in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Burosumab for treating X-linked hypophosphataemia

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam and Maastricht University

Authors
Rob Riemsma, Review Manager, KSR Ltd, UK
Isaac Corro Ramos, Health Economics Researcher, Erasmus School of Health Policy & Management, EUR, the Netherlands
Nasuh Büyükkaramikli, Health Economics Researcher, EUR
Stephanie Swift, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Remziye Zaim, Health Economics Researcher, EUR
Gimon de Graaf, Health Economics Researcher, EUR
Gill Worthy, Statistician, KSR Ltd
Lisa Stirk, Information Specialist, KSR Ltd
Maiwenn Al, Health Economics Researcher, EUR
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 23/04/2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number HST 17/40/01.

Declared competing interests of the authors
None.

Acknowledgements
We are very grateful to the following experts who offered clinical advice and comments on the draft report:
Professor Peter Selby, Consultant Physician and Honorary Clinical Professor of Metabolic Bone Disease, University of Manchester.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:
Riemsma R, Corro Ramos I, Büyükkaramikli N, Swift S, Armstrong N, Zaim R, De Graaf G, Worthy G, Stirk L, Al M, Kleijnen J. Burosumab for treating X-linked hypophosphataemia: a Highly Specialised Technology Evaluation. York: Kleijnen Systematic Reviews Ltd, 2018.

Contributions of authors
Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Isaac Corro Ramos acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Nasuh Büyükkaramikli, Remziye Zaim, Gimon de Graaf and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Stephanie Swift acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Lisa Stirk critiqued the search methods in the submission and contributed to the writing of the report. Maiwenn Al critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance on the health economics part of the project. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

ABBREVIATIONS

1,25(OH)2D	1,25-dihydroxyvitamin D
6MWT	Six minute walk test
ADHR	Autosomal dominant hypophosphataemic rickets
AE	Adverse events
ALP	Alkaline phosphatase
ANCOVA	Analysis of covariance
ARHR1	Autosomal recessive hypophosphataemic rickets type 1
ARHR2)	Autosomal recessive hypophosphataemic rickets type 2
BALP	Bone-specific alkaline phosphatase
BBS	Brittle Bone Society
BI	Budget impact
BIC	Bayesian information criterion
BNF	British National Formulary
BPABG	British Paediatric and Adolescent Bone Group
BPI	Brief Pain Inventory
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CRD	Centre for Reviews and Dissemination
CRF	Case report form
CS	Company submission
CSR	Clinical study report
DSA	Deterministic sensitivity analysis
ECG	Electrocardiogram
ECHO	Echocardiogram
EDC	Electronic data capture
EMA	European Medicines Agency
EQ-5D-5L	Euroqol 5-dimension 5-level questionnaire
ERG	Evidence Review Group
ERN-BOND	European Reference Network on Rare Bone Disorders
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
FGF23	Fibroblast growth factor 23
GEE	Generalised estimating equations
HPO	Human Phenotype Ontology
HR	Hypophosphataemic rickets
HRG	Healthcare resource group
HRQoL	Health-related quality of life
HS	Health state
HST	Highly Specialised Technologies
HTA	Health Technology Assessment
ICER	Incremental cost effectiveness ratio
IgG1	Human immunoglobulin G1
iPTH	Intact parathyroid hormone
ITT	Intent to-treat
IWRS	Interactive web response system
KSR	Kleijnen Systematic Reviews
LLN	Lower limit of normal
LS	Least squares
LVH	Left ventricular hypertrophy

MHRA	Medicines and Healthcare Products Regulatory Agency
NHSE	National Health Service England
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NSAID	Non-steroidal anti-inflammatory drug
PAS	Patient access scheme
PbR	Payments by results
PCS	Physical component summary
PD	Pharmacodynamic(s)
PDMA	Pharmaceuticals and Medical Devices Agency
PHEX	Phosphate-regulating endopeptidase homolog, X-linked (phosphate-regulating gene with homology to endopeptidases located on the X chromosome)
PIM	Promising Innovative Medicine
PK	Pharmacokinetic(s)
PODCI	Pediatric Outcomes Data Collection Instrument
POSNA	Pediatric Orthopedic Society of North America
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSM	Propensity score matching
PSS	Personal social services
PTH	Parathyroid hormone
Q2W	Biweekly, once every 2 weeks
Q4W	Monthly, once every 4 weeks
QALY(s)	Quality-adjusted life year(s)
QoL	Quality of life
RCPCH	Royal College of Paediatrics and Child Health
RCT	Randomised control trials
RGI-C	Radiographic Global Impression of Change
RSS	Rickets Severity Score
RUDY	Rare and Undiagnosed Diseases Study
SAE	Serious adverse event
SC	Subcutaneous
SD	Standard deviation
SDS	Standard deviation scores
SE	Standard error
SF-10	SF-10 Health Survey for Children
SF-36	36-Item Short Form Survey
SmPC	Summary of product characteristics
SoC	Standard of care
SOC	System organ class (for adverse events coding by MedDRA)
SPC	Summary of Product Characteristics
TEAE	Treatment emergent adverse event
TmP/GFR	Ratio of renal tubular maximum reabsorption rate of phosphate (TmP) to glomerular filtration rate (GFR)
UK	United Kingdom
ULN	Upper limit of normal
USA	United States of America
VUS	Variant of unknown significance
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index
XLH	X-linked hypophosphataemia

Table of Contents

TABLE OF TABLES

TABLE OF FIGURES

1. SUMMARY

1.1 Background

X-linked hypophosphataemia (XLH) is a rare, genetic, chronically debilitating and deforming disease that profoundly impacts the affected individual's day to day functioning and health-related quality of life (HRQoL). As a genetic disease it can affect whole families and consequently have a wide impact on the quality of life of generations of families.

In XLH, genetic mutations result in an inactive phosphate-regulating enzyme and lead to high levels of circulating fibroblast growth factor 23 (FGF23). Excess FGF23 leads to increased urinary phosphate excretion, reduced 1,25-dihydroxyvitamin D (1,25(OH)2D) synthesis, and hypophosphataemia.

1.2 Summary of submitted evidence on the nature of the condition and the impact of the new technology

XLH is characterised by dysfunction of mineral metabolism (serum phosphate, serum calcium), endocrine function and renal function. The corresponding clinical manifestations of XLH include delayed walking, waddling gait, leg bowing, enlarged cartilages, bone and/or joint pain, craniosynostosis, spontaneous dental abscesses, growth failure, fractures, mineralisation defects (rickets and osteomalacia), severe dental anomalies, hearing loss and fatigue. Rickets, the hallmark of XLH in children, is associated with substantial skeletal deformities that cause daily pain and impair physical functioning. Children may be severely limited in their daily activities, such as walking, due to deformity and antalgic gait. When these deformities become permanent, people with XLH suffer lifelong disability and pain.

Children with XLH often have trouble performing age-appropriate gross motor activities, such as walking, running, and jumping, due to bowing of the femur, tibia, and/or fibula and the rotation of the tibia that causes the feet to turn in toward each other. This impaired functionality from an early age can inhibit a child's participation in physical, educational and social activities. In adults, osteomalacia and skeletal deformities lead to development of early osteoarthritis and enthesopathy that cause pain and continue to limit physical function.

The long-term goal of therapy in children with XLH is to improve or heal rickets and prevent or correct the skeletal abnormalities associated with it, to prevent the ongoing mechanical dysfunction associated with chronic weight bearing on poorly aligned bones and joints, and to reduce the child's pain and disability.

Burosumab is a recombinant human IgG1 monoclonal antibody that binds to and inhibits the activity of FGF23. By inhibiting FGF23, burosumab increases tubular reabsorption of phosphate from the kidney and through the production of 1,25(OH)2D enhances intestinal absorption of calcium and phosphate. Burosumab improves phosphate homeostasis and its major pathologic consequences (rickets and osteomalacia), and consequently aims to resolve the skeletal and non-skeletal manifestations of XLH.

The European Medicines Agency (EMA) awarded burosumab conditional marketing authorisation on 23 February 2018. The full indication is: "Crysvita is indicated for the treatment of X-linked hypophosphataemia with radiographic evidence of bone disease in children 1 year of age and older and adolescents with growing skeletons." It is proposed that Crysvita be prescribed by physicians experienced in the management of patients with metabolic bone diseases.

1.3 Critique of the decision problem in the company's submission

Most of the evidence is presented as single arm studies including either treated patients (two studies, both with extensions that are still ongoing) or historical control patients (one study, with patients from one single centre, radiographic analysis set ([redacted])). The historical control study (CL002) included patients aged from five to 14 years and can therefore only serve as a control group for study CL201 (children aged five to 12 years). For patients with XLH aged one to four years old, the CS only presents a single arm burosumab study (CL205), no control data for this age group were provided.

A randomised controlled study comparing burosumab with active control (oral phosphate/active vitamin D therapy) in children with XLH (aged one to ≤12 years) is currently ongoing. [redacted]. These results will considerably reduce the uncertainty surrounding the clinical effectiveness of burosumab relative to conventional therapy in children with XLH aged between one and 12 years. They will provide a direct estimate of clinical effectiveness and will be more reliable and up-to-date than the current results based on propensity score matching between a single-arm and historical cohort study.

1.4 Summary of clinical effectiveness evidence submitted by the company

The studies included in the submission focus on the following populations and studies:

- Paediatric patients with XLH, five to 12 years old: Study CL201 (open-label RCT comparing different doses of burosumab (biweekly or monthly administration of burosumab doses 0.1 to 2.0 mg/kg))
- Paediatric patients with XLH, one to four years old: Study CL205 (open-label study to assess the safety, pharmacodynamics and efficacy of burosumab (biweekly administration of burosumab at a target dose of 0.8 mg/kg))
- Paediatric Patients with XLH, five to 14 years old: Study CL002 (A retrospective longitudinal study of skeletal outcomes in children with XLH. No burosumab administered; however, study inclusion required the use of conventional therapy (oral phosphate/active vitamin D))

Results from CL201 show that burosumab significantly improves rickets at week 40 and week 64, compared to baseline. The primary endpoint, the rickets severity score (RSS) was reduced from baseline by 61% at week 40 ($p<0.0001$) by 58% at week 64 ($p < 0.0001$) with biweekly burosumab. Burosumab treatment also resulted in healing of rickets as assessed by Radiographic Global Impression of Change (RGI-C) scores. The RGI-C score at Week 64 was +1.62. At Week 64, [redacted] of children treated with biweekly burosumab had healing of rickets (RGI-C global scores ≥ 1.0). Furthermore, [redacted] of children treated with burosumab had substantial healing of rickets (RGI-C global scores ≥ 2.0). Growth velocity increased by [redacted]) in children treated with burosumab every two weeks, with a corresponding least-squared (LS) mean change in standing height z-score of [redacted]). Biweekly burosumab also resulted in improved functional assessments and patient-reported outcomes in CL201. Walking ability, as assessed by LS mean distance walked in the six-minute walk test (6MWT), increased from baseline by [redacted] at week 64 [redacted]. Functional disability was assessed using the Pediatric Orthopedic Society of North America - Pediatric Outcomes Data Collection Instrument (POSNA-PODCI). Biweekly burosumab treatment increased scores for sports/physical functioning and pain/comfort into the normal range seen in healthy children; LS mean scores showed improvements of [redacted] and [redacted] at week 64, respectively.

Results from CL002 show that RSS was reduced by [redacted] (over a median period of 102 weeks) after long-term conventional therapy. The RGI-C score was [redacted] with conventional therapy in Study

CL002 (median 102 weeks). Furthermore, [redacted] of children treated with conventional therapy in Study CL002 had substantial healing of rickets (RGI-C global scores ≥ 2.0). After long-term treatment with conventional therapy in Study CL002, [redacted].

In study CL205 (13 children with XLH aged one to four years), burosumab treatment for 40 weeks significantly reduced RSS total score at week 40 by 59% (LS mean change of -1.73, $p < 0.0001$, ANCOVA model).

No patient died or discontinued from CL201 or CL205 for any reason; all patients continued treatment on study as of the data cut-off dates.

The most common adverse drug reaction reported in paediatric patients up to 64 weeks treatment with burosumab was injection site reactions (57%), headache (54%), pain in extremity (42%), vitamin D decreased (28%), rash (23%), toothache (19%), tooth abscess (14%), myalgia (14%), and dizziness (11%). Approximately 57% of the patients had an injection site reaction. The injection site reactions were generally mild in severity, occurred within one day of medicinal product administration, lasted approximately one to three days, required no treatment, and resolved in almost all instances.

In study CL201, one patient experienced serious TEAEs, and [redacted] All 52 patients (100%) experienced at least one TEAE during the study. The most frequent TEAEs (>30% incidence) in study CL201 were [redacted].

The most frequent TEAEs (> 30% incidence [four or more of 13 patients]) in study CL205 were [redacted]

Adverse events of treatment with conventional therapy have not been reported. Therefore, it is not possible to assess the relative safety and toxicity in relation to the comparator.

1.5 Summary of the ERG's critique of clinical effectiveness evidence submitted

The CS states that a systematic review search was undertaken for clinical effectiveness and adverse events evidence using a combined search for all of these areas. The company submission and response to clarification provided sufficient details for the ERG to appraise the literature searches. Of main concern to the ERG was the limited search conducted, which included few XLH synonyms and an unnecessarily restrictive use of a study design filter.

The main limitation of the efficacy data reported in the CS is the study design of the included studies. Due to the absence of a control group in most studies it is not possible to make any direct comparisons between burosumab and conventional therapy. As stated by the company, the "burosumab phase 2 studies were uncontrolled dose finding or single arm studies, therefore an indirect comparison was not feasible" (CS, page 123).

For children between one to four years old, only one study is presented in which all children received burosumab (CL205, N=13). A comparison with "established clinical management without burosumab" is not possible in this group of patients.

For children between five to 12 years old, the CS presents a study in which all children received burosumab (CL201). In addition, the CS presents a control study (CL002) in which children aged between five to 14 years received conventional therapy (i.e. oral phosphate/active vitamin D). Results of these two studies are mainly presented as a naïve comparison, simply reporting individual results from each study side by side. In addition, the company presents comparisons of 'rickets healing' with conventional therapy (Study CL002) versus burosumab (Study CL201) using propensity analysis matching.

In the CS, the company uses the terms 'healing' and 'substantial healing of rickets'. These are defined using RGI-C global scores, where scores ≥+1.0 indicate 'healing of rickets' and scores ≥+2.0 'substantial healing of rickets'. The company does explain that "Healing in this context indicates improvement in the radiographic abnormalities and does not imply that complete healing was observed" (CS, page 100). However, throughout the report the term 'healing of rickets' is used without any explanation of the degree of healing (minimal, substantial or complete). Moreover, it should be noted that RGI-C global scores and RSS scores do not capture all clinical aspects of XLH.

The naïve comparison is unreliable because there are important differences between the inclusion criteria in both studies. Inclusion criteria for patients in studies CL201 and CL002 are similar in that patients in both studies were diagnosed with XLH and were of similar age. However, children in study CL201 also had: biochemical findings associated with XLH, standing height < 50th percentile for age and gender and radiographic evidence of active bone disease including rickets in the wrists and/or knees, and/or femoral/tibial bowing, or, for expansion patients, an RSS score in the knee of at least 1.5 points as determined by central read. In other words, study CL002 included all children with XLH, while study CL201 included children with more severe symptoms of XLH. This is also reflected in the relatively [redacted] standing height and [redacted] rickets severity score for children in study CL201 when compared to children in study CL002.

The adjusted comparison, using propensity analysis matching, is unreliable because of the limitations associated with these methods, in that the matching can only include those variables measured in both studies. Randomisation in a clinical trial creates a balanced group for both measured and unmeasured variables. In observational studies, the most important factors which are predictive of the outcome may not have been measured and any treatment comparisons using observational study data may be biased. In the CS the company only included three variables in the propensity score matching (PSM): age, gender and RSS total score at baseline. The rationale for variable selection was not provided other than whether they seemed similar or not between the two study populations. No details were provided of how this similarity was judged. The ERG found no statistically significant differences in age and gender between the two groups and considered that only including three variables in the creation of the propensity scores may have been too few.

1.6 Summary of the evidence submitted to support the value for money of the treatment and cost to the NHS and PSS

The company conducted a systematic review of cost effectiveness studies of burosumab and other studies including costs, resource use and any HRQoL measure associated with XLH. A total of eight full-text studies were assessed for eligibility but none of them were deemed relevant to the economic evaluation of burosumab.

The company submission included a model-based cost-utility analysis comparing the use of burosumab with standard of care (SoC) to treat XLH patients with radiographic evidence of bone disease aged one year or older with growing skeletons.

Multiple sources of evidence were used to inform the parameters of the economic model. The proportion of males/females at baseline, the initial distribution of patients per disease severity stratified by age and the transition probabilities for burosumab were derived from the clinical studies CL201 and CL205. Transition probabilities for the SoC arm were derived from a UK chart review in the base-case analysis and from the study CL002 in a scenario analysis. General population weight data (UK growth charts) were used for the weight distribution. Mortality rates were obtained from the national life tables for England, for the period 2014 to 2016, as published by the Office of National Statistics. Utility values for the health states of the model were derived from a vignette study conducted by the company. Additionally, age specific multipliers were used based on the general population.

The price of burosumab was provided by the company. Burosumab is available in 10 mg, 20 mg and 30 mg vials. In the CS, it was stated that the Summary of Product Characteristics (SmPC) recommends dose rounding to the nearest 10 mg. Based on this assumption, annual patient costs by age and weight were estimated in the base-case analysis. Resource use for burosumab monitoring was based on expert opinion, while unit costs were taken from NHS reference costs. Standard of care treatment costs were estimated based on the dose recommended in clinical guidelines and the summary of product characteristics. Unit costs were taken from the British National Formulary (BNF). Resource use for surveillance costs was based on expert opinion and unit costs were taken from NHS reference costs. Physiotherapy resource use was based on published literature and complemented by expert opinion. Unit costs taken from PSSRU. A number of different sources were used for the estimation of orthopaedic intervention costs. Resource use was based on the prevalence observed in CL201, published literature and expert opinion. Unit costs were mostly sourced from the NHS reference costs, except the unit costs for osteotomy, which were based on published literature.

A deterministic one-way sensitivity analysis was conducted for key clinical and economic parameters in the model. A probabilistic sensitivity analysis was also conducted. A number of scenario analyses were also performed to assess the robustness of the model results to changes in structural assumptions made by the company.

The company's analysis estimated that patients treated with burosumab gained 10.304 more discounted quality adjusted life years (QALYs) compared to SoC at an additional cost of [redacted], resulting in a cost per QALY of [redacted]. When discounting was not applied, the estimated gain in QALYs was 16.891 at an additional cost of [redacted], resulting in an ICER equal to [redacted].

1.7 Summary of the ERG's critique of the value for money evidence submitted

The CS states that a systematic review search was undertaken for economic, cost and resource use and HRQoL evidence using a combined search for all of these areas. The company submission and response to clarification provided sufficient details for the ERG to appraise the literature searches. Of main concern to the ERG was the narrow search conducted, which included few XLH synonyms and an unnecessarily restrictive use of study design filters.

The ERG identified several issues in the company's analyses. The ERG main concerns were related to the method used by the company to estimate the transition probability matrices for burosumab, the source of utilities used by the company and the assumption of lifelong treatment effects of burosumab. The choice of the discount rate was also challenged by the ERG.

The results of the ERG base-case resulted in an ICER increased by [redacted] compared to the company's base-case ICER. Most of the total increase in the ICER (despite the effect of applying the 3.5% discount rate) was due to assuming a treatment effect duration for burosumab of 20 years instead of lifelong as

assumed by the company. The ERG also conducted a new probabilistic sensitivity analysis (PSA) and additional scenario analyses exploring the impact of choosing prior distributions for the burosumab transition matrices. The latter was proven to be crucial and in the several scenarios provided by the ERG, the ICER ranged from [redacted] to [redacted].

Based on the ERG results, it is expected that, from the payer perspective, the decision uncertainty related to burosumab value for money would be low, given that the ICER estimates from all ERG analyses are above the acceptable thresholds considered for orphan drugs and the burosumab cost effectiveness probability at such thresholds was [redacted].

1.8 Summary of the evidence submitted on the impact of the technology beyond direct health benefits and on the provision of specialised services

A budget impact model to estimate the costs to the NHS for a period of five years of adopting burosumab in England was also included in the CS. The results presented by the company suggested that the net budget impact of implementing burosumab (with an estimated prevalence of [redacted] patients) will be [redacted] in the first year and will rise to [redacted] in the fifth year. The cost of burosumab at year 5 amounts to [redacted]. The estimated total number of patients eligible for burosumab treatment after five years is [redacted] and the uptake of burosumab rises from 40% in year 1 to 90% in year 5.

The CS did not include any estimates of costs (savings) or benefits incurred outside of the NHS and PSS associated with of burosumab. The company indicated that at this stage this was not possible to quantify. However, the company expects significant savings to patients through healing of rickets and overall reduction or elimination of symptoms with burosumab.

1.9 Summary of the ERG's critique on the evidence submitted on the impact of the technology on non-health related benefits

The ERG considers the assumptions made in the budget impact analysis questionable. There are concerns about the theoretical population size and the expected uptake rate of burosumab in England. In the CS, it was reported that the size of the patient population ([redacted]) is not expected to change over time. This estimate is based on an assumption that the patients are only treated if they have growing skeletons. In the CS, it was stated that XLH is not associated with an increased risk of death, compared to the standard population. The potential (and theoretical) population size is assumed to remain constant.

Since real-world data suggest that there could be [redacted] XLH patients between one and 17 years of age in England, using the estimate of [redacted] children and assuming a 40% uptake in Year 1, followed by 65% uptake in Year 2 and a 90% uptake thereafter would equate to [redacted] children in year 1, [redacted] children in Year 2 and [redacted] children thereafter being treated with burosumab. The cost of burosumab at year 5 would then amount to [redacted]. The company indicated that burosumab is not expected to require additional resources to enable treatment administration, as it will be delivered via homecare. Homecare provision for XLH is being organised and funded by the company and will therefore not have any additional financial or resource impact on the NHS.

The ERG considers it inadequate that the impact of XLH on costs (savings) or benefits incurred outside of the NHS and personal social services associated with of burosumab was not identified prior to the submission to NICE.

1.10 ERG commentary on the robustness of evidence submitted including strengths, weaknesses and areas of uncertainty

The company's submission provided sufficient details for the ERG to appraise the database searches, which were generally transparent and reproducible. An adequate number of online resources were searched and a good range of additional searches were conducted for grey literature. However, the population facet for each search conducted included few synonyms, and therefore may have missed relevant literature. Given the small number of references retrieved from the search, study design filters were not essential, and may have been unnecessarily restrictive.

The main limitation of the efficacy data reported in the CS is the study design of the included studies. Due to the absence of a control group in most studies it is not possible to make any direct comparisons between burosumab and conventional therapy. As stated by the company, the "burosumab phase 2 studies were uncontrolled dose finding or single arm studies, therefore an indirect comparison was not feasible" (CS, page 123).

For children between one to four years old, only one study is presented in which all children received burosumab (CL205, N=13). A comparison with "established clinical management without burosumab" is not possible in this group of patients.

A randomised controlled study comparing burosumab with active control (oral phosphate/active vitamin D therapy) in children with XLH (aged one to ≤12 years) is currently ongoing. [redacted]. Results from this study will considerably reduce the uncertainty surrounding the clinical effectiveness of burosumab relative to conventional therapy in children with XLH aged between one and 12 years.

A range of relevant economic information was incorporated in the CS, including a QALY-based cost effectiveness model and an assessment of the expected costs to the NHS and PSS in England. However, the CS lacks information about the long-term effects of treatment with burosumab and about the treatment effects of burosumab in adults. The available evidence is limited, which makes the model results highly uncertain and sensitive to key assumptions. The CS also lacks an analysis of the wider societal (non-health) benefits associated with burosumab.

There is substantial uncertainty about the long-term effects of burosumab. The company conducted their analysis upon the assumption that these effects would be lifelong, despite treatment being stopped at the age of 16 in females and 17 in males, but there is no evidence to support that assumption. This assumption was proven to be crucial and one of the main drivers of the cost effectiveness results. Additional uncertainty is generated when translating the clinical outcomes to QALYs since the evidence on HRQoL was based on a vignette study describing the health states of the economic model that were valued by (only six) clinical experts. Since there is no direct or indirect evidence comparing burosumab to SoC, the assumed treatment effect of burosumab, as reflected by the transition probability matrices, is also very uncertain.

The ERG considers that the uncertainty around the reported ICERs is likely to be larger than suggested by the PSAs presented in this report. Given that a PSA only addresses parameter uncertainty, other sources of uncertainty, like the ones mentioned above, could not be included in the PSA.

1.11 Summary of exploratory sensitivity analyses undertaken by the ERG

The main changes made by the ERG to the company's model included the use of alternative transition probabilities for burosumab, sourcing utilities directly from the vignette study report (and not from the company submission) and the operationalisation of the treatment effect of burosumab. Minor changes

included discounting costs and health outcomes at 3.5%, although this was proven to have a major impact on the model results.

The results of the ERG base-case, before applying the 3.5% discount rate on costs and health outcomes, resulted in an ICER increased by [redacted] compared to the company's base-case ICER. After applying the 3.5% discount rate, the ICER increased by [redacted]. Although sourcing the utilities from Lloyd et al. had a substantial impact on the ICER (increased by [redacted]), most of the total increase in the ICER (before applying the 3.5% discount rate) was due to the assumption of reducing the utilities of burosumab patients 20 years after the end of treatment. Since there is uncertainty on whether this value of 20 years will be observed in real life, the ERG assessed the impact of assuming a different duration for the burosumab treatment effects on the cost effectiveness results. The difference between assuming 20 years duration of treatment effect (ERG) and lifelong treatment effects (company) was that the ICER increased by approximately [redacted] under the ERG assumption. Assuming smaller values for the duration of the burosumab treatment effect increased the ICER. In particular, when this was assumed to be five years, the ICER was [redacted].

The ERG was concerned that the PSA results presented by the company were underestimating the uncertainty associated with the transition probabilities for burosumab. For that reason, a new PSA and additional scenarios exploring the impact of choosing prior distributions for the burosumab transition matrices were conducted by the ERG. The latter was proven to be crucial and in the several scenarios provided by the ERG, the probabilistic ICER ranged from [redacted] to [redacted].

Based on the ERG results, it is expected though that, from the payer perspective, the decision uncertainty related to burosumab value for money would be low, given that the ICER estimates from all ERG analyses are above the acceptable thresholds considered for orphan drugs and the burosumab cost effectiveness probability at such thresholds was [redacted].

2 BACKGROUND

2.1 Introduction

This report provides an overview of X-linked hypophosphataemia (XLH) and its management. The content of this chapter is based on relevant literature, information provided by clinical advisors to the Evidence Review Group (ERG) and information presented in the background sections of the submission (CS),[1] with additional information provided in the company's response to clarification letter.[2] For additional information on the aetiology, epidemiology, health impact, prognosis and management of XLH, please see the CS (pages 32-57).

2.2 Description of health problem

2.2.1 Paediatric XLH

X-linked hypophosphataemia (XLH) is a rare and often genetic (hereditary) disorder. In XLH, high levels of circulating FGF23 lead to excess urinary phosphate excretion and subsequent hypophosphataemia. Since phosphate is required to build and maintain bones, patients typically develop bone deformities, defective tooth mineralisation and experience growth problems.

The major pathologic consequences of XLH in the bone are rickets (in children) and osteomalacia (in adults). Rickets, the hallmark of XLH in children, is associated with substantial skeletal deformities that cause daily pain and impair physical functioning, such that a young child may be limited in his/her daily activities and will suffer lifelong disability and pain as these deformities become irreversible when growth ceases. Children with XLH often experience difficulty performing age-appropriate gross motor activities, such as walking, running and jumping, due to bowing of the femur, tibia, and/or fibula and the tibia rotation that causes the feet to turn in toward each other. In addition, children experience muscle weakness, fatigue, and other physical functioning deficits that are likely caused by the diverse physiological impacts of hypophosphataemia, which may be independent of rickets. Bowing of the legs in children with XLH can be substantial and severe. Defects in the growth plate also lead to impairment in growth and growth potential. The combination of height loss caused by the bowing of the legs and the growth plate defects can lead to a permanent loss of growth potential and short stature which can have psychosocial consequences for the individual.[3]

Over time, symptoms may progress to include bone pain, joint pain caused by hardening (calcification) of tendons and ligaments, and dental pain. Some people with XLH may also experience hearing loss.[4, 5] In addition to the substantial impacts on skeletal disease, low serum phosphorous in XLH patients may contribute to muscle dysfunction, reduced mobility and physical functioning, and fatigue. Because XLH is a lifelong disease, bone and joint damage, osteomalacia and reduced mobility acquired during childhood, are continued into adulthood.

Rickets is typically measured using radiographs as the gold standard. The Rickets Severity Score (RSS), is a radiographic scoring method developed to assess the severity of nutritional rickets. The RSS provides the absolute score of epiphyseal/distal metaphyseal abnormalities in the wrists and knees based on the degree of metaphyseal fraying, concavity, and the proportion of the growth plate affected.[6] The RSS is a 10-point scoring method, where a score of 0 indicates no rickets and a score of 10 indicates the highest severity of rickets. The usual range of RSS total scores in XLH is between 0 and 6.5 but reflects only the epiphyseal/distal metaphyseal portion of the skeletal abnormalities that are common in affected children, as there are other aspects of XLH not fully captured in the RSS. These other findings include coxa vara (a hip deformity that causes leg length discrepancies and gait abnormalities), tibial

torsion (a twisting of the shins that causes the feet to turn inward), and genu varum (bowing) or genu valgum (knock knees).

The Radiographic Global Impression of Change (RGI-C) is an alternative radiographic scoring method for rickets. This indicates the change in abnormalities and deformities between time points. The RGI-C provides a complementary method to RSS that allows for comparison with previous radiographs. Together, both measures provide a broader insight into bone disease than any one score alone.

ERG comment: The current submission focusses on paediatric XLH, which is defined as XLH in children aged 1-17 years. Of note, the comment that "other [clinical] aspects of XLH are not fully captured in the RSS" (CS, page 41) has to be considered in the context of the economic model in the CS, which only considers RSS score alone as a clinical outcome measure. The diverse physiological impacts of hypophosphataemia, which may be independent of rickets, are not captured at all in this submission. This is acknowledged as a limitation in the CS.

Only RSS scores are used in the model; RGI-C scores are not considered, despite the company considering these to represent more sensitive readouts of rickets severity and having this information available from each of the clinical studies used to inform the economic model (CL201, CL205 and CL002).

2.2.2 Epidemiology

2.2.2.1 Prevalence of XLH

The CS contains three key references that estimate the prevalence of XLH. One published study reports on prevalence in Denmark,[7] one unpublished draft study manuscript reports on prevalence in the UK,[8] and one real-world dataset commissioned by Kyowa Kirin through the British Paediatric and Adolescent Bone Group and the European Reference Network on Rare Bone Disorders (BPABG/ERN-BOND) provides the number of XLH patients currently in selected treatment centres in the UK.

The Danish published study estimates the incidence of XLH to be 3.9 per 100,000, based on 0.57 cases being diagnosed out of 14,558 children born in Denmark in one year.[7] The estimation that this would equate to 26 new patients annually in England appears valid against a mean number of 663,157 births in England over the same incidence period (1982 to 2002).[9] Given the size of the total prevalent population [ ], this is considered by the company to be implausible.

[redacted] (based on Delmestri et al 2018[8] and a personal communication from this study's authors to Kyowa Kirin). This prevalence was applied to the general population for England in children aged between one and 17 years[10] to estimate [redacted] children with XLH (Table 2.1, below; Table 60 in the CS). However, it remains unclear how this prevalence value has been calculated (e.g. the denominator, how the 522 test cases were originally identified etc.). There is further uncertainty around this figure since, as the company have acknowledged in their clarification letter response, "[redacted]".[2] Consequently, the estimate provided from this preliminary, unpublished dataset must be interpreted with caution.

[redacted]

[redacted] Based on the information from BPABG plus information obtained through re-engaging [redacted] in England in the 1-17 age range in the company's response to clarification letter (question A5[2]). Since eligibility for treatment with burosumab requires radiographic evidence of bone disease in children and adolescents, Kyowa Kirin considers it unlikely that such patients would be undiagnosed and therefore not in treatment at one of these centres, as this degree of disease is likely to be symptomatic. According to the CS, the size of the patient population is not expected to change with time as patients are only treated if they have growing skeletons i.e. each year there may be new patients but there will also be a similar number of patients ceasing treatment.

In the company's statement in their response to clarification letter, they report that [redacted][2] [redacted] of these [redacted] patients appear to be currently treated in ERN-BOND centres. However, it is not clear if all ERN-BOND centres in England have been included in this analysis. Additional ERN-BOND centres (Oxford University Hospitals and Sheffield Teaching Hospitals[11]) do not appear in the list provided in Table 5 of the company's response to clarification letter[2]; thus, this real-world dataset may represent an underestimation of the real-world prevalence of XLH in England.

Since real-world data suggests there are [redacted] confirmed XLH patients between one and 17 years of age in England, there is a discrepancy between the Danish study's estimated values and BPABG/ERN-BOND real-world values (we would expect 26 new patients per year based on an incidence of 3.9 per 100,000,[7] but have identified an average of [redacted] new patients per year based on a real-world confirmed patient dataset). In their response to clarification letter,[2] the company questioned whether the methods used by Beck-Nielsen et al. 2009[7] may have overestimated the incidence of XLH. However, the ERG finds the methods described by Beck-Nielsen to be acceptable (patients diagnosed with rickets were identified from medical records, and the entire medical record was subsequently reviewed for biochemical and clinical parameters, similar to the methods described by Delmestri et al. 2018[8]).

Ultimately, the ERG is not confident in the data provided to support the proposed prevalence or incidence values for XLH in children aged one to 17 years the UK. This is further compounded by the suggestion that the number of cases in certain age ranges in a key study in the UK were subject to unexpected fluctuations,[8] as highlighted in the company's response to clarification letter (question A18, part IV[2]), which does not support the idea of the population of XLH remaining constant. These nuances have not been fully captured in any of the presented data.

Table 2.1: Derivation of number of XLH children on treatment in their first year

Parameter	Value	Reference
Population of females aged 1-16 years in England (2016)	5,695,613	Office for National Statistics 2016[10]
Population of males aged 1-17 years in England (2016)	5,110,255	Office for National Statistics 2016[10]
Prevalence of XLH	[redacted]	Draft abstract[8]
Number of patients eligible for burosumab per year	[redacted]	
Source: CS, Table 60		

2.2.3 Aetiology

Most XLH patients inherit their disease (i.e. have a genetic form of XLH), but a proportion (approximately 20%) develop the disease through new de novo somatic mutations.[12, 13]

The genetic form of the disease is an X-linked disorder caused by a defect in the phosphate-regulating gene with homologies to endopeptidases on the X chromosome (PHEX) peptidase which is part of the phosphate sensing system in osteocytes. Only one mutated copy of the gene is enough to cause the condition in both males and females, therefore a female with XLH has a 50% chance of passing along a mutation to each of her children. Since males only have one X-chromosome, a male with XLH will pass along the condition to all of his daughters, but to none of his sons.

ERG comment: The described aetiology of the disease is in line with the description in the literature.

2.2.4 Pathogenesis

The aetiology and pathophysiological mechanisms behind XLH remain largely unknown. Patients with XLH carry mutations in the PHEX gene, which leads to an erroneous signal in the phosphate sensing control system and an inappropriate excess of FGF23. However, the mechanism through which PHEX disruption results in elevated FGF23 is still unclear.

Excess FGF23 drives the pathophysiology of XLH, leading to impaired conservation of phosphate by the kidney and consequent hypophosphataemia.[14, 15] FGF23 also suppresses 1,25(OH)2D production,[16] resulting in decreased intestinal absorption of calcium and phosphate, further impairing the body's phosphorus supply.[17] As a consequence, patients with XLH have defective bone mineralisation, resulting in low bone turnover and poor quality bone.[18] In addition, many patients have muscle function deficits[19, 20] that may be related to insufficient quantities of adenosine triphosphate (ATP) as a consequence of chronically low concentrations of extracellular phosphate.[19, 21] The musculoskeletal effects of chronic hypophosphataemia further lead to the clinical manifestations and morbidities seen in both children and adults with XLH.

XLH is characterised by biochemical imbalance, in particular regarding:

- Measures of mineral metabolism (serum phosphate, serum calcium)
- Measures of endocrine function (serum values of FGF23, 1,25(OH)2D, insulin-like growth factor I, alkaline phosphatase (ALP), osteocalcin, growth hormone)
- Measures of renal function (urinary calcium to creatinine ratio, maximum rate of renal tubular reabsorption of phosphate normalised to the glomerular filtration rate (TmP/GFR)).

Serum ALP activity is elevated in children with XLH, to two to three times the upper limit of normal.[22] The magnitude of total and bone-specific ALP elevation correlates with the magnitude of rickets.[3] These parameters are commonly used as indicators of the presence and severity of rickets and is one of the primary methods used by physicians managing conventional therapy of XLH as a tool to assess results, since repeated X-rays are not advisable for children. Healing rickets by normalising ALP is the primary objective in children.[23]

ERG comment: In terms of normalising serum ALP, which is indicated throughout the CS to represent a primary objective towards healing rickets in children, it is important to note that only a proportion of children with XLH appear to present with elevated serum ALP while some remain within the normal reference range.[22]

The CS states that, "the magnitude of total and bone-specific ALP elevation correlates with the magnitude of rickets" and provides Carpenter 2011 as a reference.[3] However, this study does not describe a proportional relationship between ALP and rickets severity. Since normalising ALP is defined as the primary therapeutic objective in children with XLH, it would be of clear clinical relevance to include ALP as a clinical outcome in the economic model. Currently, RSS is the only clinical outcome that is used to inform the economic model. It is important to note that there is no evidence presented in the CS that rickets severity is a useful proxy marker that correlates with serum ALP; therefore, its relevance to the stated primary therapeutic objective in XLH patients remains unsupported.

2.2.5 Clinical features

The most important clinical features of paediatric XLH are reported to include: skeletal deformities, growth defects and dental issues.

Skeletal abnormalities include bowing of the femur, tibia/fibula, gait disturbance, joint pain, bone pain and restricted range of motion. Such deformities are severe enough to require at least one surgery in approximately 30% of paediatric XLH patients.[24] Skeletal abnormities, including bowing of the legs, and the associated misaligned joints, disproportionate growth and difficulty walking, persist despite treatment from an early age with conventional therapy (oral phosphate and active vitamin D).[25]

Growth failure appears frequently in children with XLH. The combination of height loss caused by the bowing of the legs and growth plate defects can lead to a permanent loss of growth potential despite the fact that children with XLH experience a normal pubertal growth spurt.[3] In the burden of illness study, CL001, diminished height was reported for (57/71 [80%]) of children.

Children with XLH who are on conventional treatment with alfacalcidol or calcitriol and phosphate show progressive stunting and body disproportion during childhood that is mainly due to diminished growth capacity in legs.[26] 25–40% of patients with well-controlled XLH show linear growth failure despite optimal treatment and have a final height under -2 standard deviation scores (SDS).[27-35] In a study of 28 XLH patients from 1971 to 2011, a significant difference was found between the initial stature and the final stature in only six patients who were treated with vitamin D and phosphate.[36]

Dental disease includes delayed dentition and dental abscesses, which are thought to arise from the limited mineralisation of the dentine compartment of the tooth. In study CL001, [redacted] of children and adolescents had previously had dental surgery.[37] Oral findings in 10 young patients with XLH and an average age of nine years have been enamel and dentine abnormalities, high pulp horns, large pulp chambers, and some cases of periapical abscesses related to teeth without caries or traumatic injuries.[38] A further study of 53 patients (adults and children) with confirmed hypophosphataemic rickets (HR) found that endodontically affected teeth are common, and the number of affected teeth increased significantly with age.[21] Hence, the need for endodontic treatment among HR patients is comprehensive.

Other studies were included in the CS to describe dental disease in XLH patients, but only focussed on adult patients alone, and therefore were not relevant to this appraisal.[4]

Clinical heterogeneity among XLH child and adult patients has been frequently reported.[3, 23] The clinical expression of the disease is widely variable, ranging from a mild abnormality, the apparent isolated occurrence of hypophosphataemia, to severe bone disease.[22] Varied clinical findings are reported even among siblings with the condition.[39]

ERG comment: Growth failure is reported in 25–40% of patients with well-controlled XLH despite optimal treatment, resulting in a final height under -2 standard deviation scores (SDS).[27-35] It is

presumed from this value that the remaining 60-75% of patients with well-controlled XLH achieve normal growth rates with conventional therapy. Other research has indicated that height velocity commonly increases during the first year of conventional therapy, and after two years of successful treatment, can be restored to its maximal potential in the majority of patients, although adult height usually remains compromised.[3, 23]

Dental disease in XLH patients is highlighted in the CS with a study by Anderson 2012 that assesses 53 patients with hypophosphataemic rickets.[21] Sixteen out of 53 patients were <18 years of age and therefore represent the population of interest for the burosumab indication described in the CS. Of these 16 patients, the mean number of endodontically affected teeth was 0.3 (standard deviation (SD) 0.9), while the median number was 0 (first and third quartile: 0.0 and 0.0). No comparisons were provided either in the referenced study, in the CS[1] or in the company's response to clarification letter (question A17[2]) for the number of endodontically affected teeth that would be expected in a healthy age-matched population. Based on the current information, the need for endodontic treatment among paediatric HR patients cannot be considered comprehensive, although it appears clear that dental issues are prevalent in adult XLH patients.

Clinical heterogeneity, which the CS highlights has been frequently reported for XLH patients, is a core issue that may impact burosumab treatment. Some patients with a PHEX mutation who are diagnosed with XLH retain residual gene activity.[17] In practical terms, this may mean that further dose-titrations are necessary that take into consideration not just weight but also residual gene activity. It is unclear if there is a validated test available to determine PHEX activity.

2.2.6 Diagnosis

Diagnosis of XLH is typically based on clinical findings, radiographic findings, biochemical testing and family history. Family history remains critically important to the early recognition of inherited forms. Although, genetic testing is increasingly used to confirm the diagnosis of XLH, radiographs have been the gold standard for the diagnosis and evaluation of rickets for several decades.[18, 40-42] The radiographic characteristics of rickets include lucency in the metaphyses, physeal widening, fraying and cupping.[6, 42] These diagnostic radiographic features of rickets typically reflect the impaired mineralisation and ossification affecting the growth plate. Bone manifestations are best seen in the metaphyses of rapidly growing bones, including the distal radius and ulna, distal femur, proximal and distal tibia and proximal humerus.[6, 42]

Paediatric patients with XLH are managed by paediatric endocrinologists and paediatric nephrologists. There are a limited number of expert clinicians with the necessary training and experience in rare metabolic bone diseases to appropriately manage children with XLH. It is anticipated that treatment would be initiated and monitored by specialist centres and clinicians.

2.2.7 Prognosis

As an update from the CS, which stated that no empirical evidence documenting the impact of XLH on mortality has been identified and that XLH is not thought to have an impact on the life expectancy of patients, a new analysis provided in the company's response to clarification letter stated that

[redacted]

[redacted].[2]

ERG comment: The original statement (that XLH had no impact on life expectancy) was unlikely to be accurate given the extensive pathological manifestations associated with the disease. The updated information that mortality is impacted in XLH patients has been updated in the company's economic model and [redacted].

2.2.8 Impact on patients' health-related quality of life (HRQoL)

2.2.8.1 Impact on paediatric HRQoL

As a rare, orphan disease area, XLH has not been the subject of extensive quality of life studies. Systematic reviews have identified very few studies including empirical evidence documenting the impact of XLH on quality of life; such studies are predominantly conducted in adult XLH patients.

From a young age, XLH has a detrimental impact on the quality of life of patients and families which continues throughout aging to adulthood. Familial cases are particularly burdensome since many members of the family may have the condition, such that a patient may also be a caregiver and vice versa.

As children grow up, they may notice the ways in which they are different from their peers; this can become more apparent to them when they go to school and can result in teasing and bullying by their peers. These differences could be associated with physical appearance, as their legs may develop 'bowing,' or their ability to join in with sports or at playtime. Even if physical appearance is not an issue, the child may begin to question why they have to take regular medication when their peers do not.[43] Difficulties may also be experienced in gross motor skills such as walking, running and jumping, due to symptoms such as bowing of the femur/tibia and/or fibula and the rotation of the tibia which causes the feet to turn inwards.

In an online survey to characterise the burden of illness in people with XLH (CL001), high levels of pain and limitations in mobility were reported by paediatric respondents with POSNA-PODCI scores for the Sports and Physical Function and Pain and Comfort domains below the normative healthy population mean. In CL001, the mean SF-10 physical health score of 35.5 was 1.5 standard deviations below the general population norm of 50. Similarly, in the phase 2 burosumab study (CL201), in children five to 12 years of age who received conventional therapy for an average of seven years, 55% had substantial functional impairment at baseline, defined as the POSNA-PODCI Global Functioning score <40, with particular functional impairments in the Sports/Physical Functioning and Pain/Comfort domains.[44] In Study CL201, the mean SF-10 physical health score at baseline was ([redacted]), below the population norm of 50. In particular, children with more severe rickets at baseline [redacted].

A further online survey, carried out in January 2018, collected background data regarding the impact of XLH and treatments that the child had received to help manage their condition.[45] [redacted].

ERG comment: It is clear that there is a paucity of data available to inform the question regarding HRQoL in the paediatric population. In the interim analysis of [redacted] in the UK, it is not clear how many children are being analysed. In the absence of comparative data with healthy age-matched children, it is also not clear if the number of missed school days or the number of days when patients could not take part in sports or other events is higher than the population norm or is directly related to their illness. HRQoL, as assessed in studies CL201 and CL205, was not used in the economic model.

2.2.8.2 Impact on family and carer HRQoL

Having a child with medical needs such as XLH requires full attention, with families and carers providing support and reassurance through the child's life progression. Frequent medication, hospital visits and tests can be overwhelming not only for the patient but for their carer as well. Regular blood tests, ultrasound scans to monitor kidneys, X-rays to check the development and condition of bones, frequent dentist visits and even orthopaedic surgery and osteotomies are required from an early age. Only the family and carers of a patient with XLH can assist with these issues. Emergency situations may also occur periodically as bone fractures or increases in pain severity are common between patients with XLH.[43] Parents of children with XLH often suffer from the condition themselves. In a UK survey, [redacted].[45]

2.3 Current service provision

Kyowa Kirin is not aware of any published NICE, NHS England, other national or expert guidelines for the diagnosis, treatment or management of XLH. XLH is listed amongst Rare Metabolic, Sclerosing and Dysplastic Bone Diseases in the National Health Services England (NHSE) document entitled "A13/S/a 2013/14 NHS STANDARD CONTRACT FOR SPECIALISED RHEUMATOLOGY SERVICES (ADULT)." There is no specialised service specification for children.

Guidelines on the diagnosis and management of XLH have been produced by a group of clinical experts in the USA.[3, 46] These guidelines provide specific recommendations for management of XLH in children and adults. This guidance also aligns with the proposals of an expert panel of the Japanese Society for Bone and Mineral Research,[47] as well as a review by UK clinicians that provides guidelines on diagnosis and management of rickets, including a short section on XLH.[48]

The CS states that most children with XLH currently receive conventional therapy, consisting of oral phosphate (divided in aliquots every four to six hours due to rapid excretion by the kidneys) and active vitamin D analogues (usually alfacalcidol in the UK, once daily).[23] Use of an active vitamin D analogue helps prevent secondary hyperparathyroidism that can be induced by phosphate administration. Calcitriol is an alternative; however, it requires multiple dosing and is only available as a capsule, making it less suitable for infants and young children.

The goal of therapy with oral phosphate and active vitamin D analogues in children is to provide just sufficient phosphorous to allow partially improved mineralisation of bone and improve skeletal outcomes, without providing so much that there is ectopic calcification. This approach aims to alleviate bone or joint pain, preventing skeletal deformities caused by rickets and improving growth. For the majority of paediatric patients with XLH (98.6%), treatment with conventional therapy (phosphate and vitamin D metabolites) does not adequately heal rickets, and improvements in serum phosphorous following administration of oral phosphate are transient, with a peak in serum phosphorus after each administration and then a return to baseline levels.[37, 49]

For children, treatment is initiated at the time of diagnosis and continued until long bone growth is complete. Almost all children with XLH require therapy until growth is complete, although the effectiveness on the skeleton is variable, and surgery is often necessary to correct lower extremity deformities. In Study CL001, over 30% of the children surveyed had already undergone at least one surgical procedure[24] and the majority (80%) had reportedly experienced bone or joint pain in the previous year.

Conventional therapy requires individualised dosing adjustment based on tolerability, evidence of secondary complications, changes in body size, growth velocity, and skeletal mineralisation.[3, 23] Frequent monitoring of height, serum calcium, alkaline phosphatase, parathyroid hormone, phosphate serum concentrations, and urinary calcium and creatinine is necessary to prevent tertiary hyperparathyroidism, induced by phosphate overdose and hypercalciuria with nephrocalcinosis and renal insufficiency, resulting from vitamin D metabolite overtreatment.[23]

UK clinicians stated that the following monitoring is required with conventional therapy:

- Monitor serum calcium, phosphorus, potassium and creatinine levels monthly until stable and thereafter every three months
- Monitor ALP, PTH and urine calcium and creatinine levels every three months.
- Perform renal ultrasonograms (to monitor nephrocalcinosis) every one to two years.

Frequent daily dosing and gastrointestinal distress and diarrhoea may compromise treatment persistence/compliance,[46] and as a result the therapeutic benefit of conventional therapy. Suboptimal therapy in childhood can result in lifelong disability. In adults, the reduced bone quality from chronic osteomalacia increases the risk for non-traumatic pseudofractures and causes bone and joint pain,[18] while ongoing skeletal deformities lead to the development of early osteoarthritis and stiffness that cause pain and continue to limit mobility and physical function.

Conventional therapy fails to address the underlying mechanism of the disease, as these supplements do not enhance proximal tubular phosphate reabsorption.

ERG comment: In describing a Japanese national survey conducted in 2010, the CS reports mean serum phosphate levels in a genetic hypophosphataemia group, and states, “Improvements in serum phosphorous following administration of oral phosphate are transient, with a peak in serum phosphorus after each administration and then a return to baseline levels”.[49] However, these values are derived from a mixed patient population that includes not only XLH but also autosomal dominant hypophosphataemic rickets (ADHR), autosomal recessive hypophosphataemic rickets type 1 (ARHR1) and type 2 (ARHR2) patients. These values therefore cannot be considered representative of XLH patients. Thus, the ERG considers that this statement is not accurate, and simply highlights the heterogeneity of the disease.

2.4 Description of the technology under assessment

2.4.1 Burosumab (KRN23, Crysvita™)

Burosumab (tradename: Crysvita™) is a recombinant human immunoglobulin G1 (IgG1) monoclonal antibody manufactured by Kyowa Kirin that binds to and inhibits the activity of fibroblast growth factor 23 (FGF23), which is produced in excess in most XLH patients. The inhibition of FGF23 is reported to improve tubular reabsorption of phosphate from the kidney and increase levels of 1,25 dihydroxy-vitamin D (1,25(OH)2D) in the serum (leading to enhanced intestinal absorption of calcium and phosphate). Normalising phosphate levels is reported to ameliorate the bone-related symptoms (e.g. rickets) associated with XLH.

Since the aetiology and pathophysiological mechanisms behind XLH remain largely unknown, the mechanism-of-action of burosumab must be considered as ameliorating the symptoms rather than treating the underlying cause.

2.5 Current usage in the NHS

Burosumab is not currently in use in the NHS. The MHRA granted burosumab a 'Promising Innovative Medicine' (PIM) designation on 31 January 2017, and the EMA awarded burosumab conditional marketing authorisation on 23 February 2018. Burosumab is expected to be used in line with the anticipated marketing authorisation in children and adolescents with XLH from the age of one year old who have radiographic evidence of bone disease.

Burosumab is a monotherapy, meaning oral phosphate and vitamin D analogue therapy should be discontinued one week prior to initiation of treatment. Concurrent use of oral phosphate and vitamin D analogues is contraindicated with burosumab. Burosumab is administered every two weeks by subcutaneous injection.

Clinical expert opinion has suggested that patients responding well to burosumab treatment are likely to have a diminishing frequency of consultant visits over the longer term. In addition, burosumab will either prevent or improve skeletal abnormalities, and reduce the need for corrective surgery. Routine treatment with burosumab should also remove the need for additional supplementation with growth hormone in a small subset of patients where this is required.

The following ongoing monitoring is recommended with burosumab (Summary of Product Characteristics (Crysvita), 2017):[50]

- Monitoring for signs and symptoms of nephrocalcinosis, e.g. by renal ultrasonography, is recommended at the start of treatment and every six months for the first 12 months of treatment, and annually thereafter.
- Monitoring of plasma alkaline phosphatases, calcium, PTH and creatinine is recommended every six months (every three months for children 1- 2 years) or as indicated. Monitoring of urine calcium and phosphate is suggested every three months. Patient's fasting serum phosphate level should be monitored due to the risk of hyperphosphataemia. To decrease the risk for ectopic mineralisation, it is recommended that fasting serum phosphate is targeted in the lower end of the normal reference range for age. Dose interruption and/or dose reduction may be required.
- Increases in serum parathyroid hormone have been observed in some XLH patients during treatment with burosumab. Periodic measurement of serum parathyroid hormone is advised.

The high burden of frequent monitoring when the drug is first introduced will tail off once the patient is on a stable dose, and the overall burden of monitoring is expected to be reduced compared with that required for conventional therapy.

ERG comment: Kyowa Kirin aim to treat a paediatric and adolescent population of XLH patients from 1-17 years of age who have radiographic evidence of bone disease. After the age of approximately 17, when growth plates fuse, it is indicated that burosumab will be discontinued as it will no longer be required to stabilise rickets symptoms. Based on the therapeutic target of burosumab (FGF23) and the largely unknown pathological mechanisms of XLH, there is no evidence presented that burosumab therapy in childhood has long-term therapeutic consequences in adulthood following treatment cessation. Bone metabolism is an ongoing and dynamic process that will continue to be subject to the pathological consequences of hypophosphataemia. Thus, the ERG considers it unlikely that the diverse pathologic and phenotypic consequences of XLH will be ameliorated without therapeutic intervention beyond the age of ~17 years, particularly with respect to progressive bone weakness. It is likely that it will continue to be necessary to treat and manage XLH patients who have received burosumab during childhood.

The economic model assumes that patients who receive burosumab and transition to the healed rickets state will remain healed. However, there is some suggestion in the literature that long-term treatment of XLH with FGF23 neutralising antibodies (in mouse models) incompletely rescues the mineralisation defect.[51]

As per the company's response, which was informed by UK-based clinical experts, growth hormone is not licensed and is not used in the UK for the treatment of XLH patients. Consequently, the statement, "Routine treatment with burosumab should also remove the need for additional supplementation with growth hormone in a small subset of patients where this is required" (CS, page 54) should be disregarded and not considered in the case for burosumab.

3 CRITIQUE OF THE COMPANY'S INTERPRETATION OF THE DECISION PROBLEM

3.1 Introduction

The remit of this appraisal, as defined in the final NICE scope, is to evaluate the benefits and costs of burosumab within its licensed indication for treating X-linked hypophosphataemia for national commissioning by NHS England.[52] The final NICE scope outlines the agreed population, intervention, comparators and outcomes for the appraisal. The NICE scope also sets out wider considerations relating to the impact of the technology beyond direct health benefits and on the delivery of the specialised service, the nature of the condition, costs to the NHS and PSS and value for money.

On 14 December 2017, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a conditional marketing authorisation for the medicinal product Crysvita (burosumab), intended for the treatment of X-linked hypophosphataemia. A conditional marketing authorisation is granted to a medicinal product that fulfils an unmet medical need when the benefit to public health of immediate availability outweighs the risk inherent in the fact that additional data are still required. The marketing authorisation holder is likely to provide comprehensive clinical data at a later stage. The EMA awarded burosumab conditional marketing authorisation on 23 February 2018.

The full indication is: "Crysvita is indicated for the treatment of X-linked hypophosphataemia with radiographic evidence of bone disease in children 1 year of age and older and adolescents with growing skeletons."[53] It is proposed that Crysvita be prescribed by physicians experienced in the management of patients with metabolic bone diseases.

3.2 Adherence to the decision problem

Table 3.1 presents a summary of the decision problem as set out in the NICE scope[52] and the company's adherence to this (based on information presented on pages 20-21 of the CS[1]).

Table 3.1: Adherence of the CS to the agreed decision problem

	Final scope issued by NICE	Deviations of submission from the scope
Population	Children and young people with X-linked hypophosphataemia	The population is in line with the licence indication: X-linked hypophosphataemia with radiographic evidence of bone disease in children one year of age and older and adolescents with growing skeletons
Intervention	Burosumab	The intervention is in line with scope
Comparator(s)	Established clinical management without burosumab	The comparator is in line with scope
Outcomes	The outcome measures to be considered include: • fractures • severity of rickets • pain (including bone pain, joint pain and joint stiffness) • motor skills • growth (including height) • tooth loss and pain • skull and spinal deformities • neurological complications (including increased intracranial pressure, craniosynostosis, problems with hearing and balance, and spinal cord compression) • radiographic response • renal function • parathyroid hormone levels • alkaline phosphatase levels • mortality • adverse effects of treatment • health-related quality of life (for patients and carers)	The following outcomes could not be accounted for: • fractures • tooth loss and pain • skull and spinal deformities • neurological complications (including increased intracranial pressure, craniosynostosis, problems with hearing and balance, and spinal cord compression) • mortality These outcomes were not captured in clinical studies. Quality of life data collected in the studies (POSNA-PODCI and SF-10) could not be used to derive utility data for the health economic modelling because there is no valuation set according to the company. Therefore, the company derived utility values from a UK study.
Subgroups to be considered	N/A	

	Final scope issued by NICE	Deviations of submission from the scope
Nature of the condition	• disease morbidity and patient clinical disability with current standard of care • impact of the disease on carer's quality of life • extent and nature of current treatment options	
Cost to the NHS and PSS, and Value for Money	• cost effectiveness using incremental cost per quality-adjusted life year • patient access schemes and other commercial agreements • the nature and extent of the resources needed to enable the new technology to be used	
Impact of the technology beyond direct health benefits, and on the delivery of the specialised service	• whether there are significant benefits other than health • whether a substantial proportion of the costs (savings) or benefits are incurred outside of the NHS and personal and social services • the potential for long-term benefits to the NHS of research and innovation • the impact of the technology on the overall delivery of the specialised service • staffing and infrastructure requirements, including training and planning for expertise	
Special considerations, including issues related to equality	Guidance will only be issued in accordance with the marketing authorisation Guidance will take into account any Managed Access Arrangements	

3.3 ERG critique of the company's adherence to the decision problem as set out in the NICE scope

3.3.1 Population

The population included in the submission relates to X-linked hypophosphataemia with radiographic evidence of bone disease in children one year of age and older and adolescents with growing skeletons. This is in line with the licence indication.

The studies included in the submission focus on the following populations and studies:

- Paediatric patients with XLH, five to 12 years old: Study CL201 (open-label RCT comparing different doses of burosumab biweekly or monthly administration of burosumab (doses 0.1 to 2.0 mg/kg))
- Paediatric patients with XLH, one to four years old: Study CL205 (open-label study to assess the safety, pharmacodynamics and efficacy of burosumab biweekly administration of burosumab at a target dose of 0.8 mg/kg))
- Paediatric Patients with XLH, five to 14 years old: Study CL002 (A retrospective longitudinal study of skeletal outcomes in children with XLH. No burosumab administered; however, study inclusion required the use of conventional therapy (oral phosphate/active vitamin D))

In addition, the CS mentions the following studies for which no data have been presented:

- A randomised, open-label, phase 3 study to assess the efficacy and safety of burosumab versus oral phosphate and active vitamin D treatment in paediatric patients with XLH, one to ≤ 12 years old with open growth plates (study CL301). Data are not yet available according to the company; although, the CS states that the primary efficacy and safety analysis for study CL301 is expected to be available [redacted].[1] Completion of this study is also a post-authorisation requirement for the conditional marketing authorisation. We asked the company to provide a precise date when data are available and whether any interim data are available.[2] The company responded that "[redacted]." Although they stress that these timelines remain provisional. The company stated they "[redacted]".
- An open-label, phase 3 study to assess the safety, pharmacodynamics and efficacy of burosumab (no control), in paediatric patients under the age of one year with XLH (study CL207). This study is planned, but no data are available. In addition, it is not relevant to the scope (children under the age of one year are outside the indication).
- XLH Disease Monitoring Program (study CL401), observing disease progression and associated side effects for up to 250 children and adults with XLH. This study is planned, but no data are available.
- A natural history survey via online questionnaire to characterise the burden of illness in adults and children with XLH (No burosumab administered). This study was used in the background section of the CS (Chapter 6 of the CS), but not as part of the clinical evidence (Chapter 9 of the CS[1]).

3.3.2 Interventions

The intervention included within the CS relates to burosumab in line with its licensed indication.

In the CS (page 12 and 31) the recommended dosage regimens of burosumab are described as: The recommended starting dose is 0.4 mg/kg of body weight and the normal maintenance dose is 0.8 mg/kg, given every two weeks. The maximum dose is 90 mg. All doses should be rounded to the nearest 10 mg. Burosumab may be initiated from one year old until end of skeletal growth. Based on UK growth data, in the cost effectiveness model, girls are assumed to remain on treatment up to 16 years of age (inclusive) and boys are assumed to remain on treatment until 17 years of age (inclusive) (CS, chapter 10.1.16, page 148).

3.3.3 Comparators

The comparator is described in the CS as "established clinical management without burosumab", this is in line with the scope.

All patients in the control study (Study CL002: A retrospective longitudinal study of skeletal outcomes in children with XLH aged five to 14 years old) received conventional therapy (i.e. oral phosphate/active vitamin D)).

3.3.4 Outcomes

As specified in the Table with the Statement of the decision problem (CS, Table 1, page 20), the studies do not provide data on the following outcomes:

- fractures
- tooth loss and pain
- skull and spinal deformities
- neurological complications (including increased intracranial pressure, craniosynostosis, problems with hearing and balance, and spinal cord compression)
- mortality

These outcomes were not captured in the clinical studies.

In the CS, the company uses the term 'healing' and 'substantial healing of rickets'. This is defined using RGI-C global scores, where scores ≥ +1.0 indicate 'healing of rickets' and scores ≥ +2.0 'substantial healing of rickets'. The company does explain that "Healing in this context indicates improvement in the radiographic abnormalities and does not imply that complete healing was observed" (CS, page 100). However, throughout the report the term 'healing of rickets' is used without any explanation of the degree of healing (minimal, substantial or complete). Moreover, RGI-C global scores and RSS scores do not capture all clinical aspects of XLH. That is of particular importance in the context of the economic model, which only considers RSS score alone as a clinical outcome measure. The diverse physiological impacts of hypophosphataemia, which may be independent of rickets, are therefore not captured as outcomes in the economic model.

In the response to the clarification letter the company described the vignettes for the various health states that informed the economic model in detail (Clarification Letter Response Question B7, Table 10). However, each health state was defined in such a way that there appears to be a perfect association between the RSS score and other clinical descriptors of the health state. For example, as the RSS score decreases so does the risk of fracture and the presence of deformity. However, this does not appear to be realistic in that it seems likely that there might be some resolution of the bone disorder such that the

RSS score decreases, but that this resolution only occurs after incurring deformity, which cannot be completely resolved and with some continued increased risk of fracture.

In addition, the model currently assumed that the effect of burosumab, although stopped at age 16 (women) or 17 (men) lasts for the rest of their lives. This also seems unrealistic, the effects of burosumab on stature, bowing of the legs, joint deformity etc. are likely to persist fairly long but may wane as osteomalacia itself and the resulting fractures may lead to associated problems in later life. Effects on bone strength will wane quicker, therefore repeated fractures and badly healing fractures after 10 or 20 years are likely to occur. Effects of burosumab on symptoms caused by hypophosphatemia itself will disappear as soon as therapy is stopped. Therefore, we have assumed in the ERG base-case that patients will experience a decline in quality of life 20 years after the end of treatment, which was operationalised by moving to the utility value of the next worse health state (see section 6.2.3 in this report).

In addition, quality of life data collected in the studies (POSNA-PODCI and SF-10) could not be used to derive utility data for the health economic modelling because there is no valuation set according to the company. Therefore, the company derived utility values from a UK study.

3.3.5 Cost to the NHS and PSS, and value for money

The CS includes a cost-consequence model in which the primary health outcome is valued in terms of incremental QALYs gained. In general, the scope was followed when assessing the costs of burosumab to the NHS and the value for money it provides.

4 IMPACT OF THE NEW TECHNOLOGY – CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

4.1.1 Searches

The ERG has presented only the major limitations of the search strategies in the main report. Further minor criticisms can be found in Appendix 1 of this report.

Section 9.1.1 of the CS states that MEDLINE, Embase and the Cochrane Central Register of Controlled Trials were searched for the identification of clinical effectiveness evidence. Search strategies were reported in detail in Appendix 17.1 of the CS and in the response to clarification. MEDLINE and Embase were searched using the Ovid interface from the earliest date available for each database until the end of October 2017. CENTRAL was searched for all available years until January 2018. The searches were also intended to identify studies on adverse events not already known to the company.

A wide range of additional searches were conducted, including the EU Clinical Trials Register, ClinicalTrials.gov, online patient organisations, online case reports and clinical studies. Three main journals in the field were hand-searched, and reference checking was carried out. Experts and clinical specialists were also consulted.

Following a request for clarification, full search strategies were provided for MEDLINE, Embase and CENTRAL. Strategies were not included for the trials register searches.

ERG comment:

- The selection of databases searched was adequate and searches were clearly reported and reproducible. The database name, host, date range and date searched were provided for the majority of the searches. A good range of additional resources were included.
- The main concern of the ERG is that the search terms used for the population facet of the strategy were insufficient. Only one indexing (MeSH/EMTREE) term was used, combined with one free-text term. Numerous synonyms are available for X-linked hypophosphataemia and use of these terms would have increased the retrieval of potentially relevant records.
- Given the small number of papers retrieved for this topic, the ERG believes that use of study design filters in the searches was unnecessarily restrictive. The ERG suggests that a single-facet search for XLH (and additional synonyms) without a study design filter would have adequately addressed all areas of interest, including clinical effectiveness, adverse events, cost-effectiveness, HRQL and resource use without retrieving unmanageably high numbers of records. See Appendix 1 for example MEDLINE, Embase and CENTRAL searches run by the ERG.

4.1.2 Inclusion criteria

The eligibility criteria for the review are described in Table 4.1 (CS, Table 7, page 60).

Table 4.1: Eligibility criteria

Inclusion criteria	
Population	Children or adults with XLH.
Interventions	Any
Outcomes	Reported statistical findings on clinical outcomes (either benefits or adverse effects).
Study design	Studies with a quantitative analytical approach and a study design of case comparison or interventional design (experimental or observational), including: Randomised Control Trials (RCTs), cluster RCTs, non-randomised controlled studies (including controlled before and after studies) and interrupted time series studies (with time points before and after the intervention to establish an underlying trend in the outcome).
Language restrictions	English
Search dates	Database inception to October 31st 2017 (Embase and Medline) and to December 2017 (Cochrane Register of Controlled Trials)
Exclusion criteria	
Population	None
Interventions	None
Outcomes	None
Study design	Animal studies or biochemical or cellular level investigations. Studies with a qualitative design, review articles or articles that investigate the genetic characteristics of XLH.
Language restrictions	Languages other than English.
Search dates	None
Source: CS, Table 7, page 60 XLH = X-linked hypophosphataemia	

ERG comment: The only criticism regarding the inclusion criteria is the language restriction used by only including English language studies.

4.1.3 Critique of data extraction

Methods for the systematic review process have not been reported. Therefore, there is no information regarding the number of reviewers involved in the study selection process and the data extraction process. It is common practice in systematic reviews that every step in the review is performed by at least two reviewers to minimise bias and to prevent mistakes. In this case there is no guarantee that the data extraction process was correct.

The CS does mention that "Data was extracted from included studies using a specially designed data extraction form" (CS, page 59); however, the form used was not presented.

4.1.4 Quality assessment

The risk of bias of included studies was evaluated using an adapted version of the Centre for Reviews and Dissemination (CRD) checklist for CL201,[54] and an adapted version of the Critical Appraisal Skills Programme (CASP) checklist for CL002 and CL205.[55] It was not reported how many reviewers were involved in the risk of bias assessment.

ERG comment: The company used appropriate risk of bias tools for different study types. However, the process of quality assessment was not fully described.

4.1.5 Evidence synthesis

As stated by the company, the "burosumab phase 2 studies were uncontrolled dose finding or single arm studies, therefore an indirect comparison was not feasible" (CS, page 123).[1]

For children between one to four years old, only one study is presented in which all children received burosumab (CL205). A comparison with "established clinical management without burosumab"[52] is not possible in this group of patients.

For children between five to 12 years old, the CS presents a study in which all children received burosumab (CL201). In addition, the CS presents a control study (CL002) in which children aged between five to 14 years received conventional therapy (i.e. oral phosphate/active vitamin D). Results of these two studies are presented as a naïve comparison, simply reporting individual results from each study side by side (See CS, Table 17, page 94). In addition, the company presents comparisons of 'rickets healing' with conventional therapy (Study CL002) versus burosumab (Study CL201) using propensity score matching. Further details of the methods and results of the naïve and propensity score matched comparisons are provided in section 4.3. As there were no controlled studies of burosumab meta-analysis was not performed.

ERG comment: Full details of the numbers of reviewers involved in the study selection, data extraction and quality assessment stages of the systematic review were not reported. Due to a lack of comparative studies meta-analyses were not possible. The lack of detail about the review methods means it is not possible to judge if appropriate steps were used to reduce the risk of reviewer error and bias. Restricting the review to studies only published in English means that some studies may have been missed, although this is unlikely due to the small amount of evidence available for burosumab.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

4.2.1 Studies included in/excluded from the submission

The CS includes two studies of burosumab in children aged 5-12 years (Study CL201) and in children aged 1-4 years (Study CL205). Study CL201 is an ongoing, multicentre, dose-finding Phase 2 study which included 52 children (10 from three clinical trial sites in the UK) with XLH aged 5-12 years and compared two dosing frequencies of burosumab: once every two weeks (n=26) or once every four weeks (n=26). Study CL205 is an ongoing, multicentre, single-arm, Phase 2 study in 13 children from one to four years old with XLH who are naive to therapy or have previously received conventional therapy with oral phosphate and active vitamin D to assess the safety, pharmacodynamics, pharmacokinetics, and efficacy of burosumab administered via subcutaneous (SC) injection once every two weeks (Q2W) for a total of 64 weeks.

In addition, the CS includes a historical control study. Study CL002 is a retrospective radiographic and medical chart review study designed to evaluate the long-term safety and efficacy of oral phosphate/active vitamin D therapy. The children in CL002, aged five to 14 years old, had received long-term (approximately eight years) conventional therapy with oral phosphate and active vitamin D (n=[redacted] in the Radiographic Analysis Set). All [redacted] patients who contributed the radiographs for RSS and RGI-C analyses were enrolled at a single US site, Shriners Hospital in St. Louis, Missouri. The study is ongoing and additional data from three other sites in the United States, France, and Canada are anticipated to add to the body of evidence. Historical images will be collected from up to 100 children.

A total of [redacted] children had been enrolled in the CL002 study at the time of the latest data cut (August 2016). One child had not received conventional therapy and was not included in the analysis. The remaining [redacted] children (98%) who met the study inclusion/exclusion criteria and had been treated with conventional therapy were included in the Full Analysis Set. The mean duration between baseline and post-baseline radiographs was [redacted]).

Since CL002 was a US study, Kyowa Kirin also commissioned a longitudinal review of patient records from three expert UK centres to provide additional data (n=43). However, results from this UK review are not included in the CS. We asked the company in the clarification letter, and the company responded that this case review was commissioned specifically for NICE, and that the data were only made available just prior to submission. For this reason, no CSR was constructed as the data has not been submitted to regulatory agencies. Instead the company provided a synopsis with details on the rationale, methodology and results as part of the response to the clarification letter.[2] A summary and critique of these data are provided in section 4.5 of this report (Additional work on clinical effectiveness undertaken by the ERG).

Table 4.2: Included studies

Study ID	Study Title	Patient Population (Type/ Number of patients)	Intervention
UX023-CL201 Clinical Study report – week 64 Analysis, May 2017 (ongoing)	Randomised, Open-Label, Dose Finding, Phase 2 Study to Assess the Pharmacodynamics and Safety of the anti-FGF23 antibody, burosumab, in Paediatric Patients with XLH	Paediatric patients with XLH, 5 to12 years old 52 initiated treatment	Multi-dose burosumab Biweekly or monthly administration of burosumab (doses 0.1 to 2.0 mg/kg) Repeat dose, up to 64 weeks
UX023-CL205 Clinical Study report – week 40 (Primary) Analysis, Oct 2017 (ongoing)	An Open-Label, Phase 2 Study to Assess the Safety, Pharmacodynamics and Efficacy of burosumab in Children from 1 to 4 Years Old with XLH	Paediatric patients with XLH, 1 to 4 years old 13 patients enrolled	Multi-dose burosumab Biweekly administration of burosumab at a target dose of 0.8 mg/kg. Repeat dose, up to 64 weeks
UX023-CL002 Clinical Study report, Nov 2016	A retrospective longitudinal study of skeletal outcomes in children with XLH	Paediatric Patients with XLH, 5 – 14 years old. Images will be collected from up to 100 children	This was not an interventional study; however, study inclusion required the use of conventional therapy (oral phosphate/ active vitamin D)
Source: CS, Tables 8 and 9, pages 63-64 XLH = X-linked hypophosphataemia			

The methodology of the three included studies is described in Tables 4.3 and 4.4, and demographic and baseline characteristics are described in Table 4.5.

ERG comment: As can be seen from Table 4.3, inclusion criteria for patients in studies CL201 and CL002 are similar in that patients in both studies were diagnosed with XLH and were of similar age. However, there are important differences between the inclusion criteria in both studies. Children in study CL201also had: biochemical findings associated with XLH, standing height < 50th percentile for

age and gender and radiographic evidence of active bone disease including rickets in the wrists and/or knees, and/or femoral/tibial bowing, or, for expansion patients, an RSS score in the knee of at least 1.5 points as determined by central read. In other words, study CL002 included all children with XLH, while study CL201 included children with more severe symptoms of XLH. This is also reflected in the relatively [redacted] standing height and [redacted] rickets severity score for children in study CL201 when compared to children in study CL002 (see Table 4.5).

Study CL205 enrolled children with XLH aged between one and four years old. In this study children had to have clinical findings consistent with XLH, including hypophosphataemia and radiographic evidence of rickets (at least five patients were required to have a Rickets Severity Score [RSS] at the knee of ≥ 1.5 points at Screening), and a confirmed PHEX mutation or variant of uncertain significance (VUS). Only 13 children were enrolled. Therefore, results in this age group are very uncertain (see Table 4.4).

Table 4.3: Summary of methodology for Studies CL201 and CL002

Study name	**UX023-CL201**	**UX023- CL002**
Objectives	• Identify a dose and dosing regimen of burosumab, based on safety and PD effect in paediatric XLH patients • Establish the safety profile of burosumab for the treatment of children with XLH including ectopic mineralisation risk, cardiovascular effects, and immunogenicity profile	To characterise change in rickets severity over time with conventional therapy (oral phosphate/active vitamin D) in children with XLH ages 5 to 14 years.
Location	This study is being conducted at a total of nine centres: four in the United States, three in the United Kingdom, one in France, and one in the Netherlands	Two sites in the USA.
Design	Randomised, multicentre, open-label, dose-finding Phase 2 study assesses the PD, efficacy, and safety of burosumab in prepubescent children (5 to 12 years old) with XLH. The study consists of two Screening Visits, a 16-week Titration Period, a 48-week Treatment Period, and a 96-week Treatment Extension Period.	Retrospective radiographic and medical chart review of patients with XLH who had longitudinal historical radiographs of the wrist, knee, or long leg taken between the ages of 5 and 14 years (inclusive).
Duration of study	The planned study duration is 160 weeks (approximately 3 years): 16 weeks in the Titration Period, 48 weeks in the Treatment Period, and 96 weeks in the Treatment Extension Period.	This is a retrospective study. The mean duration between baseline and post-baseline radiographs was [redacted] weeks]).
Sample size and Patient population	Approximately 30 paediatric patients with XLH and radiographic evidence of bone disease (“pre-expansion patients”) were planned for enrolment under the original study protocol. The study was expanded per amendment 3 of the protocol to include additional patients (“expansion patients”) who were required to have rickets severity of at least 1.5 at the knee (per the Rickets Severity Score [RSS] method), for a total of approximately 50 patients planned overall.	[redacted] paired wrist and knee images) Children with a confirmed diagnosis of XLH who have radiographic images for at least two time points taken between the ages of 5 and 14 years.

Study name	**UX023-CL201**	**UX023- CL002**

Inclusion criteria	• Male or female, aged 5 – 12 years, inclusive, with open growth plates
	• Tanner stage of 2 or less based on breast and testicular development
	• Diagnosis of XLH supported by ONE of the following:
	  o Confirmed PHEX mutation in the patient or a directly related family member with appropriate X-linked inheritance
	  o Serum FGF23 level > 30 pg/mL by Kainos assay
	• Biochemical findings (based on overnight fasting [minimum 4 hours] values collected at Screening Visit 2) associated with XLH including:
	  o Serum phosphorus ≤ 2.8 mg/dL (0.904 mmol/L)
	  o Serum creatinine within age-adjusted normal range
	• Standing height < 50th percentile for age and gender using local normative data. (Criterion was changed to "< 50th percentile" [from "< 25th percentile"] per Protocol Amendment 1)
	• Radiographic evidence of active bone disease including rickets in the wrists and/or knees, AND/OR femoral/tibial bowing, OR, for expansion patients, an RSS score in the knee of at least 1.5 points as determined by central read (The inclusion criterion of RSS ≥ 1.5 for patients enrolled with the expansion of the study was added per Protocol Amendment 3)
	• Willing to provide access to prior medical records for the collection of historical growth, biochemical and radiographic data, and disease history
	• Provide written or verbal assent (if possible) and written informed consent by a legally authorised representative after the nature of the study has been explained, and prior to any research-related procedures
	• Must, in the opinion of the investigator, be willing and able to complete all aspects of the study, adhere to the study visit schedule and comply with the assessments
	• Females who have reached menarche must have a negative pregnancy test at Screening and undergo additional pregnancy testing during the study. If sexually active, male and female patients must be willing to use an acceptable method of contraception for the duration of the study. (This inclusion criterion added per Protocol Amendment 1)
		• Male or female, with radiographic images from at least two time points taken between the ages of 5 and 14 years, inclusive
		• Diagnosis of XLH based on a confirmed PHEX mutation in the patient or a directly related family member with appropriate X-linked inheritance, or a clinical diagnosis of XLH based on biochemical profile and clinical symptoms

Exclusion criteria	• Use of a pharmacologic vitamin D metabolite or analog (eg, calcitriol, doxercalciferol, alfacalcidol, and paricalcitol) within 14 days prior to Screening Visit 2; washout took place during the Screening Period • Use of oral phosphate within 7 days prior to Screening Visit 2; washout took place during the Screening Period • Use of calcimimetics, aluminium hydroxide antacids, systemic corticosteroids, and thiazides within 7 days prior to Screening Visit 1 • Use of growth hormone therapy within 3 months before Screening Visit 1. (Criterion was changed to "within 3 months" [from "within 12 months"] per Protocol Amendment 2 • Use of bisphosphonates for 6 months or more in the 2 years prior to Screening Visit 1 • Presence of nephrocalcinosis on renal ultrasound graded ≥ 3 based on the following scale: o 0 = Normal o 1 = Faint hyperechogenic rim around the medullary pyramids o 2 = More intense echogenic rim with echoes faintly filling the entire pyramid o 3 = Uniformly intense echoes throughout the pyramid o 4 = Stone formation: solitary focus of echoes at the tip of the pyramid • Planned or recommended orthopaedic surgery, including staples, 8-plates or osteotomy, within the clinical trial period • Hypocalcaemia or hypercalcemia, defined as serum calcium levels outside the age-adjusted normal limits (based on overnight fasting [minimum 4 hours] values collected at Screening Visit 2) • Evidence of tertiary hyperparathyroidism as determined by the Investigator • Use of medication to suppress parathyroid hormone (PTH) within 2 months prior to Screening Visit 1 • Presence or history of any condition that, in the view of the investigator, places the patient at high risk of poor treatment compliance or of not completing the study • Presence of a concurrent disease or condition that would interfere with study participation or affect safety • Previously diagnosed with human immunodeficiency virus antibody, hepatitis B surface antigen, and/or hepatitis C antibody • History of recurrent infection or predisposition to infection, or of known immunodeficiency	• Currently or previously treated with burosumab in Ultragenyx protocol UX023-CL201 (images and data from patients in the current study were collected as a part of UX023-CL201)

Study name	**UX023-CL201**	**UX023- CL002**
	• Use of a therapeutic monoclonal antibody within 90 days prior to Screening Visit 1 or history of allergic or anaphylactic reactions to any monoclonal antibody • Presence or history of any hypersensitivity to burosumab excipients that, in the judgment of the investigator, places the patient at increased risk for adverse effects • Use of any investigational product or investigational medical device within 30 days prior to screening, or requirement for any investigational agent prior to completion of all scheduled study assessments	
Interventio n(s) (n =) and comparator (s) (n =)	Burosumab, n=52: Pre-expansion Patients • Dose Cohort 1, [redacted] (0.1 mg/kg Q2W [redacted] 0.2 mg/kg Q4W [n[redacted]) • Dose Cohort 2, n[redacted] (0.2 mg/kg Q2W [redacted] 0.4 mg/kg Q4W [n[redacted]) • Dose Cohort 3, n[redacted](0.3 mg/kg Q2W [redacted] 0.6 mg/kg Q4W [n[redacted] Expansion Patients • Dose Cohort 3, [redacted] (0.3 mg/kg Q2W [redacted] 0.6 mg/kg Q4W [redacted]	Not applicable (patients had been on conventional therapy for approximately 6 years prior to study enrolment).
Baseline differences	Demographic characteristics were similar for patients randomised to the Q2W and to the Q4W dose regimens.	Not applicable
Duration of follow-up, lost to follow-up information	All patients completed at least 64 weeks on study. No patient discontinued from the study, and all patients are continuing in the study as of the data cut-off date.	Patients were not followed up as this was a retrospective study. The mean duration between baseline and post-baseline radiographs was [redacted])
Statistical tests	No formal hypothesis was tested to compare treatment groups (Q2W and Q4W) in this study. Changes from baseline in efficacy parameters were tested. Statistical analyses were reported using summary tables, figures, and data listings. Statistical tests were 2-sided at the alpha=0.05 significance level, and 2-sided 95% confidence intervals (CIs) were used. All p-values were presented as nominal p-values. No adjustment on multiplicity was made. For the primary efficacy endpoint of change in RSS total score, the difference between the two dose regimens (Q2W and Q4W) was summarised with 95% CIs. For repeated measures, the generalised estimating equation (GEE) approach was used for assessing the change over time. The GEE model included regimen, study visit and interaction between regimen and study visit as categorical variables. Model-based estimates of changes from baseline and corresponding 95% CIs were provided along with	Retrospective radiographic, biochemical, growth, and conventional therapy data collected from all patients in this historical cohort were summarised by both event incidence and patient incidence. No formal hypothesis was tested in this study. The primary evaluation in the current study was the change in rickets severity, as evaluated by 2 different methods (RSS and RGI-C). Rickets was assessed based on radiographic changes from radiograph pairs that were 1 to 2 years apart, with the earlier pair considered the baseline radiograph. For each radiograph pair, growth and

Study name	**UX023-CL201**	**UX023- CL002**
	P-values for assessing statistical significance. As exploratory analyses, covariates such as baseline measures, gender, and age were considered for adjustment within GEE models. Continuous variables were summarised with means, standard deviations (SD), standard errors (SE), medians, interquartile ranges (Q1, Q3), minimums, and maximums. Categorical variables were summarised by counts and by percentages of patients in corresponding categories. No imputation on missing data was made, unless stated otherwise. All data obtained from the Case Report Forms (CRFs) as well as any derived data were included in data listings. Efficacy results were analysed by subgroups defined by RSS total score at baseline. The "higher RSS" subgroup consisted of patients with RSS total scores at baseline ≥ 1.5; the "lower RSS" subgroup consisted of patients with RSS total scores at baseline < 1.5. The value of 1.5 was based on the median RSS total score of the study population at the interim analysis of the first 12 patients. Results also were analysed by subgroups defined by degree of functional impairment: for 6MWT results by percentage of predicted 6MWT (abnormal: < 80%, or normal range: ≥ 80%) at baseline, and for the POSNA-PODCI questionnaire by Global Functioning scale score (abnormal: < 40, or normal range: ≥ 40) at baseline.	biochemical data were linked to baseline and post-baseline radiographs by time of measurement and changes in growth and biochemical parameters were summarised. RSS, growth, and biochemical data were also summarised by event incidence in addition to paired incidence; the details of assessment plan for each endpoint are provided in. Subgroups were also prespecified based on rickets severity of the baseline radiographs: baseline radiographs with RSS total score ≥1.5 were referred to as the Higher RSS subgroup and those with RSS total score <1.5 were referred to as the Lower RSS subgroup. For continuous variables, the mean, standard deviation, median, quartiles, minimum, and maximum are provided; 95% confidence intervals (95% CI) on change from baseline were calculated for paired radiographs by one sample T test. For discrete data, frequency and percent distributions are used. Analysis was performed on the analysis sets by patient incidence, by radiograph incidence, or by paired radiographs.
Primary outcomes	Primary efficacy endpoint: Change from baseline in severity of rickets as measured by Rickets Severity Score (RSS) total score The primary efficacy analysis was at week 40. Additional efficacy analysis was carried out at week 64.	Conventional therapy endpoints include the following information: • Age at the time of initiating conventional therapy • Total duration of conventional therapy • Conventional therapy treatment status at time of radiographic imaging (Yes/No) • Conventional therapy regimen at time of radiographic image taken, including medication
Secondary outcomes (including scoring methods and timings	Secondary efficacy endpoints	

Study name	**UX023-CL201**	**UX023- CL002**
of assessments)	• Change from baseline in severity of rickets as measured by RSS knee and wrist scores • Change from baseline in the radiographic appearance of rickets and bowing as measured by Radiographic Global Impression of Change (RGI-C) global, knee, wrist and long leg scores • Growth (standing height, sitting height, arm length, and leg length) • Walking Ability (Six-minute Walk Test [6MWT]) • Functional disability and pain (Pediatric Orthopedic Society of North America – Pediatric Outcomes Data Collection Instrument [POSNA-PODCI])	• names, dose and frequency of administration for both phosphate and active vitamin D • Interruptions in conventional therapy of 3 months or more and reason for interruption Radiographic measures of rickets severity were assessed by Rickets Severity Scale (RSS) and Radiographic Global Impression of Change (RGI-C. Growth endpoints include standing height (length) in cm, z-score and percentile (adjusted by gender and age). Biochemical endpoints include change over time in serum or plasma phosphorus, calcium, iPTH, 1,25(OH)2D, and ALP corresponding to dates close to the date radiographic imaging was collected, where available.
Source: CS, Tables 10 and 12, pages 66-70 and 75-77		

Table 4.4: Summary of methodology for Study CL205

Study name	**UX023- CL205**
Objectives	Primary objectives: • Establish the safety profile of burosumab for the treatment of XLH in children between 1 and 4 years old • Determine the pharmacodynamic (PD) effects of burosumab treatment on serum phosphorus and other PD markers that reflect the status of phosphate homeostasis in children between 1 and 4 years old with XLH Additional study objectives are to assess the following in children between 1 and 4 years old with XLH: • Effects of burosumab on rickets • Effects of burosumab on growth and lower extremity deformity • Pre-dose burosumab drug concentration levels
Location	This study is being conducted at 3 centres in the USA.

Study name	**UX023- CL205**
Design	Multi-centre, open-label, single-arm, Phase 2 study in children from 1 to 4 years old with XLH who are naive to therapy or have previously received conventional therapy with oral phosphate and active vitamin D to assess the safety, PD, PK, and efficacy of burosumab administered via subcutaneous (SC) injection Q2W for a total of 64 weeks.
Duration of study	The planned duration of treatment in this study is 64 weeks. Patients who complete the study may continue into an extension study.
Sample size and Patient population	Approximately 10 paediatric patients were planned for enrolment and 13 patients were enrolled. This submission summarises the planned, primary analyses of data to week 40 for all 13 patients and additional safety data available through the data cut-off date. Patients were between 1 and 4 years old, inclusive, with clinical findings consistent with XLH, including hypophosphataemia and radiographic evidence of rickets (at least 5 patients were required to have a Rickets Severity Score [RSS] at the knee of ≥ 1.5 points at Screening), and a confirmed PHEX mutation or variant of uncertain significance (VUS).
Inclusion criteria	• Male or female, aged ≥ 1 year and < 5 years • PHEX mutation or VUS in either the patient or a directly related family member with appropriate X-linked inheritance • Biochemical findings associated with XLH including serum phosphorus < 3.0 mg/dL (0.97 mmol/L) and serum creatinine within age-adjusted normal range. (Criteria to be determined based on fasting [minimum 4 hours] values collected at baseline.) • Radiographic evidence of rickets; at least 5 patients will be required to have a RSS at the knee of at least 1.5 points as determined by central read • Willing to provide access to prior medical records for the collection of historical growth, biochemical, and radiographic data and disease history • Provide written informed consent by a legally authorised representative after the nature of the study has been explained, and prior to any research-related procedures • Must, in the opinion of the Investigator, be willing and able to complete all aspects of the study, adhere to the study visit schedule, and comply with the assessments

Study name	**UX023- CL205**
Exclusion criteria	• Unwilling to stop treatment with oral phosphate and/or pharmacologic vitamin D metabolite or analog (eg, calcitriol, alfacalcidol) during the screening period and for the duration of the study • Presence of nephrocalcinosis on renal ultrasound grade 4 based on the following scale: o 0 = Normal o 1 = Faint hyperechogenic rim around the medullary pyramids o 2 = More intense echogenic rim with echoes faintly filling the entire pyramid o 3 = Uniformly intense echoes throughout the pyramid o 4 = Stone formation: solitary focus of echoes at the tip of the pyramid • Planned or recommended orthopaedic surgery, including staples, 8-plates or osteotomy, within the clinical trial period • Hypocalcaemia or hypercalcaemia, defined as serum calcium levels outside the age-adjusted normal limits. (Criteria to be determined based on fasting [minimum 4 hours] values collected at baseline.) • Presence or history of any condition that, in the view of the Investigator, places the patient at high risk of poor treatment compliance or of not completing the study • Presence of a concurrent disease or condition that would interfere with study participation or affect safety • History of recurrent infection or predisposition to infection, or of known immunodeficiency • Use of any investigational product or investigational medical device within 30 days prior to screening, or requirement for any investigational agent prior to completion of all scheduled study assessments
Intervention	Burosumab, n=13
Baseline differences	Not applicable
Duration of follow-up, lost to follow-up information	All 13 patients were included in each analysis set (Efficacy Analysis Set, PK/PD Analysis Set, and Safety Analysis Set). As of the data cut-off date (20 April 2017), all patients completed week 40, no patient had discontinued from treatment or from the study, and all patients continue in the study. Additionally, 9, 7, and 4 patients have received burosumab through weeks 42, 44, and 46, respectively, as of the data cut-off date.

Study name	**UX023- CL205**
Statistical tests	The planned sample size for this study of approximately 10 patients was considered appropriate to evaluate the burosumab dose and PK/PD relationship in children aged 1 to 4 years to confirm if that relationship is similar to that observed in older children (aged 5–12 years; N=52) in Study UX023-CL201. Analyses groups included: the Safety Analysis Set (all patients who received at least one dose of study drug), the Efficacy Analysis Set (all patients who received at least one dose of study drug and have at least one post-study drug measurement), and the PK/PD Analysis Set (all patients who received at least one dose of study drug and have evaluable blood samples). Continuous variables were summarised with means, standard deviations (SDs), standard errors (SEs), medians, interquartile range, minimums, and maximums. Categorical variables were summarised by counts and by percentages of patients in corresponding categories. No imputation on missing data was made, unless stated otherwise. All data obtained from the case report forms (CRFs) as well as any derived data were included in data listings. Changes from baseline to post-baseline time points in PD and efficacy parameters were tested for statistical significance. Statistical tests were 2-sided at the alpha = 0.05 significance level and 2-sided 95% confidence intervals (CIs) were used. All p-values were presented as nominal p-values. No adjustment for multiplicity was made. An analysis of covariance (ANCOVA) model was applied to each RGI-C score (wrist, knee, global and lower limb deformity) and change from baseline in each RSS score (wrist, knee and total). The ANCOVA model for RSS scores included the change from baseline in RSS score as the dependent variable and age and RSS score at baseline as covariates. The ANCOVA model for RGI-C scores included the RGI-C score as the dependent variable and age and RSS at baseline as covariates. By-visit analyses using the Generalised Estimating Equations (GEE) model was applied for all PD parameters; the GEE model included change from baseline as the dependent variable, time as the categorical variable and adjusted for baseline measurement, with exchangeable covariance structure. By-visit analyses using the GEE model also was applied to recumbent length/standing height; the GEE model included the change from baseline as the dependent variable, visit and gender as factor, age and recumbent length/standing height z-score at baseline as covariates, with exchangeable covariance structure.
Primary outcomes	The primary efficacy endpoint is the change from baseline in serum phosphorus.

Study name	**UX023- CL205**
Secondary outcomes (including scoring methods and timings of assessments)	• Change in rickets as assessed by the Radiographic Global Impression of Change (RGI-C) global score at weeks 40 and 64 • Change from baseline in RSS total score at weeks 40 and 64 • Change in lower extremity skeletal abnormalities, including genu varum and genu valgus, as determined by the RGI-C long leg score at weeks 40 and 64 • Change in recumbent length/standing height from baseline to post-treatment study time points in cm, height-for-age z-scores, and percentiles based on age and gender. • Historical growth records may be used to evaluate change in growth velocity • Change and percentage change from baseline over time in serum alkaline phosphatase (ALP)
Source: CS, Table 11, pages 71-73	

Table 4.5: Demographic and baseline characteristics in studies CL201, CL002 and CL205

	CL201	Study CL002	CL205
	Q2W (n=26)	Radiographic analysis set ([redacted])	(n=13)
Age (years), mean (SD)	8.7 (1.72)	[redacted]	2.9 (1.15)
Sex, male n (%)	12 (46.2%)	[redacted]	9 (69.2%)
Race White Black/ African-American Other	 23 (88.5%) 2 (7.7%) 1 (3.8%)	[redacted]	 12 (92.3%) 1 (7.7%) 0
Weight (kg), mean (SD)	31.87 (7.92)	[redacted]	12.92 (1.81)
Height (percentile for age and gender), mean (SD)	[redacted]	[redacted]	[redacted]
Standing Height (z-score), mean (SD)	-1.72, 1.03	[redacted]	-1.38 (1.19)
Renal ultrasound score, (0 – 5 scale) – n (%) 0 1 2	[redacted]	[redacted]	NR
Number (%) of Patients Who Received Prior Conventional Therapy	24 (92.3%)	[redacted]	13 (100%)
Duration of Prior Conventional Therapy, mean (SD)	7.02 (2.14) years	[redacted]	16.7 (14.39) months
Age When Conventional Therapy Was Initiated (years), mean (SD)	[redacted]	[redacted]	[redacted]
Pharmacodynamic parameters, mean (SD) Serum Phosphorus, mg/dL TmP/GFR (mg/dL) Serum 1,25(OH)2 D (pg/mL) ALP (U/L	 [redacted]	 [redacted]	 [redacted]
Rickets Severity RSS Total Score, mean (SD)	1.92 (1.17)	[redacted]	2.92 (1.37)

Source: CS, Table 13, page 82.
a) At baseline paired radiograph (the earlier radiograph pair)

4.2.2 Details of relevant studies not included in the submission

CL301 is a multi-centre, randomised, open-label, Phase 3 study comparing the efficacy and safety of burosumab with active control (oral phosphate/active vitamin D therapy) in children with XLH (aged one to ≤12 years) who have radiographic evidence of rickets, open epiphyses, and have received oral phosphate/active vitamin D therapy for ≥ 6-12 consecutive months prior to screening. Approximately 60 patients will be randomised 1:1 to receive open-label burosumab administered by subcutaneous injection or oral phosphate and active vitamin D therapy for a total of 64 weeks.

The CS does not present any results for this study. Instead the CS mentions that: "The primary efficacy and safety analysis from study UX023-CL301 is expected to be available [redacted]".[1] According to clinicaltrials.gov,[56] the estimated primary completion date is July 2018. We asked the company whether or when any (interim) results are available for the committee to look at, and the company responded that
"[redacted]".[2]

ERG comment: Results from this study will considerably reduce the uncertainty surrounding the clinical effectiveness of burosumab relative to conventional therapy in children with XLH aged between one and 12 years.

4.2.3 Summary and critique of company's analysis of validity assessment

The formal appraisal of the validity of the included studies is reported in section 9.5 of the CS (CS, Tables C7 and C8, pages 87-93).

ERG comment: The main problem with the risk of bias of included studies is that none of these studies were designed for comparison of different interventions. CL201 was a randomised controlled trial comparing two burosumab dosing regimens (Q2W versus Q4W); however, only the Q2W arm was used to compare burosumab with conventional therapy. Therefore, all comparative evidence used in the submission was derived from single arm studies. This means the risk of bias of all included studies is high.

4.2.4 Summary and critique of results

4.2.4.1 Efficacy

The CS includes two studies of burosumab in children aged 5-12 years (Study CL201) and in children aged 1-4 years (Study CL205). and one historical control study (Study CL002) in children aged five to 14 years old.

STUDY CL201 - burosumab in children aged 5-12 years

An overview of the results for CL201 are shown in Table 4.6, alongside results from the historical reference study CL002. CL201 investigated dosing every two weeks (Q2W) and every four weeks (Q4W). The Q2W regimen is the expected licensed dosing frequency and are the only results presented here. Assessments of rickets, growth, and walking ability consistently showed greater improvement with the Q2W regimen as compared with the Q4W regimen.

Table 4.6: Outcomes from CL201 and CL002

	Q2W burosumab				Conventional therapy
	Week 40 (n=26)		Week 64 (n=26)		n=[redacted]
Endpoint	**Effect Size**	**p-value**	**Effect Size**	**p-value**	**Effect Size**
RSS Total Score % mean change from baseline[a] (negative is better)	-61%	< 0.0001	-58%	< 0.0001	[redacted]
RGI-C Global Score Mean (positive is better)	+1.72	< 0.0001	+1.62	< 0.0001	[redacted]
Substantial Healing by RGI-C % with RGI-C global score ≥+2.0	[redacted]	NA	[redacted]	NA	[redacted]
Growth Velocity Mean change, comparing pre- and post-treatment[c] (cm/year)	-	-	[redacted]	[redacted]	NR
Standing Height Z-score LS mean change from baseline[b]	-	-	[redacted]	[redacted]	[redacted]
6MWT Distance LS mean change from baseline[b] (m)	[redacted]	[redacted]	[redacted]	[redacted]	NR
Sports/Physical Functioning Scale (POSNA-PODCI) LS mean change from baseline[b] (10 = 1 SD)	[redacted]	[redacted]	[redacted]	[redacted]	NR
Pain/Comfort Scale (POSNA-PODCI) LS mean change from baseline[b] (10 = 1 SD)	[redacted]	[redacted]	[redacted]	[redacted]	NR

Source: CS, Table 17, page 94

NA = Not applicable; NR = not reported; 6MWT = 6-minute walk test; GEE = generalised estimation equation; LS = least squares; POSNA-PODCI = Pediatric Orthopedic Society of North America Pediatric Outcomes Data Collection Instrument; Q2W = every 2 weeks; RGI-C = Radiographic Global Impression of Change; RSS = Rickets Severity Score

a) Percent change based on arithmetic means; p value based on GEE model.

b) LS mean and p value based on GEE model.

c) P-value based on one-sample t test on growth velocity change from baseline.

As can be seen from Table 4.6, for all outcomes that can be compared across studies, results are better for burosumab when compared to conventional treatment.

ERG comment: A naïve comparison of results from studies CL201 and CL002 is unreliable because of the differences in inclusion criteria and patient characteristics in both studies. As explained in section 4.2.1 of this report, there are important differences between the inclusion criteria in both studies. Study CL201 included children with more severe symptoms of XLH. This is also reflected in the relatively [redacted] standing height and [redacted] rickets severity score for children in study CL201 when compared to children in study CL002 (see Table 4.5).

RSS Total Score Change from baseline (Primary Efficacy Endpoint)

Table 4.7 shows the main outcomes from study CL201 for burosumab treatment at 40 weeks and 64 weeks follow up. In the Q2W group (N = 26), RSS total scores were reduced by 61% at week 40 (LS mean (SE) change: [redacted]), p < 0.0001) and by 58% at week 64 [redacted]), p < 0.0001).

In the primary analysis of the primary efficacy endpoint (overall population, N=52), RSS total score at week 40 was reduced by 50%, a statistically significant (p < 0.0001) least squares (LS) mean (SE) change of [redacted]. RSS total score at week 64 was reduced by 51%, a statistically significant (p < 0.0001) LS mean (SE) change of [redacted]. Mean (SD) RSS total scores were [redacted] at baseline, [redacted] at week 40, and [redacted] at week 64.

RSS wrist and knee scores (secondary endpoints)

[redacted] (see Table 4.7).

RGI-C Scores

Treatment for 40 weeks and 64 weeks with burosumab, resulted in healing of rickets as assessed by RGI-C scores. Mean global, wrist, and knee RGI-C scores at weeks 40 and 64 were > +1.4 in the overall group and in both treatment regimens (p < 0.0001 [GEE model]) (see Table 4.7).

Subgroup results by severity

Overall, burosumab showed better results for children with more severe baseline rickets scores. In the Q2W-treated higher RSS subgroup (baseline RSS total score ≥ 1.5; N = 17), RSS total score was reduced by 71% at week 40 (LS mean [SE] change: [redacted]], p < 0.0001) and by 62% at week 64 [redacted], p < 0.0001). In the lower RSS subgroup (baseline RSS total score < 1.5; N = 18), treatment with burosumab for 40 and 64 weeks [redacted].

In the Q2W dosing group, mean RGI-C Global Score was +2.08 (p <0.0001) in the higher RSS group and [redacted] in the lower RSS group at week 64.

Other outcomes

Walking ability, as assessed by LS mean distance walked in the six-minute walk test (6MWT), increased from baseline by [redacted] at week 64 (p [redacted]). In a subgroup with impaired walking ability (<80% of predicted normal; N = 14), the CS reported a "functionally meaningful increase in 6MWT distance of [redacted]at week 64[redacted]) to achieve normal mean values (≥ 80% of predicted normal)." Functional disability was assessed using the Pediatric Orthopedic Society of North America - Pediatric Outcomes Data Collection Instrument (POSNA-PODCI). Biweekly burosumab treatment increased scores for Sports/Physical Functioning and Pain/Comfort into the normal range seen in healthy children; LS mean scores showed improvements of [redacted] and [redacted] at week 64, respectively (see Table 4.7).

Table 4.7: Main outcomes from CL201 at weeks 40 and 64 (Q2W, ITT Analysis Set)

	Burosumab Q2W (n = 26)				
	Baseline, mean (SD)	**Week 40, mean (SD/SE*)**	**Mean change (SE), p-value[a]**	**Week 64, mean (SD/SE*)**	**Mean change (SE), p-value[a]**
RSS Wrist Score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
RSS Knee Score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
RSS Total Score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
RGI-C Wrist Score[a]	NR	[redacted]	NR	[redacted]	NR
RGI-C Knee Score[a]	NR	[redacted]	NR	[redacted]	NR
RGI-C Total Score[b]	NR	[redacted]	NR	[redacted]	NR
6MWT Distance (distance walked [m])	[redacted]	NR	NR	[redacted]	[redacted]
POSNA-PODCI-Sports/Physical Functioning Scale (Normative Score)	[redacted]	NR	NR	[redacted]	[redacted]
POSNA-PODCI-Pain/Comfort Scale (Normative Score)	[redacted]	NR	NR	[redacted]	[redacted]

Source: CS, Tables 18-21, pages 97-105

6MWT = 6-minute walk test; Q2W = every 2 weeks; POSNA-PODCI = Pediatric Orthopedic Society of North America Pediatric Outcomes Data Collection Instrument; RGI-C = Radiographic Global Impression of Change; RSS = Rickets Severity Score

*) Results are mean change from baseline with SE for RSS and mean final value with SE for RGI-C. a) LS mean and p value per GEE model, which included visit, regimen, visit by regimen as factors, and score at baseline as a covariate, with exchangeable covariance structure. b) The RGI-C score was based on a 7-point ordinal scale ranging from -3 (very much worse, or severe worsening of rickets) to +3 (very much better, or complete or near complete healing of rickets).

In the Q2W group, mean (SD) growth velocity increased, from [redacted] cm/year at baseline (i.e., the two years before study entry) to [redacted] cm/year [redacted], one sample t-test). Mean (SD) standing height z-score increased from [redacted] at baseline to [redacted] at week 64, an LS mean (SE) change of [redacted]. Mean (SD) percentile standing heights were [redacted] at baseline and [redacted] at week 64.

STUDY CL205 - burosumab in children aged 1-4 years

An overview of the results for CL205 are shown in Table 4.8. Overall, burosumab significantly improved rickets and [redacted] and [redacted]. [redacted].

Table 4.8: Overview of outcomes from Study CL205

	Week 40		
Endpoint	**N**	**Effect Size**	**p-value**
RSS Total Score % mean change from baseline[a] (negative is better)	13	-59%	< 0.0001
RGI-C Global Score LS mean[b] (positive is better)	13	+2.33	< 0.0001
Substantial Healing by RGI-C % RGI-C global score ≥ +2.0	13	[redacted]	-
ALP % mean change from baseline[c] (negative is better)	13	-36.3%	< 0.0001
RGI-C Lower Limb Deformity Score LS mean[b] (positive is better)	13	[redacted]	[redacted]
Recumbent Length/Standing Height Mean change from baseline (cm)	13	[redacted]	[redacted]
Recumbent Length/Standing Height z-score LS mean change from baseline[d]	13	[redacted]	[redacted]

Source: CS, Table 24, page 109
ALP = alkaline phosphatase; ANCOVA = analysis of covariance; LS = least squares; GEE = Generalised Estimating Equations; RGI-C = Radiographic Global Impression of Change; RSS = Rickets Severity Score
a) Percent change based on arithmetic means; p value based on ANCOVA model.
b) LS mean and p value based on ANCOVA model.
c) Percent change based on arithmetic means; p value based on GEE model.
d) LS mean and p value based on GEE model.

Impact of burosumab on bone mineral metabolism

Change in serum phosphorus (primary endpoint)

At baseline, all patients had serum phosphorus levels below normal, with a mean (SD) of 2.51 (0.284) mg/dL (0.81 [0.092] mmol/L) compared with the normal range of 3.2 to 6.1 mg/dL (1.03 to 1.97 mmol/L). Increases in serum phosphorus concentration from baseline were statistically significant at each study visit ($p < 0.0001$, GEE analysis). At week 40, mean (SD) serum phosphorus concentrations were 3.47 (0.485) mg/dL (1.12 [0.158] mmol/L); change from baseline to week 40 was 0.96 (0.439) mg/dL (0.31 [0.143] mmol/L).

Serum $1,25(OH)_2D$

Burosumab treatment increased serum $1,25(OH)_2D$ levels from [redacted]. Increases in $1,25(OH)_2D$ from baseline were statistically significant at each study visit through week 40 ($p < 0.01$, GEE analysis).

Assessment of rickets

RSS total score (secondary efficacy outcome)

Burosumab treatment for 40 weeks significantly reduced rickets severity as assessed by RSS scores. RSS total score at week 40 was reduced by 59% ($p < 0.0001$, ANCOVA model) least squares (LS) mean (SE) change of -1.73 (0.132) (see Table 4.8). Mean (SD) RSS total scores were 2.92 (1.367) at baseline and 1.19 (0.522) at week 40. Similarly, RSS wrist scores and knee scores were reduced at week 40 by [redacted]), respectively.

RGI-C global score (secondary efficacy outcome)

Burosumab treatment for 40 weeks resulted in healing of rickets as assessed by RGI-C scores. LS mean (SE) values at week 40 were +2.33 (0.080) for RGI-C global scores; +2.26 (0.110) for RGI-C wrist scores; and +2.21 (0.153) for RGI-C knee scores ($p < 0.0001$ for all, ANCOVA model) (see Table 4.8).

Other outcomes

At baseline, mean (SD) serum ALP levels were 549 (193.8) U/L, well above the upper limit of normal (ULN) for the children in this study (approximately 297 to 345 U/L, depending on the age and gender of the child). Mean (SD) serum ALP levels decreased to 389 (84.2) U/L at week 20 (mean change: -24.8%) and to 335 (87.6) U/L at week 40 (mean change: -36.3%). Changes from baseline to weeks 20 and 40 were statistically significant ($p < 0.0001$).

Burosumab treatment for 40 weeks resulted in [redacted].

Mean (SD) recumbent length/standing height [redacted].

STUDY CL002 - historical control study

Impact of conventional therapy on bone mineral metabolism

At the time of the baseline radiographs, the mean serum phosphorus level in the overall group was [redacted], below the lower limit of normal (LLN, 3.2 mg/dL [1.03 mmol/L]) for children. At the post-baseline radiographs, mean serum phosphorous level [redacted].

Effect of conventional therapy on rickets

RSS and RGI-C score change from baseline

Prolonged treatment with oral phosphate/calcitriol therapy for a median of [redacted]. Changes in RSS total scores (wrist and knee combined) showed a mean [redacted] with continued treatment with oral phosphate/calcitriol therapy.

For the higher RSS subgroup of the prespecified analysis, mean total RSS decreased (improved) from [redacted] for the baseline radiographs to [redacted] for the post-baseline radiographs. For the lower RSS subgroup, mean total RSS score [redacted].

The RGI-C global score was [redacted] post-baseline for the overall population, [redacted] for the higher RSS subgroup, and [redacted] for the lower RSS subgroup, which translate to less than minimal healing of rickets over a median period of 102 weeks.

Lower extremity deformity

After long-term treatment with conventional therapy, the mean RGI-C lower limb deformity score was [redacted] for the overall group, indicating [redacted].

Impact of conventional therapy on growth

Observational data corresponding to the [redacted] paired baseline radiographs showed that many patients in this study had decreased height for age (mean [SD] standing height z-score of [redacted] After long-term treatment with conventional therapy, [redacted].

4.2.4.2 Adverse events

In their summary of the safety profile of burosumab, the EPAR states: "The most common adverse drug reaction (ADR) reported in paediatric patients up to 64 weeks was injection site reactions (57%), headache (54%), pain in extremity (42%), vitamin D decreased (28%), rash (23%), toothache (19%), tooth abscess (14%), myalgia (14%), and dizziness (11%)".[50]

Table 4.9 gives the adverse reactions observed from clinical trials. The adverse reactions are presented by system organ class and frequency categories, defined using the following convention: very common (≥1/10); common (≥1/100 to <1/10); uncommon (≥1/1000 to <1/100); rare (≥1/10,000 to <1/1000); very rare (<1/10,000), not known (cannot be estimated from the available data). Within each frequency grouping, undesirable effects are presented in order of decreasing seriousness.[50]

Table 4.9: Adverse reactions reported in paediatric patients with XLH (N=65)

MedDRA System Organ Class	Frequency category	Adverse reaction
Infections and infestations	Very common	Tooth abscess
Nervous system disorder	Very common	Headache
	Very common	Dizziness
Gastrointestinal Disorders	Very common	Toothache
Skin and subcutaneous tissue disorder	Very common	Rash
Musculoskeletal and connective tissue disorders	Very common	Myalgia
	Very common	Pain in extremity
General disorders and administration site conditions	Very common	Injection site reaction

Investigations	Very common	Vitamin D decreased
Source: EMA - EPAR, Table 1, page 6[50]		

Injection site reactions: Local reactions (e.g. injection site urticaria, erythema, rash, swelling, bruising, pain, pruritus, and haematoma) have occurred at the site of injection. In the paediatric studies, approximately 57% of the patients had an injection site reaction. The injection site reactions were generally mild in severity, occurred within one day of medicinal product administration, lasted approximately one to three days, required no treatment, and resolved in almost all instances.

Skin reactions: In paediatric patients, the most frequent potential hypersensitivity events were rash (22%), injection site rash (6%), and urticaria (4%). The events were mild or moderate in severity.

Immunogenicity: Anti-drug antibodies (ADA) have been detected in a small percentage of patients receiving burosumab who had also tested positive for ADA prior to dosing; no adverse events or loss of efficacy was associated with these findings.[50]

In study CL201, one patient experienced serious TEAEs, and [redacted] (see Table 4.10). In study CL205, one patient experienced an SAE [redacted] considered unlikely unrelated to study drug. All 13 subjects (100%) experienced at least one TEAE during the study (see Table 4.10). [redacted].

Table 4.10: Summary of adverse events in studies CL201 and CL205 (Safety Analysis Set (SAS))

Category	Study CL201			Study CL205
	Burosumab Q2W (N = 26)	Burosumab Q4W (N = 26)	Overall (N=52)	Burosumab (N = 13)
AEs starting during screening period				4 (30.8%)
All TEAEs	26 (100%)	26 (100%)	52 (100.0%)	[redacted]
Serious TEAEs	0 (0.0%)	1 (3.8%)	1 (1.9%)	1 (7.7%)
Related TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Serious Related TEAE	0 (0.0%)	1 (3.8%)	1 (1.9%)	[redacted]
Grade 3 or 4 TEAE	[redacted]	[redacted]	[redacted]	[redacted]
TEAE leading to study discontinuation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
TEAE leading to treatment discontinuation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
TEAE leading to death	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Source: CS, Tables 26 and 28, pages 116 and 120 Q2W, every 2 weeks; TEAE, treatment-emergent adverse event				

The most frequent TEAEs (>30% incidence) in study CL201 were [redacted]

[redacted]
(see Table 4.11).

The most frequent TEAEs (> 30% incidence [four or more of 13 patients]) in study CL205 were [redacted] see Table 4.11). [redacted]

Table 4.11: Treatment-emergent adverse events* by SOC and preferred term (SAS)

System Organ Class Preferred Term	Study CL201 Q2W (N = 26)	Study CL201 Q4W (N = 26)	Study CL201 Overall (N = 52)	Study CL205
Patients with any TEAE	26 (100.0%)	26 (100.0%)	52 (100.0%)	13 (100%)
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]	
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Pharyngitis streptococcal	[redacted]	[redacted]	[redacted]	[redacted]
Tooth abscess	[redacted]	[redacted]	[redacted]	[redacted]
Gastroenteritis viral	[redacted]	[redacted]	[redacted]	
Nasopharyngitis				[redacted]
Viral upper respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Influenza	[redacted]	[redacted]	[redacted]	
Viral infection	[redacted]	[redacted]	[redacted]	
Lice infestation	[redacted]	[redacted]	[redacted]	
Gastrointestinal disorders	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Oral pain				[redacted]
Abdominal discomfort				[redacted]
Abdominal pain upper	[redacted]	[redacted]	[redacted]	[redacted]
Toothache	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	
Abdominal discomfort	[redacted]	[redacted]	[redacted]	
Abdominal pain	[redacted]	[redacted]	[redacted]	
Constipation	[redacted]	[redacted]	[redacted]	
Mouth ulceration	[redacted]	[redacted]	[redacted]	
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]	[redacted]
Injection site reaction	[redacted]	[redacted]	[redacted]	
Injection site erythema	[redacted]	[redacted]	[redacted]	
Pyrexia	[redacted]	[redacted]	[redacted]	[redacted]
Injection site pruritus	[redacted]	[redacted]	[redacted]	

System Organ Class Preferred Term	Study CL201 Q2W (N = 26)	Study CL201 Q4W (N = 26)	Study CL201 Overall (N = 52)	Study CL205
Injection site swelling	[redacted]	[redacted]	[redacted]	
Pain	[redacted]	[redacted]	[redacted]	
Fatigue	[redacted]	[redacted]	[redacted]	
Injection site pain	[redacted]	[redacted]	[redacted]	
Injection site rash	[redacted]	[redacted]	[redacted]	
Injection site bruising	[redacted]	[redacted]	[redacted]	
Malaise	[redacted]	[redacted]	[redacted]	
Respiratory thoracic and mediastinal disorders	[redacted]	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]	[redacted]
Oropharyngeal pain	[redacted]	[redacted]	[redacted]	
Nasal congestion	[redacted]	[redacted]	[redacted]	[redacted]
Rhinorrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Respiratory tract congestion				[redacted]
Epistaxis	[redacted]	[redacted]	[redacted]	
Sneezing	[redacted]	[redacted]	[redacted]	
Wheezing	[redacted]	[redacted]	[redacted]	
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]
Hypersomnia				[redacted]
Headache	[redacted]	[redacted]	[redacted]	
Dizziness	[redacted]	[redacted]	[redacted]	
Migraine	[redacted]	[redacted]	[redacted]	
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]
Pain in extremity	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]
Myalgia	[redacted]	[redacted]	[redacted]	
Back pain	[redacted]	[redacted]	[redacted]	
Bone pain	[redacted]	[redacted]	[redacted]	
Musculoskeletal pain	[redacted]	[redacted]	[redacted]	
Injury poisoning and procedural complications	[redacted]	[redacted]	[redacted]	[redacted]
Skin abrasion				[redacted]
Contusion	[redacted]	[redacted]	[redacted]	
Skin abrasion	[redacted]	[redacted]	[redacted]	
Fall	[redacted]	[redacted]	[redacted]	
Procedural pain	[redacted]	[redacted]	[redacted]	
Arthropod bite	[redacted]	[redacted]	[redacted]	

System Organ Class Preferred Term	Study CL201 Q2W (N = 26)	Study CL201 Q4W (N = 26)	Study CL201 Overall (N = 52)	Study CL205
Ligament sprain	[redacted]	[redacted]	[redacted]	
Thermal burn	[redacted]	[redacted]	[redacted]	
Skin and subcutaneous tissue disorders	[redacted]	[redacted]	[redacted]	
Rash	[redacted]	[redacted]	[redacted]	
Dry skin	[redacted]	[redacted]	[redacted]	
Investigations	[redacted]	[redacted]	[redacted]	
Vitamin D decreased	[redacted]	[redacted]	[redacted]	
Blood 25-hydroxycholecalciferol decreased	[redacted]	[redacted]	[redacted]	
Immune system disorders	[redacted]	[redacted]	[redacted]	
Seasonal allergy	[redacted]	[redacted]	[redacted]	
Ear and labyrinth disorders	[redacted]	[redacted]	[redacted]	[redacted]
Ear pain	[redacted]	[redacted]	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]	[redacted]	
Vitamin D deficiency	[redacted]	[redacted]	[redacted]	
Neoplasms benign, malignant and unspecified (incl. cysts and polyps)	[redacted]	[redacted]	[redacted]	
Skin papilloma	[redacted]	[redacted]	[redacted]	

Source: CS, Tables 27 and 29, pages 116 to 121
Q2W = every 2 weeks; TEAE, treatment-emergent adverse event
*) CL201: TEAEs occurring in ≥ 3 patients overall; CL205: TEAEs occurring in ≥ 2 patients.

ERG comment: Adverse events of treatment with conventional therapy have not been reported. Therefore, it is not possible to assess the relative safety and toxicity in relation to the comparator.

4.2.4.3 Deaths

No patient died or discontinued from CL201 or CL205 for any reason; all patients continued treatment on study as of the data cut-off dates.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

4.3.1 Methods

As stated by the company, "the burosumab phase 2 studies were uncontrolled dose finding or single arm studies, therefore an indirect comparison was not feasible" (CS, page 123).[1] However, the company provides both a naïve comparison and a matched comparison of the results from Study CL201 (burosumab in children with XLH, 5-12 years) and Study CL002 (conventional therapy in children with XLH, 5-14 years) by listing results in Table 17 (page 94) of the CS (see Table 4.6 in this report). As outlined in chapters 4.2.1 and 4.2.4 of this report, the naïve comparison is unreliable because of the differences in inclusion criteria and patient characteristics in both studies, particularly relating to rickets severity. Study CL201 included children with more severe symptoms of XLH. This is also reflected in

the relatively lower standing height and higher rickets severity score for children in study CL201 when compared to children in study CL002 (see Table 4.5).

In order to try and compensate for differences between the two studies, the company also performed a comparison of rickets severity outcomes (RSS and RGI-C) between burosumab (Study CL201) and conventional therapy (Study CL002) using propensity score matching (PSM). These analyses were carried out using the whole population of Study 201 and therefore included those who received burosumab at both doses (Q2W and Q4W). The company does mention that "the Q2W regimen showed a more stable increase in pharmacodynamic markers as compared with the Q4W regimen. Moreover, assessments of rickets, growth, and walking ability consistently showed greater improvement with the Q2W regimen as compared with the Q4W regimen, with no increase in AE's" (CS, page 93).[1] However, specific results for the Q4W regimen are not presented in the CS.

The company does acknowledge some limitations of using study CL002 as a comparator group for study CL201: "It was a retrospective radiograph and chart review study rather than a prospective natural history cohort, [redacted]." (CS, page 125).[1] There were also differences in patient characteristics between the two studies. The statistical analysis plan for the PSM provided by the company in the response to clarification stated that the two study populations were similar for race, ethnicity, and age at commencing conventional therapy but that "baseline rickets severity as measured by RSS is higher in the CL201 cohort compared to CL002. In addition, baseline age and gender for the two studies are not very comparable" (SAP, page 14).[57] However, they did not report the methods used to judge the comparability of the two studies (statistical testing or other methods). The ERG compared age and gender between the two study populations and did not find any statistically significant differences between them. For baseline age the mean was [redacted] for CL201 and [redacted] for CL002 giving a mean difference of [redacted] and for gender there were [redacted] in CL201 and [redacted] males in CL002 with a p-value = [redacted] (chi-squared test). However, the baseline total RSS score was significantly higher in CL201 (mean difference [redacted]. Therefore, the company used PSM to try and create a more comparable sample for the analysis of rickets severity between burosumab (using study CL201) and conventional therapy (using study CL002). The propensity score (PS) is the estimated conditional probability of being treated with burosumab compared to conventional therapy based on observed individual patient baseline covariates. A logistic regression model adjusting for baseline RSS total score, age and gender was used to estimate a PS value for each patient The PS values were used to adjust for differences between the patient populations of the two studies in the analyses in a number of different ways:

1. Inverse probability of treatment weighting (IPTW): in the analysis the data for each patient is weighted by their PS where patients on burosumab are given a weight of 1/PS and patients on conventional therapy are given a weight of 1/(1-PS). These weights were then included in an analysis of covariance (ANCOVA) model with change from baseline in RSS total score, of the final RGI-C global score as the outcome and adjusting for treatment group and baseline RSS total score. All subjects from both studies were included in the analysis (including Q4W burosumab).
2. Propensity score matching (PSM): patients receiving burosumab or conventional therapy where matched based on their closest PS values. Only patients who could be successfully matched were included in the analysis and the maximum tolerated difference for matching was 0.2 SD of the logit of the PS values [source SAP section 7.2.4].[57] After matching the two treatment groups were compared using the same ANCOVA model used in the IPTW analyses. Two matching methods were used:

- Matching without replacement: burosumab patients were matched one at a time to their closest control (conventional therapy patient). Once a conventional therapy patient was matched they were removed from the matching dataset and excluded from the analysis. To account for matching variability the matching was repeated 1,000 times and the order of patients in the burosumab group was randomly sorted.
- Matching with replacement: as there were fewer conventional therapy patients compared to burosumab patients matching with replacement was also used. Here a conventional therapy patient could be matched with multiple burosumab patients and they received higher weights in the analysis based on the number of times they were matched. The weights were included in the ANCOVA model.

4.3.2 Results

Details of the baseline patient characteristics of studies CL201 and CL002 before (original study data) and after PS weighting and matching are shown in Table 4.12 below. The study populations from the PSM were more comparable than those from the original studies, particularly with regards to the baseline RSS total score.

Table 4.12: Baseline characteristics in studies CL201 (burosumab) and CL002 (conventional therapy) in propensity score analysis

	Study assessment (not weighted)		Weighted by inverse probability of treatment		Propensity score matching without replacement in control		Propensity score matching with replacement in control	
	CL201	CL002	CL201	CL002	CL201	CL002	CL201	CL002
Sample size	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age at baseline (mean [SD] years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gender (% female)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age when conventional therapy initiated (mean [SD] years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline RSS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wrist score (mean [SD])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Knee score (mean [SD])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total score (mean [SD])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS, Table 30, page 127

a) Burosumab subjects (Study CL201) receive a weight equal to 1/Propensity Score, and conventional therapy subjects (Study CL002) receive a weight equal to 1/(1-Propensity Score), where the propensity score is

estimated from a logistic regression model with treatment group as response (1 = burosumab, 0 = conventional therapy), baseline RSS total score and age as covariates and sex as a categorical covariate.
b) Mean sample size and results based on 1000 iterations of PS matching without replacement.
c) A conventional therapy subject could be selected to match multiple treated subjects. Conventional therapy subjects matched multiple times received higher weights based on the number of times matched.
d) All subjects from the intent-to-treat (ITT) analysis set were selected.
e) All subjects from the radiograph analysis set were selected; when more than one radiograph pair available for a subject, the pair with the duration between two radiographs taken closest to 64 weeks is selected; radiographs that were deemed as growth plates fused or partially fused were excluded from the analysis.

Figure 4.1: Differences in RSS total scores (LS mean ± SE) between Study CL201 (burosumab treatment) and Study CL002 (conventional therapy) from propensity score analyses

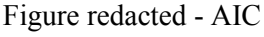
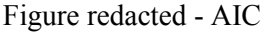


Figure 4.2: Differences in RGI-C global scores (LS mean ± SE) between Study CL201 (burosumab treatment) and Study CL002 (conventional therapy) from propensity score analyses



ERG comment: As there was no direct or indirect evidence available to compare burosumab with conventional therapy using evidence from RCTs, the evidence in the CS is based on a comparison of data from two single arm studies. Although the burosumab evidence is from a phase 2 trial, there was no control group and the randomisation was between different regimens of burosumab. The data for conventional therapy was obtained from a historical cohort study, which was different to the burosumab trial in terms of inclusion criteria and patient population. In order to try and adjust for differences between these two studies the company performed additional analyses which matched the two groups using propensity score matching. However, these analysis methods have major limitations, in that the matching can only include those variables measured in both studies. Randomisation in a clinical trial creates balanced group for both measured and unmeasured variables. In observational studies, the most important factors which are predictive of the outcome may not have been measured and any treatment comparisons using observational study data may be biased.[58]

The company only included three variables in the PSM, age, gender and RSS total score at baseline. The rationale for variable selection was not provided other than whether they seemed similar or not between the two study populations. No details were provided of how this similarity was judged. The ERG found no statistically significant differences in age and gender between the two groups and considered that only including three variables in the creation of the propensity scores may have been too few. Although the PSM groups were closer at baseline for these three variables compared to the original data, the results of the PSM analyses were very similar to those from a naïve comparison between the two study populations.

The company provided the statistical analysis programs used for the PSM analyses in the response to the clarification letter but not the data. Therefore, the ERG could not check the PSM analyses to establish that they could reproduce the results. Three different PSM methods were used and although they provided similar results it is not clear which PSM result should be considered the most reliable. The PSM analyses were only performed for rickets and not for any other relevant clinical or safety outcomes.

Due to the lack of a direct comparison between burosumab and conventional therapy and the limitations of using propensity score matching with data from two different observational studies the results of the rickets analyses presented by the company should be considered with caution. The results from CL301, a randomised controlled trial comparing the efficacy and safety of burosumab with active control (oral phosphate/active vitamin D therapy) in children with XLH (aged one to ≤12 years) are expected [redacted]. These will provide more reliable estimates for the clinical effectiveness and safety of burosumab compared to conventional therapy and should be given greater consideration than the naïve and adjusted analyses presented in the company submission.

4.4 Summary of evidence presented in other submissions

No other scientific evidence was submitted by other consultees.

4.5 Additional work on clinical effectiveness undertaken by the ERG

Additional work on clinical effectiveness undertaken by the ERG has been included in section 4.2.4 of this report. In addition, we will discuss the longitudinal review of patient records from three expert UK centres to provide additional data (n=43) commissioned by Kyowa Kirin as a UK alternative to CL002 which was a US study. The company provided a synopsis with details on the rationale, methodology and results of this UK study as part of the response to the clarification letter.[2]

The study included paediatric patients (up to age 18) with a confirmed diagnosis of XLH, as defined by radiological and clinical evidence of rickets, with documentation of a confirmed PHEX mutation. To be included in the analysis patients must have had at least two sequential radiographs. Study CL002 included paediatric patients with a confirmed diagnosis of XLH, but radiographic images from at least two time points taken between the ages of five and 14 years, inclusive, had to be available. Therefore, the UK study has a wider age range and is less comparable to study CL201in terms of age as can be seen in Table 4.13. However, the company does add that "the mean age at each RSS observation across the patients was 7.5 years, which is therefore similar to CL201 and CL002".[2]

Table 4.13: Demographic and baseline characteristics in CL201, CL205, CL002 and UK review

	CL201	Study CL002	CL205	UK Review
	Q2W (n=26)	Radiographic analysis set ([redacted])	(n=13)	Radiographic analysis (n=38)
Age (years), mean (SD)	8.7 (1.72)	[redacted]	2.9 (1.15)	[redacted]
Sex, male n (%)	12 (46.2%)	[redacted]	9 (69.2%)	[redacted]
Race White Black/ African-American Other	 23 (88.5%) 2 (7.7%) 1 (3.8%)	[redacted]	 12 (92.3%) 1 (7.7%) 0	[redacted]
Weight (kg), mean (SD)	31.87 (7.92)	[redacted]	12.92 (1.81)	[redacted]
Height (percentile for age and gender), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Standing Height (z-score), mean (SD)	-1.72, 1.03	[redacted]	-1.38 (1.19)	NR
Renal ultrasound score, (0 – 5 scale) – n (%) 0 1 2	[redacted]	[redacted]	NR	NR
Number (%) of Patients Who Received Prior Conventional Therapy	24 (92.3%)	[redacted]	13 (100%)	NR
Duration of Prior Conventional Therapy, mean (SD)	7.02 (2.14) years	[redacted]	16.7 (14.39) months	NR
Age When Conventional Therapy Was Initiated (years), mean (SD)	[redacted]	[redacted]	[redacted]	NR
Pharmacodynamic parameters, mean (SD) Serum Phosphorus, mg/dL TmP/GFR (mg/dL)	 [redacted] [redacted]	 [redacted] [redacted]	 [redacted] [redacted]	NR

	CL201	Study CL002	CL205	UK Review
	Q2W (n=26)	Radiographic analysis set ([redacted])	(n=13)	Radiographic analysis (n=38)
Serum 1,25(OH)2 D (pg/mL)	[redacted]	[redacted]	[redacted]	
ALP (U/L	[redacted]	[redacted]	[redacted]	
Rickets Severity				
RSS Total Score, mean (SD)	1.92 (1.17)	[redacted]	2.92 (1.37)	NR

Source: CS, Table 13, page 82 and Response to Clarification letter (Question A4)
a) At baseline paired radiograph (the earlier radiograph pair)

Data were collected from two participating UK expert centres (Birmingham Children's Hospital NHS Foundation Trust and Central Manchester University Hospitals NHS Foundation Trust). At the baseline visit (diagnosis) data were collected on patient demographics (age, date of diagnosis, ethnicity and gender), medical history, family history of XLH, basic parameters (weight, blood pressure, height and biochemical parameters (calcium [corrected], parathyroid hormone, phosphate and alkaline phosphatase)), current medications and rickets severity. At the follow-up visit (most recent) data were collected on significant events (for example, new comorbidities, fractures, hospitalisations, ectopic calcifications, orthopaedic surgery), basic parameters (as before), current medications and rickets severity.

Rickets severity was graded using the Rickets Severity Score (RSS; Thacher scores), as used in the burosumab clinical trial program. The same consultant radiologist based in Manchester provided RSS scores for all radiographs in the review.

Planned analyses and outcomes included the assessment of RSS at different timepoints, based on availability of radiographic data and assessment of patient weight by age and gender.

Results included data from 43 patients, diagnosed between June 1992 and August 2016. Of the 43 patient histories, data from 38 patients were included as they provided two radiographic scores.

The only results presented for the UK review are the data presented in Table 4.14 below. As such these data are not comparable to data reported in study CL002 and in the burosumab studies. It is unclear how comparable these data are to any of the burosumab data.

Table 4.14: Rickets status at x-rays from UK chart review, based on RSS

Year n+1 / Year n	Mild	Moderate	Severe	Healed	Total
Mild	12	5	4	3	24
Moderate	7	14	5	2	28
Severe	4	10	33	3	50
Healed	1	1	2	1	5
Total	24	30	44	9	107

Source: Response to clarification letter, question A4

The company states that "Due to the nature of a retrospective chart review, which provides RSS scores with varying time between visits, annualised estimates of changes in RSS score have not been analysed in detail. However, the transition matrices used in the cost-effectiveness model provide clear indication of the RSS progression amongst patients" (see Table 4.14).[2] "Nearly half of the x-rays conducted indicated that patients had severe rickets, as 50 of the 107 (47%) observations were from severe rickets. This is comparable to the baseline characteristics of the CL205 and CL201 studies, in which 43% of patients were severe. Half of the patients with mild rickets (RSS 0.5 or 1) did not have a significant change in RSS between visits, but in those that did, more deteriorated than improved (9 vs 3 patients). Few patients had healed rickets at any one time (9 of 107 x-rays) but the healed status appeared to be temporary as only one remained healed at the next x-ray".[2]

4.6 Conclusions of the clinical effectiveness section

4.6.1 Completeness of the CS with regard to relevant clinical studies and relevant data within those studies

The ERG is confident that all relevant studies (published and unpublished) of burosumab were included in the CS, including data from ongoing studies. The same applies to the historical control patients. A control study in UK patients was mentioned in the CS without any results being report in the CS. However, results were provided as part of the response to the clarification letter. The reporting of outcomes from included studies also seems complete.

A randomised controlled study comparing burosumab with active control (oral phosphate/active vitamin D therapy) in children with XLH (aged one to ≤12 years) is currently ongoing. [redacted].[2] Results from this study will considerably reduce the uncertainty surrounding the clinical effectiveness of burosumab relative to conventional therapy in children with XLH aged between one and 12 years.

4.6.2 Interpretation of treatment effects reported in the CS in relation to relevant population, interventions, comparator and outcomes

A key issue that may limit the robustness of the efficacy data reported in the CS relates to the study design of the included studies. Due to the absence of a control group in most studies, inference of treatment effects (including magnitude) may be confounded. As stated by the company, the "burosumab phase 2 studies were uncontrolled dose finding or single arm studies, therefore an indirect comparison was not feasible" (CS, page 123).[1]

For children between one to four years old, only one study is presented in which all children received burosumab (CL205). A comparison with "established clinical management without burosumab"[52] is not possible in this group of patients.

For children between five to 12 years old, the CS presents a study in which all children received burosumab (CL201). In addition, the CS presents a control study (CL002) in which children aged between five to 14 years received conventional therapy (i.e. oral phosphate/active vitamin D). Results of these two studies are mainly presented as a naïve comparison, simply reporting individual results from each study side by side. In addition, the company presents comparisons of 'rickets healing' with conventional therapy (Study CL002) versus burosumab (Study CL201) using propensity analysis matching.

In the CS, the company uses the term 'healing' and 'substantial healing of rickets'. This is defined using RGI-C global scores, where scores ≥ +1.0 indicate 'healing of rickets' and scores ≥ +2.0 'substantial

healing of rickets'. The company does explain that "Healing in this context indicates improvement in the radiographic abnormalities and does not imply that complete healing was observed" (CS, page 100).[1] However, throughout the report the term 'healing of rickets' is used without any explanation of the degree of healing (minimal, substantial or complete). Moreover, RGI-C global scores and RSS scores do not capture all clinical aspects of XLH. That is of particular importance in the context of the economic model, which only considers RSS score alone as a clinical outcome measure. The diverse physiological impacts of hypophosphataemia, which may be independent of rickets, are therefore not captured as outcomes in the economic model.

In the response to the clarification letter the company described the vignettes for the various health states that informed the economic model in detail (Clarification Letter Response Question B7, Table 10). However, each health state was defined in such a way that there appears to be a perfect association between the RSS score and other clinical descriptors of the health state. For example, as the RSS score decreases so does the risk of fracture and the presence of deformity. However, this does not appear to be realistic in that it seems likely that there might be some resolution of the bone disorder such that the RSS score decreases, but that this resolution only occurs after incurring deformity, which cannot be completely resolved and with some continued increased risk of fracture.

In addition, the model currently assumed that the effect of burosumab, although stopped at age 16 (women) or 17 (men) lasts for the rest of their lives. This also seems unrealistic, the effects of burosumab on stature, bowing of the legs, joint deformity etc. are likely to persist fairly long but may wane as osteomalacia itself and the resulting fractures may lead to associated problems in later life. Effects on bone strength will wane quicker, therefore repeated fractures and badly healing fractures after 10 or 20 years are likely to occur. Effects of burosumab on symptoms caused by hypophosphatemia itself will disappear as soon as therapy is stopped. Therefore, we have assumed in the ERG base-case that patients will experience a decline in quality of life 20 years after the end of treatment, which was operationalised by moving to the utility value of the next worse health state (see section 6.2.3 in this report).

Regarding the evidence synthesis, the naïve comparison is unreliable because there are important differences between the inclusion criteria in both studies. Inclusion criteria for patients in studies CL201 and CL002 are similar in that patients in both studies were diagnosed with XLH and were of similar age. However, children in study CL201also had: biochemical findings associated with XLH, standing height < 50th percentile for age and gender and radiographic evidence of active bone disease including rickets in the wrists and/or knees, and/or femoral/tibial bowing, or, for expansion patients, an RSS score in the knee of at least 1.5 points as determined by central read. In other words, study CL002 included all children with XLH, while study CL201 included children with more severe symptoms of XLH. This is also reflected in the relatively [redacted] standing height and [redacted] rickets severity score for children in study CL201 when compared to children in study CL002.

The adjusted comparison, using propensity analysis matching, is unreliable because of the limitations associated with these methods, in that the matching can only include those variables measured in both studies. Randomisation in a clinical trial creates balanced group for both measured and unmeasured variables. In observational studies, the most important factors which are predictive of the outcome may not have been measured and any treatment comparisons using observational study data may be biased. In the CS the company only included three variables in the PSM: age, gender and RSS total score at baseline. The rationale for variable selection was not provided other than whether they seemed similar or not between the two study populations. No details were provided of how this similarity was judged. The ERG found no statistically significant differences in age and gender between the two groups and

considered that only including three variables in the creation of the propensity scores may have been too few.

4.6.3 Uncertainties surrounding the reliability of the clinical effectiveness

The main uncertainty regarding the effectiveness evidence is the comparability of results from treated patients and historical control patients. Most of the evidence is presented as single arm studies including either treated patients (two studies, both with extensions that are still ongoing) or historical control patients (one study, with patients from one single centre, Radiographic analysis set ([redacted])). The historical control study (CL002) included patients aged from five to 14 years and can therefore only serve as a control group for study CL201 (children aged five to 12 years).

For patients with XLH aged one to four years old, the CS only presents a single arm burosumab study (CL205), no control data for this age group were provided. Only 13 children were enrolled in study CL205; therefore, results in this age group are very uncertain.

5 VALUE FOR MONEY FOR THE NHS AND PSS

5.1 Introduction

The aim of this chapter is to provide an assessment of whether or not burosumab for X-linked hypophosphatemia (XLH) represents value for money for the NHS in England. This assessment is mainly based on the evidence submitted to NICE in the company submission and in the response to the clarification letter. This includes a cost effectiveness model, a description of the methods and assumptions used to inform the input parameters of the model, and the results of economic analyses performed using the submitted cost effectiveness model. This chapter starts with a review of existing economic analyses for burosumab either from the literature or elsewhere in the public domain. Afterwards, a detailed exposition and critique of the submitted model and economic analyses is presented.

5.2 Review of existing economic analyses

The company conducted a systematic review of cost effectiveness studies of burosumab and other studies including costs, resource use and any HRQoL measure associated with XLH. The details of the search strategy were provided in Section 17.3 of the CS.[1] A summary of the search strategy and the review process leading to the selection of relevant papers is given in the remaining parts of this section.

5.2.1 Searches

Section 11.1 of the CS states that a systematic literature review of the economic and health economic evidence on XLH was undertaken. Search strategies were reported in detail in Appendix 17.3 of the CS and in the response to clarification. MEDLINE, Embase, EconLit and the NHS Economic Evaluation Database were listed as the databases searched in the identification of economic evidence. All databases were searched from the earliest date available for each database until the end of October 2017. The searches were also intended to identify studies for health-related quality of life data and for resource identification, measurement and valuation studies.

The CS states (p.150) that grey literature was identified '*provided that the foundation for the reported findings is a study with a publicly available research protocol or is a study published in full manuscript form as an academic resource*'.[1] Three main journals in the field were hand-searched, and reference checking was carried out. Experts and clinical specialists were also consulted.

The company submission and request for clarification provided full search strategies for MEDLINE, Embase and EconLit. Strategies were not provided for NHS EED, so it is not clear if this search was undertaken.

ERG comment:

- The selection of databases searched was adequate and most searches were reproducible. The database name, host, date range and date searched were provided for the majority of the searches. A good range of additional resources were included.
- The ERG only presents the major limitations of the search strategies here. Further minor criticisms can be found in Appendix 1 of this report.
- The main concern of the ERG is that the search terms used for the population facet of the strategy were insufficient. Only one indexing (MeSH/EMTREE) term was used, combined with one free-text term. Numerous synonyms are available for X-linked hypophosphatemia and use of these terms would have increased the retrieval of potentially relevant records.

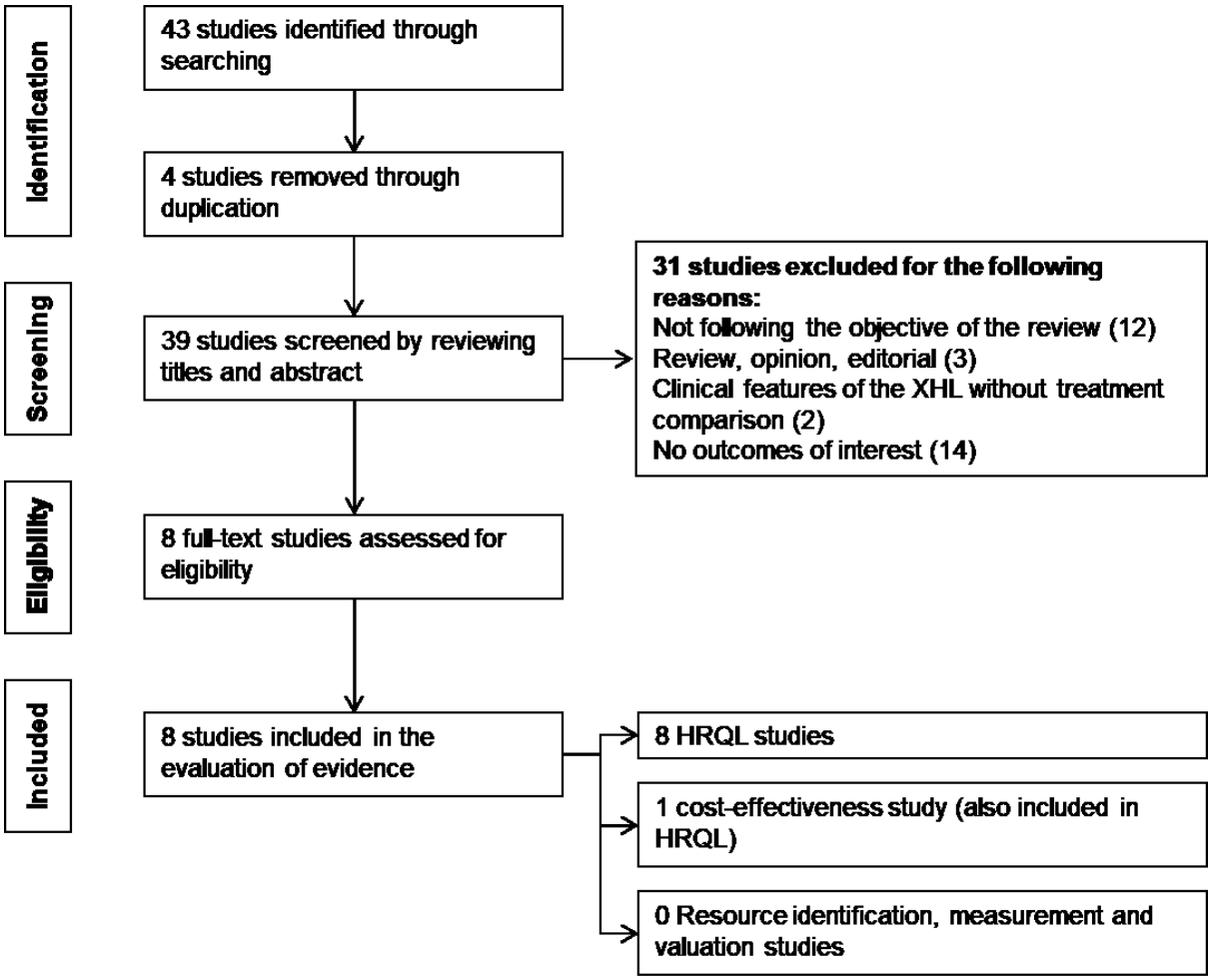
- Given the small number of papers retrieved for this topic, the ERG believes that use of study design filters in the searches was unnecessarily restrictive. The ERG suggests that a single-facet search for XLH (and additional synonyms) without a study design filter would have adequately addressed all areas of interest, including clinical effectiveness, adverse events, cost effectiveness, HRQoL and resource use, without retrieving unmanageably high numbers of records. See Appendix 1 for example MEDLINE, Embase and CENTRAL searches run by the ERG.
- The strategies provided for both MEDLINE and Embase contain repeated facets and considerable redundancy. The structure of the searches is confused; however, the final results sets appear to be correct.
- The EconLit search does not include details of the host used, database fields searched, or the number of results found. The strategy is therefore not reproducible. No strategy or results are provided for NHS EED; therefore, it is not clear whether this database was searched.

5.2.2 Review process and results

The company used broad selection criteria for the health economic evidence as reported in Table D11.1 of the company submission (CS, page 151).[1] A total of 43 publications were identified from the electronic searches. Four studies were removed due to duplication. After title and abstract screening, 31 publications were excluded as these were deemed not relevant for the research question. Thus, a total of eight full-text studies were assessed for eligibility which were included in the final evaluation of evidence. The flow of studies through the identification and selection processes is depicted in Figure 5.1.

Eight publications consisting of six studies were included in the review. An overview of the six studies is given in Section 10.1.16 of the CS.[1] The six studies were considered in terms of HRQoL but only one was related to an economic evaluation. This was the study by Forestier-Zhang et al. 2016,[59] where a cost utility simulation of 109 XLH patients (including 24 from the UK) was conducted. The paper examined various scenarios for the maximum willingness to pay threshold based on observed utility values. However, the study was not based on an economic model, considered hypothetical treatment costs, and reported only a mean EQ-5D utility (with the corresponding standard deviation), which could not be used to estimate utilities by health state in the company's model. Therefore, the study was deemed not relevant to the economic evaluation of burosumab.

Figure 5.1: PRISMA diagram for economic systematic literature review



Source: Response to clarification letter, Figure 1.[2]

ERG comment: Quality assessments, like the assessment criteria list from Drummond and Jefferson 1996,[60] for the identified studies were not included in the CS. Nevertheless, the ERG concurs that none of the identified studies are relevant to the economic evaluation of burosumab.

5.3 Exposition of the company's model

5.3.1 Economic evaluation scope

The company submission included a model-based cost-utility analysis comparing the use of burosumab with standard of care to treat patients with XLH. The patient population included in the economic evaluation were XLH patients with radiographic evidence of bone disease aged one year or older with growing skeletons. Subgroups of patients were not considered. Based on growth charts it was determined that in the UK growth is completed at the age of 16 in females and 17 in males. Therefore, treatment with burosumab was assumed to be continued until this age in the model.

Burosumab is a recombinant human IgG1 monoclonal antibody that binds to and inhibits the activity of FGF23. By inhibiting FGF23, burosumab increases tubular reabsorption of phosphate from the kidney and through the production of 1,25(OH)2D enhances intestinal absorption of calcium and phosphate. Burosumab improves phosphate homeostasis and its major pathologic consequences (rickets and osteomalacia), and consequently aims to resolve the skeletal and non-skeletal manifestations of XLH. Standard of care (SoC) treatment is the only comparator considered in the analysis and consists

of systematic oral phosphate supplements and active vitamin D analogues in the form of alfacalcidol A, or oral or injectable calcitriol.

The economic evaluation was conducted from the perspective of the NHS and PSS in England. The model estimates cost and health consequences over a lifetime time horizon for a cohort of patients with XLH aged one to 12 years at the beginning of the simulation. The cycle length of the model is one year. The outcomes of the model are the estimated incremental QALYs, the incremental costs and the incremental cost effectiveness ratio (ICER) associated with burosumab vs. SoC for treating XLH. Cost and health outcomes are discounted at a rate of 1.5%.

ERG comment: The scope of the economic evaluation is generally in line with the scope developed by NICE. Deviations in the company's decision problem were discussed in section 3.3 of this report. The adherence of the scope of the economic evaluation to the NICE reference case was also assessed by the ERG, and it is shown in Table 5.1 below.

Table 5.1: Adherence to the reference case principles relevant to highly specialised technologies

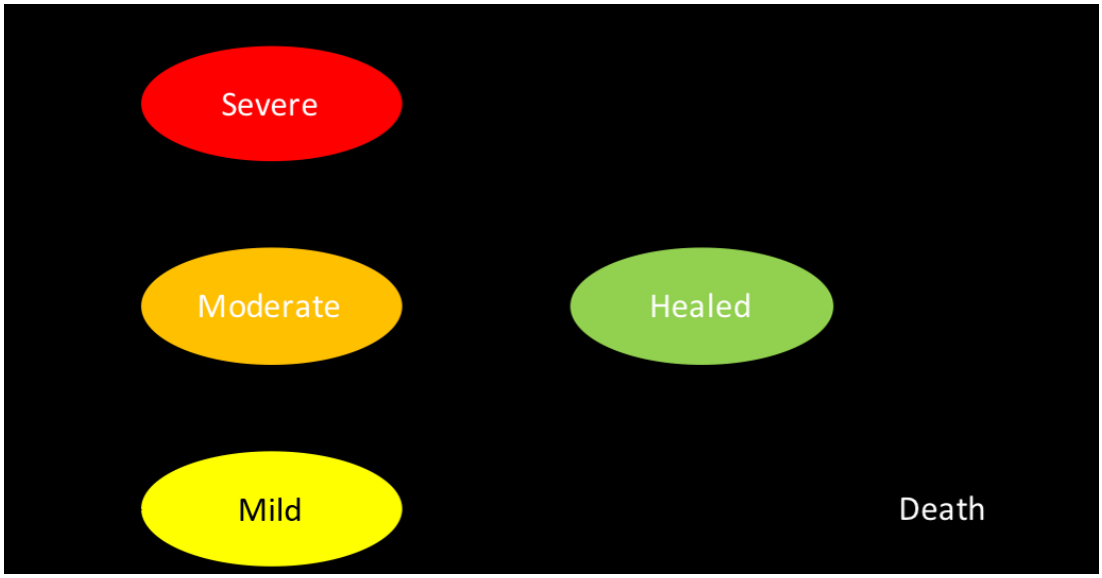
Element of economic analysis	Reference case	ERG comment
Defining the decision problem	The scope developed by NICE	The scope of the economic evaluation is generally in line with the scope developed by NICE. Deviations were discussed in Section 3.3 of this report.
Comparator	Therapies routinely used in the NHS, including technologies regarded as the current best practice	Standard of care (SoC) is the only comparator considered. It is the established clinical management without burosumab (systematic oral phosphate supplements and active vitamin D analogues in the form of alfacalcidol A, or oral or injectable calcitriol).
Perspective on costs	NHS and PSS	NHS perspective was adopted.
Perspective on outcomes	All health effects on individuals.	Patient health benefits were included in the model. Benefits to other afflicted individuals (e.g. caregivers) were not included in the model but discussed qualitatively in the company's submission (CS Chapter 14).
Type of economic evaluation	Cost-effectiveness analysis	Yes.
Time horizon	Sufficient to capture differences in costs and outcomes	Yes.
Synthesis of evidence on outcomes	Based on a systematic review	Meta-analysis was not used, as there is no direct or indirect evidence of the effectiveness of burosumab vs. SoC available. Effectiveness data was

Element of economic analysis	Reference case	ERG comment
		obtained from single-arm studies.
Measure of health effects	QALYs and life years	Health benefits are valued in terms of life years and QALYs gained.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	No, the utility values associated with the model's health states were derived from a vignette study conducted with 6 UK XLH clinical experts. The valuation was based on EQ-5D, which is the NICE standard.
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	
Discount rate	An annual rate of 3.5% on both costs and health effects.	No, costs and outcomes were discounted at 1.5%.
Equity weighting	An additional weighting can be applied for incremental QALYs above 10 years.	No additional equity weighting is applied to QALY gains.

5.3.2 Model structure

An Excel-based Markov model was developed by the company to perform the economic evaluation of burosumab for treating XLH patients in the UK. The model simulates the disease progression of XLH by using the Rickets Severity Score (RSS) as a surrogate for disease severity, which defines the different health states of the model, in patients treated with either burosumab or SoC. The impact of the disease is translated to lifetime costs and QALYs in the submitted cost effectiveness model. The model consists of four (mutually exclusive) health states representing different rickets severity levels (healed, mild, moderate, and severe) and a death state. The severity levels are defined based on the RSS, a radiographic scoring method developed to assess the severity of nutritional rickets. It scores abnormalities in the wrists and knees and is defined on a scale between 0 and 10. Healed rickets correspond to an RSS equal to 0, mild rickets correspond to an RSS between 0.5 and 1.0, moderate rickets correspond to an RSS between 1.5 and 2.0, and severe rickets correspond to an RSS larger or equal than 2.5. Transitions from every alive health state to any other alive health state are allowed in the model. Additionally, patients can move from any of the alive health states to the death state. The relation between the RSS and HRQoL and the choice of cut-offs on the RSS to define meaningful health states was based on a consensus from clinical experts. Figure 5.2 provides the graphical representation of the conceptual model as presented by the company.

Figure 5.2: Model structure as presented by the company.



Source: CS, Figure 24.[1]

It is acknowledged by the company that basing the model structure on the RSS is a limitation of the analysis because:

- Rickets and RSS do not capture all aspects of XLH symptoms and progression. Whilst rickets is the hallmark manifestation of XLH, given the heterogeneity of the condition there is a chance that someone with mild rickets may have more severe additional manifestations.
- RSS is scored independently (not compared to previous x-rays) which may result in inconsistencies in RSS scores between time points that are used to generate transition probabilities.
- The RSS can be complemented by other measures like RGI-C (as in CL201) which provides a comparison to baseline (previous x-rays). RGI-C scores are positive if there is an improvement (+3 if healed, -3 if worsening) compared to baseline. A patient showing no improvement in RSS could experience an improvement or worsening in RGI-C indicating that the patient did or did not benefit from treatment. However, this cannot be captured in the model. However, whilst the RGI-C gives an indication of change in status, it does not indicate the patient status so cannot be used to generate health states.

Despite the limitations mentioned above, the company indicated that the RSS measure provides a reasonable indication of patients' overall XLH health status because:

- Stratifying patients according to these definitions of severity reflected the reduced quality of life of the patient. Thus, the RSS is correlated with HRQoL.
- The model is built in such a way that patients in the healed rickets health state accrue costs for surveillance and drug treatment; patients in the mild rickets health state are assumed to experience additional pain and mobility problems, and associated costs; patients in the moderate and severe health states are assumed to incur orthopaedic intervention costs (in addition to costs from less severe health states). Thus, the RSS is also correlated with costs.
- Rickets severity is the primary endpoint of clinical studies as in CL201.
- In CL201 no patient's rickets worsened according to the definitions of the health states used in the model based on RSS. In addition, it was also observed that no patients' rickets worsened at Week 64 in the study, as all RGI-C scores were positive, as shown in Table 32 of the CS.[1]

Therefore, whilst the RSS is a limited measure, in CL201 it seemed to capture the treatment effect as measured by the RGI-C as well.

Transitions between the alive health states are age dependent for the burosumab arm, where two different transition probability matrices are used depending on whether the patient age is one to four years or five years and older. Transitions between the alive health states for the SoC arm are not age dependent. Only background mortality is included in the model as, according to the company, XLH is not associated with an additional mortality risk according to the available evidence. Thus, age and gender-specific background mortality risks are estimated from UK life tables. The model has a lifetime time horizon and adopted the perspective of the NHS in England. A cycle length of one year (52 weeks) with a half-cycle correction was used. The company used a discount rate of 1.5% per year for costs and effects since, according to the company, on the basis of the evidence presented, the long-term health benefits are likely to be achieved.

ERG comment: The main issues identified by the ERG within the model structure are first summarised in Box 5.1, and these issues are elaborated on afterwards.

Box 5.1: Main issues identified within the model structure in company's economic analysis

1. Appropriateness and comprehensiveness of using RSS to define health states
2. Difference of the effects of burosumab and SoC on patients younger than age five and patients older than age five.
3. Baseline weight, age and disease severity distribution
4. Appropriateness of discount factor
5. Lack of any treatment/disease related adverse events
6. Appropriateness of assuming "full recovery" in the healed rickets health state

1. Appropriateness and comprehensiveness of using RSS to define health states

The clinical rationale behind the definition of the health states in the cost effectiveness model based on the rickets severity was unclear for the ERG (healed rickets (RSS 0), mild rickets (RSS 0.5 and 1.0), moderate rickets (RSS 1.5 and 2.0) and severe rickets (RSS 2.5 or greater)). Since the RSS scale typically extends to 6.5 in a real-world XLH setting (as described on page 41 of the CS), the ERG questioned the appropriateness of allocating a RSS change of 0.5 between the first three states (healed, mild and moderate rickets) while allocating a RSS change of 4.5 (2.0 to 6.5) to the final state (severe rickets). (Question B13 in response to the CL[2]).

In their answer to the request for clarification, the company referred to pages 155-156 of the company submission[1] and the study by Mäkitie et al. 2003,[61] where rickets were graded as normal, normal/mild, mild, mild/moderate, moderate, moderate/severe, or severe rickets. Based on clinical expert opinion, the health states used in the model were simplified to healed, mild, moderate, or severe based on RSS scores. Mäkitie et al. described severe rickets as acrosteolysis, periosteal resorption, severe deformity of long bones, and/or pathological fracture. Patients with these manifestations of X-ray characteristics are most likely to be scored as 2.5 and higher. The company also indicated that resource utilisation and quality of life for patients with RSS equal to 2.5 are not expected to differ significantly compared to patients with higher RSS scores, thus yielding the definition of the severe health state in the model. Healed rickets corresponds to an RSS equal to 0. According to the RSS algorithm described in Table 6 of the CS,[1] RSS scores have intervals of 0.5.[6] Thus, the definition of mild and moderate health states had to cover the interval of RRS 0.5 to 2.0, for which an equal distribution over these health states was

assumed. Hence, the mild health state was assumed to be an RSS of 0.5 or 1 and the moderate health state was assumed to be an RSS of 1.5 or 2. Note that given this allocation, an average RSS of 1.4 would be interpreted as mild rickets, whilst an average RSS of 2.3 would be interpreted as moderate rickets.

Despite the acknowledgement by the company of the limitations of the RSS to define health states, they still assert that RSS is associated with both utility and cost, i.e. if RSS increases then so should cost and utility should decrease in a predictable way. However, as alluded to in Section 3.3.4 above and in some detail in Section 5.3.3.3 below, utilities were estimated from vignettes assuming an association between RSS and clinical characteristics that lack face validity. In particular, it is likely that RSS can improve and indeed rickets appear to be healed, but for there to be residual deformity and increased fracture risk. Since deformity and fracture risk would likely be negatively associated with utility, defining health states only by RSS is likely to overestimate any improvement due to burosumab in moving to states with a lower RSS.

2. Difference of the effects of burosumab and SoC on patients younger than age five and patients older than age five.

The health effects of burosumab are assumed to be age dependent since one set of transition probabilities was used for patients aged one to four years (CL205), whilst another set of transitions was used for patients aged between five and 12 years of age (CL201). In absence of any other source of evidence, the latter transition probabilities were also used for patients between the age of 12 and 17. From age 18 and onwards, it was assumed that patients would remain in their current health state until death occurs. For the SoC arm, the same set of transition probabilities (either the UK chart review or CL002) was used for all ages. The ERG had concerns about the different assumptions made by the company regarding the operationalisation of treatment effects in the model.

When this issue was brought up in the clarification letter (Question B17[2]), the company reiterated that transition probabilities are age dependent for burosumab but according to the ERG this answer lacked a proper justification. It seems that this assumption was made only based on the available data (CL205 for patients aged one to four and CL201 for patients aged five to 12). However, it is still unclear whether the distinction between ages 1-4 and 5-12 is due to different manifestations of the disease in those age groups or due to a different treatment effect of burosumab. If the former is correct, then a different transition probability matrix should have been used for patients 1-4 in the SoC arm as well. It should also be noted that the probabilities derived from CL205 are based on a total 13 patients only, and the probabilities derived from CL201 on a total of 26 patients. Therefore, the ERG considers that assuming such a distinction in effects between these two age groups is at least uncertain.

Transition probabilities for patients aged between five and 12 years were used for patients between the age of 12 and 17. Whilst this might be a good proxy, it is not based on any evidence. The company showed in an alternative scenario that, combining data from both age groups to estimate one set of transition probabilities for burosumab patients to be used for all ages (between one and 17), the ICER was minimally increased. Therefore, using two sets of transition probabilities for burosumab rather than one had a minimal impact on the ICER. This scenario assumed that there is no age dependent treatment effect of burosumab. However, as mentioned above, it is uncertain whether this is the case or not. Thus, a relevant additional scenario, using two different transition matrices for the SoC arm for the two age groups, could have been presented (provided that these two separate matrices could have been estimated). In such scenario, the ERG would not expect a major impact on the ICER, but the uncertainty around the model results (as presented in a PSA) would be increased.

3. Baseline weight, age and disease severity distribution

It was not clear to the ERG what the company's rationale was to select the data sources used to derive baseline weight, age and disease severity level distribution of XLH patients. Demographic parameters should be representative for the patient population expected to be treated in clinical practice, i.e. UK XLH patients. Although data from the UK chart review were available (see section 4.5 of this report), the company did not use this data source to inform the demographic parameters of the model. In the response to the clarification letter (Question B5),[2] the company indicated that due to the nature of the chart review, i.e. a retrospective study including patient histories following diagnosis, this was not considered indicative of the starting age and rickets severity distribution. Thus, combined data from CL201 and CL205 were used as proxy. Furthermore, the company compared the weights of the patients included in the UK chart review to the weights of the UK general population. Figure 2 and 3 in the response to the clarification letter suggested that the weight of XLH patients in the UK chart review was comparable to the weight of the UK general population, especially for males.[2] Females in the UK chart review seem to weigh more than females in the UK general population.

4. Appropriateness of discount factor

The ERG considers that the costs and health effects should have been discounted at a 3.5% rate, rather than at 1.5%. The NICE Technology Appraisal Methods Guide specifies that a rate of 1.5% could be considered by the Appraisal Committee if the achievement of long-term benefits is highly likely.[62] However, it is not specified that a rate of 1.5% should be applied in the base-case analysis.

The ERG considers that it is not clear from the submitted evidence that treatment with burosumab restores patients, who would otherwise die or have a very severely impaired life, to full or near full health. Throughout the CS, it is mentioned that XLH is not associated with additional mortality, and for that reason the model only considers background mortality. Thus, even though the model indicates that patients treated with burosumab will spend most of their lifetime in the healed rickets health state, it is uncertain to what extent this can be seen as full health, as discussed in section 3.3.4. More importantly, as discussed in section 3.3.4 as well, it is also uncertain whether these effects will be maintained lifelong. Therefore, the ERG will apply a 3.5% discount rate in the ERG base-case but will present a scenario analysis with a discount rate of 1.5%.

5. Lack of any treatment/disease related adverse events

Adverse events (AEs) were not included in the base-case analysis on the basis that the AEs observed in the trials are "typical for paediatric population" or frequent manifestations of the disease but not treatment related. In response to the clarification letter (Question B6 – Table 9[2]), the company presented all the treatment-emergent adverse events (TEAEs) occurring in study CL201 and classified them as typical for a paediatric population, frequent manifestation of XLH or related to treatment administration. Only "injection site reactions" were identified as related to treatment administration and were thus included as an AE in the model. AEs classified as manifestations of the disease should be captured by the model. However, the UK chart review and CL002 did not include any safety data and therefore the company did not have any evidence that could be used to model AEs in the SoC arm. Note that the AEs classified as manifestations on the disease are likely to be related to the severity of the disease. Thus, patients in more severe health states (higher RSS) are expected to experience more (or more severe) AEs. As mentioned above, only "injection site reactions" were included in the model as burosumab-related AEs, although not in the base-case analysis. Furthermore, this was considered in terms of costs only, but not in the utility calculations. The company indicated that any disutilities associated to the comparator treatments (active vitamin D and oral phosphate) are expected to be higher than those associated with burosumab (given that many children find them unpalatable). However, this statement was not based on any evidence. Furthermore, since the comparator treatments are given daily,

whereas burosumab is an injection bi-weekly, the company considers it likely that treatment-related disutilities are greater in the comparator than the burosumab arm. Finally, the company mentioned that compared to the costs and health effects currently incorporated in the model, it is likely that the inclusion of adverse events would have relatively modest impact on the model results. While the ERG agrees with the latter statement and acknowledges the challenges of incorporating AEs into the model given the available evidence, it also thinks that not incorporating AEs to the model adds an additional level of uncertainty that should be taken into account when assessing the model results.

6. Appropriateness of assuming "full recovery" in the healed rickets health state and lifelong treatment effects for burosumab

As mentioned above, the ERG considers that defining health states by RSS is likely to overestimate any improvement due to burosumab in moving to states with a lower RSS. In addition, as explained in section 4.6.2, the model currently assumed that the effect of burosumab lasts for the rest of the patients' lives, which seems to be unrealistic. Therefore, in the ERG base-case it was assumed that patients will experience a decline in quality of life 20 years after the end of treatment, which was operationalised by moving to the utility value of the next worse health state (see section 6.2.3 in this report).

5.3.3 Evidence used to inform the company's model parameters

Multiple sources of evidence were used to inform the parameters of the economic model. A summary of the evidence used to inform each group of parameters in the model is presented in Table 5.2.

Table 5.2: Summary of evidence sources used to inform key parameter groups in the company's model

Parameter group	Source of parameter values
Initial patient distribution (age, sex, weight, disease severity)	The distribution of gender and a joint distribution of age and disease severity were based on the baseline patient characteristics in the two clinical studies of burosumab (CL201 and CL205). General population weight data (UK growth charts) were used for the weight distribution.
Transition probabilities between alive states (disease severity states)	Transition probabilities for burosumab were derived from the clinical studies CL201 and CL205. SoC transition probabilities were derived from a UK chart review in the base-case and from the study CL002 in a scenario analysis. More details of these studies are shown in **Error! Reference source not found.**.
Mortality	Mortality rates were obtained from the national life tables for England, for the period 2014 to 2016, as published by the Office of National Statistics. [9]
Health related quality of life	Utility values for the health states were derived from a vignette study conducted by the company.[63] Additionally, age specific multipliers were used based on the general population.[64]
Burosumab treatment costs	The price of burosumab was provided by the company. For monitoring, resource use was based on expert opinion, while unit costs were taken from NHS reference costs.[65]

Parameter group	Source of parameter values
Standard of care treatment cost	Dosing was based on guidelines and the summary of product characteristics. Unit costs were taken from the BNF (source electronic model).[2]
Health state costs (both treatment alternatives)	For the costs of surveillance, resource use was based on expert opinion and unit costs were taken from NHS reference costs.[65] Physiotherapy resource use was based on expert opinion and a Che et al.,[66] and unit costs taken from PSSRU.[67] A number of different sources were used for the orthopaedic intervention costs. Resource use was based on prevalence observed in one of the clinical studies of burosumab (CL201), Che et al.[66] and Skrinar et al.,[68] as well as expert opinion and assumptions. Unit costs were mostly taken from NHS reference costs,[65] apart from unit costs for osteotomy, which were based on the study by Smith.[69]

An overview of the characteristics of the main clinical studies which were used to inform model parameters are listed in **Error! Reference source not found.**. No evidence from an RCT in which urosumab was compared to placebo or other relevant comparator was available. Therefore, data from separate studies were used as evidence to inform treatment effects of burosumab (two phase 2 clinical trials for different age cohorts) and standard of care (two chart review studies). These studies enrolled different populations and differed in duration of follow-up. Mortality was assumed to be the same in both treatment alternatives.

Table 5.3: Overview of studies used to inform parameters of the Markov model

Study identifier	Type of study	Evidence used in model	Number of patients	Observation interval*
CL205	Phase 2 clinical trial	Clinical effects of burosumab in children aged 1-4 y.	13	40 weeks
CL201	Phase 2 clinical trial	Clinical effects of burosumab in children aged 5-12 y.	52	64 weeks
UK chart review	Retrospective chart review	Clinical effects of standard of care.	34	Varying
CL002	Retrospective chart review	Clinical effects of standard of care.	[redacted]	2 years

* Observation interval of data used to inform model parameters
For more detailed information of the patient characteristics see Table 4.13 of this report.

5.3.3.1 Transition probabilities

Transition probabilities for standard care

A chart review study on RSS measurements conducted in the UK was used to inform transition probabilities for the standard of care alternative. Patients in this study were examined at varying time intervals. Two different approaches were employed to deal with the interval censored nature of these data. The first one assumed the last observed RSS value persisted until the next observation (i.e. if RSS=1 at Year 1 and RSS=2 at Year 3, then it was estimated that at Year 2 RSS=1), referred to as last observation carried forward (LOCF). This was used in the company's base-case. The second approach assumed a constant linear change in RSS between two time points (i.e. if RSS=1 at Year 1 and RSS=2 at Year 3, then it was estimated that at Year 2 RSS=1.5). This was included as a scenario analysis.

Observations more than three years apart were excluded from the analyses. The resulting transition probabilities for the SoC arm assuming LOCF and linear change can be seen in Table 5.4 and Table 5.5, respectively.

Table 5.4: Transition probability matrix between alive health states for standard of care treatment in base-case (estimated using last observation carried forward)

	Mild	Moderate	Severe	Healed
Mild	70%	11%	9%	9%
Moderate	18%	69%	10%	4%
Severe	5%	12%	79%	4%
Healed	7%	7%	14%	71%
Source: Table 42 in the CS.[1]				

Table 5.5: Transition probability matrix between alive health states for standard of care treatment in scenario analysis (estimated using linear change assumption)

	Mild	Moderate	Severe	Healed
Mild	51%	21%	16%	12%
Moderate	24%	52%	17%	7%
Severe	7%	19%	68%	6%
Healed	20%	20%	40%	20%
Source: Table 44 in the CS.[1]				

In a scenario analysis, the company derived transition probabilities for the SoC arm from the CL002 study.[1] This study acted as a comparison cohort for the burosumab treated population in study CL201 (thus, for patients aged five years or older).[37] During clarification, the company corrected a methodological error made when estimating this transition matrix. Therefore, the probabilities shown in Table 5.6 were obtained from the electronic model submitted by the company with the response to the clarification letter.[2]

Table 5.6: Transition probability matrix between alive health states for standard of care treatment in scenario analysis (based on CL002 data)

	Mild	Moderate	Severe	Healed
Mild	78%	7%	4%	11%
Moderate	22%	75%	4%	0%
Severe	0%	63%	37%	0%
Healed	29%	29%	0%	41%
Source: Electronic model (after clarification).[2]				

The company chose the UK chart review to derive transition probabilities for the base-case for two main reasons: it provided a better representation of the UK patient population and treatment practices (since CL002 was conducted in the US), and it was based on a longer follow-up with (on average) more observations per patient.

Transition probabilities for burosumab

Transition probabilities for the burosumab arm were estimated from two phase 2 clinical trials, one enrolling patients aged one to four years (CL201), and one enrolling patients aged five to 12 years

(CL205).[70] Since the company assumed that the treatment effect of burosumab on RSS was not the same in both trials, in the model each trial result was applied to those patients that better matched the trial population. Thus, the company assumed in the model that all patients under five would achieve the treatment effects as observed in CL205, and all patients aged five years and over would achieve the treatment effects as observed in CL201. The same methodological error mentioned above for CL002 in the SoC arm, was also corrected by the company for these transition matrices. Therefore, the probabilities shown in Table 5.7 and Table 5.8 were also obtained from the electronic model submitted by the company with the response to the clarification letter.[2]

Table 5.7: Transition probability matrix between alive states for burosumab treatment in patients aged 1one to four years

	Mild	Moderate	Severe	Healed
Mild	100%	0%	0%	0%
Moderate	59%	41%	0%	0%
Severe	50%	50%	0%	0%
Healed	0%	0%	0%	100%
Source: Electronic model (after clarification).[2]				

Table 5.8: Transition probability matrix between alive states for burosumab treatment in patients aged five years and older

	Mild	Moderate	Severe	Healed
Mild	57%	0%	0%	43%
Moderate	37%	52%	0%	12%
Severe	53%	25%	14%	8%
Healed	0%	0%	0%	100%
Source: Electronic model (after clarification).[2]				

ERG comment: The ERG does not agree with the methodology used by the company to estimate the transition probability matrices presented above. The data sources used to inform transition probabilities in the model have different observation periods (40 weeks in CL205, 64 weeks in CL201 and 104 weeks in CL002). Since the model assumed a cycle length of one year, the problem at hand was to estimate the three corresponding transition probability matrices for a different time scale (52 weeks). This was done by the company following the four steps below, as indicated on page 163 in the CS: [1]

1. Generate 40-week, 64-week, two-year and three-year transition probability matrix (based on the observe data).
2. Convert the probabilities to rates and annualise, using the formula *rate = – ln(1 – probability) / time*
3. Convert the annualised rates back to transition probabilities, using the formula *probability = 1 – exp(– annualised rate)*
4. Proportionally adjust the probabilities such that each row of the transition probability matrix equates to one.

In the recent review paper by Olariu et al. 2017,[71] the approach used by the company is summarised as well as the problem that may arise from using that approach. Thus, in order to change the time scale of a probability, the company first converted it into a rate using the formula indicated in step two above, and then calculated the re-scaled probability using the formula in step three. This is a (correct) well-known approach.[72-74] However, when a model has more than two health states, as it occurs with the

company's model, the formulae above introduces bias because these ignore competing risk between the health states of the model. This bias can have significant impact on the model results and therefore it should not be ignored.[75] A correct way to overcome this potential issue requires taking a certain root of the transition probability matrix. This method is not new, as it was described (at least) in the paper by Craig and Sendi in 2002.[76] Taking the root of a matrix is not always possible. As an alternative, Chhatwal et al. 2016 developed an algorithm to approximate such a matrix.[75] Another alternative approach would consist of choosing shorter cycle lengths in the model. That way the probability of multiple events occurring during one cycle would be reduced, thus minimising the bias.[77]

The issue described above was raised by the ERG in the clarification letter (Question B16) where the paper by Chhatwal et al. was indicated as reference.[2] However, the company did not attempt to re-estimate the transition probability matrices as suggested by the ERG. Instead of that, the company performed an exercise to quantify how large the impact of using the incorrect transition probabilities would be.

In the response to Question B16, the company made a few statements that the ERG would like to discuss. The company indicated that Chhatwal et al. "*presented an alternative approach based on finding the root of a transition probability matrix using eigendecomposition, or where that fails, a numerical approximation method*".[2] The ERG would like to emphasise that the "alternative" method of finding the root of a transition matrix is not new in the field of health economics since there is published literature on this method dating back to at least 2002.[76] The numerical approximation method seems to be indeed new. According to the company, the "*proposed methods require complex computational approaches in software such as MATLAB or Mathematica, neither of which are commonly used in economic evaluations*".[2] The ERG does not agree with this quote. Calculating the root of a matrix does not require the software mentioned by the company. In fact, the ERG has used R (as shown in Appendix 2 of this report), which is accepted by NICE. While it is true that the algorithm by Chhatwal et al. was developed in MATLAB/Mathematica, this does not mean that it cannot be translated into other language like R or VBA. In any case, "translation" was not needed because their algorithm is available online and could have been used by the company following the instructions in the link below:

http://www.mgh-ita.org/ita-tools/online-modeling-tools.html

Furthermore, the company indicated that "*despite this article being published in July 2016, no NICE appraisals have required application of this more advanced technique, rather than the commonly used method as used for the burosumab model*".[2] The ERG would like to emphasise again that this method is not new in the field of health economics. Given that it was published at least in 2002, it seems unlikely that this approach was not considered in previous NICE appraisals, although, given the time constraints, the ERG could not check this point. However, even if that would be the case and this technique was not used before in NICE appraisals, the ERG considers the company's argument still invalid since errors should be corrected at the time they are discovered independently of what has happened in the past.

Finally, the company concluded their response to Question B16 by stating that "*the approach to derive one-year transition probabilities from the trial observations seems to be valid and a multi-variate version is not required*".[2] The ERG does not agree with this statement. The company's approach is still invalid and a correct methodology, as explained above, is required. What the company has shown is that the impact of using the incorrect transition probability matrices in the model results is expected to be minor/moderate. This might be the case since the transition probability matrices are applied in the model for a relative small number of cycles.

The ERG preferred transition probability matrices are presented in section 6.2.1 of this report. The derivation and a detailed explanation of the methods used to derive these matrices can be found in Appendix 2 of this report.

5.3.3.2 Mortality

Since there is no evidence suggesting that XLH might reduce life expectancy, only age and gender specific background mortality was included in the model. Mortality rates were obtained from the national life tables for England, for the period 2014 to 2016, as published by the Office of National Statistics.[9]

ERG comment: In the clarification letter, the ERG asked the company (Question B12) about the plausibility that patients with more severe clinical manifestations of the disease were likely to have a significant reduction in life expectancy compared to an "average" UK patient. The company did not consider this implausible given the increased risk of fractures with XLH and the association between hip fractures and mortality in older healthy adults. Nevertheless, the company emphasised that there are no published articles providing evidence of this, justifying thus the assumption in the model that there is no excess mortality risk associated with XLH. In any case, the company explored an additional cost effectiveness scenario where patients in the severe health state of the model had twice the risk of dying from age 50 years and older. In that scenario, the ICER was reduced by 1% compared to the company's base-case ICER.

5.3.3.3 Health-related quality of life

The clinical trials identified by the company did not include health-related quality of life (HRQoL) measures that could be used in the economic analyses. Two studies conducted by the company included SF-36 data, but these studies did not rely on RSS (or other measures of severity). Therefore, these data could not be used to inform the company's model.[78, 79] Furthermore, as mentioned in section 5.2.2, the company did not identify any HRQoL results in the literature that could be used in the model. Thus, the company conducted a vignette study to elicit utility estimates for the health states defined in the cost effectiveness model (e.g. based on RSS). A proxy valuation of the health states with UK clinical experts was undertaken, where the experts were asked to imagine a patient as described by the vignette and to rate the impact of the health state on HRQoL by filling out the EQ-5D-5L.

Case histories (vignettes) were defined in terms of RSS and age and were created based on qualitative published studies and a series of five interviews with clinical experts. In total, 12 case histories were developed, based on four severities of rickets as defined by RSS in line with the cost effectiveness model (healed, mild, moderate and severe) and three different age categories (one to four years old, five to 12 years old and 13 years and older). The health states were validated and valuated in a series of interviews with six UK clinical experts. However, two experts did not assess the severe health state because they had no experience with patients in that condition. For each case history, the experts were asked to value the impact of the disease on different aspects of HRQoL using EQ-5D-5L. Then, the mapping algorithm developed by Van Hout et al., 2012 was used to generate EQ-5D-3L utilities.[80] Full details of the study are available in a report.[63]

The derived utilities can be seen in Table 5.9. Utility scores ranged from 0.462 (severe rickets in patients 13 years and older) to 0.969 (patients five to 12 years old). The company assumed that the utilities derived for adolescents aged 13 and over were also be applicable to adults. Moreover, it was assumed that since XLH is not associated with additional mortality, the utilities were used over the patients' lifetime, using an age decline as in the general population.[64]

Table 5.9: Utility values used in the cost effectiveness model

Health state	Utility value	Standard deviation	Source
Age 1-4			
Healed rickets	0.872	0.097*	Vignette study[63]
Mild rickets	0.774	0.094**	
Moderate rickets	0.685	0.175	
Severe rickets	0.545	0.065***	
Age 5-12			
Healed rickets	0.969	0.072*	Vignette study[63]
Mild rickets	0.757	0.119**	
Moderate rickets	0.613	0.170	
Severe rickets	0.521	0.084***	
Age 13 and over			
Healed rickets	0.862	0.105*	Vignette study[63]
Mild rickets	0.671	0.110**	
Moderate rickets	0.575	0.094	
Severe rickets	0.462	0.161***	
Utility multipliers			
Age 18-24	1.000	-	Age-decline based on the general population[64]
Age 25-34	0.992	-	
Age 35-44	0.966	-	
Age 45-54	0.930	-	
Age 55-64	0.888	-	
Age 65-74	0.851	-	
Age 75+	0.781	-	

Source: Table 31 in the CS.[1]
*This is the standard deviation around the difference between the healed and mild states. The standard error should be used in the model.
**This is the standard deviation around the difference between the mild and moderate states. The standard error should be used in the model.
***This is the standard deviation around the difference between the moderate and severe states. The standard error should be used in the model.

Given the small sample of clinical experts that valued the health states, there is significant variation around the mean values. When considering how to account for this uncertainty in probabilistic and deterministic sensitivity analysis, the company considered that using the mean and standard deviations directly would lead to implausible simulations since 'better' health states could have lower utilities than 'worse' health states. To ensure the variation was accounted for whilst generating plausible simulated utilities, the moderate health state was used as an anchor and the values for other health states were calculated based on differences to the moderate state. The moderate health state was chosen since not all clinical experts valued the healed and severe health states.

ERG comment: The ERG agrees with the company that "*the method used here to develop states and capture utilities is not the optimal source of evidence*".[1] It is a limitation that utility values were obtained

from clinical experts and not directly from XLH patients or, given that the condition affects very young children, from the parents of the patients. The latter would have been considered a more appropriate proxy for assessing HRQoL by the ERG. According to the company, "*to validate the utilities derived from the clinical experts, an ongoing study will report findings from a survey of parents of children affected by XLH. Results of this subsequent study will be reported during the NICE appraisal of burosumab and will be made available to the committee at the earliest convenience*".[1] Unfortunately, the results of this study were not available at the time this report was finished.

The utility values that the company presents in Table 31 of the CS (Table 5.9) do not match all the utility values as presented in the report about the vignette study by Lloyd et al. 2018.[63] For each age group, the value for 'healed rickets' is higher in the CS than in Lloyd et al. whereas the value for 'severe rickets' is lower in the CS than in the Lloyd et al. report. No explanation for this discrepancy was provided. In addition, it is not clear to the ERG how the standard deviations were derived that are presented in Table 5.9 for the non-moderate health states. For the three moderate health states it is unclear whether these values represent the SDs as observed from the vignette study, or the standard errors (SEs), representing the uncertainty of the mean estimate. In the electronic model, these values have been used as if they represent SEs.

At this moment, it is not possible to validate all the utility values for children reported by the company. However, the utility scores for the 'healed rickets' state can be compared to the average utility scores of the general publication. The utility values used in the model are 0.872, 0.969, and 0.862 for the 0-4, 4-13, and 13+ age-groups, respectively. In the study report by Lloyd et al. these values are substantially lower at 0.800, 0.89, and 0.811. The UK average for adults from 18-25 years, the youngest group for which a population average is available, is 0.922.[64] Thus, it appears that the utility value for 'healed rickets' in the group from four to 13 years old as used in the model is rather high, though not impossible given that the population norm is based on young adults rather than children.

However, given the rather high utility values presented in the CS compared to the report by Lloyd et al. and the lack of an explanation for the discrepancy between the two sets of utilities, the ERG considered the Lloyd-set for the ERG preferred base-case and conducted an exploratory sensitivity analysis to assess the impact of using the utility-set presented in the CS. These results will be reported in Chapter 6.

In their response to the CL (question B7), the company provided the descriptions of the vignettes that were used in the study by Lloyd et al. Per age-group, four vignettes were defined, one for each health state. The descriptions provided are strictly ordered, in that on each attribute of the vignette an equal or worse description will be given for a worst health state. For example, for the 'healed rickets' state the vignette defines five of the attributes as follows*: Patient is able to walk nearly normally for their age. They may have a slightly non-normal gait and residual bowed legs; Patient is able to complete usual activities such as dressing and playing; Patient does not experience pain associated with their XLH; Patient's mood, anxiety or sadness varies in the same way that an otherwise healthy person's would be expected to; Patient can complete school, work and many usual activities normally and doesn't have undue problems with completing tasks*.

The text for the same five attributes for 'mild rickets' reads: *Patient is able to walk nearly normally for their age. They have a slight waddling gait with some muscle weakness. They have bowed legs; Patient is able to complete usual activities such as dressing and playing. They fall over more often than other children their age; Patient does experience pain associated with their XLH, particularly in their limbs. They may need pain medication at times; Patient may be withdrawn at times and experience feelings of sadness, frustration and they may lack confidence. They may dislike the need for hospital visits. They*

may suffer teasing or bullying at school; Patient can complete school, work and many usual activities normally and doesn't have undue problems with completing tasks. They often experience quite severe tiredness or stiffness after taking part in sports.

By using experts to devise the descriptions in this very clearly ordered way, there is no possibility of improvement in one attribute with no change or even worsening of another. In contrast, more variation may be expected when patients or parents fill out an EQ-5D, as some patients with mild rickets will report e.g. moderate pain and no anxiety or depression, whereas others may report no pain and moderate anxiety or depression, thus leading to more variation in utility within one health state. Indeed, some patients with healed rickets might have considerable residual deformity, particularly if they had originally been in the severe state and still have some risk of fracture.

Treatment related adverse events were not included in the model. Whilst it is difficult to separate out some of the reported adverse events from frequent manifestations of the disease or typical for a paediatric population, this is not true for injection site reactions, erythema and swelling that were reported in [redacted], [redacted] and [redacted] of the patients, respectively. However, as indicated in section 4.2.4.2 of this report, since all injection site reactions associated with burosumab were categorised as mild in severity, the ERG agrees with the company that these are not expected to have a significant impact on the model results.

5.3.3.4 Resource use and costs included in the model

This section summarises resource use and costs presented in the CS. No studies were identified that reported resource use information. Clinical experts (Dr William G Van't Hoff and Dr Jeremy Allgrove) provided the frequencies and costs (surveillance, drugs, pain and mobility, and orthopaedic interventions) used in the CS. There is no specific healthcare resource group (HRG) or payment by results (PbR) code for XLH.

Technology and comparator costs

In the CS, it was assumed that in the first year of treatment, patients commence treatment on a recommended starting dose of 0.4 mg/kg with a stepwise increase up to 0.8 mg/kg over three months. Estimation of the treatment costs in the CS comprises a mean dose of 0.6 mg/kg for the first three months and a mean dose of 0.8 mg/kg in the subsequent nine months. The first-year dose is therefore estimated to be 0.752 mg/kg, which equates to 94% of the maintenance treatment dose. The company indicated that this assumption was in accordance with the SPC. In the CS, it was stated that the SPC recommends dose rounding to the nearest 10 mg. A scenario analysis was conducted by the company to explore the impact of rounding the dose up to the next 10-fold, rather than to the nearest as recommended in the SPC. The annual per patient cost was estimated (cost per vial) and listed in Table 5.10. Table 5.11 lists summary of acquisition costs by age and weight.

Table 5.10: Dosage and cost of burosumab

	Vial size	Cost per vial	Dose per infusion (mg per kg)
Burosumab	10 mg	£2,992	0.752mg/kg in the first 12 months of therapy, then the full dose of 0.8mg/kg
	20 mg	£5,984	
	30 mg	£8,976	
Source: CS, Table 48.[1]			

Table 5.11: Summary of acquisition treatment costs by age/weight

Age (years)	Weight (kg)	Dose (mg)	Rounded dose (mg)	Vials (10mg)	Vials (20mg)	Vials (30mg)	Annual cost
1	9.4	7.5	10	1	0	0	£77,792
2	11.8	9.4	10	1	0	0	£77,792
3	14.1	11.3	10	1	0	0	£77,792
4	16.1	12.9	10	1	0	0	£77,792
5	18.5	14.8	10	1	0	0	£77,792
6	20.7	16.5	20	0	1	0	£155,584
7	23	18.4	20	0	1	0	£155,584
8	25.9	20.7	20	0	1	0	£155,584
9	28.7	23	20	0	1	0	£155,584
10	31.8	25.4	30	0	0	1	£233,376
11	35.5	28.4	30	0	0	1	£233,376
12	39.1	31.3	30	0	0	1	£233,376
13	44	35.2	40	1	0	1	£311,168
14	49.6	39.7	40	1	0	1	£311,168
15	54.2	43.4	40	1	0	1	£311,168
16	58.2	46.6	50	0	1	1	£388,960
17	60.7	48.6	50	0	1	1	£388,960

Source: CS, Table 49.[1]

The list price of burosumab is included in the CS. [redacted]

Monitoring costs

In the CS, monitoring costs account for dose adjustments in the first year of treatment with burosumab. After initiation of treatment with burosumab, in the first month of treatment fasting serum phosphate is monitored, followed by every four weeks for the subsequent two months and thereafter as appropriate. It was indicated in the CS that if fasting serum phosphate is within the reference range for age, the same dose was maintained. In the CS, patients were assumed to require five additional blood tests and 15-minute consultations in the first year, with nurses taking blood tests to support dose titrations over the course of three months. The total monitoring cost per patient was assumed to be £126.55 (including nurse visits costs (five times for 15 minutes) of £111.25 and blood tests costs of £15.30.[65, 67]

Acquisition costs of the comparator

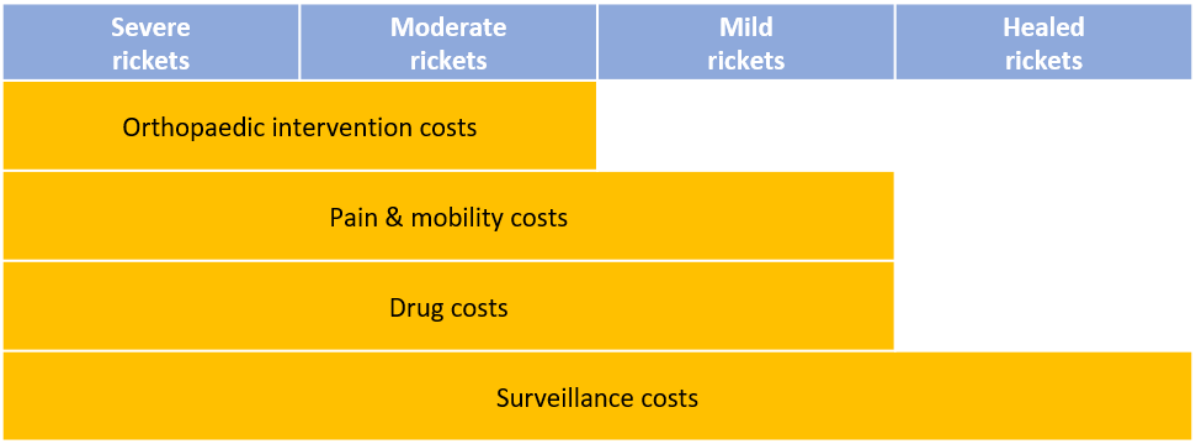
In the CS, alfacalcidol was dosed based on weight. A mean dose of 40 nanogram/kg/day was used, based on clinical expert opinion which indicates that the usual dose of alfacalcidol is 30-50 nanogram/kg/day. This is almost double the recommended dose for another vitamin D analogue, calcitriol, due to the difference in half-life between the two formulations.[3] The company indicated that the computational complexity of modelling treatment costs by age and the relatively low costs of the

comparator,[2] the mean cost of treatment across one to 17 year olds was used to estimate the average annual cost of alfacalcidol. For oral phosphate, it was assumed to be one tablet four times per day.[3]

Health state costs

In the CS, follow-up costs were categorised in four groups as shown in Figure 5.3: surveillance, pain and mobility, orthopaedic intervention and drugs (adults only). According to the CS, only patients in the moderate or severe health state are assumed to receive orthopaedic treatment. It was also assumed that patients in the mild, moderate or severe health states receive pain and mobility costs (physiotherapy). The company assumed that all patients receive the same surveillance costs regardless of health status. Only patients that have had rickets in childhood are assumed in the CS to receive the cost of vitamin D analogues and phosphate supplements in adulthood. Unit costs and resource use for all health state costs are detailed in Table 5.12.

Figure 5.3: Costs categorised by health state



Source: CS, Figure 26.[1]

Surveillance costs

In the CS, surveillance costs were assumed to be the same for all health states and in both treatment arms. Therefore, surveillance costs do not have any impact on the base-case results. In the CS, a scenario analysis was conducted in which patients who are healed at the end of childhood do not require ongoing clinical reviews in adulthood. Clinical experts could not estimate how often SoC patients would be seen in the healed health state. The details of surveillance costs are listed in Table 5.12. In the CS, surveillance costs comprise:

1. Laboratory monitoring costs, which include costs required to test serum calcium, phosphorus, potassium, and creatinine levels, ALP, PTH and urine calcium and creatinine levels.
2. A specialist consultation, which includes the costs for outpatient visits for specialist reviews.
3. Radiography, considered as the gold standard for the diagnostic and efficacy of rickets.
4. During renal ultrasonography patients are screened for signs of nephrocalcinosis, a clinical indicator for worsening XLH severity.
5. At risk of dental problems, dental outpatient appointments were assumed once every 2 years for dental examinations or minor interventions.

Drug costs

In the CS, the estimate of the costs of phosphate supplements and vitamin D analogue was based on two published studies.[66, 68] Per its SPC, the vitamin D analogue dosage was assumed to be five tablets

per day for vitamin D resistant rickets. Based on expert opinion for calcitriol a dosage of 1.125 micrograms per day was assumed.

Pain and mobility costs

In the CS, it was assumed that patients will usually use over-the-counter painkillers for pain management which would therefore not be relevant to the NHS and PSS perspective. GP visits were also excluded, as these could not be linked to specific symptoms of XLH. Thus, pain and mobility costs only consisted of physiotherapy (5% based on clinical expert opinion). It was assumed that children would receive one session (one hour) of physiotherapy per month.

Orthopaedic intervention costs

In the CS, resource use from dental abnormalities were approximated from the proportion of patients with a medical history of tooth abscess in the CL201 study.[70] The costs of the procedures were obtained from an average of dental procedures and weighted by a number of major/intermediate/minor procedures (see Table 5.12). In the CS, patients who have osteotomy procedures are assumed to require two interventions during childhood, which is applied by the company assuming that the costs occur every eight years during childhood. The same assumption was made regarding stapling of growth plates.

In the CS, it was assumed that if patients require a hip arthroplasty, the costs apply to adults only, so the cost of a hip arthroplasty was divided by 60 years to estimate an annual cost. The same calculation was used for knee arthroplasty.

Table 5.12: Summary of cost input parameters included in the model

	Age group	% of patient	Unit cost	Resource use per year	Total cost	Unit Cost Source	Resource Use Source
Surveillance costs							
Specialist Consultation	Children	100%	£249.31	4	£997.22	NHS reference costs 2016/17.[65] Using an average of consultant-led (WF01A) paediatric endocrinology (service code 252) and nephrology (service code 259) as patients are managed by both.	Clinical expert opinion
	Adults	100%	£102.33	1	£102.33	NHS reference costs 2016/17.[65] Using an average of consultant-led (WF01A) endocrinology (service code 302) and nephrology (service code 361) as patients are managed by both.	Assumption
Laboratory Monitoring	Children	100%	£4.19	4	£16.76	NHS reference costs 2016/17.[65] DAPS05 (Haematology) and DAPSS04 (Clinical biochemistry).	Clinical expert opinion
	Adults	100%	£4.19	1	£4.19		
Radiography	All	100%	£29.78	0.50	£14.89	NHS reference costs 2016/17.[65] DAPF (Direct Access Plain Film).	Clinical expert opinion
Renal Ultrasono-graphy	All	100%	£51.36	1	£51.36	NHS reference costs 2016/17.[65] IMAGDA RD40Z (Direct access ultrasound scan with duration of less than 20 minutes, without contrast).	Clinical expert opinion
Dental Check up	Children	100%	£125.39	0.50	£62.70	NHS reference costs 2016/17.[65] Outpatient attendance 142 (Paediatric dentistry).	Clinical expert opinion
	Adults	100%	£126.26	0.50	£63.13	NHS reference costs 2016/17.[65] Outpatient attendance 144 (Maxillo-facial surgery).	Clinical expert opinion
Drug costs							
Oral Phosphate	Adults	65%	£0.16 per tablet	5 tablets per day	£193.70	Cost from BNF 20th December 2017: Phosphate Sandoz effervescent tablets (100). Source electronic model.[2]	The summary of product characteristics recommends 4-6 tablets per day (using 5 average) for vitamin D resistant rickets; Che

	Age group	% of patient	Unit cost	Resource use per year	Total cost	Unit Cost Source	Resource Use Source
							et al. indicated 64.6% of adult patients receive phosphate supplements.[66]
Alfacalcidol	Adults	59%	£0.31 per 500ng capsule	Dose of 1,125 ng per day	£200.31	Cost from BNF 16th January 2018: Alfacalcitrol 500nanogram capsules (30). Source electronic model.[2]	Guidelines by Carpenter et al recommend a dose of 0.5-0.75 mcg per day for Calcitriol (another Vit D not used in UK),[22] but KOL opinion indicates that double dose is required for alfacalcidol, so a mean of 1.125 mg is used. Che et al. indicated 59.2% of adults receive a vitamin D.[66]
Pain and mobility costs							
Physiotherapy	Children	5.00%	£87 per session	1 session per month	£52.20	Cost from PSSRU 2016 (6.1).[67]	Clinical expert opinion indicated that 5% patients may request physiotherapy. Assuming one session per month.
	Adults	57.40%	£45 per hour	1 hour per month	£309.96	Cost from PSSRU 2016 (section 13). Assuming Physiotherapist specialist which is a band 8.[67]	Resource use from Che et al. [66] Assuming one hourly session per month.
Orthopaedic intervention costs							
Dental Abnormalities	Children	19.20%	£154.60	1	£29.68	NHS reference costs 2016/17.[65] Average of dental procedures in 18 years and under, weighted by the number of major/intermediate/minor procedures on the NHS (CD01B, CD02B, CD03B).	Resource use is approximated from the proportion of children with a medical history of tooth abscess in CL201 clinical study report. We assume one procedure per year.
	Adults	62.50%	£169.52	1	£271.24	NHS reference costs 2016/17.[65] Average of adult dental procedures, weighted by the number of major/intermediate/minor procedures on the NHS (CD01A, CD02A, CD03A).	The proportion of adults with dental abnormalities is sourced from Che et al.[66] The company assumed one procedure per year.
Osteotomy	Children	7.7%	£4072.99	Twice in childhood	£39.20	Smith et al.[69]	Resource use is approximated from the proportion of patients with a medical history of osteotomy in CL201 clinical study report. We assume patients have two osteotomy

	Age group	% of patient	Unit cost	Resource use per year	Total cost	Unit Cost Source	Resource Use Source
							procedures during childhood which is applied by assuming the cost occurs every 8 years as a child.
Stapling of Growth Plates	Children	17.5%	£171	Twice in childhood	£3.74	NHS reference costs 2016/17.[65] HN24E Trauma & Orthopaedics (Intermediate Knee Procedures for Non-Trauma, between 6 and 18 years, with CC Score 0).	Resource use from clinical expert opinion. In the CS, patients' growth plates are stapled twice during childhood which is applied by assuming the cost occurs every 8 years as a child.
Hip Arthroplasty	Adult	8%	£5823.53	0.017%	£7.76	Unit cost from NHS reference costs 2015-16 using the most frequent major hip procedure code (HN12F: Very Major Hip Procedures for Non-Trauma with CC Score 0-1).[65]	Resource use from Skrinar et al.[68] Once per lifetime (60 years, adulthood at approximately 20 and life expectancy approximately 80).
Knee Arthroplasty	Adult	12%	£5691.76	0.017%	£11.38	Unit cost from NHS reference costs 2015-16 using the most frequent major knee procedure code (HN22E: Very Major Knee Procedures for Non-Trauma with CC Score 0-1).[65]	Resource use from Skrinar et al.[68] Once per lifetime (60 years, adulthood at approximately 20 and life expectancy approximately 80).
Source: CS, Table 52.[1]							

Adverse event costs
No costs associated with AEs were used in the base-case analysis. In the sensitivity analysis, the impact of including costs associated with AEs (lower limit £0 and upper limit £5) were explored, using an incidence rate of 28.2% for injection site reactions based on Study CL201 and Study CL205.

ERG comment: The company indicated that all known costs and resources have been considered. The ERG requested clarification of the orthopaedic intervention costs which are only considered to occur in patients with a rickets score of 1.5 or higher, but no evidence was provided for the relevant cut-off. In the CL, it was indicated that orthopaedic interventions are only required in patients that have a need for such intervention, who are mostly likely to have more severe rickets. The assumption (confirmed by clinical experts) states that if a patient has healed or mild rickets, then it is unlikely that they would require orthopaedic interventions. The ERG also indicated that the monitoring costs are applied only in the first year of treatment (for dose adjustments). Patients up to the age of 17 are expected to see a specialist every three months, regardless of whether they receive SoC or burosumab. This is incorporated into the surveillance costs which are incurred by all patients. These consultations with clinical specialists are to monitor the disease and treatment. The company indicated that after the first three months, burosumab is not expected to require any additional monitoring over that already conducted with SoC. The ERG indicated that treatment costs of the comparator are not age specific, but an average treatment cost for all patients age one to 17 is used in the model. Given that the comparator consists of two treatments, only one of which has a cost that is age-related (alfacalcidol) and the cost of alfacalcidol is not a driver of costs, the simplification of an average cost (instead of age specific) is acceptable. These revisions have been included in the revised base-case. The revised model sent after the clarification phase comprised updated costs that reflect the same year (2016/17). Overall, the applied changes did not have an impact on the results. Surveillance costs are applicable to all patients and orthopaedic intervention costs are not drivers of the results.

In addition, the ERG had two priority questions in the CL about dosing and vial sharing of burosumab. The company indicated that vial sharing is not applied to burosumab. According to the company, if patients received their exact dose as per their weight, which could be a proxy scenario for vial sharing, the ICER would become [redacted]. Based on the SPC, if a patients' weight indicates a dose of 7.5 mg, then this will be rounded up to 10 mg. It was further stated that when patients are five years old, the calculated dose is 14.8 mg but the recommended dose to be administered is 10 mg. The recommended starting dose regimen in children, according to the CS, is based on experience in Study CL201 and Study CL205. Rounding to the nearest 10 mg was used during dose titration in Study CL201. The company indicated that when pharmacokinetic (PK) modelled dose levels were rounded to the nearest 10 mg a difference in dose of <5 mg is not expected to affect response. The maximum dose of 90 mg is recommended based on PK simulations and the practical limitation of a tolerable injection volume. It was stated that this information was presented to the EMA.

5.3.3.5 Demographic parameters included in the model

A number of demographic characteristics were considered in the model as input parameters. These included the initial distribution of patients per health state stratified by age (see Table 35 and Table 36 in the CS[1]) and the percentage of males (50.77%) at baseline. These parameters were obtained by combining the data from CL201 (all doses) and CL205. Weight by age and

gender was also included in the model as a parameter. The median weight of the general population (for each age and gender category) was assumed,[81, 82] as shown in Table 34 in the CS.[1]

ERG comment: It was not clear to the ERG what the company's rationale was to select the data sources used to derive baseline weight, age and disease severity level distribution of XLH patients. This was discussed in Box 5.1.

5.3.4 Model evaluation

The company presented the results of the health economic analyses in terms of incremental costs and incremental QALYs (combined as an ICER) for burosumab compared to standard of care. Results were obtained by performing a cohort simulation for each starting age (one to 12 years) in each treatment alternative, using the Markov model described in section 5.3.2 of this report. The results for each treatment alternative were then obtained by taking the weighted average of all the cohort simulations for that treatment alternative, using the age distribution of the treatment population. The company submission also included the results of deterministic and probabilistic sensitivity analyses (denoted by DSA and PSA, respectively), the latter consisting of 5,000 model iterations. An overview of the parameters included in the economic model is given in Table 5.13. Other parameters, like mortality or discount were not included in the sensitivity analyses. The results of a number of deterministic one-way and scenario analyses were also presented in the company submission. These are summarised in Box 5.2.

ERG comment: The company, in its response to the clarification letter, submitted an updated electronic model. The following changes were implemented to the original model in the updated version:

- For the transition probability matrices, a programming error in the original model was corrected (transition probabilities from study CL201, which has an observational interval of 64 weeks, were converted to annual probabilities as if they had an observational interval of 40 weeks). In response to Question B16 of the clarification letter, the company applied a revised method for changing the cycle length from the 40 or 64 weeks as observed in the clinical studies to the one year used in the model. As discussed in section 5.3.3.1, the company used an incorrect method to adjust cycle length, which introduced an error (by adjusting individual transition probabilities the rows of the transition matrices did not add up to one). In the original model, the error (i.e. the difference between the sum of each row of transition probabilities and 1) was resolved by dividing each element on a row by the sum of that row. In that way the error was proportionally spread over all elements. In the updated model, the error was added in full to the element on the row representing the probability of remaining in the same health state. The ERG is of the opinion that the original solution for dealing with the error introduced by the invalid method is preferred to the solution used in the updated model, because the error that is introduced is spread over multiple transition probabilities rather than just one, thereby minimizing the effects of the error. This issue has been addressed by the use of the ERG preferred transition probability matrices as discussed in section 5.3.3.1 and presented in section 6.2.1.
- An additional scenario analysis was explored, where the transition probabilities between health states for all ages was based on pooled data from both clinical studies on burosumab (CL201 and CL205).

- The adding of a factor 0.05 to the cumulative Gamma functions in the probabilistic sensitivity analysis was removed from the transition probability matrix based on the UK chart review (see section 5.4.2.3).
- Unit costs have been updated so that all costs are from 2016/17 costs/tariffs.

Table 5.13: Summary of the input parameters included in the economic model

Parameter	Mean value	Range / Distribution	Source
Baseline age and severity distribution	Table 36 in CS	Dirichlet distribution using observed values in CS Table 35.	Pooled baseline distribution from CL201 (all doses) and CL205
Percentage male	50.77%	In one-way sensitivity analysis the range is 0-100%.	Pooled data from CL201 (all doses) and CL205
Weight	Median weight of the general population in CS Table 34	A lower weight at the 25% percentile (also in CS Table 34) is tested in sensitivity analysis	Royal College of Paediatrics and Child Health[81]
Transition probabilities – treated group, age 1-4 years	CS Table 38	Dirichlet distribution using observed values in CS Table 37.	CL205 study
Transition probabilities – treated group, age 5 years and older	CS Table 40	Dirichlet distribution using observed values in CS Table 39.	CL201 study
Transition probabilities – control group, all ages	CS Table 42	Dirichlet distribution using observed values in CS Table 41. An alternative approach to missing data imputation is used in a scenario analysis. A further scenario analysis uses data from Study CL002.	UK chart review
Utilities	CS Table 31	Beta and Normal distributions using values from the UK vignette study.	UK vignette study[63]
Cost of burosumab	CS Table 48 and 49	None	Proposed list price
Monitoring costs associated with burosumab	One-off cost of £126.55 per patient at treatment initiation (CS Table 50)	Gamma distribution assuming standard error is 25% of the mean	Unit costs taken from PSSRU[67] and NHS Reference Costs 2016/17[65]
Surveillance costs and resource use Including (specialist consultations, laboratory monitoring, radiography,	CS Table 54	Gamma distribution assuming standard error is 25% of the mean	Unit costs from NHS Reference costs 2016/17.[65] Resource use taken from KOL opinion

Parameter	Mean value	Range / Distribution	Source
renal ultrasonography, dental check-ups)			Detail outlined in CS Table 52
Comparator costs (oral phosphate and alfacalcidol)	£492.57 per child and £394.01 per adult (CS Table 51 and 53)	Gamma distribution assuming standard error is 25% of the mean	Unit costs from the BNF (Source electronic model[2]) and resource use taken from Carpenter et al. for children[22] and Che et al. for adults[66]
Pain and mobility costs and resource use (physiotherapy)	CS Table 52 and Table 53	Gamma distribution assuming standard error is 25% of the mean	Unit costs taken from PSSRU[67] and resource use from Che et al.[66] Detail outlined in Table 52
Orthopaedic intervention costs and resource use Including (dental abnormalities, stapling of growth plates, hip arthroplasty, knee arthroplasty)	CS Table 52 and Table 53	Gamma distribution assuming standard error is 25% of the mean	Unit costs from NHS Reference costs 2016/17[65] Resource use and further details outlined in CS Table 52
Adverse event costs (injection-site reactions)	£0 - see section 12.3.8	Range £0 - £5	Assumed unit costs Resource use outlined in studies CL201 and CL205
Source: Table 20 in the response to the clarification letter.[2]			

Box 5.2: Deterministic sensitivity analyses and scenario analyses presented in the CS

Deterministic one-way sensitivity analyses

- Ratio between genders in treatment population
- Transition probabilities for burosumab and standard of care
- Resource use
- Unit costs
- Dosing of medication in standard of care
- Age group specific utilities of health states

Scenario analyses

- Discount rate
- Uniform age distribution at start of treatment
- Age and severity distribution based only on patients treated on Q2W schedule
- Using observed (40-week) transition probabilities for patients aged one to four years
- Using observed (64-week) transition probabilities for patients aged five and over
- Using transition probabilities based on pooled data from both clinical studies on burosumab
- Using transition probabilities for standard of care based on linear interpolation of UK chart review data
- Using transition probabilities for standard of care based on CL002 study
- Treatment is stopped at age 15 for both genders
- Treatment is stopped at age 16 for both genders
- Treatment is stopped at age 17 for both genders
- Using mean dose for burosumab from study CL201 (1.05 mg/kg) as opposed to what is recommended in the summary of product characteristics
- Rounding dose of burosumab up (as opposed to rounding to nearest 10 mg)
- Using the 25th percentile weight instead of median weight for each age
- Continuing standard of care treatment in adult patients with healed rickets
- No surveillance in adulthood for patients with healed rickets

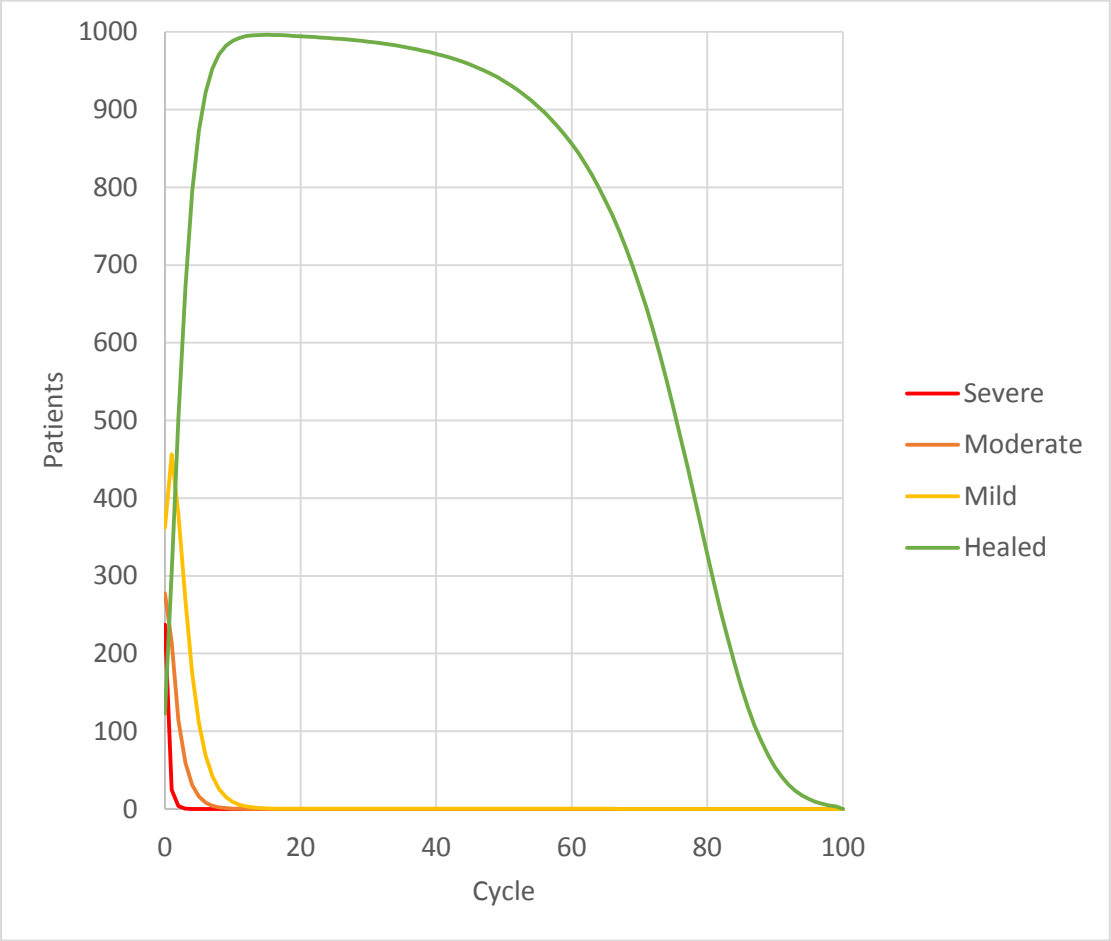
5.4 Headline results reported within the company's submission

In this section, the results of the cost consequence analysis presented by the company are summarised. During the clarification phase, the company detected and corrected two errors in the model. These are described in the response to Question B16 of the clarification letter.[2] Thus, the results described in this section are based on the version of the model submitted by the company with the response to the clarification letter. It should be emphasised that after correcting these errors the ICER increased by 1% compared to the one originally presented in the CS. Therefore, the impact on the results was minor.

The base-case Markov traces for the burosumab and SoC arms are shown in Figure 5.4 and Figure 5.5, respectively. Patients treated with burosumab are expected to spend most of their time alive in the "Healed rickets" health state. In particular, the model predicted that after six years more than 92% of the patients treated with burosumab were healed. After 13 years this was almost 100%. It is also striking that after three years of treatment with burosumab there are basically no patients in the severe health state (0.05%). In comparison, the distribution of SoC

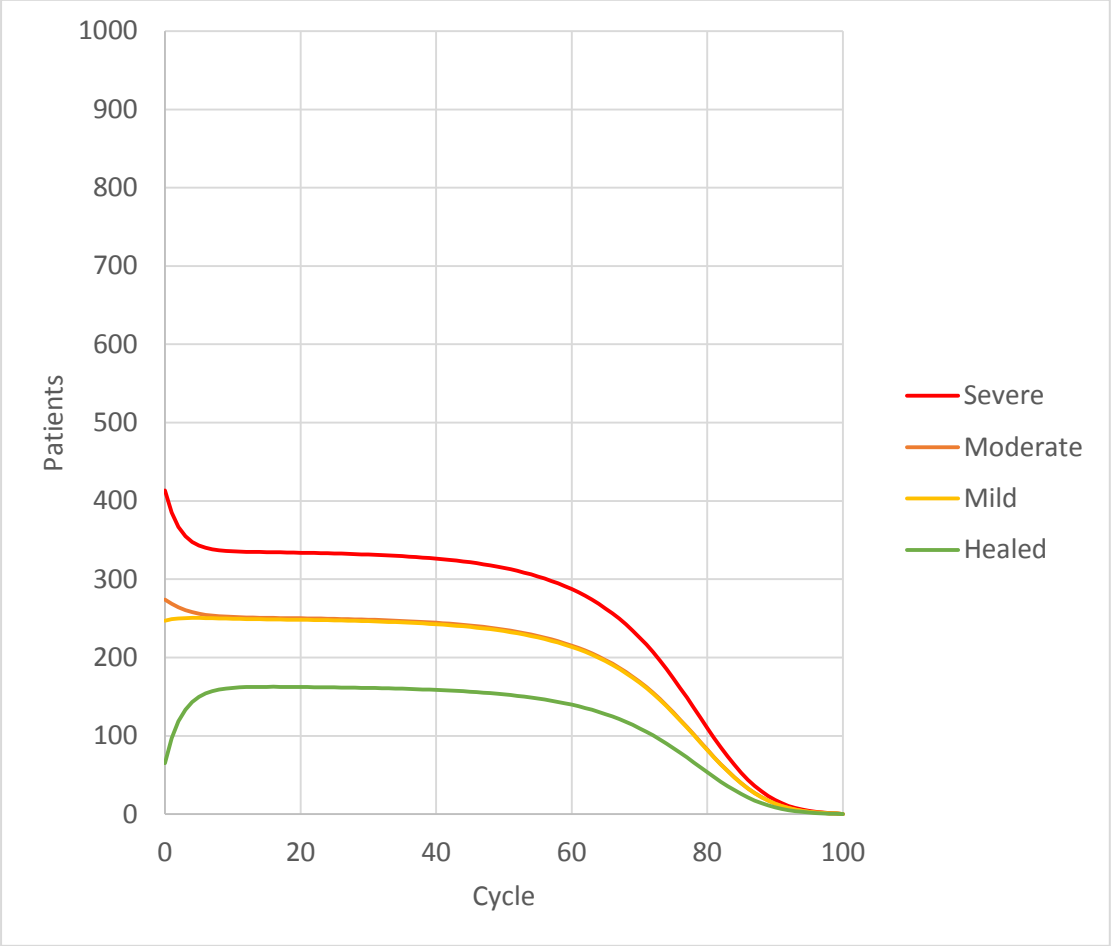
patients per health state is rather constant during most of the patient's lifetime. Approximately 35% of patients are expected to spend their time alive in the "Severe rickets" health state, 25% in the "Mild rickets" health state, another 25% in the "Moderate rickets" health state and approximately 15% in the "Healed rickets" health state. Note that there is no overall survival gain for burosumab in the base-case where the median survival is approximately 75.5 years in both arms. Differences in outcomes are thus due to the QALYs accrued over the lifetime.

Figure 5.4: Base-case: burosumab Markov trace



Source: Electronic model (after clarification).[2]

Figure 5.5: Base-case: SoC Markov trace



Source: Electronic model (after clarification).[2]

5.4.1 Headline total QALYs and total costs for burosumab versus standard care

Table 5.14 presents the results of the cost effectiveness analysis of burosumab versus SoC for the base-case scenario.

Table 5.14: Summary results of the company's base-case scenario

	Costs	QALYs	ICER	Costs	QALYs	ICER
	Discounted			Undiscounted		
SoC	[redacted]	25.989	--	[redacted]	41.786	--
Burosumab	[redacted]	36.293	[redacted]	[redacted]	58.677	[redacted]

Source: Electronic model (after clarification).[2]
Abbreviations: ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care

The company's analysis estimated that patients treated with burosumab gained 10.304 more discounted QALYs compared to SoC at an additional cost of [redacted], resulting in a cost per QALY of [redacted]. When no discounting was applied, the estimated gain in QALYs was 17.008 at an additional cost of [redacted], resulting in an ICER equal to [redacted].

Tables 5.15 and 5.16 below present a breakdown of discounted QALYs and costs for burosumab and SoC. The company's analysis suggests that under burosumab patients accrue more than 95% of the total QALYs in the "Healed rickets" health state (least severe state),

whereas for SoC, the number of QALYs accrued is similar among the four health states. This difference between the distributions of years spent in each health state, especially those spent in "Healed rickets", leads to incremental discounted QALYs of approximately 10 years.

More than 99% of the total costs for the burosumab arm are due to the treatment costs. In the SoC arm, 40% of the total costs are due to surveillance and 32% due to other medical costs. Although the burosumab arm results in cost savings in terms of pain-and-mobility ([redacted]) and orthopaedic interventions ([redacted]), the difference between burosumab and SoC is almost fully associated with burosumab treatment costs, adding up to approximately [redacted].

Table 5.15: QALY difference by health state for burosumab vs. SoC patients – base-case analysis

Health state	QALY burosumab	QALY SoC	Increment	Absolute increment	% increment
Healed rickets	34.324	5.770	28.554	28.554	61.0%
Mild rickets	1.385	7.210	-5.826	5.826	12.4%
Moderate rickets	0.444	6.230	-5.786	5.786	12.4%
Severe rickets	0.140	6.778	-6.638	6.638	14.2%
Total	36.293	25.989	10.414	46.804	100%

Source: Electronic model (after clarification).[2]
Abbreviations: QALYs = quality-adjusted life years, SoC = standard of care

Table 5.16: Costs associated with burosumab and SoC per category – base-case analysis

Cost category	Costs burosumab	Costs SoC	Increment	Absolute increment	% absolute increment
Treatment	[redacted]	[redacted]	[redacted]	[redacted]	99%
Drug (other)	[redacted]	[redacted]	[redacted]	[redacted]	1%
Monitoring	[redacted]	[redacted]	[redacted]	[redacted]	0%
Surveillance	[redacted]	[redacted]	[redacted]	[redacted]	0%
Pain and mobility	[redacted]	[redacted]	[redacted]	[redacted]	0%
Orthopaedic intervention	[redacted]	[redacted]	[redacted]	[redacted]	0%
Adverse events	[redacted]	[redacted]	[redacted]	[redacted]	0%
Total	[redacted]	[redacted]	[redacted]	[redacted]	100%

Source: Electronic model (after clarification).[2]
Abbreviations: SoC = standard of care

5.4.2 Sensitivity analyses presented within the company's submission

The company conducted sensitivity and scenario analyses. The results of these analyses are summarised below.

5.4.2.1 Deterministic sensitivity analysis

The results of the deterministic sensitivity analysis (DSA) were presented by the company as a tornado diagram where the top 20 most sensitive parameters were shown. This can be seen in Figure 5.6. It was observed that the ICER was most sensitive to changes in transition probabilities and utilities. The ICER was also sensitive to the proportion of females in the population since growth plates, and therefore treatment, stops earlier in females.

Figure 5.6: Tornado diagram illustrating results of top 20 most sensitive parameters in one-way sensitivity analysis

Source: Figure 11 in clarification letter response.[2]

ERG comment: Transition probabilities were not included in the DSA in the original version of the model submitted by the company. When this issue was raised in the clarification letter (Question B32[2]), the company included transition probabilities in the DSA, by varying the probabilities within the 90% confidence interval of a Dirichlet distribution. The results are shown in the tornado diagram above (Figure 5.6) and indicate that the model results are sensitive to the transition probabilities for patients aged five and older treated with burosumab. However, the ICER was not sensitive to changes in the transition probabilities for SoC and for burosumab patients under the age of five. In particular, the ICER increased significantly when the results were obtained at the upper limit of the 95% confidence interval, which resulted in the transition probabilities shown in in Table 5.17. These results were driven around the uncertainty in patients worsening in their rickets severity since this was not observed in the trial. In particular, remaining in the healed health state was assumed to occur with probability

one in the base-case analysis. However, the company reiterated that data from the RGI-C supported that patients had sustained improvements in rickets and therefore that it is likely that a patient would remain healed once healing has occurred. According to the company, this was also consistent with the restoration of phosphate that is associated with burosumab.

Table 5.17: Simulated upper bound of 95% confidence interval for burosumab transition matrix for patients aged 5 and over

	Mild	Moderate	Severe	Healed
Mild	53%	2%	2%	43%
Moderate	36%	46%	2%	17%
Severe	44%	26%	18%	13%
Healed	7%	7%	7%	79%
Source: Table 19 in the response to the clarification letter.[2]				

5.4.2.2 Scenario analysis

The company ran a number of scenario analyses to test the robustness of the model's results to changes in structural assumptions. The results of these analyses are summarised in Table 5.18.

The ICER was most sensitive to applying a discount rate of 3.5% for costs and effects, resulting in an ICER increased by 50% ([redacted]). Using Study CL002 data for transition probabilities in the SoC arm resulted in a 15% increase to the ICER ([redacted]), due to a 14% reduction in incremental QALYs. The ICER was also sensitive to changes in burosumab cost-relating parameters like children's weight, dosage and dose rounding, ranging from [redacted] to [redacted]. Applying a linear interpolation method for handling missing data in the UK chart review data used for SoC transition probabilities resulted in a 10% reduction in the ICER ([redacted]). Finally, the ICER was also sensitive to the age of stopping treatment (between 15 and 17 years), with ICERs ranging between [redacted] and [redacted]. For the other scenarios considered by the company, the ICER barely changed (up to a maximum of 2% increase).

ERG comment: The ERG believes that additional scenarios could have been explored, especially in terms of burosumab effectiveness. Given the low number of observations in both CL201 and CL205, scenarios showing the impact of changing the transition probabilities towards the healed and severe rickets health states could have been informative.

Furthermore, in all of the analyses, there is an underlying assumption that the treatment effect would be lifelong, since after patients reach age 18 in the model they are assumed to remain in their current health state and no deterioration in the health status of the patient occurs. However, it can be a possible that the treatment effect fades away after a certain number of years, as discussed in section 4.6.2 of this report. This was not explored by the company in the cost effectiveness analyses.

Table 5.18: Results of scenario analyses

Scenario	Total costs (£)		Total QALYs		Incremental costs (£)	Incremental QALYs	ICER (£)	Difference (%) in ICER
	Burosumab	SoC	Burosumab	SoC				
Base-case analysis	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	
Discount rate (3.5%)	[redacted]	32,626	22.318	16.121	[redacted]	6.197	[redacted]	50%
Even age distribution of cohort aged 1-12 years	[redacted]	51,284	36.580	26.215	[redacted]	10.364	[redacted]	1%
Baseline age and severity distribution: using only patients that were randomised to the bi-weekly burosumab dose	[redacted]	51,259	36.564	26.187	[redacted]	10.376	[redacted]	2%
Transition probabilities, aged 1-4 years: 40-week observations	[redacted]	50,580	36.290	25.989	[redacted]	10.301	[redacted]	0 %
Transition probabilities, aged 5 years and over: 64-week observations	[redacted]	50,580	36.403	25.989	[redacted]	10.415	[redacted]	-1%
UK chart-review data for SoC transition probabilities with missing data using linear interpolation	[redacted]	53,389	36.293	24.825	[redacted]	11.468	[redacted]	-10%
Study CL002 data for SoC transition probabilities	[redacted]	51,497	36.293	27.366	[redacted]	8.927	[redacted]	15%
Treatment stops at 15 years, both genders	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	-22%

Treatment stops at 16 years, both genders	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	-7%
Treatment stops at 17 years, both genders	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	7%
Mean burosumab dose 1.05 mg/kg	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	29%
Rounding up the dosage of burosumab required, rather than rounding to the nearest 10mg	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	12%
25th percentile children weight distribution	[redacted]	50,444	36.293	25.989	[redacted]	10.304	[redacted]	-10%
Continuing SoC drug treatment in adults with healed rickets	[redacted]	53,462	36.293	25.989	[redacted]	10.304	[redacted]	0%
Children with healed rickets no longer require surveillance in adulthood	[redacted]	48,984	36.293	25.989	[redacted]	10.304	[redacted]	-0%
Source: Electronic model (after clarification).[2] Abbreviations: ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care								

5.4.2.3 Probabilistic sensitivity analysis

A PSA was conducted using the probability distributions and parameters described throughout section 5.3.3 and summarised in Table 5.13. The average results (across 5,000 simulations) are shown in Table 5.19. The probabilistic ICER is 27% higher than the deterministic one, mostly due to the incremental QALYs, which in the PSA was approximately two QALYs smaller than in the deterministic base-case analysis.

Table 5.19: Probabilistic sensitivity analysis results

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
SoC	[redacted]	24.825			
Burosumab	[redacted]	36.293	[redacted]	8.120	[redacted]

Source: Table 17 in response to clarification letter.[2]
Abbreviations: ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care

The plot of the PSA outcomes in the cost effectiveness (CE) plane (Figure 5.7) shows that 99.9% of the simulations resulted in a gain in QALYs. The cost effectiveness acceptability curves in Figure 5.8 indicates that at a willingness to pay of £170,000, the probability of burosumab being cost-effective is [redacted].

Figure 5.7: PSA outcomes in the CE plane

Figure redacted - AIC

Source: Figure 9 in response to clarification letter.[2]

Figure 5.8: Cost effectiveness acceptability curves

Figure redacted - AIC

Source: Figure 10 in response to clarification letter.[2]

ERG comment: The PSA analyses were well-performed in general and the ERG agrees with most of the choices regarding probability distributions made by the company.

After clarification, the ERG detected an error in the model, which was using the standard deviation instead of the standard error when sampling random values for the utilities. The company used the following approach to obtain random utilities for the PSA: first a utility for the moderate health state is randomly drawn from a Beta distribution, with parameters estimated from the mean and standard deviation values obtained in the vignette study. That utility value for the moderate health state is then used as reference and the utilities for the other health states are calculated by randomly drawing the difference in utility compared to the moderate health state from a Normal distribution, with mean and standard deviation also obtained in the vignette study. For example, for patients aged 13 years and older (note that these utilities are applied in the model until patients die, thus for a large number of model cycles) the estimated mean utility in the moderate health state is 0.575 and 95% confidence interval (CI) is (0.417,0.727). In order to calculate utilities for the mild health state, a random value is drawn from a Normal distribution with mean 0.096 (the estimated mean difference in utility in the mild health state compared to the moderate health state) and standard deviation 0.11. With these parameters, a 95% confidence interval for the difference in utility in the mild health state compared to the moderate health state is (-0.085,0.277). Likewise, a 95% CI for the difference in utility in the healed and severe health states compared to the moderate health state is (0.018,0.364) and (-0.378,0.152), respectively. However, the company made a further assumption when modelling the utilities which was bounding the sampled utilities so that the health states with less severe rickets get always a higher or equal utility value compared to the next more severe health state (i.e. healed ≥ mild ≥ moderate ≥ severe). The ERG does not agree with this assumption as will be explained below. This assumption results in practice in uncertainty ranges for the difference in utility in the mild, healed and severe health states compared to the moderate health state that are (0,0.277), (0.018,0.364) and (-0.378,0), respectively. Note also that since the utility value

for the healed health state must be higher than the utility value for the mild health state, it is very likely that the lower limit for the uncertainty range of the difference in utility for the healed health state is higher than 0.018. Thus, in summary, according to the ERG the combination of using standard deviations (instead of standard errors) and the bounding condition introduced by the company implies that the model samples very large utility values for the mild and especially the healed health state, and very low for the severe health state. Since after 13 years the model predicts that alive patients in the burosumab arm have almost 100% chance of being in the healed rickets health state and that approximately 35% of the SoC patients are expected to spend their time alive in the severe rickets health state, the ERG is of the opinion that the current PSA results, as presented by the company, are biased in favour of burosumab.

As mentioned above, the ERG does not agree with the assumption of bounding the utilities so that the health states with less severe rickets always get a higher or equal utility value compared to the next more severe health state. When this issue was raised in the clarification letter (Question B9[2]), the company argued that it is common to adjust parameters associated with differing health states.[83] Otherwise, simulations may assign utilities to patients with mild rickets with values lower than those assigned to patients with severe rickets, which according to the company is implausible given the definition of the health states. The ERG disagrees with this latter statement. The company has acknowledged that rickets and RSS (and thus the model heath states) do not capture all aspects of XLH symptoms and progression and given the heterogeneity of the condition there is a chance that someone with mild rickets may have more severe additional manifestations, as mentioned above including in section 3.3.4. In fact, using the standard error instead of the standard deviation when sampling utilities for the health states, that should be very unlikely. In a less extreme case, the ERG does not consider it implausible that a patient with moderate rickets may have a lower utility than a patient with mild rickets, given the heterogeneity of XLH, the scale of the RSS (e.g. RSS = 1.49 is mild and RSS = 1.51 is moderate) and the uncertainty around the utility estimates. Nevertheless, as requested by the ERG, the company built a function into the model to enable the PSA to be run with or without bounded utilities. Unbounded utilities will be assumed in the ERG preferred base-case analysis in section 6.

As a first step for the calculation of the transition probabilities in the PSA, the model calculates "Cumulative Gamma functions" (see e.g. "Transition probabilities" sheet, cell Q9) where a factor 0.05 was added to the random draw of the Gamma distributions. It seems that this factor was added to account for non-observed transitions (empty cells in matrix) in the PSA (e.g. from Severe to Severe) as a sort of prior distribution, which in principle seems like an appropriate approach. However, the choice of 0.05 was arbitrary, as confirmed by the company in response to the clarification letter (Question B23[2]). The model results are sensitive to changes in that value and for that reason the ERG asked the company to provide a rationale for choosing 0.05 in the base-case and to perform sensitivity/scenario analyses on this factor. Unfortunately, the company simply responded that the choice of 0.05 was arbitrary but no further explanation was given. Furthermore, the ERG noted that when UK chart data were chosen for the comparator arm, this adjustment was not needed because all possible transitions were observed. The company corrected this in the model. The choice of a prior distribution for transition matrices is discussed in the paper by Briggs et al. 2003,[84] where an uninformative prior distribution over the rows of transition probability matrices is recommended to overcome the potential problem of zero observed counts in some of the cells of the matrices. This can be achieved for example by employing a minimally informative prior distribution like a Dirichlet(1, 1, 1, 1), which can

be interpreted as a uniform prior distribution expressing the belief that each transition is equally likely (i.e. this prior distribution assumes a 0.25 probability to all transitions with a high level of uncertainty). Given the low number of observations in the burosumab arm, using uninformative prior distributions for the transition matrices seems appropriate to the ERG and will be assumed in the ERG preferred base-case analysis in section 6.

The overall uncertainty associated with the PSA results is likely to be underestimated, not only for the reasons discussed above, but also because the following parameters were not included in the PSA:

- The initial distribution of patients per health state stratified by age was obtained by combining the data from CL201 and CL205. Despite being mentioned in Table 5.13 in section 5.3.4 that a Dirichlet distribution was used, these parameters seem to be fixed in the model.
- The percentage of males (50.77%) at baseline was also obtained by combining the data from CL201 and CL205. Given the limited number of observations in these trials, a Beta distribution could have been used.
- Weight by age and gender was also included in the model as a parameter. As discussed in section 5.3.2 (see e.g. Box 5.1), it is uncertain if these weights are representative for the XLH population (especially for females). Since the weight distribution per age is known, a probability distribution (e.g. Normal) could have been used to include weight in the PSA.

However, the impact of these parameters on the overall parameter uncertainty and on the decision uncertainty is expected to be minor. Because of this, and due to the time constraints associated to this assessment, the ERG did not include these parameters in the PSA conducted in section 6.

5.4.3 Validation

In the CS, there is hardly any reference to the validation efforts conducted on the model other than indicating that clinical experts validated the costs considered in the model, utilities were validated against the limited published literature and that cross-validation was not possible since there are no published cost effectiveness analyses in XLH. In the clarification letter, the ERG asked the company to provide details of the validation efforts conducted on the model. The company indicated then that the clinical experts also validated the conceptual model and supplemented information on the input parameters of the model. Furthermore, the company pointed out that "*continuous internal validation has been provided in the development of the model by two separate health economic consultancies for the absence for apparent bugs local code structure, appropriate translation of the conceptual model*".[2] Finally, an example of an extreme value test was provided. This indicated that when the treatment effect of burosumab was assumed to be zero (same transition probabilities in both arms), then the outcomes of the model were identical for both arms with the exception of drug and treatment monitoring costs.

ERG comment: While the ERG acknowledges that, due to the rarity of the disease, it might be difficult to validate many aspects of the model, it also deems the validation efforts reported in the CS insufficient. Although in the response to the clarification letter some more details were provided, it was not mentioned for example what kind of internal validation tests were conducted. A detailed discussion on the face validity of the results was missing in the CS and the response to the clarification letter. Given the lack of cost effectiveness studies on XLH, the

ERG feels that additional attention on the face validity of the results would have been helpful in this case. The ERG also asked the company to include in the response to the clarification letter the results of the ongoing external validation indicated on page 167 of the CS but these were not reported.

5.5 Discussion of available evidence relating to value for money for the NHS and PSS

Chapter 5 of this report focused on the economic evidence for burosumab submitted to NICE by the company. The company presented a QALY-based cost effectiveness model-based analysis comparing burosumab with SoC. The company's analysis estimated that patients treated with burosumab accumulated 10.304 more discounted QALYs compared to SoC at an additional cost of [redacted], resulting in a cost per QALY of [redacted]. When no discount was applied, the estimated gain in QALYs was 17.008 at an additional cost of [redacted], resulting in an ICER equal to [redacted].

The ERG identified several issues in the company's analyses. The most important concerns were related to the operationalisation of "full recovery" in the healed rickets health state and lifelong burosumab treatment effect and the choice of the utilities for the base-case. These seemed to bias the results in favour of burosumab. The choice of the discount rate also had a significant impact on the model's results, as shown by the company in one of the scenarios they conducted. The ERG was also concerned about some of the assumptions made by the company in their PSA since these also seemed to bias the results in favour of burosumab.

Other issues discussed by the ERG were the difference of the effects of burosumab on patients younger than age five and patients older than age five, the method used by the company to estimate transition probability matrices, the choice of baseline weight, age and disease severity distribution, and the lack of any treatment/disease related adverse events. However, all these were proven to have a minor impact on the model's results.

Some of the problems identified within the critical appraisal of the economic analyses were addressed by the ERG in the next chapter of this report. Thus, the next chapter outlines the additional analyses conducted by the ERG, which includes the development of a new base-case analysis (including a PSA) and several additional scenarios.

6 IMPACT ON THE COST-CONSEQUENCE ANALYSIS OF ADDITIONAL EXPLORATORY CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

6.1 Introduction

The additional analyses performed by the ERG are presented in this chapter. As described throughout Chapter 5, the ERG identified several issues in the company's analyses. Issues regarding the structure of the model were summarised in Box 5.1, whilst issues within the evidence and/or the methods used to inform the company's model parameters were discussed in section 5.3.3.1 (transition probabilities), section 5.3.3.3 (utilities) and section 5.4.2.3 (PSA). The efforts of the ERG in this chapter are focussed on solving (or partially solving) these issues. In particular, the ERG expected that the largest impact on the cost effectiveness results is caused by the choice of discount rates, the operationalisation of "full recovery" in the healed rickets health state and the lifelong treatment effects for burosumab. Furthermore, given the limited evidence in this submission, the ERG considers that great uncertainty is associated with the deterministic results and therefore, extra attention will be paid to the PSA.

6.2 Changes to the company's economic model

The changes made by the ERG to the company's model are summarised in this section. Note that the version of the model used as reference is the one submitted with the clarification letter. Compared to the original version of the model, the company made the following changes (see section 5.3.4 for details):

- Correction of a programming and a methodological error in the transition probability matrices for burosumab and SoC (CL002).
- Correction of a methodological error in the transition probability matrix for SoC (UK chart review).
- Unit costs were updated to 2016/17 costs/tariffs.

Major changes included the use of alternative annual transition probability matrices for burosumab derived from the original data, sourcing utilities directly from Lloyd et al. 2018,[63] the operationalisation of the full recovery and the lifelong treatment effects. Minor changes included discounting costs and health outcomes at 3.5% and including costs for adverse events. Based on these changes, a new ERG preferred base-case was defined in section 6.3.3.

6.2.1 Transition probabilities for burosumab

The ERG preferred transition probability matrices for burosumab are shown in Table 6.1 and Table 6.2 below. The derivation and a detailed explanation of the methods used to derive these matrices can be found in the critique to section 5.3.3.1 and Appendix 2 of this report.

Table 6.1: ERG preferred annual transition probability matrix for burosumab (patients aged one to four years)

	Mild	Moderate	Severe	Healed
Mild	100%	0%	0%	0%
Moderate	59%	41%	0%	0%
Severe	59%	41%	0%	0%
Healed	0%	0%	0%	100%
Source: Appendix 2 of this report.				

Table 6.2: ERG preferred annual transition probability matrix for burosumab (patients aged five years and older)

	Mild	Moderate	Severe	Healed
Mild	57%	0%	0%	43%
Moderate	40%	50%	0%	10%
Severe	62%	35%	0%	3%
Healed	0%	0%	0%	100%
Source: Appendix 2 of this report.				

6.2.2 Source used to estimate utilities

As mentioned in the ERG critique to section 5.3.3.3, the utility values that the company presented in Table 31 of the CS (Table 5.9) do not match all the utility values as presented in the report by Lloyd et al. 2018,[63] where the vignette study is described. It was observed that for each age group, the value for 'healed rickets' is higher in the CS than in Lloyd et al. 2018 whereas the value for 'severe rickets' is lower in the CS than in the Lloyd et al. However, no explanation for this discrepancy was provided by the company. Additionally, it was not clear to the ERG how the standard deviations that are presented in Table 5.9 for the non-moderate health states were derived. For these reason, the utilities reported in Lloyd et al. 2018, as shown in Table 6.3 below, are used in the ERG preferred base-case analysis.

Table 6.3: Mean utility values for the health states captured using EQ-5D-5L.

Health state	**Mean**	**Standard Deviation***
Age range 1-4		
Healed rickets (RSS score=0)	0.800	0.135
Mild rickets (RSS Score=0.5-1.0)	0.774	0.192
Moderate rickets (RSS Score=1.5-2.0)	0.685	0.175
Severe rickets (RSS Score>2.5)	0.610	0.184
Age range 5-12		
Healed rickets (RSS score=0)	0.890	0.113
Mild rickets (RSS Score=0.5-1.0)	0.757	0.159
Moderate rickets (RSS Score=1.5-2.0)	0.613	0.170
Severe rickets (RSS Score>2.5)	0.602	0.106
Age range 13+		
Healed rickets (RSS score=0)	0.811	0.108
Mild rickets (RSS Score=0.5-1.0)	0.671	0.154
Moderate rickets (RSS Score=1.5-2.0)	0.575	0.094
Severe rickets (RSS Score>2.5)	0.479	0.169
Source: Table 1 in Lloyd et al. 2018[63] *Standard errors should be used in the model.		

6.2.3 Operationalisation of the full recovery and lifelong treatment effects

As explained in section 4.6.2, the ERG considers that defining health states by RSS is likely to overestimate any improvement due to burosumab in moving to states with a lower RSS. In addition, the model currently assumed that the effect of burosumab lasts for the rest of the

patients' lives, which seems to be unrealistic. For that reason, the ERG assumed that the treatment effect would decline in time. Thus, it was assumed that after 20 years after the end of treatment, patients would experience a decline in quality of life which was operationalised by assuming the utility value of the next worse health state, as shown in Table 6.4.

Table 6.4: Utility values used in the ERG base-case for patients 13 years and older

Health state	Utility value (13 to 37 years)	Utility value (38 years and older)
Healed rickets	0.811	0.671
Mild rickets	0.671	0.575
Moderate rickets	0.575	0.479
Severe rickets	0.479	0.479

6.2.4 Minor changes

Minor changes included the following:

- Discounting costs and health outcomes at 3.5% (instead of 1.5% as assumed by the company).
- Including adverse events costs. These were assumed to be £0 in the base-case analysis. The CS does not report any estimation about what these costs could be. The only reference to this can be found in the electronic model where a range between £0 and £5 was used. For the ERG base-case, it was conservatively assumed £5 for the adverse event costs.

6.2.5 PSA-related changes

As discussed in the ERG critique of section 5.4.2.3, the following adjustments were made by the ERG in the PSA:

- Using the standard errors instead of the standard deviations when sampling random values for the utilities.
- Unbounding utilities so that the health states with less severe rickets do not always get a higher or equal utility value compared to the next more severe health state.
- Using a Dirichlet(1, 1, 1, 1) prior distribution for all possible transitions in the burosumab transition probability matrices.

6.3 Summary of the additional analyses undertaken by the ERG

The following analyses were undertaken using the company's model with ERG adjustments:

- ERG base-case: alternative transition probability matrices for burosumab, utilities from Lloyd et al., decline in quality of life 20 years after end of treatment, discounting costs and health outcomes at 3.5% and adverse event costs.
- ERG PSA: standard errors (instead of standard deviations) specified in utility distributions, unbound utilities with respect to next worse health state and approaches Dirichlet(1, 1, 1, 1) prior distributions for burosumab transition matrices.
- Additional scenario 1: changing the age where the decline in utilities is assumed.
- Additional scenario 2: using utilities from Table 31 in the CS.
- Additional scenario 3: rounding up the dose for burosumab.
- Additional scenario 4: running PSA with bounded utilities.

- Additional scenario 5: changing the prior distribution in transition matrices and run PSA.

6.4 Cost-consequence results produced by the ERG

6.4.1 Headline results produced by the ERG base-case analysis

The cost effectiveness results of the new ERG base-case are shown in Table 6.5. These are presented in 5 steps, showing the cumulative impact of each of the changes made by the ERG on the model results. It is clear that assuming a decline in utilities 20 years after treatment and considering a 3.5% discount rate resulted in a significant increase in the ICER. The other three changes had a minor/moderate impact on the ICER. In particular, the ERG preferred base-case analysis (Step 5 in Table 6.5) estimated that patients treated with burosumab accumulated 3.947 more discounted QALYs compared to SoC at an additional cost of [redacted], resulting in a cost per QALY of [redacted]. When the discount rate was 1.5% (Step 4 in Table 6.5), as in the company's base-case, the estimated gain in QALYs was 5.773 at an additional cost of [redacted], resulting in an ICER equal to [redacted].

Table 6.5: Comparison company base-case vs. ERG (step-by-step) base-case results

Scenario	Total costs (£)		Total QALYs		Incremental costs (£)	Incremental QALYs	ICER (£)	Difference (%) in ICER
	Burosumab	SoC	Burosumab	SoC				
Base-case (company)	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	
Step 1 – AEs costs	[redacted]	50,580	36.293	25.989	[redacted]	10.304	[redacted]	[redacted]
Step 2 – Transition matrices burosumab	[redacted]	50,580	36.301	25.989	[redacted]	10.312	[redacted]	[redacted]
Step 3 – Utilities from Lloyd et al.	[redacted]	50,580	34.232	26.007	[redacted]	8.225	[redacted]	[redacted]
Step 4 – Utilities decline 20 years after treatment	[redacted]	50,580	31.780	26.007	[redacted]	5.773	[redacted]	[redacted]
Step 5 – discount rate 3.5%	[redacted]	32,626	20.122	16.175	[redacted]	3.947	[redacted]	[redacted]

6.4.2 Probabilistic sensitivity analyses produced by the ERG

A PSA was conducted with the ERG preferred assumptions described in section 6.2.5. The average results (across 5,000 simulations) are shown in Table 6.6. The probabilistic ICER was [redacted]. This reflects the large uncertainty associated with the transition probability matrices for burosumab and the impact of choosing a prior distribution. This issue will be further discussed in section 6.4.3.5.

Table 6.6: ERG probabilistic sensitivity analysis results

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
SoC	[redacted]	16.271			
Burosumab	[redacted]	17.21	[redacted]	0.94	[redacted]
ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care					

The plot of the PSA outcomes in the cost effectiveness (CE) plane (Figure 6.1) shows that 88% of the simulations resulted in a gain in QALYs. The cost effectiveness acceptability curves in Figure 6.2 indicates that only at a high willingness to pay (approximately £500,000), the probability of burosumab being cost effective is not [redacted].

Figure 6.1: ERG PSA outcomes in the CE plane

Figure redacted - AIC

Figure 6.2: ERG-based cost effectiveness acceptability curves

Figure redacted - AIC

6.4.3 Exploratory sensitivity analyses produced by the ERG

6.4.3.1 Additional scenario 1: changing the age where the decline in utilities is assumed

In this series of scenarios, the ERG assessed the impact of assuming a different duration for the burosumab treatment effects on the cost effectiveness results. In the ERG base-case this was assumed to be 20 years after the end of treatment. Since this is unknown, the cost effectiveness results assuming a wide range of values for the burosumab treatment effect duration were calculated and summarised in Table 6.7. Note that in all these scenarios only the QALYs associated to burosumab change. Assuming five years for the duration of the burosumab treatment effects resulted in an ICER of [redacted], whilst assuming lifelong treatment effects resulted in an ICER of [redacted]. The difference between assuming 20 years duration of treatment effect (ERG) and lifelong treatment effects (company) was an ICER increased by approximately [redacted] under the ERG assumption.

Table 6.7: ERG cost effectiveness results for different durations of burosumab treatment effect

Years after treatment	Incremental costs (£)	Incremental QALYs	ICER (£)
5 years	[redacted]	3.001	[redacted]
10 years	[redacted]	3.375	[redacted]
15 years	[redacted]	3.688	[redacted]
20 years (ERG assumption)	[redacted]	3.947	[redacted]
30 years	[redacted]	4.336	[redacted]
40 years	[redacted]	4.594	[redacted]
50 years	[redacted]	4.759	[redacted]
No decline (company assumption)	[redacted]	4.906	[redacted]

6.4.3.2 Additional scenario 2: utilities from the company submission

In this scenario, the ERG explored the impact of using the utilities reported in Table 31 of the CS (Table 5.9) instead of the utility values as presented in the report about the vignette study by Lloyd et al. 2018.[63] As discussed in section 5.3.3.3, for each age group, the value for 'healed rickets' was higher in the CS than in Lloyd et al. whereas the value for 'severe rickets' was lower in the CS than in the Lloyd et al. report. The results from this scenario can be seen in Table 6.8. As expected, choosing the utilities from Table 31 in the CS, favoured the results burosumab, resulting in an ICER decreased by approximately [redacted] compared to the ERG base-case ICER.

Table 6.8: Results scenario using utilities from the company submission

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
SoC	[redacted]	16.121			
Burosumab	[redacted]	21.020	[redacted]	4.899	[redacted]
ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care					

6.4.3.3 Additional scenario 3: rounding up burosumab dose

The ERG explored in this scenario the impact on the model results of assuming that the exact dose for burosumab was given to patients. Since burosumab is available in vials of size 10 mg, 20 mg and 30 mg, it was assumed that when the calculated dose exceeded the dose of one vial, another complete vial would be needed and therefore the costs of these extra vial were added to the model's calculations. The impact of this assumption on the ICER was moderate, resulting in an ICER increased by approximately [redacted] compared to the ERG base-case ICER.

Table 6.9: Results scenario rounding up burosumab dose

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
SoC	[redacted]	16.175			
Burosumab	[redacted]	20.122	[redacted]	3.947	[redacted]
Abbreviations: ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care					

6.4.3.4 Additional scenario 4: running PSA with bounded utilities

In this scenario, the ERG tested the assumption made by the company in their base-case of bounding the utilities in such a way that the better health states were always assigned with a utility higher than or equal to the next worse health state. It should be noted that in the ERG base-case, standard errors instead of standard deviations were used to sample utilities. Therefore, the impact of this assumption was expected to be minor, as confirmed by the results shown in Table 6.10. The probabilistic ICER was [redacted]. Thus, the probabilistic ICER, the plot of the PSA outcomes in the CE plane and the cost effectiveness acceptability curves (not shown) obtained in this scenario were very similar to those obtained in the ERG PSA.

Table 6.10: Probabilistic sensitivity analysis results with bounded utilities

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
SoC	[redacted]	16.180			
Burosumab	[redacted]	17.190	[redacted]	1.01	[redacted]
Abbreviations: ICER = incremental cost-effectiveness ratio, QALYs = quality-adjusted life years, SoC = standard of care					

6.4.3.5 Additional scenario 5: prior distributions in transition probability matrices for burosumab

As mentioned in the ERG critique to section 5.4.2, an arbitrary factor 0.05 was added to the random draw of the Gamma distributions to account for non-observed transitions in the PSA. However, the model results are highly sensitive to changes in that value, as suggested by the ERG PSA results shown in section 6.4.2. The ERG asked the company to perform sensitivity/scenario analyses on this factor but unfortunately the company did not address this question (see clarification letter response to Question B23[2]).

Uninformative prior distributions over the rows of transition probability matrices are recommended by Briggs et al. 2003.[84] In particular, a prior Dirichlet(1, 1, 1, 1), in case of four health states, is suggested and this was the choice made by the ERG in their base-case. This can be interpreted as a uniform prior distribution expressing the belief that each transition is equally likely (i.e. 0.25 probability to all transitions with a high level of uncertainty). However, since the number of observations from which the transition matrices for burosumab are estimated is quite small, the choice of this prior distribution has a major impact on the PSA results as shown below. Further details on the choice and the impact of choosing prior distributions for the burosumab transition probability matrices are given in Appendix 3 of this report.

When running the PSA with the values shown in Table 6.1 and Table 6.2 (ERG preferred deterministic base-case), which should not be done because it would ignore the aforementioned uncertainty, the probabilistic ICER was [redacted], which is in line with the deterministic ICER obtained by the ERG (see Step 5 in Table 6.5), and the probability that burosumab is cost effective at thresholds smaller than or equal to £300,000 was [redacted]. When the PSA was run assuming a prior Dirichlet(0.05, 0.05, 0.05, 0.05) for all possible transitions, which was the choice made by the company, the probabilistic ICER obtained was [redacted] but the probability that burosumab is cost effective at thresholds smaller than or equal to £300,000 was still [redacted].

As the prior distribution approaches a Dirichlet(1, 1, 1, 1) , it is expected that the probabilistic ICER increases. This is because most of the cells of the observed burosumab transition probability matrices show either a probability 0 or 1 at key transitions which favour burosumab (e.g. probability of becoming severe is always 0), as shown in Table 6.1 and Table 6.2. Thus, as the prior approaches a Dirichlet(1, 1, 1, 1) , the posterior matrix deviates more from the observed matrix. Since the impact of the originally assumed 0 or 1 probabilities fades out, this has a significant impact on the model results. Thus, assuming a prior Dirichlet(0.1, 0.1, 0.1, 0.1) resulted in an ICER of [redacted] and assuming a Dirichlet(0.5, 0.5, 0.5, 0.5) resulted in an ICER of [redacted]. Finally, assuming a prior Dirichlet(1, 1, 1, 1) for all possible transitions resulted in the ERG PSA ICER of [redacted] and a [redacted] probability that burosumab is cost effective at thresholds smaller than or equal to £300,000.

6.5 Discussion

The additional analyses performed by the ERG were presented in this chapter. The main changes made by the ERG to the company's model included the use of alternative transition probabilities for burosumab, sourcing utilities directly from Lloyd et al. 2018[63] and the operationalisation of the full recovery and the lifelong treatment effects of burosumab. Minor changes included discounting costs and health outcomes at 3.5%, although this was proven to have a major impact on the model results.

The results of the ERG base-case, before applying the 3.5% discount rate on costs and health outcomes, resulted in an ICER increased by [redacted] compared to the company's base-case ICER. After applying the 3.5% discount rate, the ICER increased by [redacted]. Although sourcing the utilities from Lloyd et al. had a substantial impact on the ICER (increased by [redacted]), most of the total increase in the ICER (before applying the 3.5% discount rate) was due to the assumption of waning of treatment effect, implemented by reducing the utilities of burosumab patients 20 years after the end of treatment. Since there is uncertainty on whether this value of 20 years will be observed in real life, the ERG assessed the impact of assuming a different duration for the burosumab treatment effects on the cost effectiveness results. The difference between assuming 20 years duration of treatment effect (ERG) and lifelong treatment effects (company) was an ICER increase by approximately [redacted] under the ERG assumption. Assuming smaller values for the duration of the burosumab treatment effect increased the ICER. In particular, when this was assumed to be five years the deterministic ICER was [redacted].

The ERG was concerned that the PSA results presented by the company were underestimating the uncertainty associated with the transition probabilities for burosumab. For that reason, a new PSA and additional scenarios exploring the impact of choosing prior distributions for the burosumab transition matrices were conducted by the ERG. The latter was proven to be crucial and in the several scenarios provided by the ERG, the probabilistic ICER ranged from [redacted] to [redacted]. The ERG has concerns regarding the appropriateness of the choice of prior distribution made by the company for their PSA since this seemed to be based on matching the observed matrix and not representing prior beliefs about these transitions. The prior distribution assumed by the ERG, resulted in a more conservative approach and a more appropriate representation of the uncertainty associated to the transition probability matrices for burosumab.

Other scenarios explored by the ERG like using the utilities reported in Table 31 of the CS, rounding up the burosumab dose or bounding the utilities in the PSA were shown to have a minor to moderate impact on the model results.

Based on the ERG results, it is expected though that, from the payer perspective, the decision uncertainty related to burosumab's value for money would be low, given that the ICER estimates from all ERG analyses are above the acceptable thresholds considered for orphan drugs and the burosumab cost effectiveness probability at such thresholds was [redacted].

7 COST TO THE NHS AND PSS AND OTHER SECTORS

7.1 Summary of submitted evidence relating to the costs to the NHS and PSS

The CS includes a budget impact model to estimate the total costs to the NHS, for a period of five years, of adopting burosumab in England. [redacted].[8] This prevalence has been applied to the general population for England in children aged between one and 17 years to estimate the size of the population of [redacted] children with XLH eligible for treatment with burosumab (Table 2.1 of this report).[9] In the CS, it was reported that the number of patients eligible for burosumab [redacted] The company indicated that XLH is associated with skeletal deformations, pain and functional impairment; therefore, it is unlikely that there are undiagnosed children that would benefit from treatment with burosumab. Thus, the estimated prevalence based on primary care data is unlikely to be a significant underestimate.

In the CS, it was reported that the size of the patient population ([redacted]) is not expected to change over time as patients are only treated if they have growing skeletons i.e. each year there may be new patients but there will also be a likely similar number of patients ceasing treatment. In the CS, it was stated that XLH is not associated with an increased risk of death, compared to the standard population.[85] Therefore, the potential (and theoretical) population size is assumed to remain constant.

In the CS, based on clinical expert opinion, the yearly expected uptake rates of burosumab are calculated as follows: using the estimate of [redacted] children and assuming a 40% uptake in Year 1, followed by 65% uptake in Year 2 and a 90% uptake thereafter would equate to [redacted] children in year 1, [redacted] children in Year 2 and [redacted] children thereafter being treated with burosumab.

Table 7.1: Market update of burosumab

	Year 1	Year 2	Year 3	Year 4	Year 5
Expected uptake of burosumab	40%	65%	90%	90%	90%
Patients treated with burosumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients treated with SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 61 in the CS[1]
* The number in the CS reported here is 74; however, this should probably be 104.

[redacted]

[redacted]

The company stated that the weight, gender distribution and dosage of burosumab used to calculate treatment costs per age in the budget impact analysis is in line with the cost effectiveness model (CS Table 49).[1] The distribution of patients age was obtained from Study CL201 and CL205 (CS Table 36),[1] to estimate mean number of 10 mg, 20 mg, 30 mg vials required per patient, across the treated cohort.

[redacted] The company indicated that factoring in costs of monitoring and cost savings through displaced conventional therapy, will result in a net budget impact of [redacted] in Year 1, [redacted] in Year 2 and [redacted] per year thereafter (Table 7.2).

Table 7.2: Net budget impact of burosumab

	Year 1	Year 2	Year 3	Year 4	Year 5
Prevalent patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of patients treated with burosumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Number of new patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Number of continuing patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost of burosumab (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost offsets in drug costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Monitoring costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Net budget impact (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 62 in the CS[1]

In addition, the company reported the following information regarding resource savings associated with the use of burosumab: oral phosphate and vitamin D analogues should be discontinued one week prior to initiation of treatment with burosumab.[50] The company stated that, if a patient is treated with burosumab, there will be savings in the costs of oral phosphate and vitamin D analogues. The costs of these treatments in children are £492.57 per year (CS Table 51).[1] It was indicated in the CS that there

are also savings with regards to fewer surgical interventions, as well as reduced and/or deferred need for physiotherapy to manage the long-term consequences attributed to XLH. In the CS, these have not been factored in the budget impact analysis given its short time horizon.

7.2 ERG critique of the company's budget impact analysis

The ERG considers the assumptions made in the budget impact analysis questionable. There are concerns about the theoretical population size and the expected uptake rate of burosumab in England. In the CS, it was reported that the size of the patient population [redacted] is not expected to change over time. This estimate is based on an assumption that the patients are only treated if they have growing skeletons. In the CS, it was stated that XLH is not associated with an increased risk of death, compared to the standard population.[85] The potential (and theoretical) population size is assumed to remain constant.

Since real-world data suggests there could be [redacted] XLH patients between one and 17 years of age in England (see response to clarification letter – Question A4),[2] using the estimate of [redacted] children and assuming a 40% uptake in Year 1, followed by 65% uptake in Year 2 and a 90% uptake thereafter would equate to 77 children in year 1, 125 children in Year 2 and 174 children thereafter being treated with burosumab. The cost of burosumab at year 5 would then amount to [redacted]. The company indicated that burosumab is not expected to require additional resources to enable treatment administration, as it will be delivered via homecare. Homecare provision for XLH is being organised and funded by the company and will therefore not have any additional financial or resource impact on the NHS.

8 IMPACT OF THE TECHNOLOGY BEYOND DIRECT HEALTH BENEFITS AND ON THE DELIVERY OF THE SPECIALISED SERVICE

8.1 Summary of cost savings estimated within the CS

8.1.1 Nature of estimates presented

The CS did not include any estimates of costs (savings) or benefits incurred outside of the NHS and personal social services associated with of burosumab. The company indicated that at this stage this was not possible to quantify. However, the company expects significant savings to patients through healing of rickets and overall reduction or elimination of symptoms with burosumab.

8.1.2 Societal costs

As mentioned above, it was not possible for the company to identify and quantify costs to other government bodies. The company expects that patients treated with burosumab may be able to work more or developed further in their careers through improved education not inhibited by XLH. The company also expects that life-long disability will be avoided in people with XLH treated with burosumab. This will result in patients who will be less dependent their caregivers or on disability and other welfare payments. In the short term, the company expects that parents might not have to take time off from work to care for their child suffering with XLH.

8.1.3 Costs borne by patients

Most children experience interruptions to their schooling to attend hospital and GP appointments. Family members or caregivers may be absent from work to attend those appointments. In addition, costs of travel may be borne. Due to the limited number of specialist centres, patients and parents (or caregivers) may have to travel considerably. The results of an online survey carried out in January 2018 showed that

[redacted]

The study conducted by Berndt et al. in 1996 assessed the clinical and psychosocial aspects of XLH in 23 adults in Germany using a standardised questionnaire on pain and psychosocial rehabilitation (schooling, vocational training, employment and marital status).[28] Responders indicated that they struggled due to a lack of schooling and vocational training resulting from a lifetime of managing disease-related complications. A summary of the main findings is given below:

- Thirteen out of 20 patients were able to attend school regularly and to finish school adequately. Seven patients reported to have missed school repeatedly because of multiple hospitalisations leading to class repetition and to an inappropriate school qualification in four of them.
- Twelve out of 20 patients finished vocational training, five did not start and three attended but did not complete vocational training.
- Eight patients were employed, four were unemployed, four women were housewives, two patients received a social insurance payment because of inability to work (two patients did not answer questions on vocational training and profession).

Many adults with XLH also require surgery to correct skeletal deformities. In the study CL001,[86] [redacted]

[redacted]. In a case-note review of 59 adults with XLH, attending a single inherited metabolic disease service in the UK from 1998, 42% had had an osteotomy.4 Having surgery requires time off work not only for the surgery but also for recovery. For example, for an osteotomy, patients are not able to return to work for between two weeks and four months, depending on the nature of the work they do. Furthermore, results from CL001 indicated that most adult patents require [redacted].[86]

8.1.4 Other carer costs

As mentioned above, it was not possible for the company to quantify costs associated to caregivers.

8.1.5 Discussion of wider societal (non-health) benefits

Several issues regarding the impact of burosumab beyond direct health benefits were discussed qualitatively in the submission. However, the company were not able to provide any estimates of costs associated to inability to work or attend school, or costs borne by patients of caregivers.

The ERG considers it as inadequate that the impact of XLH on costs (savings) or benefits incurred outside of the NHS and PSS associated with burosumab was not identified prior to the submission to NICE.

8.2 Staffing and infrastructure requirements associated with the use of the technology

It was stated in the CS that, according to the draft Summary of Product Characteristics for burosumab, treatment with burosumab should be initiated by a physician experienced in the management of patients with metabolic bone diseases. The company indicated that following discussions with NHS England it was suggested that treatment with burosumab would only be initiated and prescribed by specialist centres that are members of the European Reference Network on Rare Bone Disorders (ERN-BOND). Furthermore, it is planned that burosumab will be supplied via a homecare service (to be provided and funded by the company) after patients have been established on a maintenance dose. During the initial dose titration period burosumab will be supplied directly to designated hospitals. Blood tests required for monitoring can be carried out in line with local arrangements, without visiting the specialist centre. Therefore, according to the company, no other additional facilities, technologies or infrastructures are required for the implementation of burosumab.

9 DISCUSSION

9.1 Statement of principal findings – clinical effectiveness

The studies included in the submission focus on the following populations and studies:

- Paediatric patients with XLH, five to 12 years old: Study CL201 (open-label RCT comparing different doses of burosumab biweekly or monthly administration of burosumab (doses 0.1 to 2.0 mg/kg))
- Paediatric patients with XLH, 1 to 4 years old: Study CL205 (open-label study to assess the safety, pharmacodynamics and efficacy of burosumab biweekly administration of burosumab at a target dose of 0.8 mg/kg))
- Paediatric Patients with XLH, 5 – 14 years old: Study CL002 (A retrospective longitudinal study of skeletal outcomes in children with XLH. No burosumab administered; however, study inclusion required the use of conventional therapy (oral phosphate/active vitamin D))

Results from CL201 show that burosumab significantly improves rickets at week 40 and week 64, compared to baseline. The primary endpoint, the rickets severity score (RSS) was reduced from baseline by 61% at week 40 ($p<0.0001$) by 58% at week 64 ($p < 0.0001$) with biweekly burosumab. Burosumab treatment also resulted in healing of rickets as assessed by RGI-C scores. The RGI-C score at Week 64 was +1.62. At Week 64, [redacted]% of children treated with biweekly burosumab had healing of rickets (RGI-C global scores ≥ 1.0). Furthermore, [redacted] of children treated with burosumab had substantial healing of rickets (RGI-C global scores ≥ 2.0). Growth velocity increased by [redacted]) in children treated with burosumab every two weeks, with a corresponding least-squared (LS) mean change in standing height z-score of [redacted]). Biweekly burosumab also resulted in improved functional assessments and patient-reported outcomes in CL201. Walking ability, as assessed by LS mean distance walked in the six-minute walk test (6MWT), increased from baseline by [redacted] at week 64 ([redacted]). Functional disability was assessed using the Pediatric Orthopedic Society of North America - Pediatric Outcomes Data Collection Instrument (POSNA-PODCI). Biweekly burosumab treatment increased scores for Sports/Physical Functioning and Pain/Comfort into the normal range seen in healthy children; LS mean scores showed improvements of [redacted] and [redacted] at week 64, respectively.

Results from CL002 show that RSS was reduced by [redacted] (over a median period of 102 weeks) after long-term conventional therapy. The RGI-C score was [redacted] with conventional therapy in Study CL002 (median [redacted] weeks). Furthermore, [redacted] of children treated with conventional therapy in Study CL002 had substantial healing of rickets (RGI-C global scores ≥ 2.0). After long-term treatment with conventional therapy in Study CL002, [redacted].

In study CL205 (13 children with XLH aged 1-4 years), burosumab treatment for 40 weeks significantly reduced RSS total score at week 40 by 59% (LS mean change of -1.73, $p < 0.0001$, ANCOVA model).

No patient died or discontinued from CL201 or CL205 for any reason; all patients continued treatment on study as of the data cut-off dates.

The most common adverse drug reaction reported in paediatric patients up to 64 weeks treatment with burosumab was injection site reactions (57%), headache (54%), pain in extremity (42%), vitamin D decreased (28%), rash (23%), toothache (19%), tooth abscess (14%), myalgia (14%), and dizziness (11%). Approximately 57% of the patients had an injection site reaction. The injection site reactions were generally mild in severity, occurred within one day of medicinal product administration, lasted approximately one to three days, required no treatment, and resolved in almost all instances.

In study CL201, one patient experienced serious TEAEs, and [redacted]. The most frequent TEAEs (>30% incidence) in study CL201 were [redacted].

The most frequent TEAEs (> 30% incidence [four or more of 13 patients]) in study CL205 were [redacted]

Adverse events of treatment with conventional therapy have not been reported. Therefore, it is not possible to assess the relative safety and toxicity in relation to the comparator.

9.2 Statement of principal findings – cost-consequence evaluation, NHS budget impact and societal analysis

9.2.1 Cost-consequence analysis

The company conducted a systematic review of cost effectiveness studies of burosumab and other studies including costs, resource use and any HRQoL measure associated with XLH. A total of eight full-text studies were assessed for eligibility which were included in the final evaluation of evidence. However, none of these studies were deemed relevant to the economic evaluation of burosumab.

The company's deterministic analysis estimated that patients treated with burosumab accumulated 10.304 more discounted QALYs compared to SoC at an additional cost of [redacted], resulting in a cost per QALY of [redacted]. When no discount was applied, the estimated gain in QALYs was 17.008 at an additional cost of [redacted], resulting in an ICER equal to [redacted].

The ERG identified several issues in the company's analyses. The ERG main concerns were related to the method used by the company to estimate the transition probability matrices for burosumab, the source of utilities used by the company, and the assumption of lifelong treatment effects of burosumab. The latter was expected to have a major impact on the model results. The choice of the discount rate was also challenged by the ERG. Furthermore, given the limited evidence in this submission, the ERG highlighted the extra importance of the probabilistic results. In light of these issues, the ERG performed a new base-case analysis and a number of additional scenarios.

The results of the deterministic ERG base-case resulted in an ICER increased by [redacted] compared to the company's base-case ICER. Most of the total increase in the ICER (despite the effect of applying the 3.5% discount rate) was due to assuming a treatment effect duration for burosumab of 20 years. The ERG also conducted a new PSA and additional scenario analyses exploring the impact of choosing prior distributions for the burosumab transition matrices. The latter was proven to be crucial and in the several scenarios provided by the ERG, the probabilistic ICER ranged from [redacted] to [redacted]. Other scenarios explored by the ERG like using the utilities reported in Table 31 of the CS, rounding up the burosumab dose or bounding the utilities in the PSA were shown to have a minor to moderate impact on the model results.

Based on the ERG results, it is expected that, from the payer perspective, the decision uncertainty related to burosumab value for money would be low, given that the ICER estimates from all ERG analyses are above the acceptable thresholds considered for orphan drugs and the burosumab cost effectiveness probability at such thresholds was [redacted].

9.2.2 Cost to the NHS and PSS

A budget impact model to estimate the costs to the NHS for a period of five years of adopting burosumab in England is also included in the CS. The results presented by the company suggested that the net budget impact of implementing burosumab (with an estimated prevalence of [redacted] patients) will be [redacted] in the first year and will rise to [redacted] in the fifth year. The cost of burosumab at year 5 amounts to [redacted]. The estimated total number of patients eligible for burosumab treatment after five years is [redacted] and the uptake of burosumab rises from 40% in year 1 to 90% in year 5. When a prevalence of [redacted] is considered by the ERG (with the same uptake rates), the estimated total number of patients eligible for burosumab treatment after five years reaches to [redacted]. The cost of burosumab at year 5 would then amount to [redacted].

9.2.3 Non-health benefits

The CS did not include any estimates of costs (savings) or benefits incurred outside of the NHS and personal social services associated with of burosumab. The company indicated that at this stage this was not possible to quantify. However, the company expects significant savings to patients through healing of rickets and overall reduction or elimination of symptoms with burosumab.

The ERG considers it as inadequate that the impact of XLH on costs (savings) or benefits incurred outside of the NHS and personal social services associated with of burosumab was not identified prior to the submission to NICE.

9.3 Strengths and limitations

9.3.1 Strengths of the CS

The ERG is confident that all relevant studies (published and unpublished) of burosumab were included in the CS, including data from ongoing studies. The same applies to the historical control patients. A control study in UK patients was mentioned in the CS without any results being report in the CS. However, results were provided as part of the response to the clarification letter. The reporting of outcomes from included studies also seems complete.

A range of relevant economic information was incorporated in the CS, including a QALY-based cost effectiveness model and an assessment of the expected costs to the NHS and PSS in England.

9.3.2 Weaknesses of the CS

The main limitation of the efficacy data reported in the CS is the study design of the included studies. Due to the absence of a control group in most studies it is not possible to make any direct comparisons between burosumab and conventional therapy. As stated by the company, the "burosumab phase 2 studies were uncontrolled dose finding or single arm studies, therefore an indirect comparison was not feasible" (CS, page 123).[1]

For children between one to four years old, only one study is presented in which all children received burosumab (CL205, N=13). A comparison with "established clinical management without burosumab" is not possible in this group of patients.

For children between five to 12 years old, the CS presents a study in which all children received burosumab (CL201). In addition, the CS presents a control study (CL002) in which children aged between five to 14 years received conventional therapy (i.e. oral phosphate/active vitamin D). Results of these two studies are mainly presented as a naïve comparison, simply reporting individual results from each study side by side. In addition, the company presents comparisons of 'rickets healing' with conventional therapy (Study CL002) versus burosumab (Study CL201) using propensity analysis matching.

In the CS, the company uses the term 'healing' and 'substantial healing of rickets'. This is defined using RGI-C global scores, where scores ≥ +1.0 indicate 'healing of rickets' and scores ≥ +2.0 'substantial healing of rickets'. The company does explain that "Healing in this context indicates improvement in the radiographic abnormalities and does not imply that complete healing was observed" (CS, page 100).[1] However, throughout the report the term 'healing of rickets' is used without any explanation of the degree of healing (minimal, substantial or complete). Moreover, RGI-C global scores and RSS scores do not capture all clinical aspects of XLH. That is of particular importance in the context of the economic model, which only considers RSS score alone as a clinical outcome measure. The diverse physiological impacts of hypophosphataemia, which may be independent of rickets, are therefore not captured as outcomes in the economic model.

In the response to the clarification letter the company described the vignettes for the various health states that informed the economic model in detail (Clarification Letter Response Question B7, Table 10). However, each health state was defined in such a way that there appears to be a perfect association between the RSS score and other clinical descriptors of the health state. For example, as the RSS score decreases so does the risk of fracture and the presence of deformity. However, this does not appear to be realistic in that it seems likely that there might be some resolution of the bone disorder such that the RSS score decreases, but that this resolution only occurs after incurring deformity, which cannot be completely resolved and with some continued increased risk of fracture.

In addition, the model currently assumed that the effect of burosumab, although stopped at age 16 (women) or 17 (men) lasts for the rest of their lives. This also seems unrealistic, the effects of burosumab on stature, bowing of the legs, joint deformity etc. are likely to persist fairly long but may wane as osteomalacia itself and the resulting fractures may lead to associated problems in later life. Effects on bone strength will wane quicker, therefore repeated fractures and badly healing fractures after 10 or 20 years are likely to occur. Effects of burosumab on symptoms caused by hypophosphatemia itself will disappear as soon as therapy is stopped. Therefore, we have assumed in the ERG base-case that patients will experience a decline in quality of life 20 years after the end of treatment, which was operationalised by moving to the utility value of the next worse health state (see section 6.2.3 in this report).

Regarding the evidence synthesis, the naïve comparison is unreliable because there are important differences between the inclusion criteria in both studies. Inclusion criteria for patients in studies CL201 and CL002 are similar in that patients in both studies were diagnosed with XLH and were of similar age. However, children in study CL201 also had: biochemical findings associated with XLH, standing height < 50th percentile for age and gender and radiographic evidence of active bone disease including rickets in the wrists and/or knees, and/or femoral/tibial bowing, or, for expansion patients, an RSS score in the knee of at least 1.5 points as determined by central read. In other words, study CL002 included all children with XLH, while study CL201 included children with more severe symptoms of XLH. This is also reflected in the relatively [redacted] standing height and [redacted] rickets severity score for children in study CL201 when compared to children in study CL002.

The adjusted comparison, using propensity analysis matching, is unreliable because of the limitations associated with these methods, in that the matching can only include those variables measured in both studies. Randomisation in a clinical trial creates balanced group for both measured and unmeasured variables. In observational studies, the most important factors which are predictive of the outcome may not have been measured and any treatment comparisons using observational study data may be biased. In the CS the company only included three variables in the PSM: age, gender and RSS total score at baseline. The rationale for variable selection was not provided other than whether they seemed similar or not between the two study populations. No details were provided of how this similarity was judged. The ERG found no statistically significant differences in age and gender between the two groups and considered that only including three variables in the creation of the propensity scores may have been too few.

Given the aforementioned limitations regarding the evidence presented by the company, the model results are highly uncertain and sensitive to key assumptions. Furthermore, the CS lacks an analysis of the wider societal (non-health) benefits associated with burosumab.

9.4 Uncertainties

The main uncertainty regarding the effectiveness evidence is the comparability of results from treated patients and historical control patients. Most of the evidence is presented as single arm studies including either treated patients (two studies, both with extensions that are still ongoing) or historical control patients (one study, with patients from one single centre, Radiographic analysis set ([redacted])). The historical control study (CL002) included patients aged from five to 14 years and can therefore only serve as a control group for study CL201 (children aged five to 12 years).

For patients with XLH aged one to four years old, the CS only presents a single arm burosumab study (CL205), no control data for this age group were provided. Only 13 children were enrolled in study CL205; therefore, results in this age group are very uncertain.

A randomised controlled study comparing burosumab with active control (oral phosphate/active vitamin D therapy) in children with XLH (aged one to ≤12 years) is currently ongoing. [redacted]. Results from this study will considerably reduce the uncertainty surrounding the clinical effectiveness of burosumab relative to conventional therapy in children with XLH aged between one and 12 years.

There is substantial uncertainty about the long-term effects of burosumab. The company conducted their analysis upon the assumption that these effects would be lifelong, despite treatment being stopped at the age of 16 in females and 17 in males, but there is no evidence to support that assumption. This assumption was proven to be crucial and one of the main drivers of the cost effectiveness results.

Additional uncertainty is generated when translating the clinical outcomes to QALYs since the evidence on HRQoL was based on a vignette study describing the health states of the economic model that were valued by (only six) clinical experts. Having HRQoL assessed by patients or caregivers, given that most of the patients are children, would reduce this uncertainty.

Since there is no direct or indirect evidence comparing burosumab to SoC, the assumed treatment effect of burosumab, as reflected by the transition probability matrices, is also very uncertain.

The ERG considers that the uncertainty around the reported ICERs is likely to be larger than suggested by the PSAs presented in this report. Given that PSA only addresses parameter uncertainty, other sources of uncertainty, like the ones mentioned above, could not be included in the PSA.

10 REFERENCES

[1] Kyowa Kirin. *Burosumab for treating X-linked hypophosphataemia [ID1151]. Submission to National Institute of Health and Care Excellence. Highly specialised technology evaluation (HST)*: Kyowa Kirin, 2018. 244p.

[2] Kyowa Kirin. *Burosumab for treating X-linked hypophosphataemia [ID1151]. Response to request for clarification from the ERG*: Kyowa Kirin, 2018

[3] Carpenter TO, Imel EA, Holm IA, Jan de Beur SM, Insogna KL. A clinician's guide to X-linked hypophosphatemia. *J Bone Miner Res* 2011;26(7):1381-8.

[4] Chesher D, Oddy M, Darbar U, Sayal P, Casey A, Ryan A, et al. Outcome of adult patients with X-linked hypophosphatemia caused by PHEX gene mutations. *J Inherit Metab Dis* 2018;Epub ahead of print:1-12.

[5] Ruppe MD. X-linked hypophosphatemia. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. *GeneReviews® [Internet]*. Seattle (WA): University of Washington, Seattle, 2017 [accessed 10.4.18]: 1-23. Available from: https://www.ncbi.nlm.nih.gov/books/NBK83985/

[6] Thacher TD, Fischer PR, Pettifor JM, Lawson JO, Manaster BJ, Reading JC. Radiographic scoring method for the assessment of the severity of nutritional rickets. *J Trop Pediatr* 2000;46(3):132-9.

[7] Beck-Nielsen SS, Brock-Jacobsen B, Gram J, Brixen K, Jensen TK. Incidence and prevalence of nutritional and hereditary rickets in southern Denmark. *Eur J Endocrinol* 2009;160(3):491-7.

[8] Delmestri A, Shaw N, Prieto Alhambra D, et al. *X-linked hypophosphataemia: UK prevalence using routinely collected primary care data [Unpublished report]*, 2018

[9] Office for National Statistics. Birth Summary Tables - England and Wales [Internet]. 2017 [accessed 4.4.18]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/birthsummarytables

[10] Office for National Statistics. Population estimates for UK, England and Wales, Scotland and Northern Ireland [Internet]. 2016 [accessed 10.4.18]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland

[11] NHS England. UK endorsed European Reference Network centres and the UK co-ordinating centres [Internet]. 2017 [accessed 10.4.18]. Available from: https://www.england.nhs.uk/publication/uk-endorsed-ern-centres/

[12] Dixon PH, Christie PT, Wooding C, Trump D, Grieff M, Holm I, et al. Mutational analysis of PHEX gene in X-linked hypophosphatemia. *J Clin Endocrinol Metab* 1998;83(10):3615-23.

[13] Whyte MP, Schranck FW, Armamento-Villareal R. X-linked hypophosphatemia: a search for gender, race, anticipation, or parent of origin effects on disease expression in children. *J Clin Endocrinol Metab* 1996;81(11):4075-80.

[14] Jonsson KB, Zahradnik R, Larsson T, White KE, Sugimoto T, Imanishi Y, et al. Fibroblast growth factor 23 in oncogenic osteomalacia and X-linked hypophosphatemia. *N Engl J Med* 2003;348(17):1656-63.

[15] Yamazaki Y, Okazaki R, Shibata M, Hasegawa Y, Satoh K, Tajima T, et al. Increased circulatory level of biologically active full-length FGF-23 in patients with hypophosphatemic rickets/osteomalacia. *J Clin Endocrinol Metab* 2002;87(11):4957-60.

[16] Perwad F, Azam N, Zhang MY, Yamashita T, Tenenhouse HS, Portale AA. Dietary and serum phosphorus regulate fibroblast growth factor 23 expression and 1,25-dihydroxyvitamin D metabolism in mice. *Endocrinology* 2005;146(12):5358-64.

[17] Sabbagh Y, Boileau G, Campos M, Carmona AK, Tenenhouse HS. Structure and function of disease-causing missense mutations in the PHEX gene. *J Clin Endocrinol Metab* 2003;88(5):2213-22.

[18] Shore RM, Chesney RW. Rickets: Part I. *Pediatr Radiol* 2013;43(2):140-51.

[19] Reid IR, Hardy DC, Murphy WA, Teitelbaum SL, Bergfeld MA, Whyte MP. X-linked hypophosphatemia: a clinical, biochemical, and histopathologic assessment of morbidity in adults. *Medicine (Baltim)* 1989;68(6):336-52.

[20] Veilleux LN, Cheung MS, Glorieux FH, Rauch F. The muscle-bone relationship in X-linked hypophosphatemic rickets. *J Clin Endocrinol Metab* 2013;98(5):E990-5.

[21] Andersen MG, Beck-Nielsen SS, Haubek D, Hintze H, Gjorup H, Poulsen S. Periapical and endodontic status of permanent teeth in patients with hypophosphatemic rickets. *J Oral Rehabil* 2012;39(2):144-50.

[22] Carpenter TO, Imel EA, Ruppe MD, Weber TJ, Klausner MA, Wooddell MM, et al. Randomized trial of the anti-FGF23 antibody KRN23 in X-linked hypophosphatemia. *J Clin Invest* 2014;124(4):1587-97.

[23] Linglart A, Biosse-Duplan M, Briot K, Chaussain C, Esterle L, Guillaume-Czitrom S, et al. Therapeutic management of hypophosphatemic rickets from infancy to adulthood. *Endocr Connect* 2014;3(1):R13-30.

[24] Linglart A, Dvorak-Ewell M, Marshall A, Martin JS, Skrinar A. Impaired mobility and pain significantly impact the quality of life of children with X-linked hypophosphatemia. 7th International Conference on Children's Bone Health Salzburg, Austria. 27 June 2015 - 30 June 2015. *Bone Abstracts* 2015;4:P198.

[25] Rafaelsen S, Johansson S, Raeder H, Bjerknes R. Hereditary hypophosphatemia in Norway: a retrospective population-based study of genotypes, phenotypes, and treatment complications. *Eur J Endocrinol* 2016;174(2):125-36.

[26] Zivicnjak M, Schnabel D, Staude H, Even G, Marx M, Beetz R, et al. Three-year growth hormone treatment in short children with X-linked hypophosphatemic rickets: effects on linear growth and body disproportion. *J Clin Endocrinol Metab* 2011;96(12):E2097-105.

[27] Ariceta G, Langman CB. Growth in X-linked hypophosphatemic rickets. *Eur J Pediatr* 2007;166(4):303-9.

[28] Berndt M, Ehrich JH, Lazovic D, Zimmermann J, Hillmann G, Kayser C, et al. Clinical course of hypophosphatemic rickets in 23 adults. *Clin Nephrol* 1996;45(1):33-41.

[29] Friedman NE, Lobaugh B, Drezner MK. Effects of calcitriol and phosphorus therapy on the growth of patients with X-linked hypophosphatemia. *J Clin Endocrinol Metab* 1993;76(4):839-44.

[30] Glorieux FH, Marie PJ, Pettifor JM, Delvin EE. Bone response to phosphate salts, ergocalciferol, and calcitriol in hypophosphatemic vitamin D-resistant rickets. *N Engl J Med* 1980;303(18):1023-31.

[31] Haffner D, Nissel R, Wuhl E, Mehls O. Effects of growth hormone treatment on body proportions and final height among small children with X-linked hypophosphatemic rickets. *Pediatrics* 2004;113(6):e593-6.

[32] Haffner D, Weinfurth A, Manz F, Schmidt H, Bremer HJ, Mehls O, et al. Long-term outcome of paediatric patients with hereditary tubular disorders. *Nephron* 1999;83(3):250-60.

[33] Jehan F, Gaucher C, Nguyen TM, Walrant-Debray O, Lahlou N, Sinding C, et al. Vitamin D receptor genotype in hypophosphatemic rickets as a predictor of growth and response to treatment. *J Clin Endocrinol Metab* 2008;93(12):4672-82.

[34] Steendijk R, Hauspie RC. The pattern of growth and growth retardation of patients with hypophosphataemic vitamin D-resistant rickets: a longitudinal study. *Eur J Pediatr* 1992;151(6):422-7.

[35] Verge CF, Lam A, Simpson JM, Cowell CT, Howard NJ, Silink M. Effects of therapy in X-linked hypophosphatemic rickets. *N Engl J Med* 1991;325(26):1843-8.

[36] Borghi MM, Coates V, Omar HA. Evaluation of stature development during childhood and adolescence in individuals with familial hypophosphatemic rickets. *ScientificWorldJournal* 2005;5:868-73.

[37] Ultragenyx. *UX023-CL002 clinical study interim report: a retrospective longitudinal study of skeletal outcomes in children with X-linked hypophosphatemia (XLH)*, 2016

[38] Cremonesi I, Nucci C, D'Alessandro G, Alkhamis N, Marchionni S, Piana G. X-linked hypophosphatemic rickets: enamel abnormalities and oral clinical findings. *Scanning* 2014;36(4):456-61.

[39] Surender K, Kochar IP, Ahmad A, Kapoor M. Hereditary vitamin D rickets: a case series in a family. *J Pediatr Endocrinol Metab* 2014;27(11-12):1217-22.

[40] Do TT. Clinical and radiographic evaluation of bowlegs. *Curr Opin Pediatr* 2001;13(1):42-6.

[41] Holick MF. Resurrection of vitamin D deficiency and rickets. *J Clin Invest* 2006;116(8):2062-72.

[42] Shore RM, Chesney RW. Rickets: Part II. *Pediatr Radiol* 2013;43(2):152-72.

[43] XLH Link. Support for parents [Internet]. 2018 [accessed 5.1.18]. Available from: http://xlhlink.com/en/managing-xlh/support-for-parents.html

[44] Imel EA, Carpenter TO, Linglart A, et al. Effects of burosumab (KRN23), a fully human anti-FGF23 monoclonal antibody, on functional outcomes in children with X-linked hypophosphatemia (XLH): results from a randomized, open-label phase 2 study. *International Conference on Children's Bone Health*. Würzburg, 2017: P063.

[45] Acaster Lloyd Consulting. *Parent survey [Unpublished report]*, 2018. 4p.

[46] Imel EA, Carpenter TO. A practical clinical approach to paediatric phosphate disorders. In: Allgrove J, Shaw NJ, editors. *Calcium and Bone Disorders in Children and Adolescents*. 2nd ed. Basel: Karger, 2015, 28: 134-61.

[47] Fukumoto S, Ozono K, Michigami T, Minagawa M, Okazaki R, Sugimoto T, et al. Pathogenesis and diagnostic criteria for rickets and osteomalacia - proposal by an expert panel supported by the Ministry of Health, Labour and Welfare, Japan, the Japanese Society for Bone and Mineral Research, and the Japan Endocrine Society. *J Bone Miner Metab* 2015;33(5):467-73.

[48] Pai B, Shaw N. Understanding rickets. *Paediatrics and Child Health* 2015;25(7):325-331.

[49] Endo I, Fukumoto S, Ozono K, Namba N, Inoue D, Okazaki R, et al. Nationwide survey of fibroblast growth factor 23 (FGF23)-related hypophosphatemic diseases in Japan: prevalence, biochemical data and treatment. *Endocr J* 2015;62(9):811-6.

[50] European Medicines Agency. *Crysvita -EMEA/H/C/004275. Annex I - Summary of product characteristics [Internet]*. London: European Medicines Agency, 2018 [accessed 19.3.18]. 33p. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/004275/WC500245537.pdf

[51] Xiao L, Homer-Bouthiette C, Hurley MM. FGF23 neutralizing antibody partially improves bone mineralization defect of HMWFGF2 isoforms transgenic female mice. *J Bone Miner Res* 2018;Epub ahead of print.

[52] National Institute for Health and Care Excellence. *Burosumab for treating X-linked hypophosphataemia. Final scope [Internet]*. London: NICE, 2018 [accessed 19.2.18]. 4p. Available from: https://www.nice.org.uk/guidance/gid-hst10016/documents/final-scope

[53] European Medicines Agency. *EPAR summary for the public: Crysvita - burosumab [Internet]*. London: European Medicines Agency, 2018 [accessed 19.3.18]. 8p. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/004275/WC500245541.pdf

[54] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[55] *Critical Appraisal Skills Programme (CASP): Making sense of evidence. 12 questions to help you make sense of a cohort study [Internet]*: Critical Appraisal Skills Programme (CASP), [accessed 4.4.18]. 2p. Available from: http://calder.med.miami.edu/portals/ebmfiles/UM%20CASP%20Cohort%20Assessment%20Tool.pdf

[56] Ultragenyx Pharmaceutical Inc. Efficacy and safety of KRN23 versus oral phosphate and active vitamin D treatment in pediatric patients with X linked hypophosphatemia (XLH) (PIXLES). NCT02915705 In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2017 [accessed 19.3.18]. Available from: https://clinicaltrials.gov/ct2/show/NCT02915705

[57] Mao M. *Statistical Analysis Plan: To evaluate the long-term efficacy of burosumab (UX023-CL021) compared to conventional therapy (UX023-CL002) using propensity score methodology*. Novato, CA: Ultragenyx Pharmaceutical Inc, 2017. 29p.

[58] Streiner DL, Norman GR. The pros and cons of propensity scores. *Chest* 2012;142(6):1380-1382.

[59] Forestier-Zhang L, Watts L, Turner A, Teare H, Kaye J, Barrett J, et al. Health-related quality of life and a cost-utility simulation of adults in the UK with osteogenesis imperfecta, X-linked hypophosphatemia and fibrous dysplasia. *Orphanet J Rare Dis* 2016;11(1):160.

[60] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313(7052):275-83.

[61] Mäkitie O, Doria A, Kooh SW, Cole WG, Daneman A, Sochett E. Early treatment improves growth and biochemical and radiographic outcome in X-linked hypophosphatemic rickets. *J Clin Endocrinol Metab* 2003;88(8):3591-7.

[62] National Institute for Health and Care Excellence. Guide to the processes of technology appraisal [Internet], 2014 [accessed 3.4.18]. 99p. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-sept-2014.pdf.

[63] Lloyd A, Gallop K, Piglowska N. *X-linked hypophosphatemia: its impact on HRQL as judged by clinical experts [Draft Report]*. London: Acaster Lloyd Consulting Ltd, 2018. 12p.

[64] Janssen B, Szende A. Population norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, editors. *Self-reported population health: an international perspective based on EQ-5D*: Springer, 2014: 19-30.

[65] Department of Health. *NHS reference costs 2016-2017 [Internet]*. London: Department of Health, 2017 [accessed 19.3.18] Available from: https://improvement.nhs.uk/resources/reference-costs/

[66] Che H, Roux C, Etcheto A, Rothenbuhler A, Kamenicky P, Linglart A, et al. Impaired quality of life in adults with X-linked hypophosphatemia and skeletal symptoms. *Eur J Endocrinol* 2016;174(3):325-33.

[67] Personal Social Services Research Unit. *Unit Costs of Health and Social Care 2017 [Internet]*, 2017 [accessed 19.3.18] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/

[68] Skrinar A, Marshall A, San Martin J, Dvorak-Ewell M. PP29-3: X-linked hypophosphatemia (XLH) impairs skeletal health outcomes and physical function in affected adults. In: Endocrine Society 97th Annual Meeting and Expo; 5-8 March 2015; San Diego, 2015.

[69] Smith WB, 2nd, Steinberg J, Scholtes S, McNamara IR. Medial compartment knee osteoarthritis: age-stratified cost-effectiveness of total knee arthroplasty, unicompartmental knee arthroplasty, and high tibial osteotomy. *Knee Surg Sports Traumatol Arthrosc* 2017;25(3):924-33.

[70] Ultragenyx. *UX023-CL201 Clinical Study Report*, 2017

[71] Olariu E, Cadwell KK, Hancock E, Trueman D, Chevrou-Severac H. Current recommendations on the estimation of transition probabilities in Markov cohort models for use in health care decision-making: a targeted literature review. *Clinicoecon Outcomes Res* 2017;9:537-546.

[72] Miller DK, Homan SM. Determining transition probabilities: confusion and suggestions. *Med Decis Making* 1994;14(1):52-8.

[73] Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health* 2012;15(6):812-20.

[74] Fleurence RL, Hollenbeak CS. Rates and probabilities in economic modelling: transformation, translation and appropriate application. *PharmacoEcon* 2007;25(1):3-6.

[75] Chhatwal J, Jayasuriya S, Elbasha EH. Changing cycle lengths in state-transition models: challenges and solutions. *Med Decis Making* 2016;36(8):952-64.

[76] Craig BA, Sendi PP. Estimation of the transition matrix of a discrete-time Markov chain. *Health Econ* 2002;11(1):33-42.

[77] O'Mahony JF, Newall AT, van Rosmalen J. Dealing with time in health economic evaluation: methodological issues and recommendations for practice. *PharmacoEcon* 2015;33(12):1255-68.

[78] Linglart A, Dvorak-Ewell M, Marshall A, et al. Online survey to characterise the burden of illness in children with X-linked hypophosphatemia [abstract and poster]. *ESPE 2015*. Barcelona, 2015: P-2-228.

[79] Ruppe M, Zhang X, Imel E, et al. Assessment of quality of life data after 4 monthly s.c. doses of a human monoclonal anti-fibroblast growth factor 23 antibody (KRN23) in adults with X-linked hypophosphatemia. *Horm Res Paediatr* 2014;82(Suppl 1):80-81.

[80] van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15.

[81] Royal College of Paediatrics and Child Health. *Boys: UK growth chart: 2-18 years [Internet]*: Royal College of Paediatrics and Child Health, 2013 [accessed 16.4.18] Available from: https://www.rcpch.ac.uk/system/files/protected/page/NEW%20Boys%202-18yrs%20(4TH%20JAN%202013).pdf

[82] Royal College of Paediatrics and Child Health. *Girls: UK growth chart: 2-18 years [Internet]*: Royal College of Paediatrics and Child Health, 2013 [accessed 16.4.18] Available from:

https://www.rcpch.ac.uk/system/files/protected/page/NEW%20Girls%202-18yrs(4TH%20JAN%202012).pdf

[83] Ren S, Minton J, Whyte S, Latimer NR, Stevenson M. A new approach for sampling ordered parameters in probabilistic sensitivity analysis. *PharmacoEcon* 2018;36(3):341-347.

[84] Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Med Decis Making* 2003;23(4):341-50.

[85] Nielsen LH, Rahbek ET, Beck-Nielsen SS, Christesen HT. Treatment of hypophosphataemic rickets in children remains a challenge. *Dan Med J* 2014;61(7):A4874.

[86] Ultragenyx. *UX023-CL001 online questionnaire to characterize the burden of illness in adults and children with X-linked hypophosphatemia (XLH)*, 2016

Appendix 1: ERG search strategies

The following searches were run to investigate additional population terms identified by the ERG and to identify the number of records retrieved. The ERG feels the number of references retrieved was a manageable number for the company to screen in order to identify potentially relevant clinical and cost-effectiveness studies without the use of study design filters.

MEDLINE (Ovid): 1946 to March Week 3 2018

1 exp Familial Hypophosphatemic Rickets/ (449)

2 ((familial or hereditary or genetic) adj2 (hypophosphataemi$ or hypophosphatemi$ or hypophosphatami$)).ti,ab. (269)

3 ("x linked" adj2 (hypophosphataemi$ or hypophosphatemi$ or hypophosphatami$)).ti,ab. (701)

4 (rickets adj3 (hypophosphataemi$ or hypophosphatemi$ or hypophosphatami$ or familial or hereditary or genetic or "D resistant" or "x linked")).ti,ab. (1554)

5 (XLH or HHRH or HPDR or ADHR).ti,ab. (389)

6 1 or 2 or 3 or 4 or 5 (1961)

7 limit 6 to yr="1945 - 2017" (1961)

[Records retrieved by Company searches: clinical effectiveness – 149; cost effectiveness – 10]

Embase (Ovid): 1974 to 2018 March 23

1 exp Familial Hypophosphatemic Rickets/ (742)

2 ((familial or hereditary or genetic) adj2 (hypophosphataemi$ or hypophosphatemi$ or hypophosphatami$)).ti,ab. (327)

3 ("x linked" adj2 (hypophosphataemi$ or hypophosphatemi$ or hypophosphatami$)).ti,ab. (998)

4 (rickets adj3 (hypophosphataemi$ or hypophosphatemi$ or hypophosphatami$ or familial or hereditary or genetic or "D resistant" or "x linked")).ti,ab. (2051)

5 (XLH or HHRH or HPDR or ADHR).ti,ab. (638)

6 1 or 2 or 3 or 4 or 5 (2725)

7 limit 6 to yr="1945 - 2017" (2707)

[Records retrieved by Company searches: clinical effectiveness – 200 (assuming error in reporting); cost effectiveness – 23]

CENTRAL Register of Controlled Trials (The Cochrane Library)

#1 MeSH descriptor: [Familial Hypophosphatemic Rickets] explode all trees 5

#2 (familial or hereditary or genetic) near/2 (hypophosphataemi* or hypophosphatemi* or hypophosphatami*) 15

#3 ('x linked' or 'x-linked') near/2 (hypophosphataemi* or hypophosphatemi* or hypophosphatami*) 32

#4 rickets near/3 (hypophosphataemi* or hypophosphatemi* or hypophosphatami* or familial or hereditary or genetic or 'D resistant' or 'D-resistant' or 'x linked' or 'x-linked') 43

#5 XLH or HHRH or HPDR or ADHR 23

#6 #1 or #2 or #3 or #4 or #5 in Trials 40

[Records retrieved by Company searches: clinical effectiveness – 9]

Clinical effectiveness – minor issues

- It is not clear which records in the PRISMA flow diagram were identified from database searches. The ERG assumes that flow diagram includes results from both database searches and hand-searching, as the numbers do not reflect the database searches alone.
- There appears to be an error in the documentation of the search results from Embase. In the CS, Table 2 - #24 gives the number of records retrieved as 20, but this is an unlikely reduction from the 208 records found before the date limit 1945-2017 was applied. Test searches run by the ERG suggest that this is a reporting error.
- MEDLINE In Process search strategies are not supplied separately. The ERG assumes that MEDLINE In Process is included in the MEDLINE searches, although this is not specified.

Cost effectiveness – minor issues

- It is not clear which records in the PRISMA flow diagram were identified from database searches. The ERG assumes that flow diagram includes results from both database searches and hand-searching, as the numbers do not reflect the database searches alone.
- The Embase strategy contains MEDLINE (MeSH) indexing terms (CS, Table 8 - #4)
- There are redundant lines in the MEDLINE (Table 7 - #36) and Embase (CS, Table 8 - #9, #62) strategies
- The MEDLINE strategy appears to contain unused searches (CS, Table 7 - #13, #27) on the epidemiology of XLH.

Appendix 2: Estimation of transition probability matrices

A Markov model with M health states can be characterised by the transition probability matrix P:

$$P = \begin{pmatrix} p_{1,1} & \cdots & p_{1,M} \\ \vdots & \ddots & \vdots \\ p_{M,1} & \cdots & p_{M,M} \end{pmatrix}$$

where $p_{i,j}$ denotes the transition probability from health state i to health state j (at time T) for $i, j = 1, \ldots, M$. The maximum likelihood estimate (MLE) of P, denoted by $\hat{P}$, can be obtained from the transition count matrix N

$$N = \begin{pmatrix} n_{1,1} & \cdots & n_{1,M} \\ \vdots & \ddots & \vdots \\ n_{M,1} & \cdots & n_{M,M} \end{pmatrix}$$

where $n_{i,j}$ denotes the number of event occurrences between health state i to health state j (at time T) for $i, j = 1, \ldots, M$. Then, $\hat{P}$ is the row proportions of N, so that

$$\hat{p}_{i,j} = \frac{n_{i,j}}{\sum_{m=1}^{M} n_{i,m}}$$

The company presented three transition count matrices with different observation periods (40 weeks, 64 weeks and 104 weeks). The problem at hand is to estimate the three corresponding transition probability matrices for a different time scale (52 weeks = 1 year). In general, this can be done as explained below.

Suppose the number of occurrences is obtained at time t_0, then the MLE of the transition probability matrix can be denoted by $\hat{P}_{t_0}$. If t denotes the desired time scale, then the MLE of the transition probability matrix associated with a cycle length t can be calculated as

$$\widehat{P_t} = P_{t_0}^{\left(t/t_0\right)}$$

For example, to obtain a one-year transition probability matrix from a one-month transition probability matrix, raise the one-month transition probability matrix to the twelfth power. Note that this approach works well when t is a multiple of t_0, i.e. when t/t_0 is a positive integer (as it occurs with a monthly to yearly conversion). When this is not the case, the spectral decomposition of P (eigenvalues and eigenvectors) needs to be calculated. Therefore, if we are interested in calculating $\hat{P}_t$, where t is not necessarily an integer multiple of the original scale, then $\widehat{P_t} = VD^tV^{-1}$, where

$$D^t = \begin{pmatrix} \lambda_1^t & 0 & 0 \\ 0 & \lambda_2^t & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \lambda_M^t \end{pmatrix}$$

and λ_i is the i^{th} eigenvalue of $\hat{P}_t$ and V is the matrix of eigenvectors (i^{th} column of V). Thus, in D^t the eigenvalues are raised to the power t but the eigenvectors do not change.

In practice, these calculations can be performed in *R* as shown below.

Transition probability matrix for burosumab age 1-4

The 40-week observation matrix for burosumab age 1-4 (denoted by N_40w) is the following:

```
N_40w <- matrix(c(1, 0, 0, 0, 2, 2, 0, 0, 4, 4, 0, 0, 0, 0, 0, 1), byrow = T, ncol
= 4)

rownames(N_40w) <- c("Mild","Moderate","Severe","Healed")
colnames(N_40w) <- rownames(N_40w)


N_40w

##          Mild Moderate Severe Healed
## Mild        1        0      0      0
## Moderate    2        2      0      0
## Severe      4        4      0      0
## Healed      0        0      0      1

P_40w <- matrix(nrow = 4, ncol = 4, 0)
colnames(P_40w)<-rownames(P_40w)<-colnames(N_40w)
```

The corresponding 40-week transition probabilities (denoted by P_40w) are then given below:

```
for (i in 1:4) P_40w[i, ] <- N_40w[i, ] / sum(N_40w[i, ])
P_40w

##          Mild Moderate Severe Healed
## Mild      1.0      0.0      0      0
## Moderate  0.5      0.5      0      0
## Severe    0.5      0.5      0      0
## Healed    0.0      0.0      0      1
```

Since the model's time horizon is one year (i.e. 52 weeks) the time scale of the transition matrix has to be changed. This can be done as explained above, i.e. by calculating eigenvalues and eigenvectors of the original transition matrix.

```
eig_40w <- eigen(P_40w)
eig_40w

## eigen() decomposition
## $values
## [1] 1.0 1.0 0.5 0.0
##
## $vectors
##           [,1] [,2]      [,3] [,4]
## [1,] 0.5773503    0 0.0000000    0
## [2,] 0.5773503    0 0.7071068    0
## [3,] 0.5773503    0 0.7071068    1
## [4,] 0.0000000    1 0.0000000    0

D_40w <- diag(eig_40w$values)
D_40w

##      [,1] [,2] [,3] [,4]
## [1,]    1    0  0.0    0
## [2,]    0    1  0.0    0
## [3,]    0    0  0.5    0
## [4,]    0    0  0.0    0
```

```
V_40w <- eig_40w$vectors
V_40w

##           [,1] [,2]      [,3] [,4]
## [1,] 0.5773503    0 0.0000000    0
## [2,] 0.5773503    0 0.7071068    0
## [3,] 0.5773503    0 0.7071068    1
## [4,] 0.0000000    1 0.0000000    0
```

Note that the command below should calculate the initial transition matrix (P_40w) as it occurs here.

```
V_40w %*% D_40w %*% solve(V_40w)

##      [,1] [,2] [,3] [,4]
## [1,]  1.0  0.0    0    0
## [2,]  0.5  0.5    0    0
## [3,]  0.5  0.5    0    0
## [4,]  0.0  0.0    0    1
```

We calculate first a weekly factor, since we want to obtain a transition probability matrix for one week. Then with this one-week matrix we can easily calculate the 52-week transition matrix by multiplying the one-week matrix 52 times. Note that other approaches than calculating the one-week matrix are possible but, in this case, it worked well as we will see below.

```
d_40w <- D_40w^(1/40)
d_40w

##      [,1] [,2]      [,3] [,4]
## [1,]    1    0 0.0000000    0
## [2,]    0    1 0.0000000    0
## [3,]    0    0 0.9828206    0
## [4,]    0    0 0.0000000    0
```

Thus, the one-week transition matrix is the following (P1_40w):

```
P1_40w <- V_40w %*% d_40w %*% solve(V_40w)
P1_40w

##           [,1]      [,2] [,3] [,4]
## [1,] 1.0000000 0.0000000    0    0
## [2,] 0.0171794 0.9828206    0    0
## [3,] 0.0171794 0.9828206    0    0
## [4,] 0.0000000 0.0000000    0    1
```

Note that, although it was possible to estimate the one-week transition matrix (P1_40w), some of the estimated values seem implausible, especially those regarding transitions from the severe health state (third row in P1_40w) as these values imply essentially instantaneous transition from the severe health state to either the mild or moderate health state.

As mentioned above, to obtain a one-year transition matrix we need to take the power 52 of the one-week matrix.

```
library(expm)

P1_40w %^% 52

##           [,1]      [,2] [,3] [,4]
## [1,] 1.0000000 0.0000000    0    0
## [2,] 0.5938738 0.4061262    0    0
```

```
## [3,] 0.5938738 0.4061262    0    0
## [4,] 0.0000000 0.0000000    0    1
```

As a validation step, note that by taking the power 40 of the one-week transition matrix we should obtain the original transition matrix, which is indeed happening as shown below.

```
P1_40w %^% 40

##      [,1] [,2] [,3] [,4]
## [1,]  1.0  0.0    0    0
## [2,]  0.5  0.5    0    0
## [3,]  0.5  0.5    0    0
## [4,]  0.0  0.0    0    1
```

Transition probability matrix for burosumab age 5+

The 64-week observation matrix for burosumab age 5+ (denoted by N_64w) is the following:

```
N_64w <- matrix(c(4, 0, 0, 4, 3, 3, 0, 1, 6, 3, 0, 1, 0, 0, 0, 1), byrow = T, ncol
= 4)
rownames(N_64w) <- c("Mild","Moderate","Severe","Healed")
colnames(N_64w) <- rownames(N_64w)
N_64w

##          Mild Moderate Severe Healed
## Mild        4        0      0      4
## Moderate    3        3      0      1
## Severe      6        3      0      1
## Healed      0        0      0      1

P_64w <- matrix(nrow = 4, ncol = 4, 0)
colnames(P_64w)<-rownames(P_64w)<-colnames(N_64w)
```

The corresponding 64-week transition probability matrix is then given by P_64w. We should repeat the same steps as in the 40-week case in order to obtain a one-week transition probability matrix. This is described in the `R` code below.

```
for (i in 1:4) P_64w[i, ] <- N_64w[i, ] / sum(N_64w[i, ])
round(P_64w,2)

##          Mild Moderate Severe Healed
## Mild     0.50     0.00      0   0.50
## Moderate 0.43     0.43      0   0.14
## Severe   0.60     0.30      0   0.10
## Healed   0.00     0.00      0   1.00

eig_64w <- eigen(P_64w)
eig_64w

## eigen() decomposition
## $values
## [1] 1.0000000 0.5000000 0.4285714 0.0000000
##
## $vectors
##      [,1]      [,2]      [,3] [,4]
## [1,]  0.5 0.1290564 0.0000000    0
## [2,]  0.5 0.7743386 0.8192319    0
```

```
## [3,]  0.5 0.6194709 0.5734623    1
## [4,]  0.5 0.0000000 0.0000000    0

D_64w <- diag(eig_64w$values)
D_64w

##      [,1] [,2]      [,3] [,4]
## [1,]    1  0.0 0.0000000    0
## [2,]    0  0.5 0.0000000    0
## [3,]    0  0.0 0.4285714    0
## [4,]    0  0.0 0.0000000    0

V_64w <- eig_64w$vectors
V_64w

##      [,1]      [,2]      [,3] [,4]
## [1,]  0.5 0.1290564 0.0000000    0
## [2,]  0.5 0.7743386 0.8192319    0
## [3,]  0.5 0.6194709 0.5734623    1
## [4,]  0.5 0.0000000 0.0000000    0

### This should be P
round(V_64w %*% D_64w %*% solve(V_64w),2)

##      [,1] [,2] [,3] [,4]
## [1,] 0.50 0.00    0 0.50
## [2,] 0.43 0.43    0 0.14
## [3,] 0.60 0.30    0 0.10
## [4,] 0.00 0.00    0 1.00

### Weekly factor
d_64w <- D_64w^(1/64)
d_64w

##      [,1]      [,2]      [,3] [,4]
## [1,]    1 0.000000 0.000000    0
## [2,]    0 0.989228 0.000000    0
## [3,]    0 0.000000 0.986482    0
## [4,]    0 0.000000 0.000000    0
```

However, in this case the one-week transition matrix is non-stochastic since one of its elements is negative, although the one-year transition matrix is actually stochastic, as shown below.

```
### One week transition matrix
P1_64w <- V_64w %*% d_64w %*% solve(V_64w)
round(P1_64w,2)

##      [,1] [,2] [,3]  [,4]
## [1,] 0.99 0.00    0  0.01
## [2,] 0.01 0.99    0  0.00
## [3,] 0.60 0.69    0 -0.29
## [4,] 0.00 0.00    0  1.00

### One-year transition matrix

round(P1_64w %^% 52,2)

##      [,1] [,2] [,3] [,4]
## [1,] 0.57 0.00    0 0.43
## [2,] 0.40 0.50    0 0.10
```

```
## [3,] 0.62 0.35    0 0.03
## [4,] 0.00 0.00    0 1.00
```

Furthermore, the original matrix could also be replicated.

```
round(P1_64w %^% 64,2)

##      [,1] [,2] [,3] [,4]
## [1,] 0.50 0.00    0 0.50
## [2,] 0.43 0.43    0 0.14
## [3,] 0.60 0.30    0 0.10
## [4,] 0.00 0.00    0 1.00
```

Alternatively, to overcome the issue of non-stochasticity, we propose using the approximation method described in Chhatwal et al. 2016.[75] Their algorithm is available online:

http://www.mgh-ita.org/ita-tools/online-modeling-tools.html

Using this approximation algorithm, based on the original 64-week observed matrix, the estimated stochastic four-week (note four weeks were chosen because calculating the one-week matrix was time consuming and it seemed unstable; note also that 4 is the greatest common divisor of 64 and 52, so both matrices could be estimated with the four-week matrix) matrix is the following:

```
round(matrix(c(0.963702, 0.00020038, 0., 0.036098, 0.0538478, 0.946152, 0., 0., 0.
0945621, 0.113132, 0.792306, 0., 0., 0., 0., 1.),byrow=T,nrow=4)
```

The estimated one-year matrix would be then the four-week matrix multiplied 13-times.

```
round(matrix(c(0.963702, 0.00020038, 0., 0.036098, 0.0538478, 0.946152, 0., 0., 0.
0945621, 0.113132, 0.792306, 0., 0., 0., 0., 1.),byrow=T,nrow=4)%^%13,2)

##      [,1] [,2] [,3] [,4]
## [1,] 0.62 0.00 0.00 0.38
## [2,] 0.40 0.49 0.00 0.11
## [3,] 0.48 0.32 0.05 0.15
## [4,] 0.00 0.00 0.00 1.00
```

Likewise, the 64-week matrix would be the four-week matrix multiplied 16-times, which is not the same as the observed one

```
round(matrix(c(0.963702, 0.00020038, 0., 0.036098, 0.0538478, 0.946152, 0., 0., 0.
0945621, 0.113132, 0.792306, 0., 0., 0., 0., 1.),byrow=T,nrow=4)%^%16,2)

##      [,1] [,2] [,3] [,4]
## [1,] 0.55 0.00 0.00 0.44
## [2,] 0.43 0.41 0.00 0.15
## [3,] 0.49 0.29 0.02 0.20
## [4,] 0.00 0.00 0.00 1.00
```

Appendix 3: Choice of prior distributions for transition probability matrices

Estimating reliable transition probability matrices for burosumab is challenging due to the overall low number of observed counts and the substantial number of zeroes in the matrices. *Uninformative* prior distributions over the rows of transition probability matrices are recommended by Briggs et al. 2003 to overcome this issue.[84] In particular, a prior Dirichlet(1, 1, 1, 1), in case of transition matrices having four health states, is suggested. This was the rationale for the choice made by the ERG in their base-case. Note that a Dirichlet(1, 1, 1, 1) can be interpreted as a uniform prior distribution expressing the *prior belief* that each transition is equally likely (in this case 1/4 = 0.25) but with a high level of uncertainty (since these prior estimation is only based on four counts). This prior distribution can be then be combined with the actual observed data, for example with the first row of Table A3.1 to give a *posterior* distribution Dirichlet(1+1, 1+0, 1+0, 1+0) = Dirichlet(2, 1, 1, 1), which assigns an average probability of transitioning from mild to (mild, moderate, severe, healed) equal to (2/5, 1/5, 1/5, 1/5) = (0.4, 0.2, 0.2, 0.2). Thus, the prior uninformative beliefs have been updated with the observed data and the result is a posterior probability that gives more *weight* to one transition over the others depending on the observed transitions. It is clear that, when more observed data become available, the estimated transition probabilities also become more reliable (i.e. the bias and the uncertainty in the point estimates are reduced) and the choice of the prior distribution becomes less relevant. However, since the number of observations from which the transition matrices for burosumab are estimated (in the example above just 1), the choice of this prior distribution has a major impact on the PSA results as shown in section 6.4.3.5.

An example with the transition matrix for burosumab patients aged one to four years is given below, although the same applies to the transition matrix for patients aged five to 12 years. The transition probability matrix for burosumab patients aged one to four years was estimated based on only 14 observations, which were distributed per health state as indicated in Table A3.1, although the last element of the matrix (healed, healed) was added for completeness but it was not observed in the trial (there were no healed patients).

Table A3.1. Predicted number of observations per health state at year 1 (52 weeks) for burosumab patients (one to four years old)

	Mild	Moderate	Severe	Healed	Total
Mild	1.00	0.00	0.00	0.00	1
Moderate	2.36	1.64	0.00	0.00	4
Severe	4.72	3.28	0.00	0.00	8
Healed	0.00	0.00	0.00	1.00	1
Note that non-integer observations are due to transforming the originally observed transition probability matrix from 40 weeks to 1 year (52 weeks).					

From the counts in Table A3.1, the transition probability matrix can be calculated simply by taking the proportions per row as shown in Table A3.2.

Table A3.2. ERG transition probability matrix for burosumab patients (1 to 4 years old)

	Mild	Moderate	Severe	Healed
Mild	1.00	0.00	0.00	0.00
Moderate	0.59	0.41	0.00	0.00
Severe	0.59	0.41	0.00	0.00
Healed	0.00	0.00	0.00	1.00

However, as mentioned above, due to the low number of observations, there is great uncertainty around the values shown in this transition matrix. For example, most of the cells of the matrix show either a probability 0 or 1, which have a significant impact on the model results. This issue can be overcome (or at least partially) by assuming an uninformative prior Dirichlet(1, 1, 1, 1) for all transitions. The resulting posterior distribution of the number of observations per health state at year 1 is shown in Table A3.3.

Table A3.3. Posterior distribution of the number of observations per health state at year 1 (52 weeks) for burosumab patients (one to four years old) – ERG estimate

	Mild	Moderate	Severe	Healed	Total
Mild	2.00	1.00	1.00	1.00	5
Moderate	3.36	2.64	1.00	1.00	8
Severe	5.72	4.28	1.00	1.00	12
Healed	1.00	1.00	1.00	2.00	5

The next step is then to re-estimate the transition probability matrix but now based on the 30 "observations" from Table A3.3. The resulting posterior transition probability matrix is given in Table A3.4. Note that there are significant differences between Table A3.2 and Table A3.4. Notably, Table A3.4 has no cells with a probability 0 or 1. It should be emphasised that even though the number of observations was increased from 14 to 30, the transition matrix in Table A3.4 is still surrounded by great uncertainty. This matrix was used by the ERG in their PSA.

Table A3.4. ERG transition probability matrix for burosumab patients (one to four years old) as used in the PSA

	Mild	Moderate	Severe	Healed
Mild	0.40	0.20	0.20	0.20
Moderate	0.42	0.33	0.13	0.13
Severe	0.48	0.36	0.08	0.08
Healed	0.20	0.20	0.20	0.40

It is clear that, in this case, changing the prior distribution will have a significant impact on the posterior distribution because the number of observations is very low. This is illustrated in Table A3.5 and Table A3.6, where the posterior matrices, as estimated by the company, are shown.

Note that the company chose as prior distribution a Dirichlet(0.05, 0.05, 0.05, 0.05), which can also be interpreted as a uniform prior distribution expressing the prior belief that each transition is equally likely (0.05/0.2 = 0.25) but with a very high level of uncertainty (since these prior estimation is only based on 0.2 "counts"). However, with this prior distribution, the posterior matrix in A3.6 is more similar to the original matrix in Table A3.2 than the ERG matrix in Table A3.4. Since the company indicated that the choice of this prior was arbitrary, the ERG was concerned regarding the appropriateness of this choice since it seems to be based on matching the observed matrix (which very much favours burosumab given the high number of cells with either 0 or 1) and not representing prior beliefs about these transitions.

Table A3.5. Posterior distribution of the number of observations per health state at year 1 (52 weeks) for burosumab patients (one to four years old) – company estimate

	Mild	Moderate	Severe	Healed	Total
Mild	1.05	0.05	0.05	0.05	1.2
Moderate	2.41	1.69	0.05	0.05	4.2
Severe	4.77	3.33	0.05	0.05	8.2
Healed	0.05	0.05	0.05	1.05	1.2

Table A3.6. Transition probability matrix for burosumab patients (one to four years old) as used in the company PSA

	Mild	Moderate	Severe	Healed
Mild	0.88	0.04	0.04	0.04
Moderate	0.57	0.40	0.01	0.01
Severe	0.58	0.41	0.01	0.01
Healed	0.04	0.04	0.04	0.88

In conclusion, it seems clear that running the analyses with the ERG or the company posterior transition probability matrices is expected to have a major impact on the model results. This was shown by the ERG in section 6.4.3.5. When the PSA was run with the posterior transition probability matrices estimated by the company (i.e. based on a prior Dirichlet(0.05, 0.05, 0.05, 0.05) for all possible transitions), the ICER obtained was [redacted]. As the prior distribution approached a Dirichlet(1, 1, 1, 1), the ICER increased. In particular, assuming a prior Dirichlet(1, 1, 1, 1) for all possible transitions, resulted in the ERG PSA ICER of [redacted].