,15 Prozentpunkte anstieg, war in der aktiven Kontrolle lediglich ein Anstieg um 1,88 Prozentpunkte zu verzeichnen. Die Behandlungsdifferenz von 7,27 Prozentpunkten war mit p = 0,0496 statistisch signifikant (Tabelle 4-33).

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.

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 4-34: 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. Die Serumphosphat-Konzentration wurde während der aktiv-kontrollierten Behandlungsphase zu Baseline, Woche 1, 2, 4, 8, 16, 24, 32, 40, 52 und Woche 64 gemessen. Analysepopulation: FAS Statistisches Modell: • Auswertungen zur mittleren Änderung der *Serumphosphat-Konzentration* wurden zu Woche 40 und zu Woche 64 mit einer ANCOVA bestimmt. Bei ANCOVA wurden die Behandlungsgruppe, die Baseline-RSS-Stratifizierungsfaktoren und das Baseline Alter als Faktoren berücksichtigt, die Baseline Serumphosphat-Konzentration als Kovariate. Die abhängige Variable war die Änderung zu Baseline der Serumphosphat-Konzentration. Darstellung im Dossier: • Differenz der mittleren Änderung von Baseline zu Woche 40 und 64 zwischen den Behandlungsgruppen • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier anhand von Hedges`g geprüft

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 4-35: 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.

Tabelle 4-36: Ergebnisse für *Serumphosphat* aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Serumphosphat [mg/dl]**						
(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 **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: *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 ANCOVA-Modell zu Woche 40 und zu Woche 64

b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

ANCOVA: Behandlungsgruppe, Baseline-RSS-Stratifizierungsfaktoren und Baseline Alter als Faktoren, Baseline Serumphosphat-Konzentration als 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 0,24 mg/dl zu Woche 64 gezeigt werden, hingegen konnte die Burosumab-Gruppe eine adjustierte Änderung von 0,98 mg/dl von Baseline zu Woche 64 erzielen.

Die Behandlungsdifferenz zu Woche 40 (LsMean: 0,77 mg/dl; $p < 0,0001$; H'g = 2,047 [1,427; 2,667]) und Woche 64 (LsMean: 0,74 mg/dl; $p < 0,0001$; H'g = 1,981 [1,368; 2,595]) war sowohl statistisch signifikant als auch klinisch relevant (Tabelle 4-36).

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.

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 4-37: Operationalisierung von *Schmerzerhebung* und *funktionelle Behinderung*

Studie	Operationalisierung
PROMIS Pain Interference, Physical Function und Fatigue Domain Score	
UX023-CL301	Die pädiatrische PROMIS-Fragebögen wurden von nationalen Gesundheitsinstituten entwickelt und verwenden 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. Es wurde ausschließlich die Version 2.0 des PROMIS-Fragebogens in der Studie UX023-CL301 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 [39]. 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. Der Fragebogen gliedert sich in die Domäne *Pediatric Pain Interference* mit 4 Items (Probleme beim Schlafen, Rennen, Gehen und Stehen aufgrund von Schmerzen), in die Domäne *Physical Function Mobility* mit 10 Items zu körperlichen Aktivitäten und Alltagsbewegungen sowie in die Domäne *Fatigue*, die sowohl Einschränkungen in sozialen, sportlichen und alltäglichen Aktivitäten berücksichtigt. Die Fragen wurden von den Studienteilnehmern (ab 8 Jahren) oder den Eltern (5 – 7 Jahre) anhand einer 5-Punkte-Likert-Skala bewertet. Die Antwortmöglichkeiten reichten hierbei von „nie“ bis „fast immer“. Der Bezugszeitraum war jeweils die letzte Woche. Um die jeweiligen Scores zu berechnen, wurden die Itemscores einer Domäne addiert und anschließend die Domänenscores in T-Scores transformiert. Der T-Score skaliert die Domänenscores in einen standardisierten Score mit einem Mittelwert von 50 und einer Standardabweichung von 10. Für negative Konzepte wie *Pain Interference* und *Fatigue* stellen höhere T-Scores eine stärkere Beeinträchtigung der Studienteilnehmer dar. Für die Domäne *Physical Function Mobility* korrelieren dagegen höhere T-Scores mit einer besseren körperlichen Funktion. Der Fragebogen wurde in der aktivkontrollierten Behandlungsphase zu Baseline, Woche 24, 40 und 64 erhoben. <u>Analysepopulation:</u> FAS (Baseline Alter ≥ 5 Jahre) <u>Statistisches Modell:</u> • Auswertungen zur mittleren Änderung des PROMIS wurden zu Woche 40 und zu Woche 64 mit einem GEE-Modell und einem MMRM-Modell bestimmt. Bei dem GEE-Modell bzw. MMRM-Modell wurden die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den

	Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren als Faktoren aufgenommen, sowie der Baseline-Wert des PROMIS als Kovariate. Die abhängige Variable war die Änderung von Baseline des PROMIS Scores. Darstellung im Dossier: • Hauptanalyse: Differenz der mittleren Änderung von Baseline zu Woche 40 und 64 zwischen den Behandlungsgruppen (GEE) • Sensitivitätsanalyse: Differenz der mittleren Änderung von Baseline zu Woche 40 und 64 zwischen den Behandlungsgruppen (GEE) getrennt nach Selbst- und Fremdeinschätzung • Sensitivitätsanalyse: Differenz der mittleren Änderung von Baseline zu Woche 40 und Woche 64 zwischen den Behandlungsgruppen ausgewertet mittels MMRM • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier anhand von Hedges`g geprüft
Face Pain Scale – Revised (FPS-R)	
UX023-CL301	Die überarbeitete Gesichtsschmerzskala (FPS-R) ist ein selbstberichtendes Maß zur Ermittlung des akuten Schmerzempfindens, das speziell für Kinder entwickelt wurde. 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 [39]. Es erfolgte die Befragung aller Patienten im Alter von 5 – 12 Jahren. Der Fragebogen wurde in der aktivkontrollierten Behandlungsphase zu Baseline, Woche 24, 40 und 64 erhoben. Analysepopulation: FAS (Baseline Alter ≥ 5 Jahre) Statistisches Modell: • Auswertungen zur mittleren Änderung des FPS-R wurden zu Woche 40 und zu Woche 64 mit einem GEE-Modell bestimmt. Bei dem GEE-Modell wurden die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren als Faktoren aufgenommen, sowie der Baseline-Wert des FPS-R als Kovariate. Die abhängige Variable war die Änderung von Baseline des FPS-R. Darstellung im Dossier: • Differenz der mittleren Änderung von Baseline zu Woche 40 und Woche 64 zwischen den Behandlungsgruppen (Gesamtrate in Meter und Prozentsatz) • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier anhand von Hedges`g geprüft

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 4-38: 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 ≥ 5 Jahre ist definiert als alle Patienten, die randomisiert wurden, ein Mindestalter von 5 Jahre zur Screeningvisite aufwiesen und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten der PROMIS-Erhebung mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS ≥ 5 Jahre alle randomisierten Patienten. Daher wird die FAS ≥ 5 Jahre als eine adäquate Umsetzung des ITT-Prinzips verstanden. Der Rücklauf der Fragebögen in der Studie belief sich zu allen Messzeitpunkten auf über 70 %.

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 4-39: Ergebnisse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel (GEE-Analyse)

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Pain Interference Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=20)	20 49,9 (12,05)	20 51,4 (10,57)	20 50,4 (9,51)	20 49,4 (9,52)	20 1,6 (1,52) 0,74 (1,235)	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 H'g = -0,697 [-1,390; -0,008]	-2,26 [-6,61; 2,09] 0,3091 -
Burosumab (N=15)	15 53,1 (10,95)	15 50,5 (9,17)	15 47,6 (9,84)	15 49,3 (8,07)	15 -2,7 (2,41) -2,39 (1,826)	15 -5,5 (1,78) -5,23 (1,707)	15 -3,8 (2,59) -3,55 (1,873)		
PROMIS *Pediatric Physical Function Mobility Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=20)	20 45,5 (9,86)	20 46,3 (9,69)	20 45,5 (9,71)	20 46,3 (9,63)	20 0,9 (0,81) 0,94 (0,808)	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 47,3 (8,18)	15 47,9 (8,32)	15 47,9 (9,24)	15 2,2 (1,84) 2,20 (1,625)	15 2,7 (1,49) 2,76 (1,338)	15 2,8 (1,80) 2,82 (1,648)		

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE) LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Fatigue Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=20)	20 47,0 (13,70)	20 46,8 (10,99)	20 46,6 (10,73)	20 45,0 (11,17)	20 -0,2 (1,68) -0,74 (1,188)	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 47,1 (9,55)	15 44,7 (10,49)	15 45,2 (10,69)	15 -1,6 (1,80) -1,69 (1,642)	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: *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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, Baseline-Wert des PROMIS als Kovariate

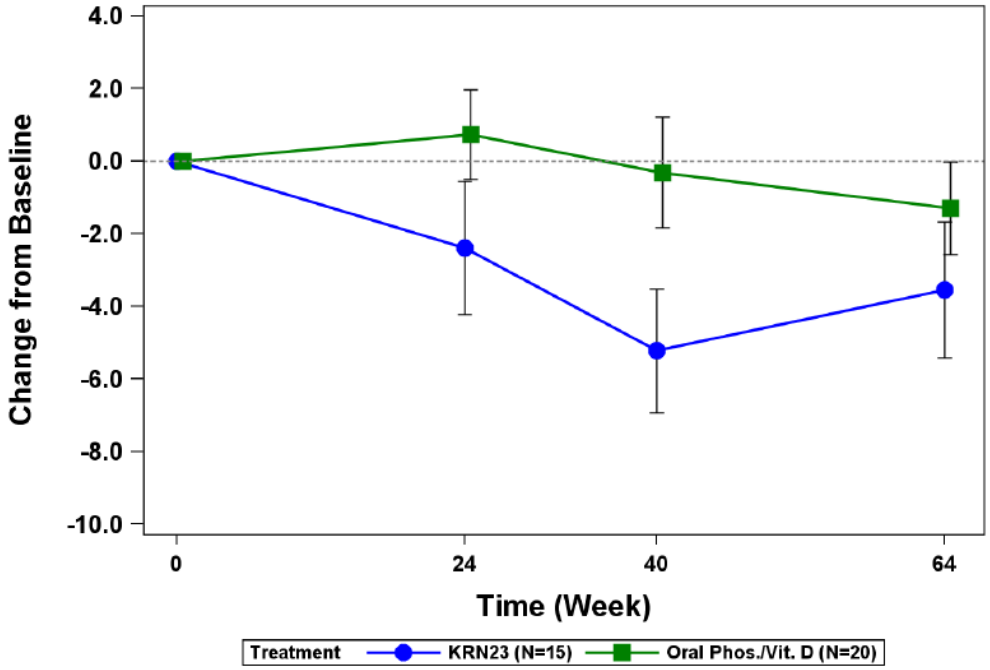

Abbildung 3: Veränderung zur Baseline der *Pediatric Pain Interference Domain* des PROMIS-Fragebogens

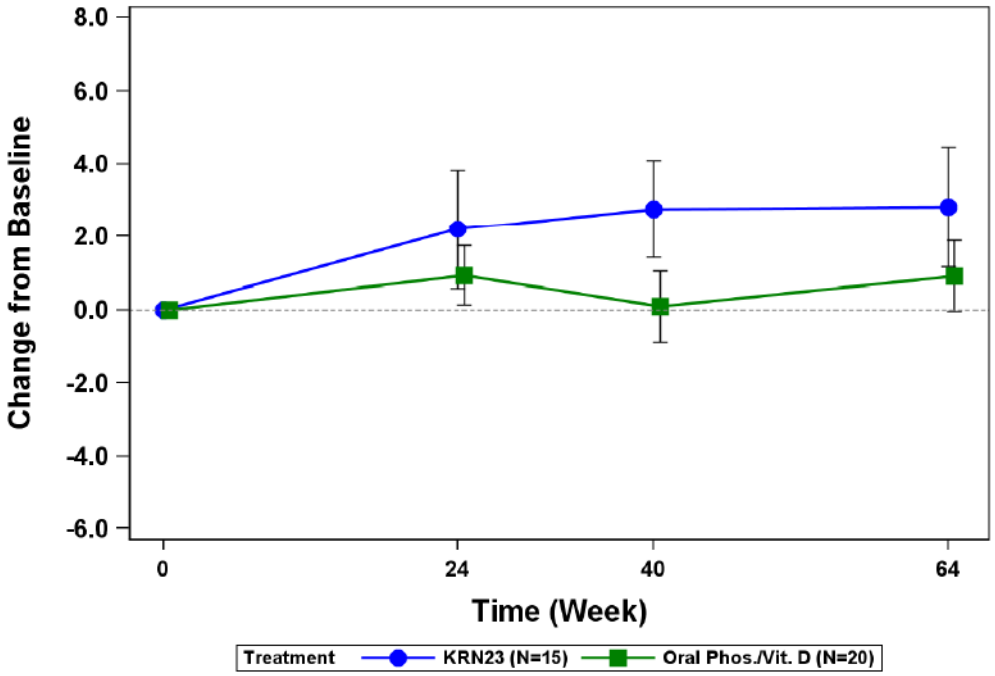

Abbildung 4: Veränderung zur Baseline der *Pediatric Physical Function Mobility Domain* des PROMIS-Fragebogens

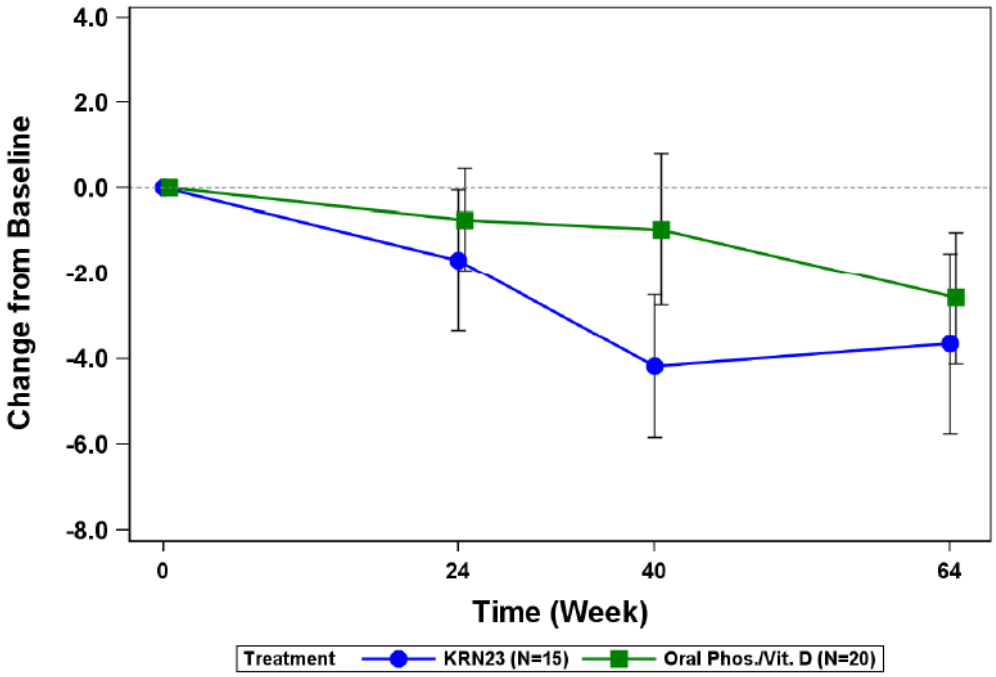

Abbildung 5: Veränderung zur Baseline der *Pediatric Fatigue Domain Score* des PROMIS-Fragebogens

Patient-Reported Outcomes Measurement Information System (PROMIS)

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. Laut der Fachexperten in der mündlichen Anhörung muss jedoch gesagt werden, dass die Schmerzsymptomatik vorwiegend adoleszente und junge Erwachsene betrifft und in ihrer Altersaktivität einschränkt. Kinder gehen, bevor die Schmerzen deutlich zu spüren sind, in eine Schonhaltung oder Pausieren während der Aktivitäten. Bei Ausflügen bzw. länger andauernden Aktivitäten wird dies sehr schnell ersichtlich. Eine weitere Erklärung hierfür ist, dass es durch die Beinfehlstellung zu einer Adaption der nachfolgenden Gelenke und später zu einer präarthrotischen Deformität kommt, die, je älter der Patient ist, erhebliche Schmerzen bereitet [13]. Im Umkehrschluss heißt dies, dass die Erhebung der Schmerzen bei Kindern im Alter von 5 – 12 Jahren aus den oben genannten Gründen noch keine gute Aussagekraft hat.

Hauptanalyse (Selbst- und Fremdeinschätzung)

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.

PROMIS Pediatric Pain Interference Domain Score

Zu Baseline wurde in der Burosumab-Gruppe (MW: 53,1) über mehr Schmerzen als in der aktiven Kontrollgruppe (MW: 49,9) berichtet. Dies änderte sich zu Woche 40. Während in der Burosumab-Gruppe zu Woche 40 eine Änderung zu Baseline von -5,23 (LsMean) beobachtet wurde, betrug die Änderung in der Kontrollgruppe lediglich -0,31 (LsMean). Der Unterschied von -4,91 zwischen den Behandlungsgruppen war statistisch signifikant (p = 0,0245; GEE-Modell). Die Abnahme von Schmerzen in der Burosumab-Gruppe (LsMean: -3,55) setzte sich auch zu Woche 64 weiter fort. Im Vergleich dazu, war nur eine geringe Abnahme von Schmerzen in der Kontrollgruppe (LsMean: -1,29) feststellbar. (Tabelle 4-39 und Abbildung 3). Die beobachtete Abnahme von Schmerzen in der Burosumab-Gruppe entsprach sowohl zu Woche 40 als auch zu Woche 64 einer MID von mindestens 3 Punkten, die Studienteilnehmer in der Kontrollgruppe erreichten dagegen keine klinisch bedeutsame Veränderung [67].

PROMIS Physical Function Mobility Domain Score

Die mittleren Werte der Mobilitätsdomäne (PROMIS *Physical Function Mobility Domain Score*) stiegen in der Burosumab-Gruppe von 45,2 (MW) zu Baseline auf 47,9 (MW) zu Woche 40 an und blieben zu Woche 64 unverändert. Die Werte der aktiven Kontrollgruppe blieben zu Woche 40 mit 45,5 (MW) unverändert und nahmen nur geringfügig zu Woche 64 zu (MW: 46,3). Der Unterschied zwischen den Behandlungsgruppen zu Woche 40 (Behandlungsdifferenz$_{W40}$: 2,68; p = 0,104) und Woche 64 (Behandlungsdifferenz$_{W64}$: 1,90; p = 0,3145) in der Mobilitätsdomäne war statistisch nicht signifikant (Tabelle 4-39 und Abbildung 4). Die Größe der minimal wichtigen Differenz war jedoch ähnlich der geschätzten MID.

PROMIS Fatigue Domain Score

In der Fatigue-Domäne (PROMIS *Fatigue Domain Score*) nahmen in der Burosumab-Gruppe die Werte von 48,8 (MW) zu Baseline auf 44,7 (MW) zu Woche 40 ab. Zu Woche 64 konnte eine Reduktion von 3,65 (LsMean) zu Baseline erzielt werden. Die aktive Kontrollgruppe konnte hingegen zu Woche 64 eine Reduktion von 2,57 (LsMean) auf 45,0 (MW) zeigen. Der Unterschied zwischen den Behandlungsgruppen zu Woche 40 (Behandlungsdifferenz$_{W40}$: -3,20; p = 0,1724) und Woche 64 (Behandlungsdifferenz$_{W64}$: -1,08; p = 0,6810) in der Fatigue-Domäne war zu keiner Zeit statistisch signifikant, jedoch war die Größe der Differenz ähnlich der geschätzten MID (Tabelle 4-39 und Abbildung 5).

Tabelle 4-40: Sensitivitätsanalyse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel unterteilt nach Alter (5 – 7 Jahre, Fremdeinschätzung)

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder 5 – 7 Jahre, Fragebogen durch Eltern ausgefüllt)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Pain Interference Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=9)	9 52,3 (12,28)	9 54,0 (11,82)	9 50,4 (10,39)	9 51,4 (10,14)	9 1,7 (1,77) 0,04 (1,540)	9 -1,8 (2,56) -3,51 (1,509)	9 -0,8 (1,73) -2,53 (1,304)	-1,15 [-7,75, 5,46] 0,7334 -	-1,52 [-7,56, 4,52] 0,6216 -
Burosumab (N=8)	8 55,9 (12,72)	8 53,5 (7,70)	8 50,5 (10,81)	7 51,6 (9,37)	8 -2,4 (4,21) -1,67 (2,434)	8 -5,4 (2,98) -4,66 (2,886)	7 -4,8 (4,94) -4,06 (2,650)		
PROMIS *Pediatric Physical Function Mobility Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=9)	9 41,9 (11,30)	9 42,7 (11,26)	9 42,9 (10,30)	9 43,0 (11,09)	9 0,8 (1,57) 1,15 (1,635)	9 1,1 (1,64) 1,43 (1,468)	9 1,1 (1,61) 1,46 (1,619)	1,52 [-3,26; 6,30] 0,5325 -	0,93 [-5,32; 7,17] 0,7711 -
Burosumab (N=8)	8 42,9 (9,51)	8 45,6 (8,15)	8 45,.6 (8,05)	7 43,4 (8,69)	8 2,7 (2,66) 2,97 (1,971)	8 2,7 (2,64) 2,95 (1,948)	7 2,6 (3,77) 2,39 (2,694)		

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder 5 – 7 Jahre, Fragebogen durch Eltern ausgefüllt)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Fatigue Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=9)	9 53,0 (16.18)	9 51,5 (13,43)	9 50,1 (12,62)	9 49,7 (13,01)	9 -1,6 (2,59) -1,98 (1,679)	9 -3,0 (2,80) -3,40 (1,728)	9 -3,3 (3,11) -3,76 (2,205)	-0,95 [-6,15, 4,26] 0,7220 -	-1,85 [-9,48; 5,77] 0,6338 -
Burosumab (N=8)	8 51,9 (10,68)	8 49,0 (10,64)	8 47,8 (12,26)	7 47,4 (13,47)	8 -2,9 (1,75) -3,13 (1,531)	8 -4,1 (1,76) -4,.35 (2,116)	7 -5,2 (3,68) -5,61 (3,224)		

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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, Baseline PROMIS als Kovariate

Tabelle 4-41: Sensitivitätsanalyse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel unterteilt nach Alter (8 – 12 Jahre, Selbsteinschätzung)

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder 8 – 12 Jahre)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Pain Interference Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=11)	11 47,9 (12,08)	11 49,3 (9,48)	11 50,3 (9,26)	11 47,7 (9,13)	11 1,4 (2,42) 1,04 (1,793)	11 2,4 (2,98) 2,03 (2,188)	11 -0,2 (2,61) -0,55 (1,886)	-7,76 [-13,10; -2,41] 0,0044 H'g = -0,853 [-1,84; 0,135]	-1,64 [-7,06; 3,79] 0,5542 -
Burosumab (N=8)	7 50,0 (8,33)	7 47,0 (10,04)	7 44,3 (8,12)	8 47,3 (6,70)	7 -3,0 (2,34) -3,06 (2,576)	7 -5,6 (1,97) -5,72 (2,138)	8 -3,0 (2,58) -2,18 (2,182)		
PROMIS *Pediatric Physical Function Mobility Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=11)	11 48,4 (7,86)	11 49,3 (7,41)	11 47,6 (9,15)	11 49,0 (7,72)	11 0,9 (0,80) 0,91 (0,825)	11 -0,8 (1,27) -0,86 (1,281)	11 0,7 (1,29) 0,63 (1,188)	3,22 [-0,09; 6,53] 0,0567 -	2,09 [-0,76; 4,94] 0,1499 -
Burosumab (N=8)	7 47,7 (8,44)	7 49,3 (8,41)	7 50,5 (8,44)	8 51,9 (8,25)	7 1,5 (2,71) 1,12 (2,421)	7 2,8 (1,32) 2,35 (1,204)	8 3,0 (1,21) 2,72 (1,178)		

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder 8 – 12 Jahre)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Fatigue Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=11)	11 42,1 (9,35)	11 42,9 (7,01)	11 43,7 (8,43)	11 41,1 (8,05)	11 0,9 (2,27) 0,10 (1,542)	11 1,6 (3,30) 0,83 (2,567)	11 -1,0 (2,34) -1,79 (1,983)	-4,81 [-10,95; 1,34] 0,1252 -	0,57 [-5,36; 6,49] 0,8517 -
Burosumab (N=8)	7 45,2 (7,31)	7 45,1 (8,44)	7 41,1 (7,34)	8 43,2 (7,99)	7 -0,1 (3,37) -0,04 (2,590)	7 -4,0 (2,66) -3,97 (2,199)	8 -2,1 (3,05) -1,22 (2,415)		

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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, Baseline PROMIS als Kovariate

Sensitivitätsanalyse: separate Auswertung der Selbst- und Fremdeinschätzung

Basierend auf der Nutzenbewertung vom 02.01.2020 wurde eine zusätzliche Sensitivitätsanalyse durchgeführt, bei der die Auswertung der Fremd- (Kinder im Alter von 5 – 7 Jahren) und Selbsteinschätzung (Kinder im Alter von 8 – 12 Jahre) getrennt voneinander erfolgte [65]. Diese Analyse wurde *post-hoc* durchgeführt und war nicht im SAP präspezifiziert.

Schmerzdomäne (PROMIS Pediatric Pain Interference Domain Score)

In der Schmerzdomäne (*PROMIS Pediatric Pain Interference Domain Score*) unterschieden sich die Baseline-Werte der Kinder, die den Fragebogen selbst ausfüllten (Burosumab: MW: 50, aktive Kontrolle: MW: 47,9), leicht von denen, bei denen eine Fremdeinschätzung erfolgte (Burosumab: MW: 55,9; aktive Kontrolle: MW: 52,3). Auch wenn bei Kindern im Alter von 8 – 12 Jahren zu Beginn der Studie über mehr Schmerzen berichtet wurde, konnte in dieser Gruppe im Vergleich zu Kindern im Alter von 5 – 7 Jahren nach 40 Wochen eine signifikante Abnahme der Schmerzbeeinträchtigung zugunsten von Burosumab festgestellt werden (LsMeanW40: -7,76; $p < 0,0044$, H`g = -0,853 [-1,84; 0,135]). Bei Kindern im Alter von 5 – 7 Jahren war zwar ebenfalls eine Abnahme von Schmerzen in der Burosumab-Gruppe feststellbar, diese erreichte jedoch keine statistische Signifikanz. Nach 64 Wochen war dagegen eine ähnliche Abnahme von Schmerzen zugunsten von Burosumab in beiden Gruppen (Kinder 5 – 7 Jahre: $LsMean_{W64}$: -1,52; Kinder 8 – 12 Jahre: $LsMean_{W64}$: -1,64) zu verzeichnen (Tabelle 4-40 und Tabelle 4-41).

Mobilitätsdomäne (PROMIS Physical Function Mobility Domain Score)

Bei der Mobilitätsdomäne war in der Gruppe 8 – 12 Jahre (Burosumab: MW: 47,7; aktive Kontrolle: MW: 48,4) analog zur Schmerzdomäne ebenfalls eine stärkere Beeinträchtigung als in der Gruppe 5 – 7 Jahre (Burosumab: MW: 42,9; aktive Kontrolle: MW: 41,9) feststellbar. Zu Woche 40 konnte im Vergleich zu Kindern im Alter von 5 – 7 Jahren (LsMean: 1,52; $p = 0,5325$) eine deutliche Verbesserung der Mobilität bei Kindern im Alter von 8 – 12 Jahren (LsMean: 3,22; $p = 0,0567$) zugunsten von Burosumab festgestellt werden. Dieser Trend setzte sich zu Woche 64 weiter fort (Tabelle 4-40 und Tabelle 4-41).

Fatigue-Domäne (PROMIS Fatigue Domain Score)

Zu Baseline wiesen Kinder im Alter von 8 – 12 Jahren (Burosumab: MW: 45,2; aktive Kontrolle: MW: 42,1) eine weniger starke Fatigue als Kinder im Alter von 5 – 7 Jahren (Burosumab: MW: 51,9; aktive Kontrolle: MW: 53,0) auf. Nach 40 Wochen war eine deutliche Abnahme der Fatigue in der Gruppe 8 – 12 Jahre zugunsten von Burosumab ($LsMean_{W40}$: -4,81) feststellbar, ohne jedoch eine statistische Signifikanz ($p = 0,1252$) zu erreichen. Das Ausmaß der Fatigue bei Kindern im Alter von 5 – 7 Jahren veränderte sich dagegen nur geringfügig zugunsten von Burosumab ($LsMean_{W40}$: -0,95; $p = 0,7220$). Zu Woche 64 war das Ausmaß der Fatigue in den Behandlungsgruppen bei Kindern im Alter von 8 – 12 Jahren ähnlich ($LsMean_{W64}$: 0,57; $p = 0,8517$), während bei Kindern im Alter von 5 – 7 Jahren sich die Fatigue in der Burosumab-Gruppe im Vergleich zur aktiven Kontrollgruppe verringerte ($LsMean_{W64}$: -1,85; $p= 0,6338$) (Tabelle 4-40 und Tabelle 4-41).

Auch wenn Unterschiede in dem Ausmaß an Schmerzen, Fatigue und Mobilität bei Kindern, die den Fragebogen selbst ausfüllten und bei Kindern, bei denen das Ausfüllen des Fragebogens durch ein Elternteil erfolgte, feststellbar waren, lässt sich anhand dieser Analyse nicht abschließend ermitteln, ob dies auf die unterschiedliche Erhebung oder auf das Alter der Kinder zurückzuführen ist. Zudem führt die getrennte Auswertung der Selbst- und Fremdeinschätzung zu einer kleineren Stichprobe und ist daher mit mehr Unsicherheiten behaftet. Aufgrund dessen wird für die Ableitung des Zusatznutzens lediglich die Auswertung der Gesamtpopulation herangezogen.

Tabelle 4-42: Ergebnisse für *Schmerzerhebung und funktionelle Behinderung* mittels PROMIS aus RCT mit dem zu bewertenden Arzneimittel (MMRM-Analyse)

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Pain Interference Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=20)	20 49,9 (12,05)	20 51,4 (10,57)	20 50,4 (9,51)	20 49,4 (9,52)	20 1,6 (1,52) 0,74 (1,506)	20 0,5 (2,01) -0,31 (1,506)	20 -0,5 (1,60) -1,29 (1,506)	-4,91 [-9,51; -0,32] 0,0365 H'g = -0,697 [-1,39; -0,008]	-2,26 [-6,85; 2,33] 0,3298 -
Burosumab (N=15)	15 53,1 (10,95)	15 50,5 (9,17)	15 47,6 (9,84)	15 49,3 (8,07)	15 -2,7 (2,41) -2,39 (1,764)	15 -5,5 (1,78) -5,23 (1,764)	15 -3,8 (2,59) -3,55 (1,764)		
PROMIS *Pediatric Physical Function Mobility Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=20)	20 45,5 (9,86)	20 46,3 (9,69)	20 45,5 (9,71)	20 46,3 (9,63)	20 0,9 (0,81) 0,94 (1,171)	20 0,0 (1,01) 0,09 (1,171)	20 0,9 (0,99) 0,92 (1,171)	2,68 [-0,87; 6,22] 0,1361 -	1,90 [-1,65; 5,44] 0,2893 -
Burosumab (N=15)	15 45,2 (9,05)	15 47,3 (8,18)	15 47,9 (8,32)	15 47,9 (9,24)	15 2,2 (1,84) 2,20 (1,366)	15 2,7 (1,49) 2,76 (1,366)	15 2,8 (1,80) 2,82 (1,366)		

Intervention	**Patient-Reported Outcomes Measurement Information System (PROMIS) (Baseline-Alter der Kinder ≥ 5 Jahren)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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]**
PROMIS *Pediatric Fatigue Domain Score*									
Gesamtrate									
Aktive Kontrolle[b] (N=20)	20 47,0 (13,70)	20 46,8 (10,99)	20 46,6 (10,73)	20 45,0 (11,17)	20 -0,2 (1,68) -0,74 (1,596)	20 -0,5 (2,21) -0,97 (1,596)	20 -2,1 (1,87) -2,57 (1,596)	-3,20 [-8,03; 1,64] 0,1915 -	-1,08 [-5,91; 3,76] 0,6581 -
Burosumab (N=15)	15 48,8 (9,60)	15 47,1 (9,55)	15 44,7 (10,49)	15 45,2 (10,69)	15 -1,6 (1,80) -1,69 (1,860)	15 -4,1 (1,50) -4,17 (1,860)	15 -3,6 (2,32) -3,65 (1,860)		

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

MMRM: Behandlungsgruppe, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, Baseline PROMIS als Kovariate Als Kovarianzstruktur wurde die *compound symmetry* verwendet.

Sensitivitätsanalyse MMRM

Laut der aktuellen Dossiervorlage wird für die Nutzenbewertung empfohlen, für die Auswertung von mit Skalen erhobenen patientenberichteten Endpunkten das MMRM zu verwenden, um die über den Studienverlauf ermittelten Informationen vollständig zu berücksichtigen. Im SAP der Studie UX023-CL301 war jedoch für die Auswertung von patientenberichteten Endpunkten die Verwendung des GEE-Modells präspezifiziert. Ähnlich wie das MMRM erlaubt auch die Verwendung des GEE-Modells die Berücksichtigung patientenspezifischer Heterogenität und stellt daher insbesondere bei Messwiederholungen ein adäquates Auswertungsmodell dar.

Um zu überprüfen, ob die Wahl des statistischen Modells gegebenenfalls einen Einfluss auf die Ergebnisse der patientenberichteten Endpunkte hat, wurde für den Endpunkt PROMIS zusätzlich zu der Auswertung mittels GEE eine Analyse mittel MMRM durchgeführt. Diese Analyse war übereinstimmend mit der Primäranalyse (Tabelle 4-42). Aufgrund der übereinstimmenden Ergebnisse, unabhängig von der Wahl des statistischen Modells (MMRM oder GEE-Modell), wird daher im Folgenden für alle patientenberichteten Endpunkte ausschließlich das GEE-Modell verwendet, da zusätzliche Analysen mittels MMRM zu keinen neuen Erkenntnissen führen würden.

Tabelle 4-43: Ergebnisse für *Schmerzerhebung* mittels FPS-R aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder 5-12 Jahre)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**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** **MW (SD)**	**n** **MWD (SE)** **LsMean[a] (SE)**	**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=20)	20 0,7 (1,17)	20 0,4 (1,05)	20 0,6 (1,60)	19 0,5 (1,12)	20 -0,3 (0,33) -0,16 (0,219)	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 0,7 (1,63)	15 0,5 (1,19)	15 0,5 (0,92)	15 0,3 (0,55) 0,17 (0,450)	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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, sowie der Baseline FPS-R als Kovariate

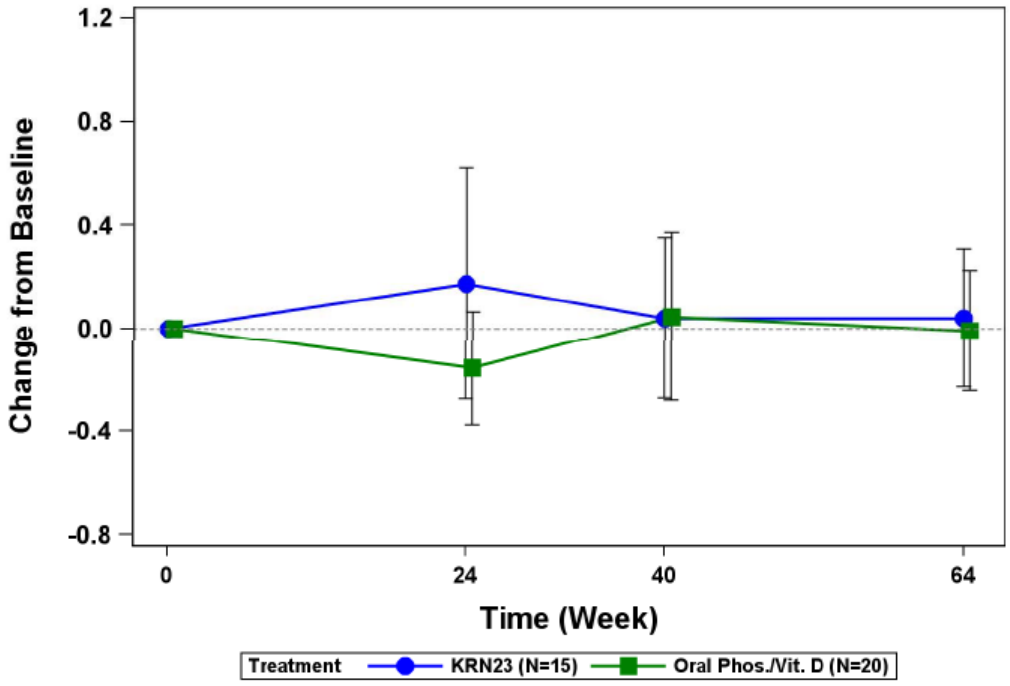

Abbildung 6: Veränderung zur Baseline des *FPS-R (Schmerzerhebung)*

Faces Pain Scale – Revised (FPS-R)

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 sowie bei allen Studienvisiten. In der mündlichen Anhörung wurde von den Fachexperten berichtet, dass die Schmerzsituation bei den pädiatrischen Patienten noch keine große Rolle spielt. Pädiatrische Patienten nehmen meist bevor es zu deutlichen Schmerzen kommt, eine Schonhaltung ein. Erst adoleszente und junge Erwachsene berichten von einer Schmerzsymptomatik, da es erst in diesem Alter durch die Beinachsenfehlstellung zu einer Adaption der nachfolgenden Gelenke und somit zur Ausbildung einer präathrotischen Deformität kommt. Diese bereitet den Patienten je nach Alter erhebliche Schmerzen.

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) (Tabelle 4-43 und Abbildung 6).

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.

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 4-44: Operationalisierung von *Lebensqualität*

Studie	Operationalisierung
UX023-CL301	Der *Short Form* 10 (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 [57]. Die 10 Items wurden aus dem *Child Health Questionnaire* (CHG) adaptiert. Hierbei werden zwei Komponenten generiert: • Physische 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 bis 9 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 [57]. Höhere globale Punktzahlen sind mit einer besseren Lebensqualität verbunden. Der Bezugszeitraum beträgt 4 Wochen. Der Fragebogen wurde in der aktivkontrollierten Behandlungsphase zu Baseline, Woche 24, 40 und 64 erhoben. <u>Analysepopulation:</u> FAS (Baseline Alter ≥ 5 Jahre) <u>Statistisches Modell:</u> • Auswertungen zur mittleren Änderung des PHS-10 bzw. PSS-10 wurden zu Woche 40 und zu Woche 64 mit einem GEE-Modell bestimmt. Bei dem GEE-Modell wurden die Behandlungsgruppe, die Behandlungsvisiten, die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und die Baseline-RSS-Stratifizierungsfaktoren als Faktoren aufgenommen, sowie der Baseline-Wert des SF-10 als Kovariate. Die abhängige Variable war die Änderung von Baseline des SF-10. <u>Darstellung im Dossier:</u> • Differenz der mittleren Änderung von Baseline zu Woche 40 und 64 zwischen den Behandlungsgruppen • Klinische Relevanz: Die klinische Relevanz der beobachteten Veränderung wird im vorliegenden Dossier mittels der SMD nach Hedges`g geprüft

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 4-45: Bewertung des Verzerrungspotenzials für *Lebensqualität* in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunktherheber**	**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 ≥ 5 Jahre ist definiert als alle Patienten, die randomisiert wurden, ein Mindestalter von 5 Jahre zur Screeningvisite aufwiesen und mindestens eine zugeteilte Studienmedikation erhalten haben. Da alle randomisierten Patienten der SF-10-Erhebung mindestens eine zugeteilte Medikation erhalten haben, umfasst die FAS ≥ 5 Jahre alle randomisierten Patienten. Daher wird die FAS ≥ 5 Jahre 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. Der Rücklauf des Fragebogens in der Studie belief sich zu allen Messzeitpunkten auf über 70 %.

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 4-46: Ergebnisse für *Lebensqualität (SF-10)* aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Gesundheitsspezifischer Fragebogen – SF-10 (Baseline-Alter der Kinder ≥ 5 Jahren)**								
(N)	**Baseline** **n** **MW (SD)**	**Woche 24** **n** **MW (SD)**	**Woche 40** **n** **MW (SD)**	**Woche 64** **n** **MW (SD)**	**Änderung von Baseline zu Woche 24** **n** **MWD (SE) LsMean[a] (SE)**	**Änderung von Baseline zu Woche 40** **n** **MWD (SE) LsMean[a] (SE)**	**Änderung von Baseline zu Woche 64** **n** **MWD (SE) LsMean[a] (SE)**	**Behandlungs-differenz von Baseline zu Woche 40** **LsMean[a] [95% KI] p-Wert falls p < 0,05: Hedges'g [95% KI]**	**Behandlungs-differenz von Baseline zu Woche 64** **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 43,30 (13,939)	20 42,27 (12,510)	20 41,07 (15,093)	20 2,56 (2,737) 2,71 (2,345)	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 43,27 (9,588)	15 46,16 (9,904)	15 46,12 (9,844)	15 3,23 (2,132) 3,14 (1,896)	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 51,36 (9,206)	20 51,85 (9,330)	20 53,95 (8,457)	20 -1,42 (1,411) -1,07 (1,262)	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 52,48 (8,153)	15 52,66 (9,225)	15 52,06 (9,137)	15 1,72 (1,891) 1,49 (1,532)	15 1,90 (1,738) 1,67 (1,640)	15 1,31 (2,112) 1,07 (1,799)		

Intervention	**Gesundheitsspezifischer Fragebogen – SF-10 (Baseline-Alter der Kinder ≥ 5 Jahren)**								
(N)	**Baseline**	**Woche 24**	**Woche 40**	**Woche 64**	**Änderung von Baseline zu Woche 24**	**Änderung von Baseline zu Woche 40**	**Änderung von Baseline zu Woche 64**	**Behandlungs-differenz von Baseline zu**	**Behandlungs-differenz von Baseline zu**
	n	**n**	**n**	**n**	**n**	**n**	**n**	**Woche 40**	**Woche 64**
	MW (SD)	**MW (SD)**	**MW (SD)**	**MW (SD)**	**MWD (SE) LsMean[a] (SE)**	**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]**

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, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Baseline-RSS-Stratifizierungsfaktoren als Faktoren, Baseline SF-10 als Kovariate.

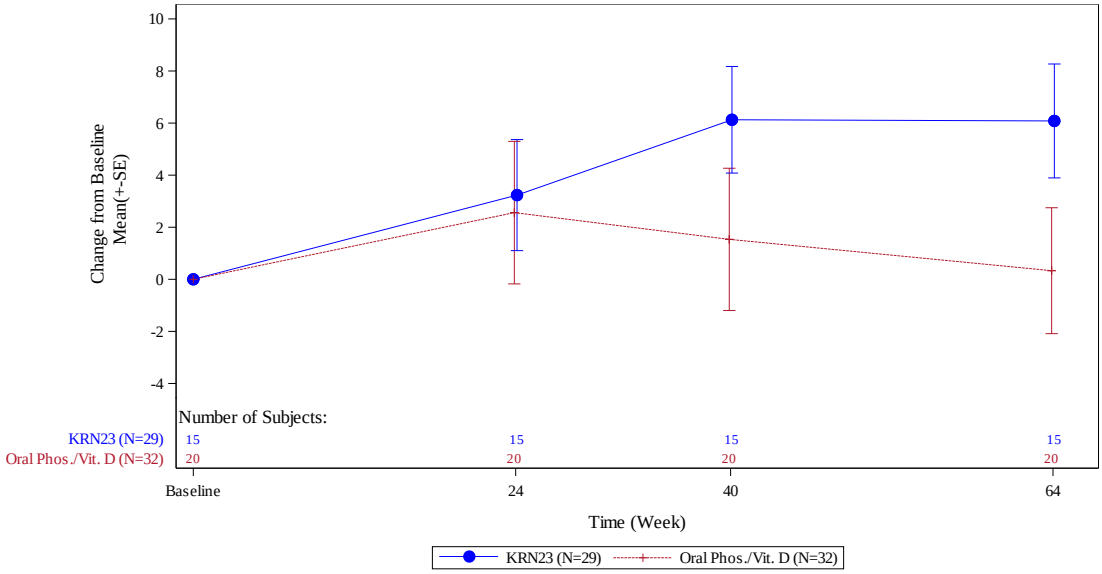

Abbildung 7: Veränderung zu Baseline für den Endpunkt *Physical Health Summary Score* über den Verlauf der 64-wöchigen aktiv-kontrollierten Phase der Studie UX023-CL301.

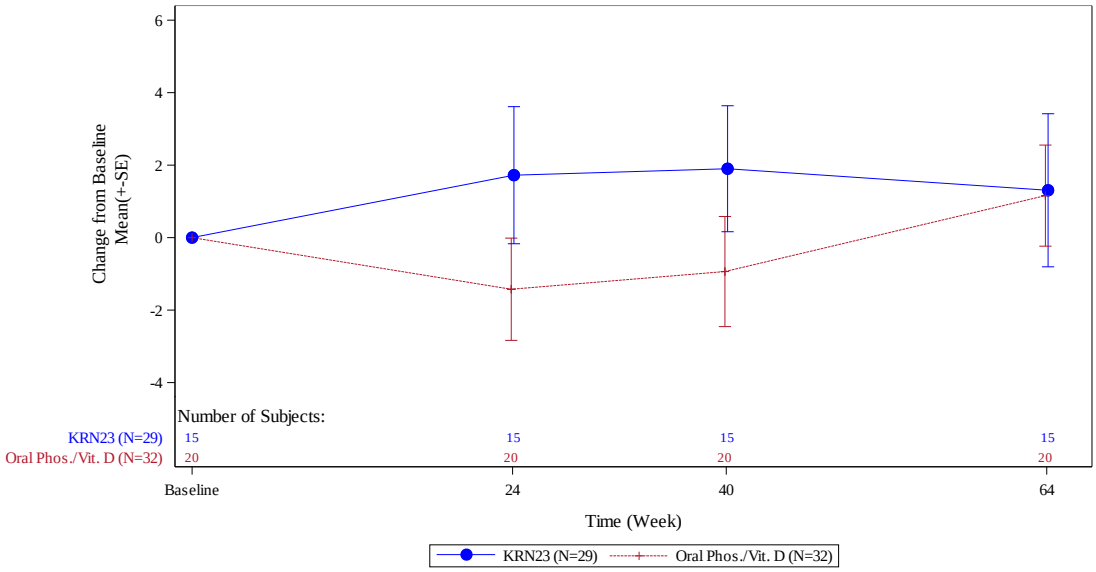

Abbildung 8: Veränderung zu Baseline für den Endpunkt *Psychosocial Summary Score* über den Verlauf der 64-wöchigen aktiv-kontrollierten Phase der Studie UX023-CL301.

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 (MW) und bei der aktiven Kontrollgruppe 40,74 (MW), 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 (MW), was eine adjustierte Veränderung von 1,69 (LsMean) bedeutet. Zu Woche 64 betrug die adjustierte mittlere Änderung in der Kontrollgruppe lediglich 0,48 (LsMean). Die Behandlungsdifferenz von 5,51 zu Woche 64 unterstreicht den positiven Effekt der Burosumab-Behandlung auf die Lebensqualität von XLH-Patienten ohne jedoch eine statistische Signifikanz zu erreichen ($p = 0,0571$) (Tabelle 4-46 und Abbildung 7).

Beim *Psychosocial Summary Score* (PSS-10) wiesen die Burosumab-Gruppe und die aktive Kontrolle ähnliche Werte nahe dem standardisierten Mittelwert von 50 auf (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 (MW) und in Woche 64 auf 52,06 (MW). Die mittleren PSS-10-Werte der aktiven Kontrollgruppe sanken in Woche 40 auf 51,85 (MW) und stiegen wieder in Woche 64 auf 53,95 (MW). Weder zu Woche 40 noch zu Woche 64 lag ein statistisch signifikanter Behandlungsunterschied zwischen den Behandlungsgruppen vor ($p_{40} = 0,2841$; $p_{64} = 0,8337$) (Tabelle 4-46 und Abbildung 8).

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.

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 4-47: Operationalisierung von *Sicherheit*

Studie	Operationalisierung
UX023-CL301	**Unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)** Ein *unerwünschtes Ereignis* (*UE*) ist eine jegliche unerwünschte medizinische Erscheinung bei einem Patienten, unabhängig davon, ob das Ereignis in kausalem Zusammenhang mit der Verabreichung der Studienmedikation steht. Alle *UE* 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. In die Auswertung wurden jedoch nur *UE* eingeschlossen, die nach der erstmaligen Verabreichung der Prüfsubstanz (Burosumab) oder Phosphat auftraten (*Treatment emergent adverse events* (*TEAE*)), um die Inzidenz, Häufigkeit und Schwere der *UE* sowie einen möglichen Zusammenhang mit der Behandlung zu bewerten. *UE*, die zwischen dem Zeitpunkt der Einwilligung und der Baseline-Visite auftraten, wurden hierbei nicht berücksichtigt, da diese *UE* in keinem Bezug zur Behandlung mit der Prüfsubstanz standen und die objektive Beurteilung des Sicherheitsprofils in beiden Behandlungsarmen beeinträchtigt hätten. Bei jedem Besuch wurden die Patienten zu neuen oder laufenden *UE* seit dem letzten Besuch befragt. Die Beurteilung der *UE* erfolgte bei jedem Besuch der Prüfstelle. Der Prüfarzt beurteilt die Schwere, die Kausalität (Zusammenhang mit der Studienmedikation) und die Bedeutung jedes *UE*. Folgende Klassifizierungen der *TEAE* wurden vorgenommen: *Schweregrad der TEAE* Der Prüfarzt schätzt den Schweregrad aller *UE* gemäß der *National Cancer Institute Common Terminology Criteria for Adverse Evenst (NCI CTCAE)* ein, die wie folgt definiert sind: • Mild: Jedes *TEAE*, das normalerweise vorübergehend ist und nur minimale Behandlung oder therapeutisches Eingreifen benötigt. Das Ereignis beeinflusst normale tägliche Aktivitäten generell nicht. • Moderat: Jedes *TEAE*, 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. • Schwer (Stufe 3): Jedes *TEAE*, 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 hat zu versterben. • Tod (Grad 5): Ereignisse, die zum Tod führen. *Kausalität der TEAE* Der Prüfarzt bewertete die potenzielle Beziehung der *TEAE* der Studie in Abhängigkeit des Auftretens durch die Medikation. Hierzu dienen folgende Kategorien: • Definitiv nicht zusammenhängend: Diese Kategorie bezieht sich auf *TEAE*, 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 *TEAE*, 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 *TEAE*, 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).

Schwerwiegendes TEAE

Ein schwerwiegendes *TEAE* ist jegliches *TEAE*, 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.

SOC und PT nach MedDRA

Alle *TEAE* wurden gemäß MedDRA kodiert und entsprechend den SOC und PT klassifiziert. Im Dossier werden auf Ebene der SOC und PT folgende Auswertungen dargestellt:

- Gesamtrate *TEAE*: Alles SOC und PT, die von ≥ 10 % der Patienten in einem Studienarm berichtet wurden
- Schwerwiegende *TEAE*: Alle SOC und PT, die von ≥ 5 % der Patienten in einem Studienarm berichtet wurden
- *TEAE* mit NCI-CTCAE-Schweregrad 3 – 4: Alle SOC und PT, die von ≥ 5 % der Patienten in einem Studienarm berichtet wurden

Analysepopulation: *Safety*

Statistisches Modell:

Im Studienprotokoll der Studie UX023-CL301 war prospektiv eine deskriptive Darstellung des Anteils der Patienten mit *TEAE* geplant. *Post-hoc* wurden für das

vorliegende Dossier sowohl für die Gesamtraten als auch auf Ebene von SOC und PT Effektschätzer berechnet, um die Anforderungen der Dossiervorlage zu erfüllen.

- Das OR wurde mittels der logistischen Regression bestimmt. Als Faktoren dienten die Baseline-Alters-Stratifizierungsfaktoren und die Behandlungsgruppe, als kontinuierliche Kovariate der RSS Total Score.
- Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation und der p-Wert mittels Fishers exaktem Test berechnet.
- Zu Woche 64 wurde das GLMM verwendet, dass die Behandlungsgruppe, die Behandlungsvisiten und die Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren berücksichtigte und den Baseline RSS Total Score als kontinuierliche Kovariate. Zusätzlich wurden patientenspezifische Effekte für die Modellkonstante als zufällige Effekte berücksichtigt.

Erkrankungsbezogene Ereignisse:

Laut der aktuellen Dossiervorlage sollen bei der Erhebung *UE* zusätzliche *UE*-Analysen durchgeführt werden, bei denen krankheitsbezogene Ereignisse unberücksichtigt bleiben. In der Nutzenbewertung vom 02.01.2020 wurde bereits darauf hingewiesen, dass zwar *UE* der Systemorganklasse *Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen* sowie die PT *Zahnschmerzen, Zahnkaries und Zahnabszess* typische Symptome der betrachteten Grunderkrankung darstellen können, einzelne Symptome wie z. B. Schmerzen in den Extremitäten sich aber nicht nachträglich aus den betreffenden Systemorganklassen separieren lassen [65]. Aufgrund dieser Problematik werden im Folgenden ausschließlich die Gesamtanalysen der jeglichen, schweren und schwerwiegenden *UE* dargestellt, bei denen nicht ausgeschlossen werden kann, dass krankheitsbezogene Ereignisse beinhaltet sind.

Darstellung im Dossier:

- Gesamtrate *TEAE* bis Woche 64
- Gesamtrate *schwerer TEAE* bis Woche 64
- Gesamtrate *schwerwiegender TEAE* bis Woche 64
- *TEAE* unterteilt nach MedDRA SOC und PT, die bei mind. ≥ 10 % der Studienteilnehmer in mindestens einer Behandlungsgruppe aufgetreten sind
- *Schwere TEAE* unterteilt nach MedDRA SOC und PT, die bei mind. ≥ 5 % der Studienteilnehmer in mindestens einer Behandlungsgruppe aufgetreten sind
- *Schwerwiegende TEAE* unterteilt nach MedDRA SOC und PT, die bei mind. ≥ 5 % der Studienteilnehmer in mindestens einer Behandlungsgruppe aufgetreten sind
- *TEAE von besonderem Interesse* (unabhängig von der Häufigkeit)

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 4-48: 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 Analyse des Endpunkts Sicherheit beruht auf dem *Safety-Set* der Studienpopulation, da dieses die Sicherheit der Interventionen untersuchte. Die *Safety*-Population umfasst alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Patienten wurden dem Behandlungsarm zugeordnet, dessen Behandlung sie tatsächlich erhalten haben. Alle randomisierten Patienten erhielten in der Studie UX023-CL301 mindestens eine zugeteilte Medikation, sodass das ITT-Prinzip adäquat umgesetzt wurde.

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.

Unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)

Tabelle 4-49: Ergebnisse für *unerwünschte Ereignisse nach Beginn der Behandlung* (*TEAE*) aus RCT mit dem zu bewertenden Arzneimittel (*Safety*-Population)

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b] [95% KI]**	**ARD[b] [95% KI]**	**p-Wert[c]**
Gesamtrate - Baseline zu Woche 40						
Jegliche TEAE						
Aktive Kontrolle[d] (32)	26 (81,3)	6,121 [0,652; 57,447]	0,1107	1,188 [0,922; 1,423]	0,153 [0,002; 0,304]	0,1064
Burosumab (29)	28 (96,6)					
Schwerwiegende TEAE						
Aktive Kontrolle[d] (32)	1 (3,1)	3,473 [0,320; 37,691]	0,3002	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					
TEAE Grad 3 - 4						
Aktive Kontrolle[d] (32)	2 (6,3)	2,832 [0,440; 18,213]	0,2673	2,207 [0,436; 11,166]	0,075 [-0,076; 0,226]	0,4107
Burosumab (29)	4 (13,8)					
Studienabbruch durch TEAE						
Aktive Kontrolle[d] (32)	0 (0)	-	-	-	-	-
Burosumab (29)	0 (0)					
Therapieabbruch durch TEAE						
Aktive Kontrolle[d] (32)	0 (0)	-	-	-	-	-
Burosumab (29)	0 (0)					
Tod durch TEAE						
Aktive Kontrolle[d] (32)	0 (0)	-	-	-	-	-
Burosumab (29)	0 (0)					

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b] [95% KI]**	**ARD[b] [95% KI]**	**p-Wert[c]**
Gesamtrate - Baseline zu Woche 64						
Jegliche TEAE						
Aktive Kontrolle[d] (32)	27 (84,4)	n.b.		1,185 [1,021; 1,376]	0,156 [0,030; 0,282]	0,0538
Burosumab (29)	29 (100)					
Schwerwiegende TEAE						
Aktive Kontrolle[d] (32)	3 (9,4)	1,133 [0,197; 6,520]	0,8868	1,115 [0,207; 6,018]	0,010 [-0,140; 0,160]	1,0000
Burosumab (29)	3 (10,3)					
TEAE Grad 3 – 4						
Aktive Kontrolle[d] (32)	3 (9,4)	1,808 [0,338; 9,664]	0,4824	1,547 [0,316; 7,582]	0,044 [-0,117; 0,205]	0,6988
Burosumab (29)	4 (13,8)					
Studienabbruch durch TEAE						
Aktive Kontrolle[d] (32)	0 (0)	-	-	-	-	-
Burosumab (29)	0 (0)					
Therapieabbruch durch TEAE						
Aktive Kontrolle[d] (32)	0 (0)	-	-	-	-	-
Burosumab (29)	0 (0)					
Tod durch TEAE						
Aktive Kontrolle[d] (32)	0 (0)	-	-	-	-	-
Burosumab (29)	0 (0)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; *TEAE*: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 40): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*. GLMM (Woche 64): Behandlungsgruppe, Behandlungsvisiten und Interaktion der Behandlungsgruppe mit den Behandlungsvisiten sowie die Baseline-Alters-Stratifizierungsfaktoren als Faktoren, Baseline RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b] [95% KI]	ARD[b] [95% KI]	p-Wert[c]
b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*. c: Der p-Wert wurde mittels Fishers exaktem Test berechnet. d: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						

Die Auswertung der *TEAE* ergab keine signifikanten Sicherheitsbedenken. Während der Studie wurde kein Kind aufgrund von *TEAE* oder aus einem anderen Grund aus der Studie ausgeschlossen. Kein Studienteilnehmer verstarb. Die meisten Studienteilnehmer hatten mindestens ein *TEAE* (Burosumab: 100 %; Aktive Kontrolle: 84 %). Bei 7 Patienten ($n_{(Burosumab)} = 4$; $n_{(aktive\ Kontrolle)} = 3$) traten *TEAE* Grad 3 auf. Die restlichen *TEAE* waren mild oder moderat (Grad 1 oder 2) und es wurden keine lebensbedrohlichen *TEAE* Grad 4 gemeldet (Tabelle 4-49).

Tabelle 4-50: Deskriptive Darstellung der *unerwünschten Ereignisse nach Beginn der Behandlung (TEAE)* aus RCT mit dem zu bewertenden Arzneimittel zu Woche 40, 64 und in der Extensionsphase

UX023-CL301 Population (N) (Safety)	Woche 40 Patienten n (%)	Woche 64 Patienten n (%)	Extensionsphase[a] Patienten n (%)
TEAE			
Aktive Kontrolle[b] (32)	26 (81,3)	27 (84,4)	13 (59,1)
Burosumab (29)	28 (96,6)	29 (100,0)	29 (100,0)
Schwerwiegende TEAE			
Aktive Kontrolle[b] (32)	1 (3,1)	3 (9,4)	0 (0,0)
Burosumab (29)	3 (10,3)	3 (10,3)	4 (13,8)
TEAE Grad 3 – 4			
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 TEAE			
Aktive Kontrolle[b] (32)	0 (0)	0 (0)	0 (0)
Burosumab (29)	0 (0)	0 (0)	0 (0)
Therapieabbruch durch TEAE			
Aktive Kontrolle[b] (32)	0 (0)	0 (0)	0 (0)
Burosumab (29)	0 (0)	0 (0)	0 (0)

UX023-CL301 **Population (N)** **(Safety)**	**Woche 40** **Patienten** **n (%)**	**Woche 64** **Patienten** **n (%)**	**Extensionsphase[a]** **Patienten** **n (%)**
Tod durch TEAE			
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; *TEAE*: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung)

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 D

Extensionsphase (Woche 66 bis 140)

Die Bewertung der Extensionsphase ergab keine Sicherheitsbedenken. Kein Patient wurde aufgrund von *TEAE* aus der Studie ausgeschlossen. Es wurde kein Todesfall berichtet. 59 % der Patienten, die von der Substitutionstherapie auf die Burosumab-Behandlung wechselten, berichteten von einem *TEAE* (Tabelle 4-50). Die meisten dieser Ereignisse waren ebenfalls Reaktionen an der Injektionsstelle. In der Burosumab-Burosumab-Gruppe trat während der Extensionsphase ein *schweres TEAE* 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 intrakraniellem Druck und ophthalmologischen Komplikationen verbunden war und das Papillenödem dadurch bedingt sein könnte. Während der Extensionsphase waren alle *TEAE* mild oder moderat mit Ausnahme von zwei *TEAE* Grad 3. Diese traten bei den beiden Patienten auf, die während der Behandlungsphase bereits *schwere TEAE* berichteten. Eines war das oben genannte bilaterale Papillenödem. Bei dem zweiten *schweren TEAE* 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 therapieassoziiert eingestuft und konnte innerhalb von zwei Tagen mit einer Zahnextraktion behoben werden.

Unerwünschte Ereignisse nach Beginn der Behandlung (TEAE) nach MedDRA SOC und PT (bei ≥ 10 % der Patienten in mindestens einer Behandlungsgruppe)

Tabelle 4-51: Ergebnisse für jegliche *unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)* nach *System Organ Class* (SOC) und *Preferred Term* (PT) (bei ≥ 10 % der Patienten in mindestens einer Behandlungsgruppe) aus RCT mit dem zu bewertenden Arzneimittel (Baseline – W 40, *Safety*-Population)

UX023-CL301 Population (N) (*Safety*)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
Baseline zu Woche 40						
SOC Kongenitale, familiäre und genetische Erkrankungen						
Aktive Kontrolle[e] (32)	1 (3,1)	3,204 [0,284, 36,151]	0,3400	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					
SOC Erkrankungen des Ohrs und des Labyrinths						
Aktive Kontrolle[e] (32)	1 (3,1)	3,999 [0,364; 43,955]	0,2517	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					
SOC Erkrankungen des Gastrointestinaltrakts						
Aktive Kontrolle[e] (32)	14 (43,8)	2,879 [0,973; 8,518]	0,0558	1,576 [0,993; 2,503]	0,252 [0,012; 0,493]	0,0709
Burosumab (29)	20 (69,0)					
PT Schmerzen Oberbauch						
Aktive Kontrolle[e] (32)	2 (6,3)	1,981 [0,285; 13,757]	0,4830	1,655 [0,297; 9,217]	0,041 [-0,098; 0,180]	0,6615
Burosumab (29)	3 (10,3)					
PT Karies						
Aktive Kontrolle[e] (32)	2 (6,3)	4,352 [0,629; 30,085]	0,1333	2,759 [0,579; 13,140]	0,110 [-0,051; 0,271]	0,2414
Burosumab (29)	5 (17,2)					
PT Diarrhoe						
Aktive Kontrolle[e] (32)	1 (3,1)	9,854 [0,971; 99,985]	0,0529	6,621 [0,847; 51,76]	0,176 [0,016; 0,335]	0,0461
Burosumab (29)	6 (20,7)					

UX023-CL301 Population (N) (*Safety*)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
PT Erbrechen						
Aktive Kontrolle[e] (32)	6 (18,8)	2,620 [0,793; 8,655]	0,1119	2,023 [0,857; 4,773]	0,192 [-0,031; 0,414]	0,1522
Burosumab (29)	11 (37,9)					
SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort						
Aktive Kontrolle[e] (32)	6 (18,8)	12,051 [3,379; 42,973]	0,0002	3,862 [1,814; 8,220]	0,537 [0,325; 0,748]	< 0,0001
Burosumab (29)	21 (72,4)					
PT Erythem an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		14,3 [0,841; 243,201]	0,207 [0,059; 0,354]	0,0086
Burosumab (29)	6 (20,7)					
PT Injektionsstelle juckend						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10.3)					
PT Ausschlag an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					
PT Reaktion an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		14,3 [0,841; 243,201]	0,207 [0,059; 0,354]	0,0086
Burosumab (29)	6 (20,7)					
PT Fieber						
Aktive Kontrolle[e] (32)	4 (12,5)	5,798 [1,478; 22,744]	0,0126	3,586 [1,317; 9,765]	0,323 [0,109; 0,537]	0,0090
Burosumab (29)	13 (44,8)					

UX023-CL301 Population (N) (*Safety*)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
SOC Erkrankungen des Immunsystems						
Aktive Kontrolle[e] (32)	2 (6,3)	2,566 [0,410; 16,075]	0,3080	2,207 [0,436; 11,166]	0,075 [-0,076; 0,226]	0,4107
Burosumab (29)	4 (13,8)					
PT Jahreszeitbedingte Allergie						
Aktive Kontrolle[e] (32)	1 (3,1)	4,024 [0,365; 44,423]	0,2504	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					
SOC Infektionen und parasitäre Erkrankungen						
Aktive Kontrolle[e] (32)	18 (56,3)	2,301 [0,731; 7,237]	0,1510	1,349 [0,933; 1,949]	0,196 [-0,036; 0,428]	0,1767
Burosumab (29)	22 (75,9)					
PT Nasopharyngitis						
Aktive Kontrolle[e] (32)	9 (28,1)	1,135 [0,366; 3,525]	0,8234	1,103 [0,508; 2,396]	0,029 [-0,200; 0,258]	1,0000
Burosumab (29)	9 (31,0)					
PT Zahnabszess						
Aktive Kontrolle[e] (32)	3 (9,4)	2,714 [0,581, 12,690]	0,2000	2,207 [0,607; 8,030]	0,113 [-0,066; 0,292]	0,2870
Burosumab (29)	6 (20,7)					
PT Infektion der oberen Atemwege						
Aktive Kontrolle[e] (32)	2 (6,3)	1,801 [0,264; 12,308]	0,5424	1,655 [0,297; 9,217]	0,041 [-0,098; 0,180]	0,6615
Burosumab (29)	3 (10,3)					
SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen						
Aktive Kontrolle[e] (32)	1 (3,1)	13,187 [1,427; 121,867]	0,0238	8,828 [1,174; 66,360]	0,245 [0,071; 0,418]	0,0101
Burosumab (29)	8 (27,6)					

UX023-CL301 Population (N) (*Safety*)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
PT Kontusion						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		9,9 [0,556; 176,29]	0,138 [0,012; 0,263]	0,0455
Burosumab (29)	4 (13,8)					
SOC Untersuchungen						
Aktive Kontrolle[e] (32)	4 (12,5)	1,551 [0,354; 6,795]	0,5546	1,379 [0,409; 4,649]	0,047 [-0,132; 0,226]	0,7242
Burosumab (29)	5 (17,2)					
SOC Stoffwechsel- und Ernährungsstörungen						
Aktive Kontrolle[e] (32)	3 (9,4)	3,602 [0,686; 18,903]	0,1272	2,207 [0,607; 8,030]	0,113 [-0,066; 0,292]	0,2870
Burosumab (29)	6 (20,7)					
PT Vitamin D-Mangel						
Aktive Kontrolle[e] (32)	1 (3,1)	9,631 [0,921; 100,728]	0,0583	5,517 [0,684; 44,491]	0,141 [-0,009; 0,291]	0,0933
Burosumab (29)	5 (17,2)					
SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen						
Aktive Kontrolle[e] (32)	11 (34,4)	2,051 [0,682; 6,168]	0,1966	1,404 [0,764; 2,583]	0,139 [-0,106; 0,384]	0,3067
Burosumab (29)	14 (48,3)					
PT Arthralgie						
Aktive Kontrolle[e] (32)	6 (18,8)	2,294 [0,643; 8,185]	0,1965	1,655 [0,671; 4,081]	0,123 [-0,093; 0,339]	0,3736
Burosumab (29)	9 (31,0)					
PT Schmerz in einer Extremität						
Aktive Kontrolle[e] (32)	8 (25,0)	1,718 [0,541; 5,456]	0,3527	1,379 [0,631; 3,016]	0,095 [-0,134; 0,324]	0,5749
Burosumab (29)	10 (34,5)					

UX023-CL301 Population (N) (*Safety*)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
SOC Erkrankungen des Nervensystems						
Aktive Kontrolle[e] (32)	8 (25,0)	1,525 [0,459, 5,062]	0,4843	1,241 [0,553; 2,788]	0,060 [-0,165; 0,286]	0,7757
Burosumab (29)	9 (31,0)					
PT Kopfschmerz						
Aktive Kontrolle[e] (32)	5 (15,6)	2,540 [0,652; 9,900]	0,1755	1,766 [0,651; 4,789]	0,120 [-0,086; 0,325]	0,3509
Burosumab (29)	8 (27,6)					
SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums						
Aktive Kontrolle[e] (32)	8 (25)	5,028 [1,594; 15,857]	0,0067	2,483 [1,278; 4,824]	0,371 [0,139; 0,602]	0,0046
Burosumab (29)	18 (62,1)					
PT Asthma						
Aktive Kontrolle[e] (32)	1 (3,1)	5,365 [0,523; 55,080]	0,1541	4,414 [0,523; 37,250]	0,107 [-0,033; 0,246]	0,1816
Burosumab (29)	4 (13,8)					
PT Husten						
Aktive Kontrolle[e] (32)	5 (15,6)	3,338 [0,952; 11,705]	0,0593	2,428 [0,958; 6,153]	0,223 [0,006; 0,440]	0,0791
Burosumab (29)	11 (37,9)					
PT Nasenverstopfung						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					
PT Schmerzen im Oropharynx						
Aktive Kontrolle[e] (32)	1 (3,1)	4,591 [0,372; 56,614]	0,2294	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					

UX023-CL301 Population (N) (*Safety*)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
PT Rhinorrhoe						
Aktive Kontrolle[e] (32)	1 (3,1)	6,947 [0,689; 70,052]	0,0985	5,517 [0,684; 44,491]	0,141 [-0,009; 0,291]	0,0933
Burosumab (29)	5 (17,2)					
SOC Erkrankungen der Haut und des Unterhautzellgewebes						
Aktive Kontrolle[e] (32)	2 (6,3)	8,112 [1,475; 44,625]	0,0170	5,517 [1,317; 23,115]	0,282 [0,090; 0,475]	0,0086
Burosumab (29)	10 (34,5)					
PT Ausschlag						
Aktive Kontrolle[e] (32)	2 (6,3)	1,317 [0,166; 10,417]	0,7909	1,655 [0,297; 9,217]	0,041 [-0,098; 0,180]	0,6615
Burosumab (29)	3 (10,3)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; TEAE: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 40): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*.

c: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

d: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.

e: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

Tabelle 4-52: Ergebnisse für jegliche *unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)* nach *System Organ Class* (SOC) und *Preferred Term* (PT) (bei ≥ 10 % der Patienten in mindestens einer Behandlungsgruppe) aus RCT mit dem zu bewertenden Arzneimittel (Baseline – Woche 64, *Safety*-Population)

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
Baseline zu Woche 64						
SOC Kongenitale, familiäre und genetische Erkrankungen						
Aktive Kontrolle[e] (32)	1 (3,1)	3,204 [0,284, 36,151]	0,3400	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
SOC Erkrankungen des Ohrs und des Labyrinths						
Aktive Kontrolle[e] (32)	3 (9,4)	2,995 [0,620; 14,462]	0,1684	2,207 [0,607; 8,030]	0,113 [-0,066; 0,292]	0,2870
Burosumab (29)	6 (20,7)					
PT Ohrschmerzen						
Aktive Kontrolle[e] (32)	1 (3,1)	5,118 [0,505; 51,873]	0,1635	4,414 [0,523; 37,250]	0,107 [-0,033; 0,246]	0,1816
Burosumab (29)	4 (13,8)					
SOC Erkrankungen des Gastrointestinaltrakts						
Aktive Kontrolle[e] (32)	17 (53,1)	3,676 [1,118; 12,081]	0,0326	1,493 [1,026; 2,172]	0,262 [0,035; 0,489]	0,0577
Burosumab (29)	23 (79,3)					
PT Schmerzen Oberbauch						
Aktive Kontrolle[e] (32)	3 (9,4)	1,248 [0,212; 7,330]	0,8032	1,103 [0,241; 5,042]	0,010 [-0,140; 0,160]	1,0000
Burosumab (29)	3 (10,3)					
PT Obstipation						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		12,1 [0,698; 209,71]	0,172 [0,035; 0,310]	0,0200
Burosumab (29)	5 (17,2)					
PT Karies						
Aktive Kontrolle[e] (32)	2 (6,3)	7,734 [1,360; 43,987]	0,0219	4,966 [1,168; 21,114]	0,248 [0,060; 0,436]	0,0183
Burosumab (29)	9 (31,0)					
PT Diarrhoe						
Aktive Kontrolle[e] (32)	2 (6,3)	4.799 [0,852; 27,043]	0,0746	3,862 [0,871; 17,119]	0.179 [0,002; 0,356]	0,0723
Burosumab (29)	7 (24,1)					

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
PT Übelkeit						
Aktive Kontrolle[e] (32)	1 (3,1)	4,579 [0,390; 53,740]	0,2212	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					
PT Zahnschmerzen						
Aktive Kontrolle[e] (32)	1 (3,1)	5,505 [0,531; 57,114]	0,1498	4,414 [0,523; 37,250]	0,107 [-0,033; 0,246]	0,1816
Burosumab (29)	4 (13,8)					
PT Erbrechen						
Aktive Kontrolle[e] (32)	8 (25,0)	2,036 [0,663; 6,253]	0,2099	1,655 [0,790; 3,470]	0,164 [-0,070; 0,398]	0,2746
Burosumab (29)	12 (41,4)					
SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort						
Aktive Kontrolle[e] (32)	8 (25,0)	20,183 [4,988; 81,660]	< 0,0001	3,448 [1,860; 6,394]	0,612 [0,416; 0,808]	< 0,0001
Burosumab (29)	25 (86,2)					
PT Erythem an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		20,9 [1,270; 343,87]	0,310 [0,142, 0.479]	0,0006
Burosumab (29)	9 (31,0)					
PT Injektionsstelle juckend						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					
PT Ausschlag an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
PT Reaktion an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		16,5 [0,984; 276,73]	0,241 [0,086, 0,397]	0,0036
Burosumab (29)	7 (24,1)					
PT Schwellung an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					
PT Fieber						
Aktive Kontrolle[e] (32)	6 (18,8)	5,225 [1,557; 17,530]	0,0083	2,943 [1,332; 6,499]	0,364 [0,138; 0,590]	0,0039
Burosumab (29)	16 (55,2)					
SOC Erkrankungen des Immunsystems						
Aktive Kontrolle[e] (32)	5 (15,6)	1,078 [0,263, 4,426]	0,9152	1,103 [0,355; 3,427]	0,016 [-0,170, 0,203]	1,0000
Burosumab (29)	5 (17,2)					
PT Jahreszeitbedingte Allergie						
Aktive Kontrolle[e] (32)	2 (6,3)	2,353 [0,374; 14,797]	0,3552	2,207 [0,436; 11,166]	0,075 [-0,076; 0,226]	0,4107
Burosumab (29)	4 (13,8)					
SOC Infektionen und parasitäre Erkrankungen						
Aktive Kontrolle[e] (32)	24 (75,0)	1,134 [0,317; 4,053]	0,8445	1,057 [0,805; 1,390]	0,043 [-0,167; 0,253]	0,7666
Burosumab (29)	23 (79,3)					
PT Influenza						
Aktive Kontrolle[e] (32)	6 (18,8)	0,652 [0,154; 2,764]	0,5557	0,736 [0,230; 2,349]	0,050 [-0,135; 0,234]	0,7348
Burosumab (29)	4 (13,8)					

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
PT Nasopharyngitis						
Aktive Kontrolle[e] (32)	14 (43,8)	0,792 [0,275; 2,281]	0,6611	0,867 [0,471; 1,594]	0,058 [-0,188; 0,305]	0,7951
Burosumab (29)	11 (37,9)					
PT Zahnabszess						
Aktive Kontrolle[e] (32)	3 (9,4)	4,259 [0,938; 19,346]	0,0602	2,943 [0,862; 10,048]	0,182 [-0,009; 0,374]	0,0964
Burosumab (29)	8 (27,6)					
PT Infektion der oberen Atemwege						
Aktive Kontrolle[e] (32)	3 (9,4)	1,176 [0,207; 6,676]	0,8523	1,103 [0,241; 5,042]	0,010 [-0,140; 0,160]	1,0000
Burosumab (29)	3 (10,3)					
SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen						
Aktive Kontrolle[e] (32)	2 (6,3)	9,023 [1,656; 49,166]	0,0119	5,517 [1,317; 23,115]	0,282 [0,090; 0,475]	0,0086
Burosumab (29)	10 (34,5)					
PT Kontusion						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		9,9 [0,556; 176,29]	0,138 [0,012; 0,263]	0,0455
Burosumab (29)	4 (13,8)					
PT Sturz						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					
SOC Untersuchungen						
Aktive Kontrolle[e] (32)	4 (12,5)	3,119 [0,768; 12,666]	0,1096	2,207 [0,742; 6,564]	0,151 [-0,048; 0,350]	0,1996
Burosumab (29)	8 (27,6)					

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
PT Vitamin D erniedrigt						
Aktive Kontrolle[e] (32)	1 (3,1)	11,595 [1,135; 118,49]	0,0391	6,621 [0,847; 51,764]	0,176 [0,016; 0,335]	0,0461
Burosumab (29)	6 (20,7)					
SOC Stoffwechsel- und Ernährungsstörungen						
Aktive Kontrolle[e] (32)	3 (9,4)	3,602 [0,686; 18,903]	0,1272	2,207 [0,607; 8,030]	0,113 [-0,066; 0,292]	0,2870
Burosumab (29)	6 (20,7)					
PT Vitamin-D-Mangel						
Aktive Kontrolle[e] (32)	1 (3,1)	9,631 [0,921; 100,73]	0,0583	5,517 [0,684; 44,491]	0,141 [-0,009; 0,291]	0,0933
Burosumab (29)	5 (17,2)					
SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen						
Aktive Kontrolle[e] (32)	15 (46,9)	1,938 [0,641; 5,860]	0,2361	1,251 [0,775; 2,019]	0,117 [-0,132; 0,367]	0,4443
Burosumab (29)	17 (58,6)					
PT Arthralgie						
Aktive Kontrolle[e] (32)	10 (31,3)	2,036 [0,667; 6,216]	0,2072	1,434 [0,746; 2,758]	0,136 [-0,106; 0,378]	0,3024
Burosumab (29)	13 (44,8)					
PT Schmerz in einer Extremität						
Aktive Kontrolle[e] (32)	10 (31,3)	1,518 [0,495; 4,656]	0,4585	1,214 [0,607; 2,428]	0,067 [-0,172; 0,306]	0,6020
Burosumab (29)	11 (37,9)					
SOC Erkrankungen des Nervensystems						
Aktive Kontrolle[e] (32)	9 (28,1)	2,090 [0,665; 6,566]	0,2023	1,471 [0,728; 2,972]	0,133 [-0,105; 0,370]	0,2963
Burosumab (29)	12 (41,4)					

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
PT Kopfschmerzen						
Aktive Kontrolle[e] (32)	6 (18,8)	3,003 [0,822; 10,972]	0,0947	1,839 [0,764; 4,427]	0,157 [-0,062; 0,377]	0,2442
Burosumab (29)	10 (34,5)					
SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums						
Aktive Kontrolle[e] (32)	9 (28,1)	7,044 [2,147; 23,108]	0,0017	2,575 [1,416; 4,681]	0,443 [0,218; 0,668]	0,0008
Burosumab (29)	21 (72,4)					
PT Asthma						
Aktive Kontrolle[e] (32)	1 (3,1)	5,365 [0,523; 55,080]	0,1541	4,414 [0,523; 37,250]	0,107 [-0,033; 0,246]	0,1816
Burosumab (29)	4 (13,8)					
PT Husten						
Aktive Kontrolle[e] (32)	6 (18,8)	4,587 [1,404; 14,982]	0,0126	2,759 [1,237; 6,154]	0,330 [0,103; 0,556]	0,0083
Burosumab (29)	15 (51,7)					
PT Nasenverstopfung						
Aktive Kontrolle[e] (32)	1 (3,1)	7,383 [0,736; 74,028]	0,0879	5,517 [0,684; 44,491]	0,141 [-0,009; 0,291]	0,0933
Burosumab (29)	5 (17,2)					
PT Schmerzen im Oropharynx						
Aktive Kontrolle[e] (32)	1 (3,1)	6,947 [0,689; 70,052]	0,0985	5,517 [0,684; 44,491]	0,141 [-0,009; 0,291]	0,0933
Burosumab (29)	5 (17,2)					
PT Rhinorrhoe						
Aktive Kontrolle[e] (32)	2 (6,3)	5,083 [0,892; 28,977]	0,0665	3,862 [0,871; 17,119]	0,179 [0,002; 0,356]	0,0723
Burosumab (29)	7 (24,1)					

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
SOC Erkrankungen der Haut und des Unterhautzellgewebes						
Aktive Kontrolle[e] (32)	4 (12,5)	4,568 [1,145; 18,220]	0,0320	3,034 [1,085; 8,484]	0,254 [0,044; 0,465]	0,0357
Burosumab (29)	11 (37,9)					
PT Ausschlag						
Aktive Kontrolle[e] (32)	2 (6,3)	1,317 [0,166; 10,417]	0,7909	1,655 [0,297; 9,217]	0,041 [-0,098; 0,180]	0,6615
Burosumab (29)	3 (10,3)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; TEAE: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 64): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*.

c: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

d: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.

e: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

Alle berichteten *TEAE* auf SOC und PT Ebene (bei ≥ 10 % der Patienten in mindestens einer Behandlungsgruppe) waren mild (Grad 1) oder moderat (Grad 2). Die häufigsten *TEAE* 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).

Auf Ebene der SOC zeigte sich ein statistisch signifikanter Unterschied für *Allgemeine Erkrankungen und Beschwerden am Verabreichungsort* (RR_{W64} [95% KI]: 3,448; [1,860; 6,394]; p < 0,0001) sowie die dazugehörigen PT *Erythem an der Injektionsstelle* (RR_{W64} [95% KI]: 20,936; [1,270; 343,87]; p = 0,0006) und PT *Reaktion an der Injektionsstelle* (RR_{W64} [95% KI]*:* 16,528; [0,984; 276,73]; p = 0,0036). Hierbei ist jedoch zu beachten, dass eine subkutan verabreichte Prüfsubstanz (Burosumab) mit einer oral verabreichten Substanz (Phosphat) verglichen wird. Eine valide Aussage zum Sicherheitsprofil von Burosumab im Vergleich zu oralem Phosphat lässt sich daher anhand der SOC *Allgemeine Erkrankungen und Beschwerden am Verabreichungsort* und den dazugehörigen PT nicht treffen (Tabelle 4-51 und Tabelle 4-52).

Weitere statistisch signifikante Unterschiede zuungunsten von Burosumab zeigten sich für die SOC *Erkrankungen der Haut und des Unterhautzellgewebes* (RR_{W64} [95% KI]: 3,034; [1,085; 8,484]; p = 0,0357). Der Unterschied besteht ausschließlich auf Ebene der SOC; statistisch signifikante Unterschiede auf Ebene der dazugehörigen PT konnten nicht festgestellt werden (Tabelle 4-51 und Tabelle 4-52).

Für die SOC *Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen* (RR_{W64} [95% KI]: 5,517; [1,317; 23,115]; p = 0,0086) und den dazugehörigen PT *Vitamin D erniedrigt* (RR_{W64} [95% KI]: 6,621; [0,847; 51,764]; p = 0,0461) zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Burosumab. Ein Vitamin D-Mangel stellt eine charakteristische Manifestation der XLH-Erkrankung dar und wird durch einen FGF23-stimulierten Abbau von Calcitriol in der Niere bewirkt. Alle Kinder erhielten vor Studienbeginn neben oralem Phosphat daher auch für mindestens 12 Monate aktives Vitamin D. Auch während der Studie wurde im Kontrollarm die Gabe von aktivem Vitamin D fortgesetzt. In der Burosumab-Gruppe dagegen war die Gabe von aktivem Vitamin D während der Studie nicht erlaubt. Es ist daher nicht auszuschließen, dass durch das Absetzen von aktivem Vitamin D vor Studienbeginn der Vitamin-D-Mangel in der Burosumab-Gruppe zunächst verstärkt wurde und sich erst im Verlauf der Burosumab-Behandlung normalisierte. Unterstützt wird dies ebenfalls durch die Beobachtung, dass in der Burosumab-Gruppe nach der 64-wöchigen Behandlungsphase keine weiteren *TEAE* der Kategorie *Vitamin D erniedrigt* berichtet wurden. Eine Assoziation des Vitamin D-Mangels mit der Burosumab-Behandlung konnte nicht festgestellt werden (Tabelle 4-51 und Tabelle 4-52).

Für die PT *Fieber* (RR_{W64} [95% KI]: 2,943; [1,332; 6,499]; p = 0,0039) und *Husten* (RR_{W64} [95% KI]: 2,759; [1,237; 6,154]; p = 0,0083) und die SOC *Erkrankungen der Atemwege des Brustraums und Mediastinums* (RR_{W64} [95% KI]: 2,575; [1,416; 4,681]; p = 0,0008) zeigte sich ebenfalls jeweils ein statistisch signifikanter Unterschied zuungunsten von Burosumab gegenüber Phosphat (Tabelle 4-51 und Tabelle 4-52). Hierbei handelt es sich um typische Kindheitssymptome. Die Gründe für die unterschiedlichen Häufigkeiten zwischen den Behandlungsgruppen sind nicht vollständig verstanden. Es ist jedoch hervorzuheben, dass keines dieser Ereignisse vom Prüfarzt als schwerwiegendes Ereignis eingestuft und auch nicht als mit der Burosumab-Behandlung zusammenhängend bewertet wurde mit der Ausnahme von 4 Patienten, die von Fieber am Tag der Burosumab-Injektion berichteten. 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 in die Studie 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 UE 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 [80]. 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 Überempfindlichkeitsereignisse (z. B. Hautausschlag), die in der Burosumab-Gruppe beobachtet wurden.

Schwerwiegende *TEAE* und *TEAE* des Grads 3 traten auf Ebene der MedDRA SOC und PT in keinem Fall bei 5 % oder mehr der Patienten in einer Behandlungsgruppe auf, sodass von einer zusätzlichen Darstellung im Abschnitt 4.3.1.3.1.7 abgesehen wird. Darüber hinaus waren keine *TEAE* des Grads 4 zu verzeichnen.

Zahnmedizinische Ereignisse (mündliche Befragung)

Tabelle 4-53: Zusammenfassung der *zahnmedizinischen Untersuchungen* nach Beginn der Behandlung

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
Baseline zu Woche 40						
Jegliche zahnmedizinische Ereignisse						
Aktive Kontrolle[e] (32)	5 (15,6)	3,391 [0,919; 12,504]	0,0661	2,207 [0,855; 5,699]	0,189 [-0,025; 0,402]	0,1363
Burosumab (29)	10 (34,5)					
Zahnextraktion (aufgrund von Abszess oder Karies)						
Aktive Kontrolle[e] (32)	2 (6,3)	2,257 [0,302; 16,865]	0,4212	1,655 [0,297; 9,217]	0,041 [-0,098; 0,180]	0,6615
Burosumab (29)	3 (10,3)					
Zahnabszess						
Aktive Kontrolle[e] (32)	2 (6,3)	7,817 [1,295; 47,179]	0,0257	4,414 [1,019; 19,115]	0,213 [0,030; 0,396]	0,0371
Burosumab (29)	8 (27,6)					
Wurzelkanal						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
Zahnkaries						
Aktive Kontrolle[e] (32)	1 (3,1)	6,027 [0,454; 80,029]	0,1697	3,310 [0,364; 30,076]	0,072 [-0,054; 0,198]	0,3385
Burosumab (29)	3 (10,3)					
Baseline zu Woche 64						
Jegliche zahnmedizinische Ereignisse						
Aktive Kontrolle[e] (32)	5 (15,6)	4,537 [1,249; 16,480]	0,0224	2,648 [1,061; 6,607]	0,258 [0,039; 0,477]	0,0438
Burosumab (29)	12 (41,4)					
Zahnextraktion (aufgrund von Abszess oder Karies)						
Aktive Kontrolle[e] (32)	3 (9,4)	2,498 [0,492; 12,679]	0,2639	1,839 [0,481; 7,025]	0,079 [-0,092; 0,249]	0,4601
Burosumab (29)	5 (17,2)					
Zahnabszess						
Aktive Kontrolle[e] (32)	2 (6,3)	11,255 [1,921; 65,954]	0,0082	5,517 [1,317; 23,115]	0,282 [0,090; 0,475]	0,0086
Burosumab (29)	10 (34,5)					
Wurzelkanal						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		7,7 [0,415; 143,00]	0,103 [-0,007; 0,214]	0,1015
Burosumab (29)	3 (10,3)					
Zahnkaries						
Aktive Kontrolle[e] (32)	1 (3,1)	11,853 [1,105; 127,107]	0,0414	6,621 [0,847; 51,764]	0,176 [0,016; 0,335]	0,0461
Burosumab (29)	6 (20,7)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; TEAE: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 40 und 64): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*.

c: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
d: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt. e: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						

Mit der XLH geht regelhaft eine schlechte Zahngesundheit einher [10]. Über Zahnabszesse, übermäßige Karies, Wurzelkanal oder Zahnextraktion (aufgrund von Karies oder Abszess) wurde 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 eines der folgenden Ereignisse aufgetreten ist: Zahnabszess, Zahnextraktion (durch Abszess oder Karies), Wurzelbehandlung und/oder Gingivitis.

Die Ergebnisse dieser mündlichen Prüfung wurden getrennt von den TEAE tabellarisch dargestellt. Der Prüfarzt beurteilte, ob es sich um Zahnbeschwerden oder ein TEAE 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. Zahnabszesse traten bei 10 Studienteilnehmern der Burosumab-Gruppe und bei 2 Studienteilnehmern der aktiven Kontrollgruppe auf (RR_{W64} [95% KI]: 5,517; [1,317; 23,115]; p = 0,0086). Zahnkaries war dagegen bei 6 Studienteilnehmern der Burosumab-Gruppe und bei einem Studienteilnehmer der aktiven Kontrollgruppe zu Woche 64 vorhanden (RR_{W64} [95% KI]: 6,621; [0,847; 51,764]; p = 0,0461) (Tabelle 4-53). Die Zahnbeschwerden traten in beiden Gruppen meist bei den 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 (Tabelle 4-53).

Dentale unerwünschte Ereignisse (TEAE)

Tabelle 4-54: Ergebnisse der *dentalen unerwünschten Ereignisse nach Beginn der Behandlung (TEAE)*

UX023-CL301 Population (N) (Safety)	Patienten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
Baseline zu Woche 40						
Jegliche dentale TEAE						
Aktive Kontrolle[e] (32)	9 (28,1)	2,138 [0,714, 6,406]	0,1710	1,594 [0,803; 3,163]	0,167 [-0,072, 0,406]	0,1943
Burosumab (29)	13 (44,8)					

UX023-CL301 Population (N) (Safety)	Patien-ten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
Zahnkaries						
Aktive Kontrolle[e] (32)	2 (6,3)	4,352 [0,629; 30,085]	0,1333	2,759 [0,579; 13,140]	0,110 [-0,051; 0,271]	0,2414
Burosumab (29)	5 (17,2)					
Zahnabszess						
Aktive Kontrolle[e] (32)	3 (9,4)	2,714 [0,581; 12,690]	0,2000	2,207 [0,607; 8,030]	0,113 [-0,066; 0,292]	0,2870
Burosumab (29)	6 (20,7)					
Zahnschmerzen						
Aktive Kontrolle[e] (32)	1 (3,1)	2,499 [0,199; 31,442]	0,4719	2,207 [0,211; 23,077]	0,038 [-0,072; 0,148]	0,6003
Burosumab (29)	2 (6,9)					
Zahnhypoplasie						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		5,5 [0,275; 110,01]	0,069 [-0,023; 0,161]	0,2219
Burosumab (29)	2 (6,9)					
Lockerer Zahn						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		3,3 [0,140; 77,953]	0,034 [-0,032; 0,101]	0,4754
Burosumab (29)	1 (3,4)					
Periodontitis						
Aktive Kontrolle[e] (32)	1 (3,1)	1,236 [0,064; 23,811]	0,8865	1,103 [0,072; 16,849]	0,003 [-0,086; 0,093]	1,0000
Burosumab (29)	1 (3,4)					
Zahnverlust						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		3,3 [0,140; 77,953]	0,034 [-0,032; 0,101]	0,4754
Burosumab (29)	1 (3,4)					

UX023-CL301 Population (N) (Safety)	Patien-ten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
Maloklusion						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		3,3 [0,140; 77,953]	0,034 [-0,032; 0,101]	0,4754
Burosumab (29)	1 (3,4)					
Zahnverfärbungen						
Aktive Kontrolle[e] (32)	1 (3,1)	n.b.		0,367 [0,016; 8,661]	-0,031 [-0,092; 0,030]	< 0,0001
Burosumab (29)	0 (0)					
Baseline zu Woche 64						
Jegliche dentale TEAE						
Aktive Kontrolle[e] (32)	10 (31,3)	2,479 [0,838; 7,336]	0,0991	1,655 [0,888; 3,085]	0,205 [-0,038; 0,447]	0,1245
Burosumab (29)	15 (51,7)					
Zahnschmerzen						
Aktive Kontrolle[e] (32)	1 (3,1)	5,505 [0,531; 57,114]	0,1498	4,414 [0,523; 37,250]	0,107 [-0,033; 0,246]	0,1816
Burosumab (29)	4 (13,8)					
Zahnabszess						
Aktive Kontrolle[e] (32)	3 (9,4)	4,259 [0,938; 19,346]	0,0602	2,943 [0,862; 10,048]	0,182 [-0,009; 0,374]	0,0964
Burosumab (29)	8 (27,6)					
Zahnhypoplasie						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		5,5 [0,275; 110,01]	0,069 [-0,023; 0,161]	0,2219
Burosumab (29)	2 (6,9)					
Zahnkaries						
Aktive Kontrolle[e] (32)	2 (6,3)	7,734 [1,360; 43,987]	0,0219	4,966 [1,168; 21,114]	0,248 [0,060; 0,436]	0,0183
Burosumab (29)	9 (31,0)					

UX023-CL301 Population (N) (Safety)	Patien-ten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
Lockerer Zahn						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		3,3 [0,140; 77,953]	0,034 [-0,032; 0,101]	0,4754
Burosumab (29)	1 (3,4)					
Periodontitis						
Aktive Kontrolle[e] (32)	1 (3,1)	1,236 [0,064; 23,811]	0,8865	1,103 [0,072; 16,849]	0,003 [-0,086; 0,093]	1,000
Burosumab (29)	1 (3,4)					
Zahnverlust						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		3,3 [0,140; 77,953]	0,034 [-0,032; 0,101]	0,4754
Burosumab (29)	1 (3,4)					
Maloklusion						
Aktive Kontrolle[e] (32)	1 (3,1)	n.b.		0,367 [0,016; 8,661]	-0,031 [-0,092; 0,030]	1,000
Burosumab (29)	0 (0)					
Zahnfehlstellung						
Aktive Kontrolle[e] (32)	1 (3,1)	n.b.		0,367 [0,016; 8,661]	-0,031 [-0,092; 0,030]	1,000
Burosumab (29)	0 (0)					
Zahnverfärbung						
Aktive Kontrolle[e] (32)	1 (3,1)	n.b.		0,367 [0,016; 8,661]	-0,031 [-0,092; 0,030]	1,000
Burosumab (29)	0 (0)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; TEAE: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 40 und 64): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*.

c: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

d: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.

UX023-CL301 Population (N) (Safety)	Patien-ten n (%)	OR[a] [95% KI]	p-Wert[a]	RR[b,d] [95% KI]	ARD[b,d] [95% KI]	p-Wert[c]
e: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D						

Dental-assoziierte-TEAE 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 während der 64-wöchigen aktiv kontrollierten Phase der Studie berichtet (Tabelle 4-54). Dentale TEAE wurden bei allen Kindern dokumentiert, die nach Aussage der mündlichen zahnmedizinischen Befragung an Zahnbeschwerden litten ($n_{(Burosumab)}$ = 12; $n_{(aktive\ Kontrolle)}$ = 5). Ebenfalls wurden dentale TEAE für Kinder notiert, die in der mündlichen zahnmedizinischen Befragung keine Zahnbeschwerden angegeben hatten ($n_{(Burosumab)}$ = 4; $n_{(aktive\ Kontrolle)}$ = 5). Nicht berücksichtigt wurde hierbei das dentale TEAE „Zahnen“, welches bei einem Kind in der Burosumab-Gruppe aufgetreten war. Von diesen Kindern mit TEAE 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.

Zahnärztliche Komplikationen traten in der Burosumab-Gruppe generell häufiger auf. Dabei handelte es sich jedoch fast ausschließlich um numerische Unterschiede. Die einzige zahnärztliche Komplikation mit einem signifikanten Unterschied zuungunsten von Burosumab stellt Zahnkaries dar. Hier betrug zu Woche 64 das RR 4,966 ([95% KI]: [1,168; 21,114]; p = 0,0183) (Tabelle 4-54). Allerdings waren hiervon lediglich 9 Teilnehmer der Burosumab-Gruppe betroffen.

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 können. Das Fehlen von TEAE bei der Substitutionstherapie sowie die höhere Inzidenz der am häufigsten gemeldeten TEAE 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.

Unerwünschte Ereignisse von besonderem Interesse nach Beginn der Behandlung

Tabelle 4-55: Ergebnisse zu *unerwünschten Ereignissen von besonderem Interesse nach Beginn der Behandlung*

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
Baseline zu Woche 40						
Reaktionen an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		27,5 [1,70; 444,66]	0,414 [0,235; 0,593]	< 0,0001
Burosumab (29)	12 (41,4)					
Hypersensitivität						
Aktive Kontrolle[e] (32)	2 (6,3)	7,076 [1,206; 41,508]	0,0308	4,966 [1,168; 21,114]	0,248 [0,060; 0,436]	0,0183
Burosumab (29)	9 (31,0)					
Baseline zu Woche 64						
Reaktionen an der Injektionsstelle						
Aktive Kontrolle[e] (32)	0 (0)	n.b.		34,1 [2,132; 545,51]	0,517 [0,335; 0,699]	< 0,0001
Burosumab (29)	15 (51,7)					
Hypersensitivität						
Aktive Kontrolle[e] (32)	6 (18,8)	2,853 [0,744; 10,946]	0,1240	2,023 [0,857; 4,773]	0,192 [-0,031; 0,414]	0,1522
Burosumab (29)	11 (37,9)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; TEAE: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 40 und 64): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*.

c: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

d: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.

e: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

Reaktionen an der Injektionsstelle und damit Ereignisse von besonderem Interesse, die aufgrund der unterschiedlichen Verabreichungsform der Prüfsubsatz und der aktiven Kontrolle nur bei Patienten in der Burosumab-Gruppe untersucht wurden, wurden zu Woche 40 von 12

(41,4 %) und zu Woche 64 von 15 Kindern (51,7 %) (RR_{W64} [95% KI]: 34,1; [2,132; 545,51]; p < 0,0001) berichtet (Tabelle 4-55). 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 %) zu Woche 64 in der Burosumab-Gruppe und bei 6/32 Patienten (19 %) in der aktiven Kontrollgruppe beobachtet wurde (RR_{W64} [95% KI]: 2,023; [0,857; 4,773]; p = 0,1522) (Tabelle 4-55).

Es traten keine schwerwiegenden TEAE von besonderem Interesse auf oder TEAE von besonderem Interesse Grad ≥ 3.

Bis zum Datenschnitt erlebte kein Kind eine Hyperphosphatämie, eine ektopische Mineralisierung oder ein Restless-Legs-Syndrom, wodurch das gute Sicherheitsprofil von Burosumab bestätigt wird.

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.

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 Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1. [16]

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

[16] unbesetzt

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.

- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.

- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.

- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.

- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.

- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.

- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.

- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.

Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-56: Matrix der durchgeführten Subgruppenanalysen

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

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.

-: Subgruppe erfüllt nicht die Kriterien in Abschnitt 4.3.1.3.2

a: Subgruppenanalysen auf SOC und PT-Ebene waren nicht a priori definiert, sondern wurden *post-hoc* durchgeführt.

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

Die Subgruppen Region und *PHEX*-Mutation waren im Studienprotokoll präspezifiziert, erfüllten aber nicht die in Abschnitt 4.3.1.3.2 aufgeführten Kriterien, sodass auf eine Berechnung und Darstellung der Subgruppen im Dossier verzichtet wird.

Der Endpunkt *Unerwünschte Ereignisse von besonderem Interesse* erfüllte nicht die in Abschnitt 4.3.1.3.2 erforderlichen Kriterien zur Darstellung von Subgruppenanalysen, sodass von einer Berechnung der Interaktions p-Werte und Subgruppen abgesehen wird.

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 ist folgendes zu berücksichtigen:

- 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 zunächst die Ergebnisse der Interaktionstests für alle untersuchten Subgruppen dargestellt. Entsprechend dem strukturierten, mehrstufigen Vorgehen zur Bewertung einer möglichen Effektmodifikation erfolgt anschließend im Abschnitt 4.3.1.3.2.1 eine ausführliche Ergebnisdarstellung für diejenigen Subgruppen, bei denen ein signifikanter Interaktionstest ($p < 0,05$) vorliegt. Die Ergebnisse der Subgruppen mit einem nicht signifikantem Interaktionstest ($p \geq 0,05$) befinden sich im Anhang 4-G.

Stellen Sie anschließend in Tabelle 4-57 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ($p < 0,05$) Interaktionsterme.

Die folgende Tabelle 4-57 zeigt in einer Übersicht die Ergebnisse sämtlicher Interaktionstests, die für Subgruppenanalysen der UX023-CL301 durchgeführt wurden.

Tabelle 4-57: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für UX023-CL301

OUTCOMES									
Subgruppe / **Endpunkt**	Rachitis-Schwere (≤ 2,5 vs. > 2,5)			Alter (< 5 Jahre vs. ≥ 5 Jahre)			Geschlecht (weiblich vs. männlich)		
Datenschnitt	W24	W40	W64	W24	W40	W64	W24	W40	W64
Morbidität									
RGI-C Global Score	-[b]	0,9828	0,7551	-[b]	0,8392	0,8777	-[b]	0,3834	0,7074
RGI-C Global Score = + 2,0	-[b]	0,9784	0,9744	-[b]	0,7035	0,8945	-[b]	0,3925	0,9685
RSS	-[b]	0,9910	0,9528	-[b]	0,2974	0,1872	-[b]	0,9794	0,8650
Körpergröße (Standhöhe/Liegelänge z-score)	-[b]	0,7451	0,9458	-[b]	0,1254	0,1816	-[b]	0,6319	0,8412
Sitzhöhe	-[b]	0,2362	0,6198	-[b]	0,2708	0,4151	-[b]	0,5377	0,6838
6MWT [m]	-[b]	0,7670	0,6173	-[c]	-[c]	-[c]	-[b]	0,6273	0,7328
6MWT [vorhergesagter Prozentsatz]	-[b]	0,9053	0,7956	-[c]	-[c]	-[c]	-[b]	0,7356	0,8572
6MWT [m][a]	-[b]	0,7985	0,9762	-[c]	-[c]	-[c]	-[b]	0,2726	0,3538
6MWT [vorhergesagter Prozentsatz][a]	-[b]	0,8130	0,9110	-[c]	-[c]	-[c]	-[b]	0,3016	0,3733
Serumphosphat	-[b]	0,1432	0,1612	-[b]	0,2802	0,2669	-[b]	0,0754	0,1412
PROMIS – Pain Domain	0,6782	0,4626	0,7860	-[c]	-[c]	-[c]	0,1427	0,7002	0,3785
PROMIS – Physical Function and Morbidity Domain	0,9201	0,5295	0,5334	-[c]	-[c]	-[c]	0,7513	0,8162	0,8180

OUTCOMES									
Subgruppe / **Endpunkt**	Rachitis-Schwere (≤ 2,5 vs. > 2,5)			Alter (< 5 Jahre vs. ≥ 5 Jahre)			Geschlecht (weiblich vs. männlich)		
Datenschnitt	W24	W40	W64	W24	W40	W64	W24	W40	W64
PROMIS – Fatigue Domain	0,1824	0,1683	0,3849	-[c]	-[c]	-[c]	0,7074	0,7523	0,7071
FPS-R	0,2650	0,2424	0,6453	-[c]	-[c]	-[c]	0,0998	**0,0043**	**0,0004**
Lebensqualität									
SF-10 – Physical Health Summary Score	0,5634	0,2775	0,5106	-[c]	-[c]	-[c]	0,3839	0,8104	0,7502
SF-10 – Psychosocial Summary Score	0,6836	0,7624	0,7028	-[c]	-[c]	-[c]	0,4879	0,2667	0,2360
Unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)									
TEAE	-[b]	0,8048	0,8184	-[b]	0,0688	0,0518	-[b]	0,8708	0,9971
Schwerwiegende TEAE	-[b]	0,9954	0,8150	-[b]	0,9965	0,9951	-[b]	0,9956	0,9000
TEAE Grad 3 – 4	-[b]	0,9958	0,9950	-[b]	0,9961	0,9957	-[b]	0,2061	0,2703
TEAE nach MedDRA SOC und PT (Inzidenz ≥ 10% in mindestens einer Behandlungsgruppe)									
SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums	-[b]	0,2821	0,8396	-[b]	0,9119	0,6248	-[b]	0,9331	0,5080
PT Husten	-[b]	-[d]	0,5564	-[b]	-[d]	0,6172	-[b]	-[d]	0,8280
SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	-[b]	0,1423	0,1349	-[b]	0,6233	0,3642	-[b]	0,0899	0,7541
PT Fieber	-[b]	0,3627	0,2400	-[b]	0,5778	0,5847	-[b]	**0,0474**	0,0685

OUTCOMES									
Subgruppe / **Endpunkt**	Rachitis-Schwere (≤ 2,5 vs. > 2,5)			Alter (< 5 Jahre vs. ≥ 5 Jahre)			Geschlecht (weiblich vs. männlich)		
Datenschnitt	W24	W40	W64	W24	W40	W64	W24	W40	W64
SOC Erkrankungen des Gastrointestinaltrakts	-[b]	-[d]	0,7599	-[b]	-[d]	0,5519	-[b]	-[d]	0,9556
Jegliche zahnmedizinische Ereignisse	-[b]	-[d]	0,9009	-[b]	-[d]	0,1642	-[b]	-[d]	0,4179
TEAE von besonderem Interesse									
Hypersensitivität	-[b]	-[d]	-[d]	-[b]	-[d]	0,4174	-[b]	-[d]	0,2808
Reaktionen an der Einstichstelle	-[b]	-[e]	-[e]	-[b]	-[e]	-[e]	-[b]	-[e]	-[e]

RGI-C: *Radiographic Global Impression of Change*; RSS: *Rickets Severity Score*; 6MWT: 6-Minuten-Gehdistanz Test; PROMIS: *Patient Reported Outcomes Measurement Information System*; FPS-R: *Face Pain Scale – Revised*; SF-10: *Short Form (SF)-10 Health Survey*; *TEAE: Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung)

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.

b: Erhebungszeitpunkt nicht relevant für die Nutzenbewertung

c: Endpunkt wurde nur bei Kindern ≥ 5 Jahre erhoben.

d: Erfüllt nicht die Anforderungen an die Subgruppenanalysen in Abschnitt 4.3.1.3.2.

e: Es traten keine Ereignisse im Kontrollarm auf.

Stellen Sie schließlich alle Subgruppenergebnisse dar.

Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich“ und „männlich“) als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig“ und „hoch“) eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.

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.

4.3.1.3.2.1 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 Patientenanzahl pro Behandlungsarm handelt.

4.3.1.3.2.1.1 Ergebnisse für präspezifizierten Subgruppenanalysen

Für keine der präspezifizierten Subgruppenanalysen zeigte sich im Interaktionstest eine Effektmodifikation (alle $p > 0,05$). Aufgrund dessen werden die präspezifizierten Subgruppenanalysen ausschließlich in Anhang 4-G dargestellt.

4.3.1.3.2.1.2 Ergebnisse für post-hoc durchgeführte Subgruppenanalysen

4.3.1.3.2.1.2.1 Endpunkt: FPS-R

Tabelle 4-58: Subgruppenanalysen – Ergebnisse für *FPS-R* aus RCT mit dem zu bewertenden Arzneimittel

Intervention	**Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder ≥ 5 Jahre)**							
(N)	**Baseline**	**Woche 24**	**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 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=11)	11 0,7 (1,35)	11 0,2 (0,60)	11 0,0 (0,0)	11 0,0 (0,0)	11 -0,7 (0,41) -0,50 (0,024)	11 -0,7 (0,41) -0,50 (0,024)	1,19 [-0,14; 2,52] 0,0793 -	1,19 [0,43; 1,95] 0,0022 H'g = 1,330 [0,177; 2,484]
Burosumab (N=5)	5 0,0 (0,0)	5 1,6 (2,61)	5 1,2 (1,79)	5 1,2 (1,10)	5 1,2 (0,80) 0,69 (0,691)	5 1,2 (0,49) 0,69 (0,400)		
Männliche Patienten								
Aktive Kontrolle[b] (N=9)	9 0,7 (1,00)	9 0,7 (1,41)	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=10)	10 0,6 (1,35)	10 0,2 (0,63)	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)		

Intervention	**Face Pain Scale – Revised (FPS-R) (Baseline-Alter der Kinder ≥ 5 Jahre)**							
(N)	**Baseline**	**Woche 24**	**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** **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]**

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 GEE-Modell zu Woche 40 und Woche 64.

b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

GEE: Behandlungsgruppe, Behandlungsvisiten, Interaktion der Behandlungsgruppe mit den Behandlungsvisiten und Geschlecht als Faktoren, sowie der Baseline FPS-R als Kovariate

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_{W40} = 0,0043 bzw. p_{W64} = 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: $\text{Differenz}_{\text{weiblich}}$: 1,19; p = 0,0022 und $\text{Differenz}_{\text{männlich}}$ -1,05; p = 0,0284) (Tabelle 4-58) (Abbildung 9).

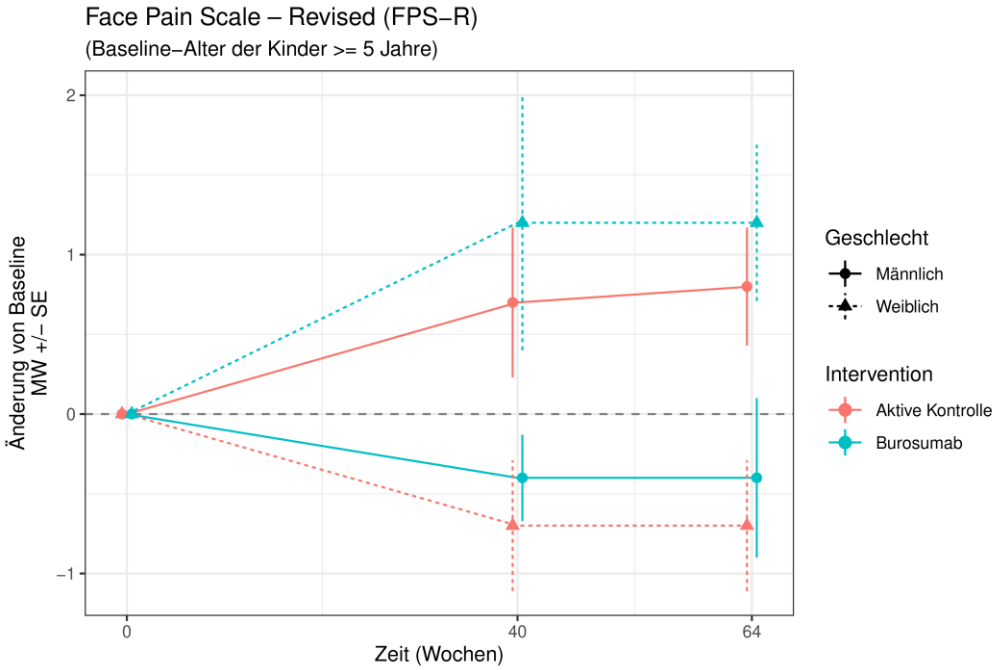

Abbildung 9: Subgruppenanalyse *FPS-R*

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) und PT Fieber 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.

4.3.1.3.2.1.2.2 Endpunkt: PT Fieber

Tabelle 4-59: Subgruppenanalysen – Ergebnisse für *PT Fieber* aus RCT mit dem zu bewertenden Arzneimittel

UX023-CL301 Population (N) (Safety)	**Patienten n (%)**	**OR[a] [95% KI]**	**p-Wert[a]**	**RR[b,d] [95% KI]**	**ARD[b,d] [95% KI]**	**p-Wert[c]**
Subgruppe: Geschlecht - Baseline zu Woche 40						
Männliche Patienten						
Aktive Kontrolle[e] (14)	0 (0)	n.b.		16,071 [1,009; 256,10]	0,538 [0,267; 0,809]	0,0019
Burosumab (13)	7 (53,8)					
Weibliche Patienten						
Aktive Kontrolle[e] (18)	4 (22,2)	1,372 [0,238; 7,909]	0,7152	1,688 [0,578; 4,925]	0,153 [-0,152; 0,458]	0,4569
Burosumab (16)	6 (37,5)					

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; [95% KI]: 95%-iges Konfidenzintervall; OR: *Odds Ratio*; RR: *Risk Ratio* (relatives Risiko); ARD: *Absolute Risk Difference* (Absolute Risikodifferenz); n.b.: nicht berechenbar

Ein RR bzw. OR < 1 zeigt einen Vorteil für Burosumab an und ein RR bzw. OR > 1 einen Vorteil für die aktive Kontrolle.

a: Logistische Regression (Woche 40 und 64): Baseline-Alter-Stratifizierungsfaktoren und Behandlungsgruppe als Faktoren, RSS Total Score als kontinuierliche Kovariate. Die Berechnung erfolgte *post-hoc*.

b: Das RR und die ARD wurden unadjustiert aus den zum jeweiligen Endpunkt gehörenden Vierfeldertafeln berechnet. Die zugehörigen 95% KI wurden mittels Normalverteilungsapproximation berechnet. Die Berechnung erfolgte *post-hoc*.

c: Der p-Wert wurde mittels Fishers exaktem Test berechnet.

d: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.

e: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D

Für das PT *Fieber* lässt sich eine potenzielle Effektmodifikation in der Subgruppe Geschlecht männliche Patienten zu Woche 40 feststellen (männliche Patienten: RR: 16,071; weibliche Patienten: RR: 1,688). Die Effektschätzer sind gleichgerichtet, unterscheiden sich aber in der Signifikanz (Männlich: p = 0,0019; Weiblich: p = 0,4569) (Tabelle 4-59). Der beobachtete Effekt ist zu Woche 64 am Ende der aktiv kontrollierten Behandlungsphase nicht mehr zu beobachten, sodass die Ableitung des Zusatznutzens ausschließlich aus den Ergebnissen für die Gesamtpopulation erfolgt.

4.3.1.4 Liste der eingeschlossenen Studien - RCT

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 4-60: Liste der zur Bewertung des Zusatznutzens eingeschlossenen Studien - RCT

Studie	Datenquellen
UX023-CL301 RCT	Studienberichte [26, 78]
	Studienprotokoll [25]
	SAP [27]
	Studienregistereinträge [68-71]
	Studienpublikation [66, 67]
	Dokumente von der Internetseite des G-BA [17, 65, 72-77]

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 sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der Internetseite des G-BA*
- *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 identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. ***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 4-61: 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 drei 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.**

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

Tabelle 4-62: 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 4-63: 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 4-64: 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 4-65: 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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

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.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

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).

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/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *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 4-66: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	**Zeitliche Parallelität der Gruppen**	**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**	**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

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

Tabelle 4-67: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

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

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 4-68: 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 4-69: 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.

4.3.2.2.3.2 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.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende 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).

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 sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *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

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

Tabelle 4-70: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

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

Nicht zutreffend.

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 4-71: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	**Operationalisierung**
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. 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.

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.

4.3.2.3.3.2 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.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

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).

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 [81]. Der Orphan-Drug-Status wird Arzneimitteln verliehen, die zur Behandlung einer seltenen chronischen oder lebensbedrohlichen Erkrankung eingesetzt werden, für die keine zufriedenstellende Therapie existiert. Nachweise gemäß § 35a Abs. 1 Satz 3 Nummer 2 und 3 SGB V über den medizinischen Nutzen und Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie (zVT) 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).Übersteigt der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten zwölf Kalendermonaten einen Betrag von 50 Millionen Euro, hat der pharmazeutische Unternehmer gemäß § 35a Absatz 1 Satz 12 SGB V innerhalb von drei Monaten nach Aufforderung durch den G-BA Nachweise gemäß 5. Kapitel § 5 Absatz 1 bis 6 VerfO des G-BA zu übermitteln.

Der G-BA hat mit Schreiben vom 21.10.2021 die Kyowa Kirin GmbH darüber informiert, dass der Umsatz mit Burosumab (Crysvita®) zu Lasten der GKV einen Betrag von 50 Millionen Euro überschritten hat. Maßgeblicher Zeitpunkt für die Einreichung der Nachweise gemäß § 35a Abs. 1 Satz 12 SGB V ist damit der 01.02.2022. Im vorliegenden Dossier werden nun die erforderlichen Nachweise des Zusatznutzens von Burosumab gegenüber der zVT dargestellt.

Für die erneute Bewertung von Burosumab nach Überschreitung der Umsatzgrenze von 50 Millionen Euro werden die Studiendaten zur abgeschlossenen Behandlungsphase (Woche 40 und Woche 64) der Phase-III-Zulassungsstudie UX023-CL301 mit einer Evidenzstufe Ib (RCT) herangezogen, die bereits in dem Nutzenbewertungsverfahren 2019-10-01-D-492 nach Fristablauf bewertet wurden [65].

In der Studie UX023-CL301 wurden die Patienten 1:1 mittels *Interactive Web Response System* (IWRS) auf die Behandlungsarme randomisiert. Die Stratifizierung erfolgte nach Rachitis-Schwere, Alter und Region. Eine Verblindung des klinischen Personals und der Studienteilnehmer war aufgrund der unterschiedlichen Verabreichungsmethoden und der individuellen und häufigen Dosisanpassung, die bei einer oralen Phosphat- und aktiven Vitamin-D-Therapie erforderlich sind, nicht möglich. Ausschließlich die Auswerter der Röntgenbildaufnahmen, die den RSS Total Score und RGI-C Score ermittelten, waren verblindet. Ebenso stellt die Serumphosphat-Konzentration durch die Messung am Gerät sowie die Körpergröße/Sitzhöhe ein geringes Verzerrungspotenzial dar. Sowohl bei der Messung des Endpunktes 6-Minuten-Gehdistanz Test (6MWT) als auch bei 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. Es fand zu allen Zeitpunkten eine ergebnisunabhängige Berichterstattung statt. Aufgrund dessen ist trotz des offenen Studiendesigns von einem niedrigen Verzerrungspotenzial und einer hohen Aussagesicherheit der Ergebnisse auszugehen. Die Validität und Patientenrelevanz der Nachweise sind ausführlich im Abschnitt 4.2.5.2 dargestellt.

Anhand des Methodenpapiers 6.0 des IQWiG sind die in diesem Dossier dargestellten Nachweise zum medizinischen Zusatznutzen von Burosumab als **Beleg** einzustufen [63].

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 einem Jahr und Jugendlichen mit XLH und röntgenologischem Nachweis einer Knochenerkrankung zugelassen [1]. Bis zur Zulassung von Burosumab gab 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 multizentrischen, randomisierten, offenen und aktiv kontrollierten Phase-III-Studie UX023-CL301 der Evidenzstufe 1b herangezogen [73].

Die medizinischen Experten in der mündlichen Anhörung (21.08.2018) bestätigten bereits damals den hohen medizinischen Stellenwert von Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren:

- *Prof. Dr. Partsch (DGKED):* „[…] Dieses Problem wird jetzt **kausal** durch **Burosumab** gelöst, und Sie können tatsächlich nach **wenigen Injektionen** sehen, dass die Kinder **deutlich besser laufen**, ein normaleres Gangbild haben. […] Nach wenigen Monaten ist dieses Gangbild deutlich verbessert. […] Insofern ist das auch funktionell ein **Quantensprung** in der **Therapie** der XLH.“
- *Prof. Dr. Schnabel (DGKED): „*Bei Burosumab ist die Applikation eine andere, man gibt es nicht oral, mehrfach täglich, sondern es wird einmal alle vierzehn Tage subkutan appliziert. Man sieht schon nach **ein, zwei subkutanen Injektionen** einen **Anstieg** der **tubulären Phosphatrückresorption**, die die Basis dafür ist, dass überhaupt dieses Medikament im Organismus wirkt. Ich verliere nicht mehr 30 – 40 % meines Phosphats, sondern nur noch 4 – 5 %, so wie es normalerweise auch physiologisch ist. Mit der verminderten Ausscheidung von Phosphat kommt es im Blut zum Anstieg und zum **Erreichen eines Serumphosphatwerts in dem unteren Normbereich**. Dieses Phosphat geht dann verstärkt in den Knochen und sorgt dafür, dass sich der **Knochenumsatz normalisiert**.*“*

Da die Studie UX023-CL301 noch nicht abgeschlossen 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 patientenrelevanten Endpunkten Serumphosphat und Rachitissymptomatik einen **beträchtlichen Zusatznutzen** sehe.

Nach Ablauf der Befristung wurden die Ergebnisse aus der abgeschlossenen Behandlungsphase (Woche 64) der multizentrischen, randomisierten, offenen und aktiv kontrollierten Phase-III-Studie UX023-CL301 eingereicht und nutzenbewertet (2019-10-01-D-492). Diese bestätigten die Erkenntnisse aus dem ersten Nutzenbewertungsverfahren und zeigten eindrucksvoll den beträchtlichen Zusatznutzen von Burosumab bei der Behandlung der XLH bei Kindern.

Als Grundlage für das dritte Nutzenbewertungsverfahren von Burosumab zur Behandlung von Kindern mit XLH, nach Überschreitung der Umsatzgrenze von 50 Millionen Euro, werden die

Daten dieser multizentrischen, randomisierten, offenen und aktiv kontrollierten Phase III-Studie UX023-CL301 abermals im Dossier aufgeführt und zur Beurteilung des Zusatznutzens von Burosumab zur Behandlung von Kindern ab einem Jahr und Jugendlichen mit XLH und röntgenologischem Nachweis einer Knochenerkrankung herangezogen.

Interpretation der Ergebnisse zu Morbidität

Die folgende Tabelle 4-72 zeigt in einer Übersicht die zugrunde liegenden Ergebnisse der Endpunktkategorie *Morbidität* bezüglich des Vergleichs der Behandlungsgruppen, mit denen das Ausmaß des Zusatznutzens der Burosumab-Behandlung abgeleitet werden konnte.

Tabelle 4-72: Zusammenfassung der Ergebnisse der Kategorie *Morbidität* zum medizinischen Zusatznutzen von Burosumab

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab** **Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab** **Woche 64**
Rachitissymptomatik		
Radiographic Global Impression of Change (RGI-C) Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	1,14 [0,83; 1,45] < 0,0001 H'g = 1,846 [1,246; 2,446]	1,02 [0,72; 1,33] < 0,0001 H'g = 1,600 [1,023; 2,178]
Rickets Severity Score (RSS)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-1,34 [-1,74; -0,94] < 0,0001 H'g = -1,399 [-1,964; -0,833]	-1,21 [-1,59; -0,83] < 0,0001 H'g = -1,119 [-1,710; -0,529]
Körpergröße		
Stehhöhe/Liegelänge [Z-Score]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	0,12 [0,01; 0,24] 0,0408 H'g = 0,537 [0,021; 1,053]	0,14 [0,00; 0,29] 0,0490 H'g = 0,490 [-0,025; 1,005]
Sitzhöhe (cm)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-0,79 [-1,98; 0,39] 0,1901 -	0,08 [-1,58; 1,75] 0,9213 -

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64**
Gehfähigkeit[c]		
6-Minuten-Gehdistanz Test[c] (6MWT) [in Meter]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	38,30 [-2,38; 78,98] 0,0650 -	45,55 [2,09; 89,02] 0,0399 H'g = 0,679 [-0,038; 1,396]
6-Minuten-Gehdistanz Test[c] (6MWT) Gesamtrate [Prozentsatz]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	6,19 [-0,68; 13,06] 0,0775 -	7,27 [0,01; 14,52] 0,0496 H'g = 0,703 [-0,015; 1,422]
Serumphosphat		
Serumphosphat-Konzentration [mg/dl]		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	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]
Schmerzerhebung		
Patient Reported Outcomes Measurement Information System (PROMIS) Pediatric Pain Interference Domain Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-4,91 [-9,19; -0,63] 0,0245 H'g = -0,697 [-1,390; -0,008]	-2,26 [-6,61; 2,09] 0,3091 -
Patient Reported Outcomes Measurement Information System (PROMIS) Physical Function Mobility Domain Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	2,68 [-0,53; 5,88] 0,1014 -	1,90 [-1,80; 5,59] 0,3145 -
Patient Reported Outcomes Measurement Information System (PROMIS) Fatigue Domain Score		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-3,20 [-7,79; 1,39] 0,1724 -	-1,08 [-6,21; 4,06] 0,6810 -
Face Pain Scale – Revised (FPS-R)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	-0,00 [-0,80; 0,79] 0,9905 -	0,05 [-0,58; 0,68] 0,8786 -

Patientenrelevante Endpunkte	UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40	UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64
(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 4.2.5.2 und 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 unvollständigen Datensatz aufwiesen (zu Baseline und zu Woche 40). Eine detaillierte Beschreibung erfolgt in den Abschnitten 4.2.5.4 und 4.3.1.3.1.3.		

Mortalität

Während der 64-wöchigen aktiv kontrollierten Phase der Studie UX023-CL301 traten keine Todesfälle in den Behandlungsgruppen (Burosumab-Gruppe und aktiven Kontrollgruppe) auf.

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 die 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) (Tabelle 4-72). Somit 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 und Sitzhöhe

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 (Tabelle 4-72).

Zusätzlich wurde in der Studie UX023-301 die absolute Sitzhöhe in Zentimetern analysiert. Im Gegensatz zur Körpergröße erfolgte keine Adjustierung der Sitzhöhe für Alter und Geschlecht, sodass hier eine potenzielle Verzerrung der Ergebnisse nicht auszuschließen ist. Der Vollständigkeit halber werden die Analysen dennoch im vorliegenden Dossier dargestellt. Die Studienteilnehmer in der Burosumab-Gruppe und in der aktiven Kontrollgruppe wiesen eine ähnliche Zunahme der Sitzhöhe über die 64-wöchige randomisierte aktiv kontrollierte Phase der Studie auf (Behandlungsdifferenz$_{W64}$: 0,08; p = 0,9213) (Tabelle 4-72). Hierbei ist jedoch anzumerken, dass sich die durch die chronische Hypophosphatämie bedingten Skelettanomalien bei Kindern mit XLH überwiegend anhand von Verformungen der unteren Extremitäten zeigen und zur Ausbildung von O- (*Genua vara*) oder X-Beinen (*Genua valga*) führen. Das Messen der Sitzhöhe berücksichtigt diese Skelettanomalien nur unzureichend und führt daher möglicherweise zu einer Unterschätzung des beobachteten Effekts von Burosumab auf das Skelettwachstum.

Aufgrund dessen stellt insbesondere bei XLH-Patienten die Stehhöhe bzw. die Liegelänge die sensitivere Messmethode dar, um das Ausmaß der Burosumab-Behandlung auf das Wachstum von XLH-Patienten zu analysieren.

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 4-72 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 ($p_{absolut} = 0{,}0399$; $p_{relativ} = 0{,}0496$) auf (Tabelle 4-72). 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 zu Baseline eine schlechtere Gehfähigkeit im Vergleich zur Kontrollgruppe zeigte.

Serumphosphat

Mit den zwei Phase-II-Studien (UX023-CL201 und UX023-CL205) konnte bereits im ersten Nutzenbewertungsverfahren eindrucksvoll gezeigt werden, dass sich durch die Burosumab-Behandlung bereits innerhalb kurzer Zeit die Serumphosphat-Homöostase wiederherstellen lässt. Mit den Ergebnissen der abgeschlossenen Behandlungsphasen der aktiv kontrollierten Phase-III-Studie UX023-CL301 werden diese Erkenntnisse aus den pivotalen Studien 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-Studie konnte innerhalb der Burosumab-Gruppe ein schneller Anstieg der Serumphosphat-Konzentration in den altersspezifischen Normbereich erreicht und langfristig bis Woche 64 gehalten werden. Bei der aktiven Kontrollgruppe dagegen wurde lediglich ein geringer 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 von Burosumab bei der Behandlung der XLH (Tabelle 4-72). Burosumab ist damit die erste kausale Therapie zur Behandlung der XLH, die die zugrunde liegende Pathophysiologie der FGF23-induzierten Hypophosphatämie signifikant 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 der 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 (Tabelle 4-72). Die Behandlungsdifferenz von -4,91 war mit p = 0,0245 statistisch signifikant. Der positive Effekt auf das Schmerzempfinden in der Burosumab-Gruppe setzte sich auch zu Woche 64 (LsMean: -2,26) weiter fort, ohne jedoch eine statistische Signifikanz zu erreichen (p = 0,3091).

Die beobachtete Abnahme von Schmerzen in der Burosumab-Gruppe erreichte sowohl zu Woche 40 als auch zu Woche 64 eine MID von mindestens 3 Punkten und bestätigt die spürbare Veränderung von XLH-assoziierten Schmerzen unter der Burosumab-Therapie. In der Kontrollgruppe wurde dagegen weder zu Woche 40 noch zu Woche 64 eine klinisch spürbare Veränderung von Schmerzen festgestellt [67].

Bei den Domänen Körperfunktion sowie Fatigue konnte kein signifikanter Behandlungsunterschied nachgewiesen werden. Doch ist eine geringe Verbesserung der Körperfunktion im Studienverlauf unter der Burosumab-Behandlung zu erwähnen (Tabelle 4-72). 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 (Crysvita®) ist die erste **kausale Therapie** zur Behandlung der XLH. Im Gegensatz zur Phosphatsubstitution, die lediglich die Symptomatik verbessern kann, nicht aber kausal in den Pathomechanismus eingreift, setzt die Behandlung mit Burosumab direkt am pathophysiologischen Defekt der XLH an. Bereits innerhalb kurzer Zeit wird dadurch eine **Normalisierung** der **Serumphosphat-Homöostase** erreicht und auch langfristig aufrechterhalten. Infolgedessen wird die **Knochenmineralisierung** gefördert und das **Heilen der Rachitis** stimuliert. Diese positiven Effekte haben direkt Auswirkungen auf das Wohlbefinden und die physische Funktion und führen zu einer signifikanten Zunahme des Wachstums und der Gehfähigkeit sowie zu einer Abnahme von Schmerzen. Durch Burosumab steht nun eine **zielgerichtete Therapie** zur Verfügung, durch die eine **bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens** bei der Behandlung der XLH erzielt werden kann.

Interpretation der Ergebnisse zur Lebensqualität

Die folgende Tabelle Tabelle 4-73 zeigt in einer Übersicht die zugrunde liegenden Ergebnisse der Endpunktkategorie *Lebensqualität* bezüglich des Vergleichs der Behandlungsgruppen, mit denen das Ausmaß des Zusatznutzens der Burosumab-Behandlung abgeleitet werden konnte.

Tabelle 4-73: Zusammenfassung der Ergebnisse der Kategorie *Lebensqualität* zum medizinischen Zusatznutzen von Burosumab

Patientenrelevante Endpunkte	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 40**	**UX023-CL301 Behandlungsdifferenz (B-P) aktive Kontrolle vs. Burosumab Woche 64**
Lebensqualität		
Short Form (SF)-10 Health Survey Physical Heath Score (PHS-10)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	4,35 [-1,15; 9,85] 0,1213 -	5,51 [-0,17; 11,18] 0,0571 -
Short Form (SF)-10 Health Survey Psychosocial Summary Score (PSS-10)		
LsMean[a] [95% KI] p-Wert Hedges'g[b] [95% KI]	2,24 [-1,86; 6,35] 0,2841 -	-0,44 [-4,59; 3,70] 0,8337 -

(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 4.2.5.2 und 4.3.1.3.1

b: Hedges'g-Berechnung erfolgt nur, falls $p < 0,05$

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. In der Studie UX023-C301 zeigten die Patienten der aktiven Kontrollgruppe weder eine Verbesserung der Körperfunktion noch der psychosozialen Domäne. Unter der Burosumab-Behandlung dagegen ist eine Verbesserung der Körperfunktion sowohl zu Woche 40 als auch Woche 64 gegeben (Tabelle 4-73). Die psychosoziale Domäne weist dagegen nur auf eine geringfügige Verbesserung bei der Burosumab-Gruppe hin (Tabelle 4-73). 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 4-74 zeigt in einer Übersicht die zugrunde liegenden Ergebnisse der Endpunktkategorie *Sicherheit* bezüglich des Vergleichs der Behandlungsgruppen, mit denen das Ausmaß des Zusatznutzens der Burosumab-Behandlung abgeleitet werden konnte.

Tabelle 4-74: Zusammenfassung der Ergebnisse für *unerwünschte Ereignisse nach Beginn der Behandlung (TEAE)* aus RCT mit dem zu bewertenden Arzneimittel

UX023-CL301 *Safety*-Population	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 40**	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 64**
Gesamtrate		
Jegliche TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	28 (96,6) vs. 26 (81,3)	29 (100,0) vs. 27 (84,4)
RR[a,c] [95% KI] p-Wert	1,188 [0,922; 1,423] 0,1064	1,185 [1,021; 1,376] 0,0538
Schwerwiegende TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	3 (10,3) vs. 1 (3,1)	3 (10,3) vs. 3 (9,4)
RR[a,c] [95% KI] p-Wert	3,310 [0,364; 30,076] 0,3385	1,115 [0,207; 6,018] 1,0000
Schwere TEAE Grad 3 – 4		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	4 (13,8) vs. 2 (6,3)	4 (13,8) vs. 3 (9,4)
RR[a,c] [95% KI] p-Wert	2,207 [0,436; 11,166] 0,4107	1,547 [0,316; 7,582] 0,6988
Therapieabbruch durch TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	0 (0) vs. 0 (0)	0 (0) vs. 0 (0)
RR[a,c] [95% KI] p-Wert	-	-
Tod durch TEAE		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	0 (0) vs. 0 (0)	0 (0) vs. 0 (0)
RR[a,c] [95% KI] p-Wert	-	-
TEAE von besonderem Interesse		
Reaktionen an der Injektionsstelle		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	12 (41,4) vs. 0 (0)	15 (51,7) vs. 0 (0)
RR[a,c] [95% KI] p-Wert	27,5 [1,70; 444,66] < 0,0001	34,1 [2,132; 545,51] < 0,0001
Hypersensitivität		
Burosumab (n (%)) vs. aktive Kontrolle[b] (n (%))	9 (31,0) vs. 2 (6,3)	11 (37,9) vs. 6 (18,8)
RR[a,c] [95% KI] p-Wert	4,966 [1,168; 21,114] 0,0183	2,023 [0,857; 4,773] 0,1522

N: Patientenzahl der entsprechend definierten Population; n: Anzahl der Patienten mit mindestens einem Ereignis in der entsprechenden Kategorie; %: Anteil von N; *TEAE*: *Treatment emergent adverse event* (unerwünschtes Ereignis nach Beginn der Behandlung); [95% KI]: 95%-iges Konfidenzintervall; RR: *Risk Ratio* (relatives Risiko)

UX023-CL301 *Safety*-Population	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 40**	**Burosumab (N = 29) vs. aktive Kontrolle (32) Woche 64**
Ein RR < 1 zeigt einen Vorteil für Burosumab an und ein RR > 1 einen Vorteil für die aktive Kontrolle. a: RR: Eine detaillierte Beschreibung des statistischen Modells ist in den Abschnitten 4.2.5.2 und 4.3.1.3.1.7 zu finden. b: Aktive Kontrolle: Orales Phosphat und aktives Vitamin D c: Treten keine Ereignisse auf, wird ein Korrekturwert von 0,5 zu den Zellen der Vierfeldertafel hinzugefügt.		

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 den *unerwünschten Ereignissen nach Beginn der Behandlung* (*TEAE*) keine signifikanten Sicherheitsbedenken (Tabelle 4-74). In allen drei Studien wurde weder ein Therapie-/Studienabbruch noch Todesfall aufgrund von *TEAE* berichtet. Demnach erweist sich Burosumab insgesamt als wirksames, gut verträgliches und sichereres Arzneimittel zur Behandlung von Kindern ab einem Jahr und Jugendlichen mit XLH und röntgenologischen Nachweis einer Knochenerkrankung.

In der Behandlungsphase (Baseline bis Woche 64) berichteten die meisten Patienten, sowohl in der Burosumab-Gruppe (100 %) als auch in der aktiven Kontrollgruppe (84 %), mindestens ein *TEAE*. Der Großteil der berichteten *TEAE* wurde als mild oder moderat eingestuft. Bei 3 Patienten der aktiven Kontrollgruppe sowie 4 Patienten der Burosumab-Gruppe traten *TEAE* vom Grad 3 auf. Diese Vorkommnisse wurden von den Prüfärzten als nicht therapieassoziiert 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 beobachtet.

Statistisch signifikante Unterschiede in der Sicherheit lagen lediglich bezogen auf die *TEAE*-Kategorie jegliche *TEAE* auf Ebene der SOC und PT vor. Hierbei zeigte sich bei der SOC *Allgemeine Erkrankungen und Beschwerden am Verabreichungsort* sowie den PT *Erythem an der Injektionsstelle* und *Reaktion an der Injektionsstelle* ein signifikanter Unterschied zuungunsten von Burosumab (Tabelle 4-52). Eine Reaktion an der Injektionsstelle ist eine bekannte und zu erwartende Folge einer subkutanen Verabreichung von Proteintherapeutika (Antikörpertherapeutika). Die aufgetretenen Reaktionen wurden vom Prüfarzt als allgemeine lokale Reizung und nicht als schwere Überempfindlichkeitsreaktion angesehen. Die Dauer der meisten Reaktionen an der Injektionsstelle betrug etwa ein bis zwei Tage. Da lediglich Burosumab, nicht aber Phosphat, subkutan verabreicht wurde, lässt sich anhand dieser SOC und PT keine valide Aussage zum Sicherheitsprofils von Burosumab im Vergleich zu oralem Phosphat machen.

Weitere statistisch signifikante Unterschiede zuungunsten von Burosumab zeigten sich für die SOC *Erkrankungen der Haut und des Unterhautzellgewebes*. Der Unterschied besteht

ausschließlich auf Ebene der SOC; statistisch signifikante Unterschiede auf Ebene der dazugehörigen PT konnten nicht festgestellt werden (Tabelle 4-52).

Für die SOC *Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen* und den dazugehörigen PT *Vitamin D erniedrigt* zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Burosumab (Tabelle 4-52). Ein Vitamin D-Mangel stellt eine charakteristische Manifestation der XLH-Erkrankung dar und wird durch einen FGF23-stimulierten Abbau von Calcitriol in der Niere bewirkt. Alle Kinder erhielten vor Studienbeginn neben oralem Phosphat daher auch für mindestens 12 Monate aktives Vitamin D. Auch während der Studie wurde im Kontrollarm die Gabe von aktivem Vitamin D fortgesetzt. In der Burosumab-Gruppe dagegen war die Gabe von aktivem Vitamin D während der Studie nicht erlaubt. Es ist daher nicht auszuschließen, dass durch das Absetzen von aktivem Vitamin D vor Studienbeginn der Vitamin D -Mangel in der Burosumab-Gruppe zunächst verstärkt wurde und sich erst im Verlauf der Burosumab-Behandlung normalisierte. Unterstützt wird dies ebenfalls durch die Beobachtung, dass in der Burosumab-Gruppe nach der 64-wöchigen Behandlung keine weiteren *TEAE* der Kategorie Vitamin D erniedrigt berichtet wurden. Eine Assoziation des Vitamin D-Mangels mit der Burosumab-Behandlung konnte nicht festgestellt werden.

Für die PT *Fieber* und *Husten* und die SOC *Erkrankungen der Atemwege, des Brustraums und Mediastinums* zeigte sich ebenfalls ein statistisch signifikanter Unterschied zuungunsten von Busorumab gegenüber Phosphat (Tabelle 4-52). Hierbei handelt es sich um typische Kindheitssymptome. Die Gründe für die unterschiedlichen Häufigkeiten zwischen den Behandlungsgruppen sind nicht vollständig verstanden. Es ist jedoch hervorzuheben, dass keines dieser Ereignisse vom Prüfarzt als schwerwiegendes Ereignis eingestuft und auch nicht als mit der Burosumab-Behandlung zusammenhängend bewertet wurde, mit der Ausnahme von 4 Patienten, die von Fieber am Tag der Burosumab-Injektion berichteten. 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 *TEAE* 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 [80]. 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 Überempfindlichkeitsereignisse (z. B. Hautausschlag), die in der Burosumab-Gruppe beobachtet wurden.

Schwerwiegende TEAE und *TEAE* des Grads 3 und 4 traten auf Ebene der MedDRA SOC und PT in keinem Fall bei 5 % oder mehr der Patienten in einer Behandlungsgruppe auf.

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, ausbleiben.

Zusammenfassung Sicherheit

Die Studiendaten der Studie UX023-CL301 belegen die gute Verträglichkeit von Burosumab bei Kindern mit XLH. Die meisten TEAE waren mild oder moderat, schwere oder schwerwiegende TEAE traten nur bei wenigen Studienteilnehmern auf. Hyperphosphatämien waren während der gesamten Studiendauer nicht zu verzeichnen. Nephrokalzinosen oder auch ein Hyperparathyreoidismus, die zu den typischen therapieassoziierten Folgeerkrankungen einer langjährigen Phosphatgabe zählen, waren nicht zu beobachten. Zudem führte kein TEAE zum Tod oder vorzeitigen Therapieabbruch der Studienteilnehmer. Insgesamt ist eine Behandlung mit Burosumab somit durch ein **ausgewogenes Nutzen-Risiko-Profil** gekennzeichnet.

Subgruppenanalysen

In den Subgruppenanalysen ergaben sich für keinen Endpunkt bedeutsame Effektmodifikationen. Aufgrund dessen wird der Zusatznutzen ausschließlich von der Gesamtpopulation abgeleitet.

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.

Übertragbarkeit auf die Altersgruppe 13 – 17 Jahre

Auch wenn in der Studie UX023-CL301 keine Studienteilnehmer im Alter von 13 – 17 Jahren eingeschlossen waren, ist davon auszugehen, dass der beobachtete Effekt auf die Serumphosphat-Homöostase und den Knochenstoffwechsel bei dieser Altersgruppe bestehen bleibt. Zum einen ist der Anteil an Kindern mit einem abgeschlossenen Skelettwachstum in dieser Altersgruppe gering, zum anderen konnte anhand weiterer Daten gezeigt werden, dass sich der positive Effekt einer Burosumab-Behandlung auch nach dem Schließen der Wachstumsfugen weiter fortsetzt. Dies belegen Erkenntnisse aus der einarmigen Studie UX023-CL201, die die Wirksamkeit und Sicherheit von Burosumab bei Kindern mit XLH untersuchte [82]. 11 der 52 Kinder mit XLH, die an dieser Studie teilnahmen, erreichten im Verlauf der Studie das Jugendalter, definiert als Studienteilnehmer, deren distale femorale- und proximale tibiale Wachstumsfugen sich während der Studie ganz oder teilweise schlossen. Bei

diesen Studienteilnehmern zeigte sich anhand des Serumphosphats und der „Heilung“ der Rachitis ebenfalls ein Ansprechen auf die Burosumab-Therapie analog zur Gesamtpopulation. Unterstützt wird dies ebenfalls durch ein *Early Access Programm* aus Großbritannien, dass explizit die Wirksamkeit von Burosumab bei Jugendlichen mit geschlossenen Wachstumsfugen untersuchte [83]. Die Übertragbarkeit der Evidenz auf die Altersgruppe 13 – 17 Jahre wurde auch im Rahmen der Zulassung durch die EMA bestätigt [84].

Abschließende Bewertung

Die 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 Akkumulation der Knochendeformationen und starken körperlichen Beeinträchtigungen zu verzögern bzw. zu verhindern und eine maximal mögliche Lebensqualität der Patienten zu gewährleisten.

Die Ergebnisse der Studie UX023-CL301 zeigen eindrucksvoll, dass durch den innovativen Wirkmechanismus von Burosumab die **Serumphosphat-Homöostase** bereits innerhalb kurzer Zeit **normalisiert** und auch langfristig aufrechterhalten wird. Folglich werden die damit verbundenen weiteren Stoffwechselkreisläufe, wie die des Calciums und Vitamin D, ebenfalls positiv beeinflusst. Die ausreichende Versorgung des Körpers mit Phosphat stimuliert den Knochenstoffwechsel und resultiert in einer “**Heilung**“ der **Rachitis**. Diese positiven Effekte haben direkt Auswirkungen auf das Wohlbefinden und die physische Funktion und führen zu einer **signifikanten Zunahme** des **Wachstums** und der **Gehfähigkeit** sowie zu einer **Abnahme** von **Schmerzen**. Darüber hinaus belegt die Studie UX023-CL301 das **ausgewogene Nutzen-Risiko-Profil** der Burosumab-Therapie. Die schweren therapieassoziierten Folgeerkrankungen, die häufig unter einer Phosphatsubstitution zu beobachten sind, treten unter einer Burosumab-Therapie nicht auf. Die Ergebnisse der Studie UX023-CL301 belegen somit den Zusatznutzen von Burosumab, der ersten den Pathomechanismus-adressierenden Therapieoption zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren mit röntgenologischem Nachweis einer Knochenerkrankung.

Gesamtfazit zum medizinischen Zusatznutzen von Burosumab

Die X-chromosomale Hypophosphatämie ist eine schwere genetisch bedingte Phosphatverlustkrankheit, die zu erheblichen Beeinträchtigungen in der Lebensqualität und im Alltag führt. Bis zur Zulassung von Burosumab im Jahr 2018 war keine spezifische Therapie zur Behandlung von XLH-Patienten verfügbar. Mangels fehlender Alternativen wurde daher versucht, den chronischen Phosphatverlust durch eine orale Phosphatgabe zu kompensieren. Auch wenn bei Kindern mit XLH dadurch in manchen Fällen eine Verbesserung der Symptomatik erzielt werden kann, lässt sich durch die Substitution mit Phosphat der chronische Phosphatverlust nicht beheben. Hinzu kommen schwere therapieassoziierte Folgeerkrankungen, die mit einer jahrelangen Phosphatsubstitution einhergehen. Die Folgen

dieser unzureichenden Versorgung des Körpers mit Phosphat sind schwer und führen zu weitreichenden Behinderungen der Patienten.

Bei Kindern und Erwachsenen mit XLH treten häufig Knochen- und Gelenkschmerzen sowie verminderte Gelenkbeweglichkeit auf, die einen erheblichen Einfluss auf das tägliche Leben nehmen. Viele Patienten leiden unter funktionellen Einschränkungen bei Tätigkeiten im Alltag. Zudem können die Patienten die körperlichen und psychischen Belastungen teilweise nur schlecht bewältigen, die vor allem auf Knochenschmerzen, Gelenkerkrankungen mit eingeschränkter Mobilität und Krankenhausaufenthalte zurückzuführen sind und wiederum zu einer unzureichenden Schul- und Berufsausbildung führen können. Die offensichtlichen Fehlbildungen und Gangauffälligkeiten führen immer wieder zu einer Stigmatisierung im öffentlichen Alltag. Insbesondere die Ganganomalien sind erheblich und führen zu einer starken Beeinträchtigung der XLH-Patienten [85]. Charakteristisch bei der XLH ist hierbei der „watschelnde“, breitbeinige Gang, der mit einer deutlich reduzierten Gehgeschwindigkeit einhergeht. Gerade für Kinder können solche Merkmale immer wieder Anlass für soziale Ausgrenzungen sein und unter Umständen schwere psychische Erkrankungen nach sich ziehen. Der therapeutische Bedarf ist daher hoch.

Erst seit der Zulassung von Burosumab steht diesen Patienten eine zielgerichtete Therapieoption zur Verfügung, die direkt in den Pathomechanismus eingreift und die zugrundeliegende Dysfunktion normalisiert. Im Gegensatz zur Phosphatsubstitution zeichnet sich die Burosumab-Behandlung durch eine gute Wirksamkeit und ein ausgewogenes Sicherheitsprofil aus. Die Verleihung des *Orphan-Drug*-Status und die Durchführung eines Arzneimittel-Härtefallprogramms durch das Paul-Ehrlich-Institut (PEI) vor der Zulassung von Burosumab unterstreichen die bis dahin fehlenden therapeutischen Alternativen im vorliegenden Anwendungsgebiet.

Die hier dargestellte Studie UX023-CL301 mit der Evidenzstufe 1b bestätigt eindrucksvoll die Ergebnisse der Studien UX023-CL201 und UX023-CL205 aus dem ersten Nutzenbewertungsverfahren von Burosumab [15].

Aufgrund der raschen Wiederherstellung der Mineral- und Vitamin-Dysregulation und der damit verbundenen Normalisierung des Knochenstoffwechsels kommt es unter der Burosumab-Therapie zu einer „Heilung“ der Rachitis und zur Rückbildung skelettaler Veränderungen, die zu einer Reduktion der Knochen- und Gelenkschmerzen führen. Die für die Patienten spürbare Linderung der Erkrankung führt zu einem deutlichen Anstieg ihrer Lebensqualität. Die ausreichende Versorgung des Körpers mit Phosphat wirkt sich zudem auch auf die Zusammensetzung der Körpermasse aus, fördert die Bildung von Muskelmasse und steigert die Mobilität der XLH-Patienten [86]. Die häufig bei einer Phosphatsubstitution beobachteten schweren Folgeerkrankungen, wie z. B. ein sekundärer oder tertiärer Hyperparathyreoidismus, treten bei der Behandlung mit Burosumab nicht auf. Der Wechsel von der Phosphatsubstitution auf die Burosumab-Therapie führt sogar nachweislich zu einer Abnahme der renalen Calcium-Ausscheidung und hat dadurch das Potenzial, langfristig die Entstehung von Nephrokalzinosen

zu verhindern [14]. Eine Behandlung mit Burosumab sollte aufgrund dieser vielfältigen Vorteile so früh wie möglich initiiert werden.

Burosumab ist mittlerweile seit mehr als drei Jahren in Deutschland auf dem Markt und hat sich im Versorgungsalltag etabliert.

In der Zusammenschau ergibt sich für Burosumab nach § 5 Abs. 7 der AM-NutzenV eine bisher noch nicht erreichte deutliche Verbesserung der Erkrankung basierend insbesondere auf einer **Abschwächung schwerwiegender Symptomatik** (Kleinwuchs, stark eingeschränkte Gehfähigkeit, erhebliche Mobilitäts- und Funktionsstörungen, damit einhergehende starke Schmerzen, geringere Lebensqualität).

Unter Berücksichtigung des sehr guten Wirksamkeits- und Sicherheitsprofils liegt für Burosumab zur Behandlung der XLH bei Kindern im Alter von 1 – 17 Jahren mit röntgenologischem Nachweis einer Knochenerkrankung somit ein **Beleg** auf einen **beträchtlichen Zusatznutzen** vor.

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 4-75: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Kinder ab 1 Jahr und Jugendliche mit XLH und röntgenologischem 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 2005[17], Molenberghs 2010[18]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende 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 2006[19]) 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

[17] Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.
[18] 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.
[19] Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006[20]) 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

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 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), 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: 05/2021 [Zugriff: 22.12.2021]. 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_Questionaire.
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.

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

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: 07.12.2021]. URL: https://www.g-ba.de/downloads/91-1031-359/2018-08-21_Wortprotokoll_Burosumab_D-349.pdf.
13. Gemeinsamer Bundesausschuss (G-BA) (2020): Mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses. Wirkstoff Burosumab. [Zugriff: 07.12.2021]. URL: https://www.g-ba.de/downloads/91-1031-497/2020-02-10_Wortprotokoll_Burosumab_D-492.pdf.
14. Harada D, Ueyama K, Oriyama K, Ishiura Y, Kashiwagi H, Yamada H, et al. (2021): Switching from conventional therapy to burosumab injection has the potential to prevent nephrocalcinosis in patients with X-linked hypophosphatemic rickets. Journal of Pediatric Endocrinology and Metabolism; 34(6):791-8.
15. Gemeinsamer Bundesausschuss (G-BA) (2018): Nutzenbewertung nach § 35a SGB V-Nutzenbewertungsverfahren zum Wirkstoff Burosumab. [Zugriff: 27.12.2021]. URL: https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/359/.
16. Linglart A, Imel EA, Whyte MP, Portale AA, Högler W, Boot AM, et al. (2021): Sustained efficacy and safety of burosumab, a monoclonal antibody to FGF23, in children with X-linked hypophosphatemia. The Journal of Clinical Endocrinology & Metabolism;
17. Gemeinsamer Bundesausschuss (G-BA) (2018): 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 – Burosumab. [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/40-268-5295/2018-10-04_AM-RL-XII_Burosumab_D-349_TrG.pdf.
18. Deutsche Gesellschaft für Kinderendokrinologie und -diabetologie (DGKED) e. V. (2016): S1-Leitlinie 174-008: Hereditäre hypophosphatämische Rachitiden; Stand: 03.2016. [Zugriff: 15.12.2021]. URL: https://www.awmf.org/uploads/tx_szleitlinien/174-008l_S1_Hereditaere_hypophosphataemische_Rachitiden_2016-05-abgelaufen.pdf.
19. Gemeinsamer Bundesausschuss (G-BA) (2017): Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 AM-NutzenV Beratungsanforderung 2017-B-176.
20. Gemeinsamer Bundesausschuss (G-BA) (2020): Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV Beratungsanforderung 2020-B-003.
21. 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: 27.12.2021]. URL: https://www.g-ba.de/downloads/40-268-3450/2015-11-27_AM-RL-XII_Secukinumab_2015-06-01-D167_TrG.pdf.

22. International Conference on Harmonization (ICH) (1995): Guideline for industry - The extent of population exposure to assess clinical safets: for drugs intended for longterm treatment of non-life-threatening conditions E1. [Zugriff: 27.12.2021]. URL: https://www.fda.gov/media/71180/download.
23. 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.
24. Ultragenyx pharmaceutical (2017): Ultragenyx RSS Variability Assessment for Study UX023-CL201 Version 1.0 (11 July 2017).
25. 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).
26. 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).
27. 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).
28. 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.
29. Lim R, Shailam R, Hulett R, Skrinar A, Nixon A, Williams A, et al. (2021): Validation of the Radiographic Global Impression of Change (RGI-C) score to assess healing of rickets in pediatric X-linked hypophosphatemia (XLH). Bone; 148:115964.
30. Ultragenyx pharmaceutical (2016): WhitePaper: Rickets in X-linked Hypophosphatemia Biology, Clinical Features, and Radiographic Findings.
31. 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: 27.12.2021]. URL: https://www.g-ba.de/downloads/40-268-3662/2016-03-17_AM-RL-XII_Asfotase-alfa_D-188_TrG.pdf.
32. 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).
33. 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: 27.12.2021]. URL:

https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/KiGGS_Referenzperzentile.pdf?__blob=publicationFile.
34. 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: 27.12.2021]. URL: https://www.cdc.gov/growthcharts/data_tables.htm.
35. 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: 27.12.2021]. URL: https://www.g-ba.de/downloads/92-975-1192/2015-12-23_Nutzenbewertung-G-BA_Asfotase%20alfa.pdf.
36. 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.
37. Carpenter TO, Whyte MP, Imel EA, Boot AM, Hogler W, Linglart A, et al. (2018): Burosumab Therapy in Children with X-Linked Hypophosphatemia. N Engl J Med; 378(21):1987-98.
38. American Thoracic Society (ATS) (2002): American Thoracic Society (ATS) Statement: Guidelines for the Six-Minute Walk Test Am J Respir Crit Care Med; 166:111-7.
39. 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).
40. 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.
41. Schnabel D (2013): Genetische Störungen der Knochenmineralisation. Osteologie 22:100-7.
42. 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.
43. Health Measures (2021): PROMIS® Reference Populations. [Zugriff: 27.12.2021]. URL: http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/reference-populations.
44. Ultragenyx pharmaceutical Javier San Martin (12. September 2016): UX023-CL301 Protocol Clarification Memo 1.
45. Ultragenyx pharmaceutical Javier San Martin (20. März 2017): UX023-CL301 Protocol Clarification Memo 3.
46. Ultragenyx pharmaceutical Javier San Martin (28. November 2016): UX023-CL301 Protocol Clarification Memo 2.
47. Health Measures (2021): List of Pedrisatric Measures: Available PROMIS® Measures for Pediatric Self-Report (ages 8-17) and Parent Proxy Report (ages 5-17) [Zugriff: 27.12.2021]. URL: http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-pediatric-measures.

48. 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.
49. 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.
50. 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.
51. 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.
52. 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.
53. 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.
54. 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.
55. European Medicinal Agnecy (EMA) (2008): CHMP Assessment report for sovrima, International Nonproprietary Name: idebenone, Procedure No. EMEA/H/C/000908 [Zugriff: 27.12.2021]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/000908/WC500070576.pdf.
56. European Medicinal Agnecy (EMA) (2014): CHMP assessment report Tracleer, International non-proprietary name: Bosentan, Procedure No. EMEA/H/C/000401/II/0066. [Zugriff: 27.12.2021]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000401/WC500184279.pdf.
57. 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.
58. 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.
59. Sabandal MM, Robotta P, Burklein S, Schafer E (2015): Review of the dental implications of X-linked hypophosphataemic rickets (XLHR). Clin Oral Investig; 19(4):759-68.
60. Foster BL, Ramnitz MS, Gafni RI, Burke AB, Boyce AM, Lee JS, et al. (2014): Rare bone diseases and their dental, oral, and craniofacial manifestations. J Dent Res; 93(7 Suppl):7S-19S.
61. Slade GD (1997): Derivation and validation of a short-form oral health impact profile. Community Dent Oral Epidemiol; 25(4):284-90.

62. Hanisch M, Bohner L, Sabandal MMI, Kleinheinz J, Jung S (2019): Oral symptoms and oral health-related quality of life of individuals with x-linked hypophosphatemia. Head Face Med; 15(1):8.
63. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2022): Allgemeine Methoden. [Zugriff: 24.01.2022]. URL: https://www.iqwig.de/methoden/allgemeine-methoden-v6-1.pdf.
64. Zhang J, Kai FY (1998): What's the relative risk?: A method of correcting the odds ratio in cohort studies of common outcomes. Jama; 280(19):1690-1.
65. Gemeinsamer Bundesausschuss (G-BA) (2020): Nutzenbewertung nach § 35a SGB V-Nutzenbewertungsverfahren zum Wirkstoff Burosumab. [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/92-975-3331/2019-10-01_Nutzenbewertung-G-BA_Burosumab_D-492.pdf.
66. 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.
67. Padidela R, Whyte MP, Glorieux FH, Munns CF, Ward LM, Nilsson O, et al. (2021): Patient-Reported Outcomes from a Randomized, Active-Controlled, Open-Label, Phase 3 Trial of Burosumab Versus Conventional Therapy in Children with X-Linked Hypophosphatemia. Calcified tissue international; 108(5):622-33.
68. Ultragenyx Pharmaceutical Inc (2016): UX023-CL301 - Efficacy and Safety of Burosumab (KRN23) Versus Oral Phosphate and Active Vitamin D Treatment in Pediatric Patients With X linked Hypophosphatemia (XLH) - ClinicalTrials.gov (NCT02915705). Stand des Eintrags: 23.01.2020. [Zugriff: 02.11.2021]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT02915705
69. Ultragenyx Pharmaceuticals Inc (2017): 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) (2016-000600-29). [Zugriff: 02.11.2021]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000600-29
70. Ultragenyx Pharmaceuticals Inc (2016): EMA/902676 UX023-CL301 - A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 - WHO ICTRP (EUCTR2016-000600-29-DK). Stand des Eintrags: 03.07.2017. [Zugriff: 02.11.2021]. URL: https://trialsearch.who.int/Trial2.aspx?TrialID=EUCTR2016-000600-29-DK
71. Ultragenyx Pharmaceutical Inc (2016): (2016-000600-29) - 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) Stand des Eintrags: 2016. [Zugriff: 02.11.2021]. URL: https://portal.dimdi.de/clinical-trials/servlet/FlowController/DisplayDocuments#__DEFANCHOR__
72. Gemeinsamer Bundesausschuss (G-BA) (2018): Amendment 2/2 zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO Wirkstoff: Burosumab. [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/92-975-2502/2018-10-04_Amendment-2_Burosumab_D-349.pdf.

73. Gemeinsamer Bundesausschuss (G-BA) (2018): 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 – Burosumab. [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/39-261-3515/2018-10-04_AM-RL-XII_Burosumab_D-349_BAnz.pdf.
74. Gemeinsamer Bundesausschuss (G-BA) (2018): Zusammenfassende Dokumentation über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Burosumab. [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/40-268-5583/2018-10-04_AM-RL-XII_Burosumab_D-349_ZD.pdf.
75. Gemeinsamer Bundesausschuss (G-BA) (2020): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Burosumab (Neubewertung nach Fristablauf: Hypophosphatämie). [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/40-268-6476/2020-04-02_AM-RL-XII_Burosumab_D-492_TrG.pdf.
76. Kyowa Kirin GmbH (2019): Dossier zur Nutzenbewertung gemäß § 35a SGB V Burosumab (Crysvita) Modul 4A. [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/92-975-3330/2019-09-23_Modul4A_Burosumab.pdf.
77. Gemeinsamer Bundesausschuss (G-BA) (2020): Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Burosumab (Neubewertung nach Fristablauf: Hypophosphatämie). [Zugriff: 05.11.2021]. URL: https://www.g-ba.de/downloads/39-261-4239/2020-04-02_AM-RL-XII_Burosumab_D-492_BAnz.pdf.
78. Ultragenyx Pharmaceutical Inc (2018): 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).
79. Geiger R, Strasak A, Treml B, Gasser K, Kleinsasser A, Fischer V, et al. (2007): Six-minute walk test in children and adolescents. J Pediatr; 150(4):395-9, 9 e1-2.
80. 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.
81. European Medicinal Agnecy (EMA) (2014): Public summary of opinion on orphan designation - Recombinant human monoclonal IgG1 antibody for fibroblast growth factor 23 for the treatment of X-linked hypophosphataemia. [Zugriff: 16.12.2021]. URL: https://www.ema.europa.eu/en/documents/orphan-designation/eu/3/14/1351-public-summary-opinion-orphan-designation-recombinant-human-monoclonal-igg1-antibody-fibroblast_en.pdf.
82. Ultragenyx pharmaceutical (2019): 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).
83. Dharmaraj P, Burren C, Cheung M, Padidela R, Mughal Z, Shaw N, et al. (2019): SUN-525 Burosumab Experience In A UK Adolescent Population. Journal of the Endocrine Society; 3(Supplement_1):SUN-525.

84. European Medicinal Agnecy (EMA) (2020): Assessment Report. [Zugriff: 16.12.2021]. URL: https://www.ema.europa.eu/en/documents/variation-report/crysvita-h-c-4275-ii-010-g-epar-assessment-report-variation_en.pdf.
85. Mindler GT, Kranzl A, Stauffer A, Kocijan R, Ganger R, Radler C, et al. (2021): Lower Limb Deformity and Gait Deviations Among Adolescents and Adults With X-Linked Hypophosphatemia. Frontiers in Endocrinology:1170.
86. Brener A, Lebenthal Y, Cleper R, Kapusta L, Zeitlin L (2021): Body composition and cardiometabolic health of pediatric patients with X-linked hypophosphatemia (XLH) under burosumab therapy. Therapeutic Advances in Endocrinology and Metabolism; 12:20420188211001150.

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. DIMDI, 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	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[21]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

[21] 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 4-76: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EMBASE

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	02.11.2021	
Zeitsegment	1974 to 2021 November 01	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	exp burosumab/	257
2	krn23.mp.[a]	44
3	"krn 23".mp.[a]	19
4	burosumab.mp.	275
5	anti-FGF23.mp.	122
6	"anti-FGF 23".mp.	18
7	crysvita.mp.	28
8	1 or 2 or 3 or 4 or 5 or 6 or 7	354
9	random*.tw. or placebo*.mp. or double-blind*.tw.	1984382
10	8 und 9	75
a: für krn 23 bzw. krn23 lag kein MeSH-Term vor.		

Tabelle 4-77: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in MEDLINE

Datenbankname	MEDLINE	
Suchoberfläche	Ovid	
Datum der Suche	02.11.2021	
Zeitsegment	Ovid MEDLINE(R) and In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to November 01, 2021	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[22]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	burosumab.mp.	103
2	krn23.mp.	16
3	"krn 23".mp.	1
4	anti-FGF23.mp.	57
5	"anti-FGF 23".mp.	8
6	crysvita.mp.	6
7	1 or 2 or 3 or 4 or 5 or 6	142
8	randomized controlled trial.pt. or randomized.mp. or placebo.mp.	979763
9	7 and 8	11

Tabelle 4-78: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in Cochrane

Datenbankname	Cochrane	
Suchoberfläche	Cochrane Library	
Datum der Suche	02.11.2021	
Zeitsegment	Keine Einschränkung	
Suchfilter	Keine Suchfilter	
#	**Suchbegriffe**	**Ergebnis**
1	krn23	21
2	"krn 23"	2
3	anti-FGF23	27
4	"anti-FGF 23"	2
5	crysvita	0
6	burosumab	54
7	#1 or #2 or #3 or #4 or #5 or #6	69

Die Recherche in bibliographischen Datenbanken wurde am 02.11.2021 durchgeführt. Treffer in Studienregistern wurden dabei nicht berücksichtigt, da diese bereits über die Studienregistersuche abgedeckt sind. Insgesamt ergab diese Recherche 155 Treffer von denen

80 als Duplikate ausgeschlossen wurden. Die verbleibenden 75 Treffer wurden anhand von *Titel* und *Abstract* mit den in Abschnitt 4.2.2 genannten Kriterien bewertet. 73 konnten so ausgeschlossen werden. Zwei Publikationen wurden als relevant für die Neubewertung des medizinischen Zusatznutzens von Burosumab nach Überschreitung der Umsatzgrenze von 50 Millionen Euro erachtet und eingeschlossen. Beide Publikationen beziehen sich auf die pivotale Phase-III-Studie UX023-CL301 mit Datenschnitt zu Woche 64 und dienen somit als Evidenz für das vorliegende Anwendungsgebiet [66, 67].

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.

[22] 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-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken 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 jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank 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/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-79: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in clinicaltrials.gov

Studienregister/ Studienergebnisdatenbank	clinicaltrials.gov
Internetadresse	https://clinicaltrials.gov/
Datum der Suche	02.11.2021
Suchstrategie	[OTHER TERMS] KRN23 OR "KRN 23" OR burosumab OR "anti-FGF 23" OR anti-FGF23 OR crysvita
Treffer	31

Die Suche in clinicaltrials.gov ergibt insgesamt 31 Treffer. Ein Treffer erfüllte die Einschlusskriterien aus der Tabelle 4-6. Dabei handelt es sich um die pivotale Studie UX023-CL301, die in Abschnitt 4.3.1 dargestellt ist [68].

Tabelle 4-80: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in EU Clinical Trials Register (EU-CTR)

Studienregister/ Studienergebnisdatenbank	EU Clinical Trials Register (EU-CTR)
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	02.11.2021
Suchstrategie	KRN23 OR "KRN 23" OR burosumab OR "anti-FGF 23" OR anti-FGF23 OR crysvita
Treffer	8

Die Suche in EU-CTR ergibt insgesamt 8 Treffer. Ein Treffer erfüllte die Einschlusskriterien aus der Tabelle 4-6. Dabei handelt es sich um die pivotale Studie UX023-CL301, die in Abschnitt 4.3.1 dargestellt ist [69].

Tabelle 4-81: Dokumentation der Suchstrategie für randomisierte Studien mit Burosumab (Crysvita®) zur Behandlung der XLH in *WHO International Clinical Trials Registry Platform* (WHO ICTRP)

Studienregister/ Studienergebnisdatenbank	WHO International Clinical Trials Registry Platform (WHO ICTRP)
Internetadresse	https://trialsearch.who.int/Default.aspx
Datum der Suche	02.11.2021
Suchstrategie	KRN23 OR burosumab OR anti-FGF 23 OR anti-FGF23 OR crysvita
Treffer	36

Die Suche im *International Clinical Trials Registry Platform Search Portal* der WHO ergab 58 Einträge für 36 Studien. Unter den Treffern befindet sich auch die relevante pivotale Studie UX023-CL301, die in Abschnitt 4.3.1 dargestellt ist [70].

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 4-7 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/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträgen 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

Tabelle 4-82: Übersicht der Treffer in den Studienregistern

Studienregister	Einträge	Ausgeschlossen	Eingeschlossen
clinicaltrials.gov	31	30	1
EU-CTR	8	7	1
ICTRP	36	35	1
Gesamt	75	72	3

Tabelle 4-83: Ausgeschlossene Studien ClincialTrials.gov

Nr.	Titel	Register-nummer	Ausschluss-grund
1	An Open-label, Multi-center, Single-cohort, Post-marketing Phase 4 Study to Evaluate the Efficacy, Pharmacodynamics, and Safety of the Anti-FGF23 Antibody, KRN23, in Pediatric Chinese Patients With X-linked Hypophosphatemic Rickets/Osteomalacia (XLH) - ClinicalTrials.gov	NCT04842032	A5
2	An Open-label, Multi-center, Single-cohort, Post-marketing Phase 4 Study to Evaluate the Efficacy, Pharmacodynamics, and Safety of the Anti-FGF23 Antibody, KRN23, in Adult Chinese Patients With X-linked Hypophosphatemic Rickets/Osteomalacia - ClinicalTrials.gov	NCT04842019	A1
3	An Investigator-sponsored Phase 3b Open-label Study of Anti-FGF23 Antibody Burosumab (KRN23) in Adult Patients With Xlinked Hypophosphatemia (XLH) in GERmany - BurGER - ClinicalTrials.gov	NCT04695860	A1
4	A Phase 3 Long-term Extension Study of KRN23 in Patients With X-linked Hypophosphatemic Rickets/Osteomalacia - ClinicalTrials.gov	NCT04308096	A5
5	An Open Label Trial to Assess the Safety and Efficacy of KRN23, an Investigational Antibody to FGF23, in a Single Pediatric Patient With Epidermal Nevus Syndrome (ENS) and Associated Hypophosphatemic Rickets - ClinicalTrials.gov	NCT04320316	A5

Nr.	Titel	Register-nummer	Ausschluss-grund
6	Examining the Effect of Burosumab on Muscle Function Using MR Spectroscopy - ClinicalTrials.gov	NCT04146935	A5
7	A Phase 1/2, Open-label, Multicenter, Non-randomized Study to Assess the Safety, Tolerability, Pharmacokinetics and Efficacy of Burosumab in Pediatric Patients From Birth to Less Than 1 Year of Age With X-linked Hypophosphatemia (XLH) - ClinicalTrials.gov	NCT04188964	A5
8	An Open Label Trial to Assess the Safety and Efficacy of Burosumab in a Single Patient With Cutaneous Skeletal Hypophosphatemia Syndrome (CSHS) - ClinicalTrials.gov	NCT03993821	A5
9	A Phase 3b Open-label Study of the Anti-FGF23 Antibody, Burosumab (KRN23) in Adult Patients With X-linked Hypophosphatemia (XLH) - ClinicalTrials.gov	NCT03920072	A1
10	12-months of Treatment With Burosumab in Children and Adolescents With X-linked Hypophosphatemia: a Prospective Longitudinal Cohort Study - ClinicalTrials.gov	NCT04419363	A5
11	An Open Label Trial to Assess the Safety and Efficacy of Burosumab (KRN23), an Investigational Antibody to FGF23, in a Single Pediatric Patient With Epidermal Nevus Syndrome (ENS) and Associated Hypophosphatemic Rickets - ClinicalTrials.gov	NCT03581591	A1
12	A Phase 3 Open-Label Trial to Assess the Efficacy and Safety of KRN23 in Pediatric Patients With X-linked Hypophosphatemic Rickets/Osteomalacia - ClinicalTrials.gov	NCT03233126	A5
13	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) - ClinicalTrials.gov	NCT02915705	E
14	A Phase 2 Open-Label Trial to Assess the Efficacy and Safety of KRN23 in Patients With Tumor-Induced Osteomalacia or Epidermal Nevus Syndrome - ClinicalTrials.gov	NCT02722798	A1
15	An Open-Label, Phase 2 Study to Assess the Safety, Pharmacodynamics, and Efficacy of KRN23 in Children From 1 to 4 Years Old With X-linked Hypophosphatemia (XLH) - ClinicalTrials.gov	NCT02750618	A5
16	An Open-Label, Single-Arm, Phase 3 Study to Evaluate the Effects of KRN23 on Osteomalacia in Adults With X-linked Hypophosphatemia (XLH) - ClinicalTrials.gov	NCT02537431	A1
17	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) - ClinicalTrials.gov	NCT02526160	A1
18	A Phase 2b, Open-Label, Long-Term Extension Study to Evaluate the Safety and Pharmacodynamics of KRN23 in Adult Subjects With X-Linked Hypophosphatemia (XLH) - ClinicalTrials.gov	NCT02312687	A1
19	A Phase 2 Open-Label Trial to Assess the Efficacy and Safety of KRN23, an Antibody to FGF23, in Subjects With Tumor-Induced Osteomalacia (TIO) or Epidermal Nevus Syndrome (ENS)-Associated Osteomalacia - ClinicalTrials.gov	NCT02304367	A1
20	A Randomized, Open-Label, Dose Finding, Phase 2 Study to Assess the Pharmacodynamics and Safety of the Anti-FGF23	NCT02163577	A3

Nr.	Titel	Register-nummer	Ausschluss-grund
	Antibody, KRN23, in Pediatric Patients With X-linked Hypophosphatemia (XLH) - ClinicalTrials.gov		
21	A Phase 1, Multicenter, Open-label, Sequential Dose-escalation, Single-dose Study to Assess the Safety and Tolerability of KRN23 in Subjects With X-linked Hypophosphatemic Rickets/Osteomalacia. - ClinicalTrials.gov	NCT02181764	A5
22	An Open-Label, Long-Term, Extension Study to Evaluate the Safety and Efficacy of KRN23 in Adult Subjects With X-Linked Hypophosphatemia - ClinicalTrials.gov	NCT01571596	A1
23	A Phase I/II, Open-Label, Repeat-Dose, Dose-Escalation Study of KRN23 in Adult Subjects With X-Linked Hypophosphatemia - ClinicalTrials.gov	NCT01340482	A1
24	Expanded Access to Burosumab	NCT03775187	A5
25	Effect of Burosumab and 1-25 (OH) Vitamin D on Human Osteoblasts From Patients Requiring Craniosynostosis Surgery for Idiopathic Reason or Due to Hypophosphatemic Rickets (HR) - ClinicalTrials.gov	NCT04159675	A5
26	A Phase I, Double-blind, Randomized, Placebo-controlled, Single-dose, Dose-escalation Study of KRN23 in X-linked Hypophosphatemia - ClinicalTrials.gov	NCT00830674	A3
27	Effect of Burosumab and 1-25 (OH) Vitamin D on Human Osteoclasts From Patients With Hypophosphatemic Rickets (HR) - ClinicalTrials.gov	NCT04184661	A5
28	X-linked Hypophosphatemia Disease Monitoring Program (XLH-DMP) - ClinicalTrials.gov	NCT03651505	A5
29	Tumor-induced Osteomalacia Disease Monitoring Program (TIO DMP) - ClinicalTrials.gov	NCT04783428	A1
30	A UK Multicentre, Non-interventional, Observational, Health-Related Quality of Life Study for Children and Adolescents With X-linked Hypophosphataemia - ClinicalTrials.gov	NCT04819490	A5
31	An International, Multicentre, Prospective, Non-interventional Observational Registry for Patients With X-Linked Hypophosphatemia (XLH) - ClinicalTrials.gov	NCT03193476	A5

Tabelle 4-84: Ausgeschlossene Studien EU-CTR

Nr.	Titel	Register-nummer	Ausschluss-grund
32	Hypophosphatemia (XLH) - EU-CTR	2018-000202-37	A1
33	An investigator-sponsored Phase 3b Open-label Study of Anti-FGF23 Antibody Burosumab (KRN23) in Adult Patients with X-linked Hypophosphatemia (XLH) in GERmany - BurGER - EU-CTR	2019-003190-26	A1
34	A Phase 1/2, Open-label, Multicenter, Non-randomized Study to Assess the Safety, Tolerability, Pharmacokinetics and Efficacy of Burosumab in Pediatric Patients from Birth to Less than 1 Year of Age with X-linked Hypophosphatemia (XLH) - EU-CTR	2019-000469-19	A5
35	A Randomized, open Label, Dose Finding, Phase 2 Study to Assess the Pharmacodynamics and Safety of the anti-FGF23	2014-000406-35	A3

Nr.	Titel	Register-nummer	Ausschluss-grund
	antibody, KRN23, in Pediatric Patients with X-linked Hypophosphatemia (XLH) - EU-CTR		
36	An Open-Label, Phase 2 Study to Assess the Safety, Pharmacodynamics, and Efficacy of KRN23 in Children from 1 to 4 Years Old with X-linked Hypophosphatemia (XLH) - EU-CTR	2018-001983-49	A5
37	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)	2016-000600-29	E
38	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) - EU-CTR	2014-005529-11	A1
39	An Open-Label, Single-Arm, Phase 3 Study to Evaluate the Effects of KRN23 on Osteomalacia in Adults with X-linked Hypophosphatemia (XLH) - EU-CTR	2015-001775-41	A1

Tabelle 4-85: Ausgeschlossene Studien WHO ICTRP

Nr.	Titel	Register-nummer	Ausschluss-grund
40	An Open-label, Multi-center, Single-cohort, Post-marketing Phase 4 Study to Evaluate the Efficacy, Pharmacodynamics, and Safety of the Anti-FGF23 Antibody, KRN23, in Adult Chinese Patients With X-linked Hypophosphatemic Rickets/Osteomalacia - WHO ICTRP	NCT04842019	A1
41	An Open-label, Multi-center, Single-cohort, Post-marketing Phase 4 Study to Evaluate the Efficacy, Pharmacodynamics, and Safety of the Anti-FGF23 Antibody, KRN23, in Pediatric Chinese Patients With X-linked Hypophosphatemic Rickets/Osteomalacia (XLH) - WHO ICTRP	NCT04842032	A5
42	An Investigator-sponsored Phase 3b Open-label Study of Anti- FGF23 Antibody Burosumab (KRN23) in Adult Patients With Xlinked Hypophosphatemia (XLH) in GERmany - BurGER - WHO ICTRP	NCT04695860	A1
43	An investigator-sponsored Phase 3b Open-label Study of Anti-FGF23 Antibody Burosumab (KRN23) in Adult Patients with X-linked Hypophosphatemia (XLH) in GERmany - BurGER - WHO ICTRP	2019-003190-26	A1
44	12-months of Treatment With Burosumab in Children and Adolescents With X-linked Hypophosphatemia: a Prospective Longitudinal Cohort Study - WHO ICTRP	NCT04419363	A5
45	Phase 3 Long-term Extension Study of KRN23 in Patients with X-linked Hypophosphatemic Rickets/Osteomalacia - WHO ICTRP	JPRN-JapicCTI-205284	A5
46	An Open Label Trial to Assess the Safety and Efficacy of KRN23, an Investigational Antibody to FGF23, in a Single Pediatric Patient With Epidermal Nevus Syndrome (ENS) and Associated Hypophosphatemic Rickets - WHO ICTRP	NCT04320316	A5
47	A Phase 3 Long-term Extension Study of KRN23 in Patients With X-linked Hypophosphatemic Rickets/Osteomalacia - WHO ICTRP	NCT04308096	A5
48	Effect of Burosumab and 1-25 (OH) Vitamin D on Human Osteoclasts From Patients With Hypophosphatemic Rickets (HR) - WHO ICTRP	NCT04184661	A5

Nr.	Titel	Register-nummer	Ausschluss-grund
49	A Phase 1/2, Open-label, Multicenter, Non-randomized Study to Assess the Safety, Tolerability, Pharmacokinetics and Efficacy of Burosumab in Pediatric Patients From Birth to Less Than 1 Year of Age With X-linked Hypophosphatemia (XLH) - WHO ICTRP	NCT04188964	A5
50	Effect of Burosumab and 1-25 (OH) Vitamin D on Human Osteoblasts From Patients Requiring Craniosynostosis Surgery for Idiopathic Reason or Due to Hypophosphatemic Rickets (HR) - WHO ICTRP	NCT04159675	A5
51	A Phase 1/2, Open-label, Multicenter, Non-randomized Study to Assess the Safety, Tolerability, Pharmacokinetics and Efficacy of Burosumab in Pediatric Patients from Birth to Less than 1 Year of Age with X-linked Hypophosphatemia (XLH) - WHO ICTRP	2019-000469-19	A5
52	Examining the Effect of Burosumab on Muscle Function Using MR Spectroscopy - WHO ICTRP	NCT04146935	A5
53	A Phase 3b Open-label Study of the Anti-FGF23 Antibody, - WHO ICTRP	2018-000202-37	A1
54	An Open Label Trial to Assess the Safety and Efficacy of Burosumab in a Single Patient With Cutaneous Skeletal Hypophosphatemia Syndrome (CSHS) - WHO ICTRP	NCT03993821	A5
55	Expanded Access to Burosumab - WHO ICTRP	NCT03775187	A5
56	A Phase 3b Open-label Study of the Anti-FGF23 Antibody, Burosumab (KRN23) in Adult Patients With X-linked Hypophosphatemia (XLH) - WHO ICTRP	NCT03920072	A1
57	A Phase 3b Open-label Study of the Anti-FGF23 Antibody, Burosumab (KRN23) in Adult Patients with X-linked Hypophosphatemia (XLH) - WHO ICTRP	2018-000202-37	A1
58	An Open Label Trial to Assess the Safety and Efficacy of Burosumab (KRN23), an Investigational Antibody to FGF23, in a Single Pediatric Patient With Epidermal Nevus Syndrome (ENS) and Associated Hypophosphatemic Rickets - WHO ICTRP	NCT03581591	A1
59	A Phase 3 Open-Label Trial to Assess the Efficacy and Safety of KRN23 in Pediatric Patients with X-linked Hypophosphatemic Rickets/Osteomalacia - WHO ICTRP	JPRN-JapicCTI-173614	A5
60	A Randomized, Open-Label, Phase 3 Study to Assess the Efficacy and Safety of KRN23 - WHO ICTRP	2016-000600-29	E
61	A Phase 2 Open-Label Trial to Assess the Efficacy and Safety of KRN23 in Patients With Tumor-Induced Osteomalacia or Epidermal Nevus Syndrome - WHO ICTRP	NCT02722798	A1
62	An Open-Label, Phase 2 Study to Assess the Safety, Pharmacodynamics, and Efficacy of KRN23 in Children From 1 to 4 Years Old With X-linked Hypophosphatemia (XLH) - WHO ICTRP	NCT02750618	A5
63	A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Assess the Efficacy and Safety of KRN23 in Adults with X-linked Hypophosphatemia (XLH) - WHO ICTRP	2014-005529-11	A1
64	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) - WHO ICTRP	2014-005529-11	A1
65	An Open-Label, Single-Arm, Phase 3 Study to Evaluate the Effects of KRN23 on Osteomalacia in Adults With X-linked Hypophosphatemia (XLH) - WHO ICTRP	NCT02537431	A1

Nr.	Titel	Register-nummer	Ausschluss-grund
66	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) - WHO ICTRP	NCT02526160	A1
67	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) - WHO ICTRP	2014-000406-35	A3
68	A Phase 2b, Open-Label, Long-Term Extension Study to Evaluate the Safety and Pharmacodynamics of KRN23 in Adult Subjects With X-Linked Hypophosphatemia (XLH) - WHO ICTRP	NCT02312687	A1
69	A Phase 2 Open-Label Trial to Assess the Efficacy and Safety of KRN23, an Antibody to FGF23, in Subjects With Tumor-Induced Osteomalacia (TIO) or Epidermal Nevus Syndrome (ENS)-Associated Osteomalacia - WHO ICTRP	NCT02304367	A1
70	A Phase 1, Multicenter, Open-label, Sequential Dose-escalation, Single-dose Study to Assess the Safety and Tolerability of KRN23 in Subjects With X-linked Hypophosphatemic Rickets/Osteomalacia. - WHO ICTRP	NCT02181764	A5
71	A Phase 1 study of KRN23 in Subjects with X-linked Hypophosphatemic Rickets/Osteomalacia - WHO ICTRP	JPRN-JapicCTI-142601	A5
72	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) - WHO ICTRP	2014-000406-35	A3
73	An Open-Label, Long-Term, Extension Study to Evaluate the Safety and Efficacy of KRN23 in Adult Subjects With X-Linked Hypophosphatemia - WHO ICTRP	NCT01571596	A1
74	A Phase I/II, Open-Label, Repeat-Dose, Dose-Escalation Study of KRN23 in Adult Subjects With X-Linked Hypophosphatemia - WHO ICTRP	NCT01340482	A1
75	A Phase I, Double-blind, Randomized, Placebo-controlled, Single-dose, Dose-escalation Study of KRN23 in X-linked Hypophosphatemia - WHO ICTRP	NCT00830674	A5

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 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-86 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 4-86 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-86: Studiendesign und - methodik für Studien 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 Rachitissymptomatik bei Kindern im Alter von 1 – 12 Jahren mit XLH.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	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. 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 abbrachen, einer 64-wöchigen (Woche 0 - 64) Behandlungsphase und einer 76-wöchigen Extensionsphase (Woche 140). <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. <u>Behandlungsphase aktive Kontrolle (Woche 0 – 64):</u> Mehrere tägliche Dosen von oralem Phosphat/aktivem Vitamin D. Aufgrund der Variabilität der Dosen und Einnahmezeiten

Item[a]	Charakteristikum	Studieninformation
		erfolgte 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 Die Behandlung mit Alfacalcidol oder Calcitriol erfolgt basierend auf klinischen Parametern.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Das Original-Studienprotokoll ist auf den 04. April 2016 datiert. <u>Amendment 1 (Global) datiert auf den 03. November 2017:</u> • 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, Kanada und Australien): Nach Abschluss der 64-wöchigen Behandlungsphase werden die Patienten, die zu Burosumab randomisiert wurden, die

Item[a]	Charakteristikum	Studieninformation
		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 Phosphat/aktivem Vitamin D) konnte die erste Dosis von Burosumab zu Woche 66 appliziert werden. • Das maximale Geschlechterverhältnis wurde von 60 % auf 70 % weiblicher Patienten erhöht, um das reale Geschlechterverhältnis der XLH-Patientenpopulation in der Studie besser widerzuspiegeln. Diese Änderung wurde aus Gründen der Genauigkeit zur exakteren Erfassung des Krankheitsbilds vorgenommen, hatte aber letztendlich keinen Einfluss auf den Anteil der weiblichen Studienteilnehmer, der in der Studie UX023-CL301 55,7 % betrug. • Es wurde klargestellt, dass „mindestens“ und nicht „in etwa“ 20 Studienteilnehmer im Alter von 1 bis < 5 Jahre sein müssen. Auswahl der Studienpopulation: • 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

Item[a]	Charakteristikum	Studieninformation
		von Klammernahtgeräten (staples), 8-Platten oder andere orthopädische Hardware“ korrigiert. • 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. Sicherheitsanalysen • 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. Verbotene Medikation • Es wurde klargestellt, dass in der Burosumab-Gruppe aktives Vitamin D, nicht die Vitamin D Supplementation, verboten ist. Studienmedikation • 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.

Item[a]	Charakteristikum	Studieninformation
4	Probanden / Patienten	Es wurden 32 Kinder in den aktiven Kontrollarm und 29 Kinder in den Burosumab-Arm im Alter zwischen 1 – 12 Jahren mit X-chromosomaler Hypophosphatämie randomisiert.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** Kinder, die in die Studie aufgenommen werden sollten, mussten alle Kriterien erfüllen: - 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 - *PHEX*-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

Item[a]	Charakteristikum	Studieninformation
		unterziehen. Frauen im gebärfähigen Alter müssen bereit sein eine hochwirksame 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. **Ausschlusskriterien:** Kinder, die eines der folgenden Ausschlusskriterien erfüllten, waren nicht zur Teilnahme an der Studie berechtigt: - 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: 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
		altersabhängigen Normalwerte (basierend auf Nüchternwerten [Mindestens 4 Stunden] der Screeningvisite und/oder Baseline - Nachweis eines Hyperparathyreoidismus (iPTH-Konzentration 2,5X ULN) - Einnahme/ Verwendung von Medikamenten zur Unterdrückung von PTH (z. B. Cinacalcet, Kalzimimetika) innerhalb von 2 Monaten vor der Screeingvisite - 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 oder in Japan die Verwendung eines Prüfpräparats oder eines Prüfmedikaments innerhalb von 4 Monaten vor der Screeningvisite oder die Notwendigkeit des Prüfpräparates vor Abschluss aller Studienbewertungen.
4b	Studienorganisation und Ort der Studiendurchführung	Sponsor war Ultragenyx Pharmaceutical Inc. Novato, CA, USA. 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)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Burosumab (KNR23) ist ein rekombinanter, vollständig humaner, monoklonaler Antikörper (IgG1), der die überschüssige biologische Aktivität des Fibroblasten-Wachstumsfaktors 23

Item[a]	Charakteristikum	Studieninformation
		(FGF23) bindet und hemmt. 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. <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. <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.
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	**<u>Wirksamkeit:</u>** **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 o Erhebungszeitpunkte: Screeningvisite, W 40, W 64 o Schulung und Verblindung der 3 Auswerter sowie Mehrfachbeurteilung **Sekundärer Wirksamkeitsendpunkt:** Evaluation des Effekts von Burosumab verglichen mit der aktiven Kontrolle: <u>Wesentliche sekundäre Endpunkte:</u> • Wachstum, Wachstumsgeschwindigkeit und Deformation der unteren Extremitäten (gemessen durch RGI-C-Score des gestreckten Beins, sowie Z-Scores der Liegelänge/Stehhöhe) o Erhebungszeitpunkte: Screeningvisite, W 40, W 64 (RGI-C) bzw. Baseline, W 24, W 40, W 64 (Stehhöhe/Liegelänge) o Schulung und Verblindung der Auswerter sowie

		Mehrfachbeurteilung von 3 Auswertern (RGI-C)
		• Änderung des RSS Total Score
		  ○ Erhebungszeitpunkte: Screeningvisite, W 40, W 64
		  ○ Schulung und Verblindung des einzelnen Auswerters
		• Änderung der Blut-Parameter Serum-Phosphat und Alkalische Phosphatase (ALP) zur Bewertung der Knochenmineralisierung und daraus folgend der Rachitis-Schwere
		<u>Weitere sekundäre Endpunkte:</u>
		• Beurteilung der Rachitissymptomatik anhand von RGI-C- und RSS-Handgelenk- sowie Kniescores
		  ○ Erhebungszeitpunkte: Screeningvisite, W 40, W 64
		  ○ Schulung und Verblindung des einzelnen Auswerters
		• 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 D_3, TmP/GFR und TRP)
		• Gehfähigkeit (6MWT), Patienten-/Elternberichte zu Schmerz, Fatigue und körperliche Funktion/Mobilität (PROMIS; FPS-R)
		  ○ Erhebungszeitpunkte: Baseline, W 24, W 40, W 64
		  ○ Schulungen des klinischen Personals auf Basis der ATS Guidelines (6MWT)
		  ○ Validierung des PROMIS-Fragebogens durch Interviewstudie
		  ○ Schulungen des klinischen Personals (FPS-R)
		Explorative Endpunkte:
		• Gesundheitsbezogene Lebensqualität mittels Short Form 10 (SF-10) for Children Health Survey (SF-10: für Kinder ab 5 Jahren zur Screeningvisite; Domänen: *Physical Health Summary Score* (PHS-10) und *Psychosocial Summary Score* (PSS-10))
		  ○ Erhebungszeitpunkte: Baseline, W 24, W 40, W 64
		  ○ 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
		Sicherheitsendpunkte:
		• Frequenz und Schwere der unerwünschten

Item[a]	Charakteristikum	Studieninformation
		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 urinanalytische 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) **Pharmakokinetik Sicherheitsendpunkte:** • Serum Burosumab
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	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. 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. Das Alter der teilnahmeberechtigten Personen wird auf Altersverteilung überwacht. Etwa 20 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

Item[a]	Charakteristikum	Studieninformation
		Patienten werden aufgenommen. 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.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Die Primäranalyse erfolgte zu Woche 40. Eine zusätzliche Analyse wird zu Woche 64 durchgeführt, um Langzeitbehandlungseffekte, zusätzliche Wirksamkeitsergebnisse und langfristige Sicherheit zu bewerten. Es kam zu keinen vorzeitigen Studienabbrüchen.
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 ≤ 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?	Zu a) und b) 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 Therapie (oralen Phosphat/aktives Vitamin D) erforderlich sind, problematisch wäre. Aus diesem Grund war dies eine offene Studie, weshalb sowohl das Studienpersonal als auch die teilnehmenden Patienten wussten zu welcher Behandlungsgruppe (Burosumab vs. aktive

Item[a]	Charakteristikum	Studieninformation
		Kontrolle) sie zugeteilt wurden. Zu c) 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.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	**Wirksamkeits-Analyse:** Bei wiederholten Messungen der klinischen Endpunkte wurde ein *Generalized Estimating Equations* (GEE)-Modell zur Bewertung der Veränderung im Laufe der Zeit verwendet. Das GEE-Modell beinhaltet die Behandlungsgruppe, die Studienvisite und die Interaktion zwischen Behandlungsgruppe und Studienbesuch als kategorische Variablen. Die Stratifizierungsfaktoren des Baseline RSS Total Score und das Alter zu Baseline 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. Für die Kovarianz wurde zunächst eine austauschbare (exchangeable) Struktur gewählt. Für den Fall, dass diese zu einer Nichtkonvergenz führte, wird eine unabhängige Kovarianzstruktur gewählt. **Analyse des primären Wirksamkeitsendpunktes:** Der primäre Wirksamkeitsendpunkt wurde getestet, um den mittleren RGI-C-Global-Score zwischen der Burosumab-Gruppe und der aktiven Kontrollgruppe zu vergleichen. Bei der Analyse der Woche 40 wurde eine Kovarizanzanalyse (ANCOVA-Modell) durchgeführt und beinhaltet den RGI-C-Global-Score als abhängige Variable, die Behandlungsgruppe und die Baseline-Alters-Stratifizierungsfaktoren als Faktoren, und den Baseline RSS Total Score als kontinuierliche Kovariate. Der Stratifizierungsfaktor des Baseline RSS Total Scores wurde nicht in das Modell aufgenommen. Zu Woche 64 wurde ein GEE-Modell, welches die Behandlungsgruppe, Behandlungsvisite, Interaktion zwischen

Item[a]	Charakteristikum	Studieninformation
		Behandlungsgruppe und Studienvisite und die Baseline-Alters-Stratifizierungsfaktoren als unabhängige Variablen beinhaltete und den Baseline RSS Total Score als kontinuierliche Kovariate, verwendet. Der RSS-Gesamtscore wurde analog zum RGI-C analysiert. **Analyse der sekundären Wirksamkeitsendpunkte:** 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 *GEE-Modell*, welches die Faktoren Behandlungsgruppe, Studienvisite, Interaktion zwischen Behandlungsgruppe und Behandlungsvisite und Baseline-RSS-Stratifizierungsfaktoren als unabhängige Variablen beinhaltete. und Alter beinhaltet. Zusätzlich wurde bei den Endpunkten Stehhöhe/Liegelänge/Sitzhöhe das Baseline Alter als unabhängige Variable berücksichtigt. Weitere Baseline-Kovariaten werden gegebenenfalls zur Korrektur berücksichtigt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	**Responderanalysen:** Die RGI-C Global Score Responderanalyse wurde für Woche 40 und ebenfalls für Woche 64 durchgeführt. RGI-C-Responder sind Patienten mit einem mittleren RGI-C Global Score ≥ + 2,0. Die Woche 40 Analyse wurde mithilfe eines logistischen Regressionsmodells mit der Behandlungsgruppe und den Baseline-Alters-Stratifizierungsfaktoren als unabhängige Variablen und der Baseline RSS Total Score als kontinuierliche Kovariate durchgeführt. Der Stratifizierungsfaktor RSS Total Score wurde nicht in das Modell aufgenommen. In der Woche 64 Analyse wurde ein ähnliches logistisches Regressionsmodell angewendet (GLMM), das sowohl die Daten der Woche 40 als auch der Woche 64 als abhängige Variable, Behandlungsgruppe, Behandlungsvisite, Interaktion zwischen Behandlungsgruppe und Studienvisite und Baseline-Alters-Stratifizierungsfaktoren als unabhängige Variablen enthält, sowie den Baseline RSS Total Score als kontinuierliche Kovariate. Der Stratifizierungsfaktor Baseline RSS Total Score wurde nicht in das Modell aufgenommen. Patientenspezifische Heterogenität wurde mit Hilfe von zufälligen Effekten für die Modellkonstante auf Patientenebene

Item[a]	Charakteristikum	Studieninformation
		berücksichtigt. 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. **Subgruppenanalysen:** Für die Wirksamkeitsdaten wurden für folgende Endpunkte Subgruppen definiert: - ***Rickets Severity Score* RSS:** Subgruppen definiert nach RSS Total Score zu Baseline. o Baseline RSS Total Score ≥ 2,5 o Baseline RSS Total Score < 2,5 - ***Alter*:** Subgruppen definiert nach Alter zu Baseline. o Alter zu Baseline ≥ 5 o Alter zu Baseline < 5 - ***Herkunft:*** Subgruppen definiert nach Herkunft o Japan o Rest der Welt - ***Geschlecht:*** Subgruppen definiert nach Geschlecht. o männlich o weiblich - ***PHEX-Mutation:*** o Positiv o Nicht-positiv **Sensitivitätsanalysen** Zur Beurteilung der Robustheit des primären Analyseergebnisses der RSS- und RGI-C-Endpunkte für fehlende Werte wurden Sensitivitätsanalysen durchgeführt. Diese können folgende beinhalten, sind aber nicht beschränkt auf: • *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

Item[a]	Charakteristikum	Studieninformation
		imputieren. Die imputierten Datensätze wurden mit einem ANCOVA-Modell in Woche 40 Analyse oder einem GEE-Modell in Woche 64 analysiert. Die Ergebnisse wurden mit der PROC MIANALYZE-Prozedur von SAS kombiniert. Dabei wurden der Baseline RSS Total Score, Behandlungsgruppe und Baseline-Alters-Stratifizierungsfaktoren als Prädiktionsvariablen verwendet. Andere Variablen wie auch die Änderung von ALP können ebenfalls als Prädiktionsvariable betrachtet werden.
Resultate		
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, einschließlich der Extensionsphase, ist beendet. Für Studienteilnehmer aus Europa, Australien und den USA endete die Behandlungsfortsetzungsphase im September 2018, nachdem Burosumab auch kommerziell erwerblich war. Für Studienteilnehmer aus Kanada endete die Behandlungsfortsetzungsphase im Juni 2019, nachdem dort ebenfalls Burosumab kommerziell erhältlich war. Für Studienteilnehmer aus Japan und Korea endete die Studie dagegen nach der 64-wöchigen Behandlungshase.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

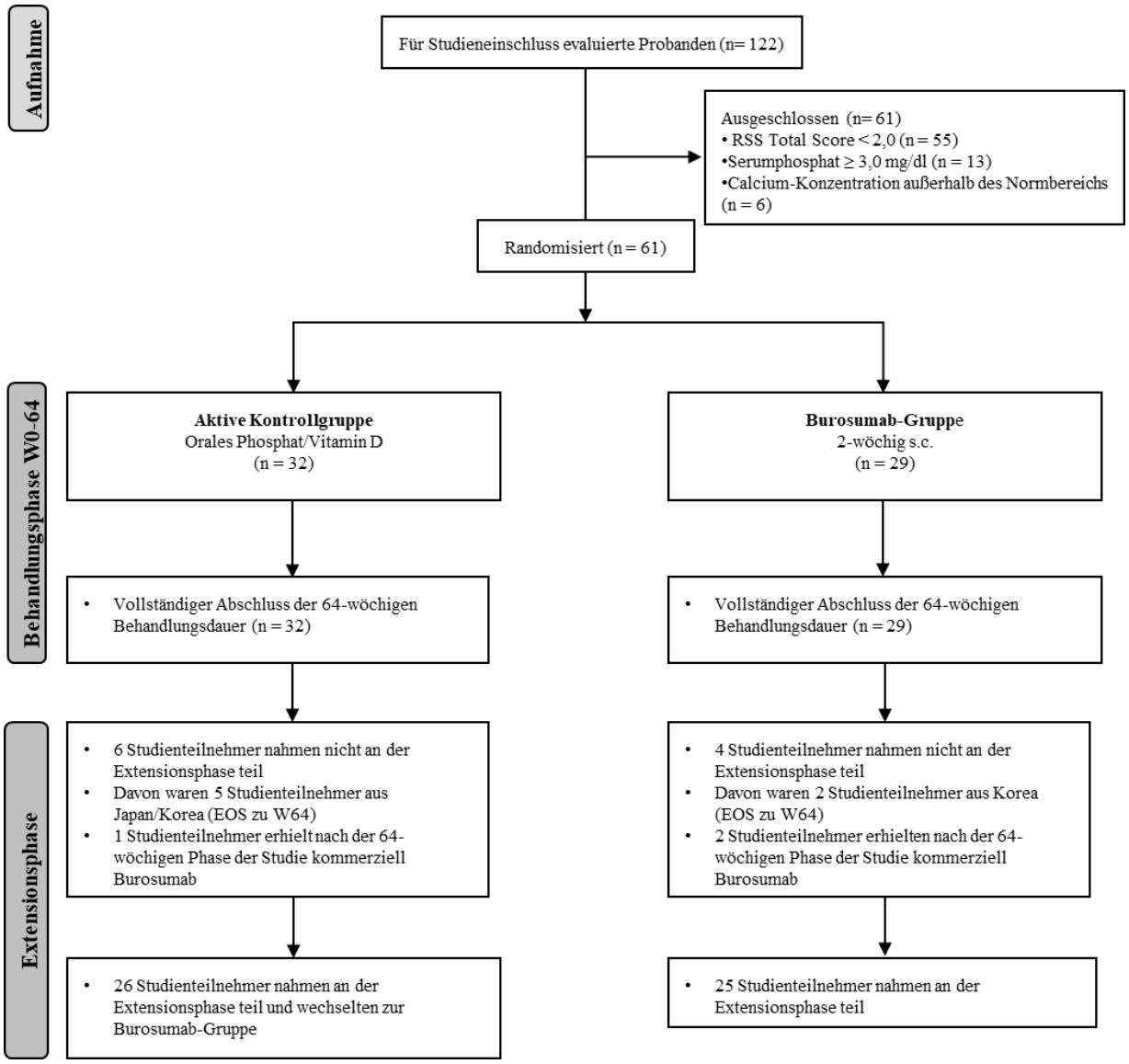

Abbildung 10: Patientenfluss der Studie UX023-CL301 gemäß CONSORT

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 4-87: (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UX023-CL301

Studie: UX023-CL301

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
UX023-CL301 Protocol (week 64) + Amendment	
UX023-CL301 Statistical Analysis Plan	**UX023-CL301 SAP**
UX023-CL301 Clinical Study Report	**UX023-CL301 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. <u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> 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.

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

Nicht zutreffend.

2. <u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> 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, nummerischer Parameter sowie der verblindeten Auswertung der Röntgenbilder eine objektive Beurteilung.

5. Keine sonstigen (endpunktübergreifenden) 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 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** ☐ **nein**

Angaben 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 pädiatrische 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** ☐ **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** ☐ **nein**

Angaben 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** ☐ **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:

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 und Sitzhöhe)

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** ☐ **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. Aufgrund der Verwendung des gleichen Radiometers zur Messung der Stehhöhe/Sitzhöhe 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** ☒ **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. 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** ☐ **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** ☐ **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. Beim Serumphosphat handelt es sich um einen objektiven nummerischen 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** ☒ **nein**

Angaben 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** ☐ **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. 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 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** ☐ **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. 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** ☐ **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. 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** ☒ **nein**

Angaben 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** ☐ **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 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** ☐ **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** ☐ **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. 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 Evenst (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-Cl301 wird der Endpunkt **zahnärztliche Untersuchungen** unter dem Aspekt **Sicherheit** berücksichtigt.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☐ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein**: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.
→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.
für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

☐ **ja**: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

☐ **unklar**: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

☐ **nein**: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

☐ **ja**: Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein**: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar**: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

☐ **nein**: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

☐ **unklar**: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

☐ **nein**: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja:** Die Patienten waren verblindet.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:
- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter
- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:
- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
  Zulässige Gründe sind:
  - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
  - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
  - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.
Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.
- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
  - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
  - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.*
  - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
  - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

☐ **ja**

☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig**: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch**: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

_____

_____

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. **Verblindung der Endpunkterheber**

   Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
   In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

   ☐ **ja:** Der Endpunkt wurde verblindet erhoben.

   ☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

   ☐ **nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

   Angaben zum Kriterium; obligate Begründung für die Einstufung:

   _____

   _____

2. **Adäquate Umsetzung des ITT-Prinzips**

   Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

   ☐ **ja:** Eines der folgenden Merkmale trifft zu:
   - Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

☐ **unklar**: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

☐ **nein**: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.
- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

☐ **ja**

☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig**: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch**: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

Anhang 4-G: Ergänzende Analysen

Ergänzende *post-hoc* durchgeführte Analysen zu den Abschnitten 4.3.1.3 und 4.3.1.3.2 sind separat im Dokument 2022-01-31_Modul4A_Burosumab_Anhang4-G.pdf aufgeführt.