Bone 122 (2019) 76–81

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

Rickets severity predicts clinical outcomes in children with X-linked hypophosphatemia: Utility of the radiographic Rickets Severity Score

Tom D. Thacher[a,*], John M. Pettifor[b], Peter J. Tebben[a], Ana L. Creo[a], Alison Skrinar[c], Meng Mao[c], Chao-Yin Chen[c], Ting Chang[c], Javier San Martin[c], Thomas O. Carpenter[d]

[a] Mayo Clinic, Rochester, MN, USA
[b] MRC/Wits Developmental Pathways for Health Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[c] Ultragenyx Pharmaceutical Inc., Novato, CA, USA
[d] Yale University School of Medicine, New Haven, CT, USA

ARTICLE INFO

Keywords:
Metabolic bone disease
Clinical trial
Radiology
Children
Validation

ABSTRACT

The Rickets Severity Score (RSS) was used to evaluate X-linked hypophosphatemic rickets (XLH), a genetic disorder mediated by increased circulating FGF23. The reliability of the RSS was assessed using data from a randomized, phase 2 clinical trial that evaluated the effects of burosumab, a fully human anti-FGF23 monoclonal antibody, in 52 children with XLH ages 5 to 12 years. Bilateral knee and wrist radiographs were obtained at baseline, week 40, and week 64. We evaluated the relationships of the RSS to the Radiographic Global Impression of Change (RGI-C), serum alkaline phosphatase (ALP), height Z-score, 6-minute walk test (6MWT) percent predicted, and the Pediatric Orthopedic Society of North America Pediatric Outcomes Data Collection Instrument (POSNA-PODCI). The RSS showed moderate-to-substantial inter-rater reliability (weighted kappa, 0.45–0.65; Pearson correlation coefficient (r), 0.83–0.89) and substantial intra-rater reliability (weighted Kappa, 0.66; r = 0.91). Baseline RSS correlated with serum ALP (r = 0.47). Baseline RSS identified two subgroups (higher [RSS ≥1.5] and lower RSS [RSS < 1.5]) that discriminated between subjects with greater and lesser rachitic disease. Higher RSS was associated with more severe clinical features, including impaired growth (Z-score, −2.12 vs −1.44) and walking ability (6MWT percent predicted, 77% vs 86%), more severe self-reported pain (29.9 [more severe] vs 45.3 [less severe]) and less physical function (29.6 [more severe] vs 40.9 [less severe]). During burosumab treatment, greater reductions in RSS corresponded to higher RGI-C global scores (r = −0.65). Improvements in RSS correlated with decreased serum ALP (r = 0.47). These results show the reliability of the RSS in XLH, and demonstrate that higher RSS values are associated with greater biochemical, clinical, and functional impairments in children with XLH.

1. Introduction

Rickets is a disorder characterized by impaired mineralization and ossification of the growth plates. Radiographs remain the gold standard for accurately diagnosing and evaluating the severity of rickets [1–4]. Radiographic features of rickets at the growth plate reflect the perturbations in cartilage maturation and apoptosis and in matrix mineralization and ossification. These manifestations are best seen at the physes of rapidly growing bones, including the distal radius and ulna, distal femur, proximal and distal tibia, and proximal humerus [4,5].

The Rickets Severity Score (RSS) is a quantitative method, validated in nutritional rickets, to assess rickets severity in the wrists and knees based on the degree of metaphyseal fraying, concavity, and the proportion of the growth plate affected [5]. It is a 10-point scale, where 10 represents the most extreme degree of rickets severity and 0 represents the absence of radiographic changes of rickets. The radiographic response following treatment of nutritional rickets can be assessed by the RSS, and RSS values correlate with values of serum alkaline phosphatase (ALP), a biochemical measure of rachitic activity. Originally used to assess nutritional rickets in Nigerian children, the RSS has been utilized in similar studies in other countries and independently validated in the US, India, Egypt, and Turkey [6–16]. The RSS has also been used to assess treatment response in hypophosphatasia, a rare hereditary form of rickets [17,18].

Unlike nutritional rickets, which is due to deficiencies of vitamin D or calcium, X-linked hypophosphatemic rickets (XLH) is a hereditary

* Corresponding author at: Mayo Clinic, 200 First Street S.W., Rochester, MN, USA.
E-mail address: Thacher.Thomas@mayo.edu (T.D. Thacher).

https://doi.org/10.1016/j.bone.2019.02.010
Received 13 November 2018; Received in revised form 23 January 2019; Accepted 10 February 2019
Available online 14 February 2019

form of rickets caused by loss-of-function mutations in the *PHEX* gene, resulting in renal phosphate wasting mediated by increased circulating FGF23 [19]. The radiographic features of rickets at the growth plate are similar in XLH and nutritional rickets, although the severity and duration of rickets vary considerably between the two disorders. Radiographic features in XLH are slower to correct with conventional therapy (oral phosphate and active vitamin D) than nutritional rickets and some degree of growth plate abnormality usually persists despite conventional therapy [20,21].

Because of the similar radiographic features observed in nutritional rickets and XLH, the RSS was used as the primary endpoint to evaluate the effect of burosumab in 52 children with XLH in a randomized, open-label, phase 2 clinical trial, UX023-CL201. In this study, treatment with burosumab provided a sustained increase in serum phosphorus levels to normal or near normal levels. No clinically significant safety findings with burosumab were observed during the trial, and the most commonly reported adverse events were common manifestations of XLH [22]. Burosumab, which has been approved by the United States Food and Drug Administration and the European Medicines Agency (EMA) for treatment of patients with XLH based in part on the results of this trial, is a fully human monoclonal antibody that inhibits FGF23. In the phase 2 study, there was significant improvement in rickets severity after 40 and 64 weeks of burosumab treatment, as assessed by two different methodologies, the RSS and the Radiographic Global Impression of Change (RGI-C) [22]. Here, our objective was to assess the validity of the RSS in XLH by examining the relationship between the RSS and severity of clinical and biochemical abnormalities at baseline, radiographic changes after burosumab treatment, and inter- and intra-rater reliability of the RSS.

2. Materials and methods

2.1. Radiographs

Study UX023-CL201 enrolled 52 children with XLH ranging from 5 to 12 years old (NCT02163577). The study initially enrolled 36 subjects with XLH and radiographic evidence of bone disease, and was subsequently expanded to include additional subjects who were required to have an RSS ≥ 1.5 points at the knee. The institutional review board at each participating site approved the protocol. Additional details regarding study design and treatment outcomes can be found in Carpenter et al. [22]. Bilateral anteroposterior (AP) knee radiographs and bilateral posteroanterior (PA) wrist radiographs were obtained at screening (serving as baseline radiographs) and at weeks 40 and 64. A manual was developed to standardize wrist and knee radiographs taken at each site. The majority of enrolled patients had been on conventional therapy for XLH prior to enrollment.

2.2. Rickets Severity Score method

The severity of rickets was measured using the RSS based on the degree of metaphyseal fraying, concavity, and the proportion of the growth plate affected [5]. Each radiograph was scored individually without the need to reference any prior images for comparison, thereby allowing the rater to remain blinded to radiograph sequence and treatment status. Radiographs of the wrists and knees from individual subjects were presented in random order and evaluated separately with the reader blinded to the treatment status of the subject and the timing of the radiograph.

Although radiographs of both knees and wrists at each assessment time were available, the RSS methodology involves the unilateral assessment of the more severely affected wrist and knee based on the clinical judgment of the reader. The reader assigns separate scores for the wrist (maximum score 4) and knee (maximum score 6) that are summed to create a total RSS (0 − 10), where higher scores represent greater severity of rickets (Table 1) [5]. The original RSS was slightly modified to allow for 0.5 point increments at the wrist rather than only 1.0 point increments. For the knee, the method assigns a multiplier factor to the score at the femur and tibia: 0.5 if one condyle or plateau is affected, and 1.0 if both condyles or plateaus are affected. The score of each image was entered into an electronic data capture system and transferred electronically to a central imaging facility. The scores could not be retrieved from the system by the reader after submission.

2.3. RSS inter- and intra-reliability assessments

The inter-rater reliability of the RSS was evaluated from a total of 104 radiographs (both wrists and knees) from the phase 2 study, 52 from baseline and 52 from Week 40. Each of the 104 radiographs was evaluated independently by four readers (TDT, PJT, ALC, and JMP). Two of these readers (TDT and JMP) had participated in the original validation of the RSS in nutritional rickets [5]. A rater manual was developed by Ultragenyx and each reader was provided with training to standardize the rating procedure. One of the four readers (TDT) had already read all the radiographs in his role as the primary reader for the phase 2 study. Therefore, this reader read each image twice and was the only reader used in an intra-reader variability assessment.

2.4. Radiographic Global Impression of Change (RGI-C) method

Another radiographic scoring method, the RGI-C, validated and compared with the RSS in hypophosphatasia [23], was used in the same data set from the phase 2 study to provide complementary assessment of the effect of treatment on rickets severity. The RGI-C utilizes a 7-point ordinal scale to evaluate the extent of change between radiographs taken at two different time points. Unlike the RSS, where each radiograph receives a score, the RGI-C represents the reader's impression of differences between earlier (baseline) and later (post-baseline) time points and clinical significance is explicit. Three readers, none of whom were readers for the RSS, conducted a side-by-side comparison of radiographs of bilateral wrist images with the baseline image on the left and post-treatment image on the right, and assigned a wrist score. RGI-C scores of −3, −2, and −1 between the two radiographs measured indicate severe, moderate, and minimal worsening, respectively; a score of zero indicates no change; and scores of +1, +2 and +3, indicate minimal healing, substantial healing, and complete/near complete healing, respectively. Bilateral knee images were scored similarly. The rater provided an RGI-C global score based on the overall impression of changes in the wrist and knee images, using the same 7-point scale. The average of the three readers' scores was calculated to generate wrist, knee, and global scores. While the RSS is a quantitative measure, the RGI-C provides a complementary qualitative evaluation of radiographic changes. As a measure of the validity of the RSS, changes in RSS with treatment were compared with the RGI-C score.

The relationship of the RSS with biochemical (serum ALP) and clinical measures were also assessed. Clinical assessments included height-for-age z-score and growth velocity, mobility (as evaluated with the 6-minute walk test [6MWT]), and patient-reported pain and functional disability (as measured with the use of the Pediatric Orthopedic Society of North America Pediatric Outcomes Data Collection Instrument [24]; normative scores range from 0 to 100, with higher scores indicating better functioning).

To determine growth velocity, two regression lines were fitted for each subject during the pre-treatment and treatment periods separately; the slope of the regression line was used to estimate the growth velocity in cm/year. Baseline growth velocity was estimated based on pre-burosumab treatment data collected from patients' historical medical records from the 2 years prior to baseline. We did not use the standard auxological method due to the limitation of historical growth data collection.

Table 1
Wrist and knee Rickets Severity Score. Total possible score is 10, with 4 for the wrist and 6 for the knee.

Grade definitions for radius & ulna					
0	Normal growth plate without changes of rickets				
0.5	Lucency of metaphyseal margin without fraying or irregularity				
1	Widened growth plate, irregularity of metaphyseal margin, but without concave cupping				
1.5	Partial metaphyseal concavity or incomplete fraying of metaphyseal margin				
2	Metaphyseal concavity with fraying of margins				
Grade radius	0	0.5	1	1.5	2
Grade ulna	0	0.5	1	1.5	2
Radius grade __ + Ulna grade__ = Total wrist__					

Grade definitions for femur & tibia				
0	Normal growth plate without changes of rickets			
1	Partial lucency, smooth margin of metaphysis visible			
2	Partial lucency, smooth margin of metaphysis NOT visible			
3	Complete lucency, epiphysis appears widely separated from distal metaphysis			
Grade femur	0	1	2	3
Grade tibia	0	1	2	3
Femur & tibia multiplier	0.5 if ≤ one condyle or plateau affected		1 if two condyles or plateaus affected	
(Femur grade __ × multiplier ___) + (Tibia grade__ × multiplier ___) = Total knee__				

2.5. Statistical analysis

The relationships between total RSS and biochemical and functional continuous variables were assessed with the Pearson correlation coefficient (r). Total RSS at baseline was dichotomized into two subgroups: higher RSS (≥1.5) and lower RSS (< 1.5). Baseline impairments were compared between the high and low RSS groups using two-sample *t*-test. A cut-off score of 1.5 was chosen to divide the groups based on the median total RSS of the study population at the time of the first interim analysis with 36 subjects. After the initial 36 subjects were enrolled, another 16 subjects were enrolled to provide additional safety and efficacy data, and to enrich for the rickets endpoint; the 16 additional subjects were required to have an RSS ≥1.5 at the knee. Least-squares mean values were used to assess change in the RSS between baseline and week 64 radiographs. Least-squares mean values were calculated with generalized estimating equation models (SAS 9.4) and included visit, regimen, visit by regimen as factors, and total RSS at baseline as a covariate, with exchangeable covariance structure. The differences in baseline impairments were tested by ANOVA between RSS < 1.5 vs ≥1.5. All p-values were presented as nominal p-values. No adjustment for multiplicity was made.

For RSS inter- and intra-rater reliability assessments, the inter-rater agreement was assessed for all pairs of readers, and for the one reader who read the radiographs twice, RSS of the first read was compared with RSS of the second read for the intra-rater assessment. Inter- and intra-rater agreement of the RSS was assessed using Pearson's and Spearman's Correlation Coefficients, agreement rate, and Weighted Kappa (Cicchetti-Allison weights). Only Pearson's Correlation Coefficients are reported, but were similar to analyses using Spearman's Correlation Coefficients (data not shown). The interpretation of agreement for the weighted Kappa was categorized as follows: almost perfect (0.80–1.00), substantial (0.60–0.79), moderate (0.40–0.59), fair (0.20–0.39), slight (0.00–0.19), and poor (< 0.00) [25].

3. Results

3.1. Study population and RSS

Primary results up to week 64 were previously reported in Carpenter et al. [22]. Briefly, a total of 52 subjects were included in this analysis. Mean (SD) age was 8.5 (1.9) years. Overall, mean (SD) RSS at baseline was 1.80 (1.09), ranging from 0 to 4.5. Baseline RSS from the phase 2 study were used to identify two subgroups (higher and lower RSS subgroups) to compare subjects with greater and lesser degrees of rachitic disease. At baseline, 34 subjects had a higher RSS (≥1.5) and 18 subjects had a lower RSS (< 1.5). In the higher and lower RSS subgroups, the mean (SD) baseline RSS was 2.46 (0.69) and 0.56 (0.34), respectively. Three subjects had a baseline RSS of zero; nevertheless, these subjects all had positive *PHEX* mutations and hypophosphatemia, and showed other skeletal abnormalities, including bowing of the legs.

After 64 weeks of treatment with burosumab, the mean (SD) RSS had improved overall to 0.88 (0.56). The mean (SD) RSS in the higher and lower RSS subgroups at week 64 was 1.01 (0.56) and 0.61 (0.47), respectively. The least squares mean change in RSS from baseline to week 64 was −0.92 overall ($p < 0.0001$). In the higher and lower RSS subgroups, least squares mean change from baseline to week 64 was −1.44 ($p < 0.0001$) and +0.06 ($p = 0.54$), respectively [22].

3.2. Reliability of the RSS in XLH

The inter-rater RSS reliability between four independent readers of 104 radiographs from the phase 2 study is presented in Table 2. Exact agreement and agreement within ± 0.5 point are shown in Table 2 for the RSS; weighted kappa values were 0.45 to 0.65 and Spearman

Table 2
RSS inter- and intra-rater reliability in subjects with XLH.

Analysis	Reader	Identical agreement (%)	Agreement within ± 0.5 (%)	Weighted kappa (95% CI)	Pearson correlation
Inter-rater	2 vs 3	21/103 (20.4%)	55/103 (53.4%)	0.57 (0.48, 0.65)	0.826
	2 vs 4	34/103 (33.0%)	71/103 (68.9%)	0.65 (0.57, 0.72)	0.864
	2 vs 1	20/103 (19.4%)	60/103 (58.3%)	0.50 (0.43, 0.58)	0.863
	3 vs 4	19/104 (18.3%)	62/104 (59.6%)	0.58 (0.51, 0.65)	0.871
	3 vs 1	21/104 (20.2%)	54/104 (51.9%)	0.45 (0.38, 0.52)	0.866
	4 vs 1	25/104 (24.0%)	68/104 (65.4%)	0.56 (0.49, 0.63)	0.887
Intra-rater	1 vs 1	38/104 (36.5%)	88/104 (84.6%)	0.66 (0.60, 0.72)	0.911

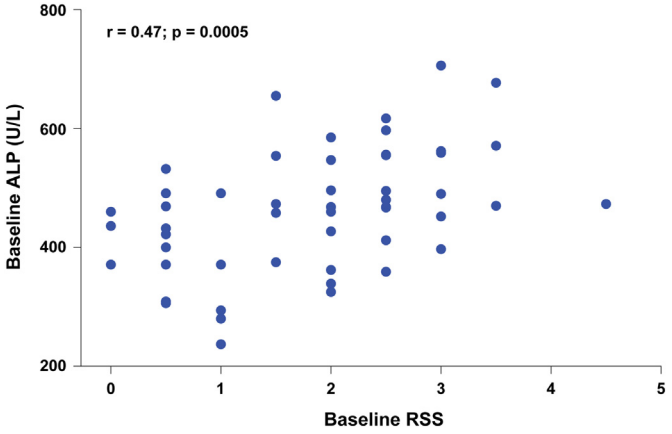

Fig. 1. Relationship between baseline ALP and RSS.
Note: normal ranges for ALP vary based on the age and gender of a child. The approximate upper limit of normal for healthy children between 5 and 12 years of age is 297 to 385 U/L.

correlation values were 0.69 to 0.82.

Intra-rater RSS reliability from radiographs in the phase 2 study was superior to or similar to inter-rater reliability (Table 2). A weighted Kappa of 0.66 and Pearson and Spearman correlation coefficients (0.91 and 0.90, respectively) showed substantial agreement and a high level of consistency.

3.3. RSS correlates with alkaline phosphatase in XLH

At baseline, higher serum ALP levels were associated with higher RSS (Fig. 1, r = 0.47; p = 0.0005). At baseline, mean (SD) serum ALP levels were greater in the higher RSS subgroup than in the lower RSS subgroup (497 [95] vs 388 [85] U/L) (p = 0.0002).

3.4. RSS correlates with clinical features of XLH

RSS subgroup analysis also discriminated between subjects with greater or lesser impairments in growth, mobility, and other clinical features at baseline. Higher baseline RSS was associated with more severe clinical features and poorer functional assessments, including growth, walking ability, and self-reported pain and physical function (Table 3). Subjects in the higher RSS subgroup had a mean standing height z-score of −2.1 compared to −1.4 in the lower RSS group. Growth velocity was also lower for the higher RSS subgroup by nearly 1 cm/year (5.1 vs 5.9 cm/year).

Walking ability, as measured by the 6-minute walking test (6MWT), also showed greater impairment in the higher RSS subgroup with a mean percentage of predicted normal walking distance of 77%. Nearly 60% of subjects in the higher RSS subgroup had baseline 6MWT values < 80% of predicted normal (defined as impaired) compared with ~20% for the lower RSS subgroup.

Finally, patient/caregiver perception of impairment was greater for the higher RSS subgroup with lower mean (SD) scores for the Sports/Physical Functioning and Pain Comfort scales of the POSNA-PODCI. Mean scores for these scales were within 1 SD of the norms for a healthy population (Mean = 50, SD = 10) at baseline for the lower RSS subgroup. For the higher RSS subgroup, mean baseline scores for the Sports/Physical Functioning and Pain Comfort scales were 2 SD below the norm, 29.6 (17.0) and 29.9 (14.6), respectively, indicating limited ability to participate in age-appropriate gross motor activities and greater pain associated with more severe rickets.

Table 3
Comparison of baseline growth and functional measures by RSS subgroup.

	Higher RSS (≥1.5)	Lower RSS (< 1.5)	p value
	(N = 34)	(N = 18)	
Standing height			
Z score, mean (SD)	−2.1 (1.0)	−1.4 (0.9)	0.0183
Pre-treatment growth velocity, cm/year, mean (SD)	5.1 (1.3)	5.9 (1.2)	0.0271
6MWT			
Distance (m), mean (SD)	468.2 (94.3)	511.3 (96.9)	0.1272
% predicted distance, mean (SD)	77.4 (13.7)	86.0 (13.5)	0.0349
Subjects with 6MWT < 80% predicted distance at baseline, n (%)	20 (58.8%)	4 (22.2%)	0.6052
POSNA-PODCI			
Sports/physical functioning scale (score), mean (SD)	29.6 (17.0)	40.9 (16.2)	0.0275
Pain/comfort scale (score), mean (SD)	29.9 (14.6)	45.3 (13.4)	0.0006

6MWT, 6-minute walk test; POSNA-PODCI, Pediatric Orthopedic Society of North America Pediatric Outcomes Data Collection Instrument.
p values are based on two-sample *t*-test.

3.5. Sensitivity of RSS to changes in rickets severity

Improvement in rickets was related to decline in serum ALP. At Week 64, mean (SD) serum ALP levels decreased from 459 (105) U/L at baseline to 369 (76) U/L at Week 64, which approaches the upper limit of normal range for healthy children between 5 and 12 years. Changes in ALP correlated with changes in RSS after 64 weeks of burosumab treatment (r = 0.47; p = 0.0008; Fig. 2), with greater reductions in ALP associated with greater reductions in rickets severity as assessed by the RSS. In the higher RSS subgroup, mean (SD) serum ALP levels decreased from 497 (95) U/L at baseline to 380 (76) U/L at week 64 and in the lower RSS group from 388 (85) U/L at baseline to 345 (73) U/L.

Subjects with greater improvement in the RSS also showed greater improvement in the RGI-C. During burosumab treatment, greater reductions in total RSS corresponded to higher RGI-C global scores (Fig. 3; r = −0.65; p < 0.0001). Concordance between the RSS and RGI-C was also confirmed by a large overlap between RSS and RGI-C responders. Of the 27 subjects with a baseline RSS ≥1.0 and an RGI-C global score ≥ +2.0 at Week 64, 21 (78%) had a reduction in RSS of ≥1.0 at Week 64. The mean (SD) change in RSS for subjects with RGI-C global scores < +2.0 was −0.22 (0.64), compared with −1.55 (0.99) for subjects with RGI-C global scores ≥ +2.0.

4. Discussion

Our results demonstrate that greater rickets severity, as demonstrated by RSS, in children with XLH was associated with higher serum

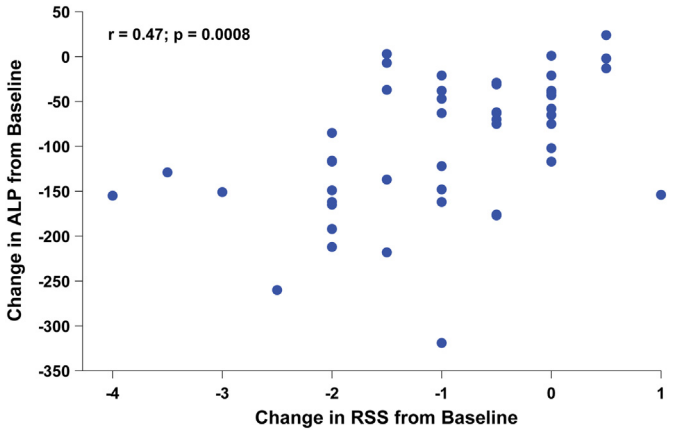

Fig. 2. Relationship between change in ALP and change in RSS after 64 weeks of burosumab treatment.

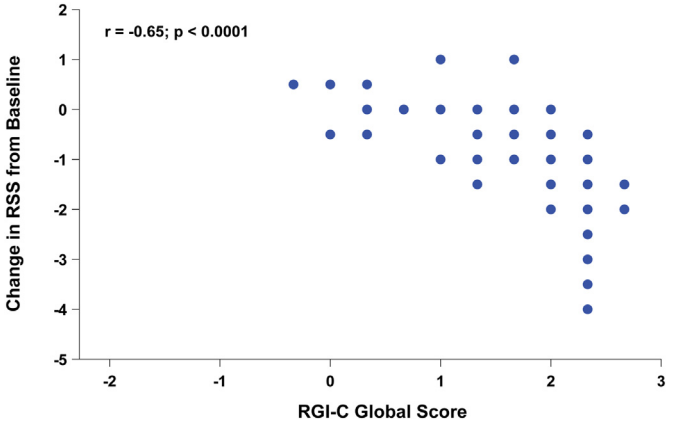

Fig. 3. Relationship of change in RSS with RGI-C global score after 64 weeks of burosumab treatment.
BL = baseline; RGI-C = Radiographic Global Impression of Change; RSS = Rickets Severity Score. Some dots may represent more than one subject.

ALP levels, greater growth impairments, increased pain, and worse physical function at baseline. The RSS was a reliable and valid instrument for evaluating the severity of radiographic rickets in XLH.

Assessment of rickets severity using the RSS showed that XLH exhibits similar radiographic abnormalities at the growth plate as nutritional rickets, reflecting the lack of mineralization and/or delayed growth plate ossification. However, there are notable differences between nutritional rickets and XLH. Although the RSS in nutritional rickets may range from 0.5 to 10, RSS values in XLH in this study were more restricted, ranging from 0 to 4.5. The lower RSS range in the phase 2 study possibly reflects the fact that most (96%) subjects had previously received conventional therapy for a mean of almost 7 years and thus were likely to have partially responded to treatment. Moreover, while the radiographic XLH features may not be as pronounced as those of nutritional rickets, the disease's impact is frequently greater because of the chronicity and lifelong effects leading to debilitating musculoskeletal impairments. Despite the lower range of RSS values in XLH, the RSS remained sensitive to changes in the severity of radiographic abnormalities following burosumab treatment.

A previous study that assessed radiographic changes with clinical outcomes in XLH also noted incomplete healing of rickets with conventional treatment, which was more pronounced with initiation of treatment after the age of 1 year [21]. In that study, a 6-point radiographic severity assessment was used. Radiographs improved but residual skeletal changes were evident with conventional treatment. Advantages of the RSS over the previous scoring method include its validation in multiple etiologies of rickets and the correlation with functional and biochemical measures of severity.

Our subgroup analyses showed that subjects who had a baseline RSS ≥ 1.5 had severely compromised bone health, with higher levels of ALP, and greater impairments in clinical outcomes, compared with subjects who had a baseline RSS < 1.5. These results are consistent with expert reviews that rickets in children with XLH is associated with severe long-term developmental abnormalities including stunted growth and skeletal deformities, particularly in the lower extremities, that often require surgical correction to improve or maintain mobility [19,26]. The relationship between RSS and growth is expected given the impact of deficient mineralization of cartilage and osteoid and delayed endochondral ossification on the growth plate, and are consistent with a previous study showing that the degree of severity of rickets is negatively correlated with height [21]. The RSS was also sensitive to changes in disease activity following burosumab treatment as seen by its concordance with the RGI-C and reductions in serum ALP levels.

At the lower end of the scale (e.g., a patient with an RSS of 0.5), the RSS scoring method may not have adequate sensitivity to detect possible small changes in patients with XLH due to the 0.5-point increments in the RSS scoring method. The RSS does not assess gradations or size of lucency, but only if the lucency is complete, partial, or absent. Therefore, the utility of the RSS is greater when a higher dynamic range of values are present in subjects with rickets, and greater improvements in RSS were observed in subjects with greater disease severity at baseline. However, even with the lower range of RSS values seen in patients with XLH, compared with those with nutritional rickets or hypophosphatasia, the RSS proved to be useful for assessing radiographic rickets severity.

Despite the smaller RSS range observed in XLH, inter- and intra-rater reliability was similar to those reported for nutritional rickets [5]. In the original evaluation of RSS in nutritional rickets, inter-rater correlation was 0.84 or greater and intra-rater correlation was 0.89 or greater for each observer. A study in India using the RSS in 178 children with nutritional rickets found that the initial RSS was linearly related with time to radiographic resolution (adjusted $R^2 = 0.90$) [12]. In 98 children with nutritional rickets in Egypt and Turkey, RSS was significantly related with serum phosphorus ($r = -0.30$), bone alkaline phosphatase ($r = 0.31$), and C-terminal telopeptide of type I collagen (CTX; $r = 0.22$) [10]. Validation of the RSS in different forms of rickets and the high inter- and intra-rater reliability support the validity and utility of this assessment in clinical research. The RSS allows readers to be blinded to timing of each radiograph, thereby reducing potential bias. Finally, the greater statistical power of a scaled ordinal variable is important for assessing treatment effects in clinical trials of rare diseases like XLH where the sample size is small.

The RSS is limited by a lack of comparison with previous radiographs, which may prevent detecting small changes; however, even assessments performed at isolated time points remain useful. The RGI-C can be utilized as a complementary method to the RSS to assess change in a side-by-side comparison of radiographs. Variable thickness of lucency and the proportion of growth plate involvement are given the same score with the RSS, and the software used in this study did not allow for < 0.5 increments in the RSS. The RSS does not evaluate the degree of leg deformity, which may also have an impact on function.

5. Conclusions

We demonstrated that higher RSS is associated with greater biochemical, clinical, and functional impairments in children with XLH. We further show the reliability of the RSS to assess rickets severity in XLH and radiographic changes following treatment.

Disclosures

A Skrinar, M Mao, C-Y Chen, T Chang, and J San Martin are employees of Ultragenyx Pharmaceutical Inc. T.D. Thacher, J.M. Pettifor, P. J. Tebben, A. L. Creo, and T. O. Carpenter have consulted for Ultragenyx Pharmaceutical Inc. T. D. Thacher has consulted for Kyowa Kirin Pharmaceutical Development, Inc. T.O. Carpenter and P. J. Tebben have received grant support and served as investigator for studies sponsored by Ultragenyx. This study was sponsored and funded by Ultragenyx Pharmaceutical Inc. in partnership with Kyowa Kirin International plc.

Acknowledgments

We thank the patients, parents, caregivers, and the healthcare professionals who participated in Study UX023-CL201, and Catherine Woods, PhD, and Minna Balbas, PhD, of Ultragenyx Pharmaceutical Inc. for editorial support. Funding for the study was supported by Ultragenyx Pharmaceutical Inc.

Authors' roles

Study design: AS, JSM, and TOC. Study conduct: TDT, PJT, ALC, JMP, AS, JSM, and TOC. Data collection: MM, C-Y C. Data analysis: AS, MM, C-Y C, JSM, and TOC. Data interpretation: TDT, PJT, ALC, JMP, AS, JSM, TOC. Drafting manuscript: TDT and TC. Revising manuscript content: TDT, PJT, ALC, JMP, AS, MM, C-Y C, JSM, and TOC. All authors approved the final version of the manuscript and take responsibility for the integrity of the data analysis.

References

[1] M.F. Holick, Resurrection of vitamin D deficiency and rickets, J. Clin. Invest. 116 (8) (2006) 2062–2072, https://doi.org/10.1172/JCI29449 (PubMed PMID: 16886050; PubMed Central PMCID: PMC1523417).

[2] T.T. Do, Clinical and radiographic evaluation of bowlegs, Curr. Opin. Pediatr. 13 (1) (2001) 42–46 (PubMed PMID: 11176242).

[3] R.M. Shore, R.W. Chesney, Rickets: part I, Pediatr. Radiol. 43 (2) (2013) 140–151, https://doi.org/10.1007/s00247-012-2532-x (PubMed PMID: 23208530).

[4] R.M. Shore, R.W. Chesney, Rickets: part II, Pediatr. Radiol. 43 (2) (2013) 152–172, https://doi.org/10.1007/s00247-012-2536-6 (PubMed PMID: 23179485).

[5] T.D. Thacher, P.R. Fischer, J.M. Pettifor, J.O. Lawson, B.J. Manaster, J.C. Reading, Radiographic scoring method for the assessment of the severity of nutritional rickets, J. Trop. Pediatr. 46 (3) (2000) 132–139 (PubMed PMID: 10893912).

[6] T.D. Thacher, P.R. Fischer, J.M. Pettifor, J.O. Lawson, C.O. Isichei, J.C. Reading, et al., A comparison of calcium, vitamin D, or both for nutritional rickets in Nigerian children, N. Engl. J. Med. 341 (8) (1999) 563–568, https://doi.org/10.1056/NEJM199908193410803 (PubMed PMID: 10451461. Epub 1999/08/19).

[7] T.D. Thacher, P.R. Fischer, J.M. Pettifor, Vitamin D treatment in calcium-deficiency rickets: a randomised controlled trial, Arch. Dis. Child. 99 (9) (2014) 807–811, https://doi.org/10.1136/archdischild-2013-305275 (PubMed PMID: 24748637; PubMed Central PMCID: PMCPMC4145444. Epub 2014/04/22).

[8] T.D. Thacher, T.J. Bommersbach, J.M. Pettifor, C.O. Isichei, P.R. Fischer, Comparison of limestone and ground fish for treatment of nutritional rickets in children in Nigeria, J. Pediatr. 167 (1) (2015) 148 e1–54 e1, https://doi.org/10.1016/j.jpeds.2015.02.008 (PubMed PMID: 25799193. Epub 2015/03/24).

[9] H.L. Jones, L. Jammeh, S. Owens, A.J. Fulford, S.E. Moore, J.M. Pettifor, et al., Prevalence of rickets-like bone deformities in rural Gambian children, Bone 77 (2015) 1–5, https://doi.org/10.1016/j.bone.2015.04.011 (PubMed PMID: 25871880; PubMed Central PMCID: PMCPMC4456426. Epub 2015/04/15).

[10] G.I. Baroncelli, A. Bereket, M. El Kholy, L. Audi, Y. Cesur, B. Ozkan, et al., Rickets in the Middle East: role of environment and genetic predisposition, J. Clin. Endocrinol. Metab. 93 (5) (2008) 1743–1750, https://doi.org/10.1210/jc.2007-1413 (PubMed PMID: 18285415. Epub 2008/02/21).

[11] K. Mondal, A. Seth, R.K. Marwaha, D. Dhanwal, S. Aneja, R. Singh, et al., A randomized controlled trial on safety and efficacy of single intramuscular versus staggered oral dose of 600 000IU vitamin D in treatment of nutritional rickets, J. Trop. Pediatr. 60 (3) (2014) 203–210, https://doi.org/10.1093/tropej/fmt105 (PubMed PMID: 24401754. Epub 2014/01/10).

[12] D. Chatterjee, V. Gupta, V. Sharma, B. Sinha, S. Samanta, A reliable and cost effective approach for radiographic monitoring in nutritional rickets, Br. J. Radiol. 87 (1036) (2014) 20130648, , https://doi.org/10.1259/bjr.20130648 (PubMed PMID: 24593231; PubMed Central PMCID: PMCPMC4067015. Epub 2014/03/07).

[13] V. Aggarwal, A. Seth, R.K. Marwaha, B. Sharma, P. Sonkar, S. Singh, et al., Management of nutritional rickets in Indian children: a randomized controlled trial, J. Trop. Pediatr. 59 (2) (2013) 127–133, https://doi.org/10.1093/tropej/fms058 (PubMed PMID: 23104564. Epub 2012/10/30).

[14] G.F. Combs Jr., N. Hassan, N. Dellagana, D. Staab, P. Fischer, C. Hunt, et al., Apparent efficacy of food-based calcium supplementation in preventing rickets in Bangladesh, Biol. Trace Elem. Res. 121 (3) (2008) 193–204, https://doi.org/10.1007/s12011-007-8053-z (PubMed PMID: 18180882. Epub 2008/01/09).

[15] C.M. Gordon, H.A. Feldman, L. Sinclair, A.L. Williams, P.K. Kleinman, J. Perez-Rossello, et al., Prevalence of vitamin D deficiency among healthy infants and toddlers, Arch. Pediatr. Adolesc. Med. 162 (6) (2008) 505–512, https://doi.org/10.1001/archpedi.162.6.505 (PubMed PMID: 18524739; PubMed Central PMCID: PMCPMC3206624. Epub 2008/06/06).

[16] J.M. Perez-Rossello, H.A. Feldman, P.K. Kleinman, S.A. Connolly, R.A. Fair, R.M. Myers, et al., Rachitic changes, demineralization, and fracture risk in healthy infants and toddlers with vitamin D deficiency, Radiology 262 (1) (2012) 234–241, https://doi.org/10.1148/radiol.11110358 (PubMed PMID: 22106354; PubMed Central PMCID: PMCPMC3244670. Epub 2011/11/23).

[17] M.P. Whyte, C.R. Greenberg, N.J. Salman, M.B. Bober, W.H. McAlister, D. Wenkert, et al., Enzyme-replacement therapy in life-threatening hypophosphatasia, N. Engl. J. Med. 366 (10) (2012) 904–913, https://doi.org/10.1056/NEJMoa1106173 (PubMed PMID: 22397652).

[18] T. Kitaoka, T. Tajima, K. Nagasaki, T. Kikuchi, K. Yamamoto, T. Michigami, et al., Safety and efficacy of treatment with asfotase alfa in patients with hypophosphatasia: results from a Japanese clinical trial, Clin. Endocrinol. 87 (1) (2017) 10–19, https://doi.org/10.1111/cen.13343 (PubMed PMID: 28374482. Epub 2017/04/05).

[19] T.O. Carpenter, E.A. Imel, I.A. Holm, S.M. Jan de Beur, K.L. Insogna, A clinician's guide to X-linked hypophosphatemia, J. Bone Miner. Res. 26 (7) (2011) 1381–1388, https://doi.org/10.1002/jbmr.340 (PubMed PMID: 21538511; PubMed Central PMCID: PMC3157040).

[20] M. Lempicki, A. Rothenbuhler, V. Merzoug, S. Franchi-Abella, C. Chaussain, C. Adamsbaum, et al., Magnetic resonance imaging features as surrogate markers of X-linked hypophosphatemic rickets activity, Horm. Res. Paediatr. 87 (4) (2017) 244–253, https://doi.org/10.1159/000464142 (PubMed PMID: 28376474).

[21] O. Makitie, A. Doria, S.W. Kooh, W.G. Cole, A. Daneman, E. Sochett, Early treatment improves growth and biochemical and radiographic outcome in X-linked hypophosphatemic rickets, J. Clin. Endocrinol. Metab. 88 (8) (2003) 3591–3597.

[22] T.O. Carpenter, M.P. Whyte, E.A. Imel, A.M. Boot, W. Hogler, A. Linglart, et al., Burosumab therapy in children with X-linked hypophosphatemia, N. Engl. J. Med. 378 (21) (2018) 1987–1998, https://doi.org/10.1056/NEJMoa1714641 (PubMed PMID: 29791829. Epub 2018/05/24).

[23] M.P. Whyte, K.P. Fujita, S. Moseley, D.D. Thompson, W.H. McAlister, Validation of a novel scoring system for changes in skeletal manifestations of hypophosphatasia in newborns, infants, and children: the radiographic global impression of change scale, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 33 (5) (2018) 868–874, https://doi.org/10.1002/jbmr.3377 (PubMed PMID: 29297597. Epub 2018/01/04).

[24] L.H. Daltroy, M.H. Liang, A.H. Fossel, M.J. Goldberg, The POSNA pediatric musculoskeletal functional health questionnaire: report on reliability, validity, and sensitivity to change. Pediatric Outcomes Instrument Development Group. Pediatric Orthopaedic Society of North America, J. Pediatr. Orthop. 18 (5) (1998) 561–571 (PubMed PMID: 9746401).

[25] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1) (1977) 159–174 (Epub 1977/03/01. PubMed PMID: 843571).

[26] A. Linglart, M. Biosse-Duplan, K. Briot, C. Chaussain, L. Esterle, S. Guillaume-Czitrom, et al., Therapeutic management of hypophosphatemic rickets from infancy to adulthood, Endocr. Connect. 3 (1) (2014), https://doi.org/10.1530/EC-13-0103 (PubMed Central PMCID: PMCPMC3959730. Epub Mar 15, 2014).