Inter-rater reliability of the radiographic assessment of simple bone cysts

S. Cho1
R. Yankanah1
P. Babyn2
J. Stimec3
A. S. Doria3
D. Stephens4
J. G. Wright1

Abstract

Purpose To develop and evaluate the reliability of an explicit set of parameters and criteria for simple bone cysts (SBCs) and evaluate the reliability of single versus serial chronological reading methods.

Methods Radiographic criteria were developed based on the literature and expert consensus. A single anteroposterior/lateral radiograph from 32 subjects with SBC were evaluated by three radiologists. A second reading was then conducted using revised criteria including a visual schematic. In the third reading the same images were assessed but radiologists had access to images from two additional time points. Inter-rater reliability was assessed after each reading using kappa (κ) and percentage agreement for categorical and binary parameters and intra-class correlation coefficient (ICC) for continuous parameters.

Results Parameters that were revised with more explicit definitions including the visual schematic demonstrated consistent or improved inter-rater reliability with the exception of continuous cortical rim present and cyst location in the metaphysis and mid-diaphysis. Cortical rim displayed only slight reliability throughout (κ = -0.008 to 0.16). All other categorical parameters had a percentage agreement above 0.8 or a moderate (κ = 0.41 to 0.60), substantial (κ = 0.61 to 0.80) or almost perfect inter-rater reliability (κ = 0.81 to 1.0) in at least one reading. All continuous parameters demonstrated excellent inter-rater reliability (ICC > 0.75) in at least one reading with the exception of scalloping (ICC = 0.37 to 0.70). Inter-rater reliability values did not indicate an obviously superior method of assessment between single and serial chronological readings.

Conclusion Explicit criteria for SBC parameters used in their assessment demonstrated improved and substantial inter-rater reliability. Inter-rater reliability did not differ between single and serial chronological readings.

Level of Evidence: Not Applicable

Cite this article: Cho S, Yankanah R, Babyn P, Stimec J, Doria AS, Stephens D, Wright JG. Inter-rater reliability of the radiographic assessment of simple bone cysts. J Child Orthop 2019;13:226-235. DOI: 10.1302/1863-2548.13.180140

Keywords: simple bone cyst; unicameral bone cyst; bone cyst; inter-rater reliability; radiography

Introduction

Simple bone cysts (SBC) otherwise known as unicameral bone cysts are the most common benign bone lesions in children and adolescents.1,2 Patients with SBCs usually present with fracture or pain in the affected limb.3-5 SBCs have a high probability of recurrence and repeat fracture after treatment,3,6-10 and may even cause discrepant limb lengths or angular deformity.1,4,7,11 While benign, SBCs can have a life-long impact on children through risk of fracture and activity restriction, thus, clinicians need reliable measures for their assessment.8,10,12

The radiographic features of SBC can include those used for diagnosis such as the location of the cyst, prognosis such as the cyst index,13 or as an outcome of treatment, for instance, the Neer classification.14 Other SBC features are also consistent with various benign bone lesions, such as cortical reaction. However, few of these parameters have been shown to be reliable and the absence of specific criteria may partially explain the conflicting findings as to which characteristics can be used to predict healing or fracture.12,13

Beyond the criteria used to evaluate an evolving disease such as SBC, whether radiographs should either be interpreted individually without review of prior images or assessed in a chronological series has received little...
attention in the orthopaedic literature. While serial chronological readings may introduce bias from prior examinations, it may be advantageous to have an illustration of the chronological progression of the disease which could lead to assessments more sensitive to change. The purpose of this study was: a) to evaluate the reliability of radiographic criteria used in diagnosis, prognosis or outcome of SBC as a typical benign bone lesion; and b) to evaluate the reliability of a single versus serial chronological evaluation of radiographs for SBCs.

Materials and methods

Patient sample

The patient sample consisted of 32 patients aged four to 14 years (median ten years) with a radiographic diagnosis of a SBC. All participants were previously enrolled in a multi-centre trial for the treatment of a SBC. For each study subject, an anteroposterior (AP) and lateral (Lat) radiograph of the cyst at baseline, one year and two years post-treatment were obtained and de-identified. Approval for this study was granted by the institutional research ethics board of the Hospital for Sick Children, Toronto, Ontario, Canada.

Development of radiographic parameters

A literature search using MEDLINE was conducted to collate radiographic characteristics of SBCs. The following search string was used: bone cysts/and (unicameral or multcameral or simple), radiography, treatment outcome, treatment failure and aneurysmal bone cysts. The search was limited to the English language and filtered to include humans and children 0 to 18 years old. Redundant parameters describing similar characteristics were consolidated or eliminated. Parameters for characterization of bone cysts were defined either by consensus of authors or by modified criteria from existing definitions in the literature (Table 1). In some cases, variables were assigned reference images (Fig. 1) and those parameters related to continuous data such as cortical thinning were divided into categorical scales such as 0% to 25%, 25% to 50%, etc. Factors were grouped into those used for diagnosis, prognosis and outcome assessment of SBC.

Radiographic assessment

The readers involved were three board-certified paediatric musculoskeletal radiologists with 29, seven and 13 years of experience after training, respectively (PB, ASD, JS). Each radiograph was assigned a subject ID and images were read using picture archiving and communications system workstations with the default calibration settings and measurement tools available in clinical practice. The radiologists had no prior discussion regarding measurement or specific instructions other than the criteria provided. Data was recorded on a paper case report form and later inputted into a secure REDCap database.

First reading

AP/Lat images from one of the three time points available for each of the 32 patients was randomly selected and reviewed by the three radiologists. After an initial analysis, with the assistance of one of the radiologists (ASD) more precise definitions were added to the criteria with particular attention to those with poor reliability (Table 1). A visual schematic illustrating how to classify various cystic features was also created (Fig. 2).

Second reading

Upwards of one year after the first reading, the two radiologists (PB, JS) not involved in the revisions to the parameters evaluated the same 32 AP/Lat radiographs using the revised criteria, case report form and the visual schematic.

Third reading

At least two weeks after the second reading, the two radiologists (PB, JS) evaluated the same 32 radiographs using the criteria. However, each radiologist had the ability to view all three images for the corresponding patient simultaneously and had knowledge of their chronological order (serial chronological method).

Statistical analysis

Based on the number of radiologists per reading, sample size calculations yielded estimates of 32 patients required for the first reading and 19 required for the second and third reading. All inter-reader reliability was determined using the intra-class correlation coefficient (ICC) for continuous parameters and the kappa statistic ($\kappa$) and percentage agreement for categorical or binary parameters. ICC values range from 0 to 1 and values above 0.75 were considered to have excellent reliability. Kappa values range from -1 to 1 and were categorized as poor ($\kappa < 0$), slight ($\kappa = 0$ to 0.20), fair ($\kappa = 0.21$ to 0.40), moderate ($\kappa = 0.41$ to 0.60), substantial ($\kappa = 0.61$ to 0.80) and almost perfect ($\kappa = 0.81$ to 1). The kappa statistic tends to penalize the reliability of data with an imbalanced distribution that is heavily skewed towards positive agreement, as agreement attributed to chance is then assumed to be high which then lowers the kappa value. Therefore, percentage agreement was also used as a primary measure of reliability in this study.
Results

Out of the 32 patients, 22 (69%) were male and ten (31%) were female. There were 24 (75%) and eight (25%) SBCs located in the humerus and femur, respectively. A summary of the resulting reliability values for bone measurements and each of the parameters can be found in Table 2. All categorical parameters demonstrated moderate to almost perfect kappa values during at least one of the readings with the exception of the percentage of continuous cortical rim and cyst location, which consistently had a poor or slight kappa value. The percentage of scalloping present around the cyst border was the only continuous parameter to consistently demonstrate less than excellent reliability.

Within the diagnostic parameters, cyst location and periosteal reaction demonstrated variability in kappa values, however, both had percentage agreement values above 0.8 in at least one of the readings. Loculation had fair to substantial reliability across readings. Prognostic parameters based on continuous data demonstrating consistently excellent reliability included cyst dimensions, cyst volume and tubulation measurements. Cyst activity and indices demonstrated less than excellent reliability in the first reading only, otherwise they also showed excellent reliability. Prognostic parameters based on categorical data including the presence and absence of tubulation, whether the tubulation caused expansion to the bone and the presence and absence of scalloping demonstrated slight to substantial kappa values but had percentage agreement values above 0.8 across the second and third readings. The outcome parameter of fracture had fair to substantial kappa values but percentage agreement values above 0.8. Cyst grade had fair to moderate reliability.

Fig. 1 Product scale developed to represent the cyst grade spectrum.
### Table 1 Diagnostic, prognostic and outcome parameters and their definitions, criteria and interpretation when present in simple bone cysts

| Parameter                  | Definition and interpretation                                                                 | Initial radiographic criteria                           | Revised radiographic criteria                           |
|----------------------------|------------------------------------------------------------------------------------------------|--------------------------------------------------------|--------------------------------------------------------|
| **Diagnostic parameters**  |                                                                                                |                                                        |                                                        |
| Cyst location*             | Location of cyst within the bone.                                                             | Check all that apply:                                   | Check all that apply:                                   |
|                            | Are usually located in the proximal humerus or femur.*                                       | 1) metaphyseal (in flare of bone)                       | 1) metaphysis                                          |
|                            |                                                                                               | 2) epiphyseal (across physis)                           | 2) epiphysis                                           |
|                            |                                                                                               | 3) diaphyseal (below flare of bone)                     | 3) proximal diaphysis                                  |
|                            |                                                                                               | 4) proximal                                            | 4) middle diaphysis                                    |
|                            |                                                                                               | 5) middle                                              | 5) distal                                              |
|                            |                                                                                               | 6) distal                                              |                                                        |
| Loculation*                | Presence of well corticated cavities.                                                         | Number of loculations (well corticated cavities) present: | Number of intra-cystic loculations with a continuous contour present: |
|                            | Unilocular: having a single cavity.                                                            | 1) unilocular                                          | 1) unilocular                                          |
|                            | Multiloculated: having many small cavities, a greater indication of recurrence.7,4,27          | 2) 2 to 4                                              | 2) 2 to 4                                              |
|                            |                                                                                               | 3) 5 to 10                                             | 3) 5 to 10                                             |
|                            |                                                                                               | 4) > 10.                                               | 4) > 10.                                               |
| Cortical rim*              | Cortical bone surrounding the cyst.                                                           | Continuous cortical rim around cyst:                   | Continuous cortical rim around the cyst (% of total circumference): |
|                            | A thickened continuous border is indicative of healing.                                       | 1) absent                                              | 1) absent                                              |
|                            | An osteolytic rim is indicative of a higher risk of recurrence.23,18                           | 2) 1% to 24%                                           | 2) 1% to 24%                                           |
|                            | Incorporates radiodensity, zone of transition and geographic or permeative borders. A continuous cortical rim implies a greater percentage of radiodense rim, a sharper zone of transition and a geographic border. | 3) 25% to 49%                                         | 3) 25% to 49%                                         |
|                            |                                                                                               | 4) 50% to 74%                                         | 4) 50% to 74%                                         |
|                            |                                                                                               | 5) 75% to 99%                                         | 5) 75% to 99%                                         |
|                            |                                                                                               | 6) 100%.                                              | 6) 100%.                                              |
| Periosteal reaction*       | Callus formation of thick, dense, wavy uniform new bone resulting from irritation or injury.19 | 1) Absent                                             | 1) Absent                                             |
|                            |                                                                                               | 2) present.                                           |                                                        |
| Bone measurements*         | Dimensions of the bone in which the cyst is located.                                         | 1) Length of bone (middle of joint to middle of joint (AP)) | 1) Length of bone (from level of proximal extent of joint to level of distal extent of joint (AP)) |
|                            |                                                                                               | 2) width of nearest physis (AP).                      | 2) thickness/width of nearest physis to cyst (AP) |
| Prognostic parameters      |                                                                                                |                                                        |                                                        |
| Cyst activity*             | Distance of cyst to physis.                                                                   | Smallest distance from middle of physis to top or bottom of cyst | Shortest distance from middle of the nearest physis to the outermost top or bottom margin of the cyst |
|                            | Active cysts may increase fracture risk1,17.                                                  |                                                        |                                                        |
|                            | Active: within 1.0 cm to physis19                                                             |                                                        |                                                        |
|                            | Inactive: > 1.0 cm to physis19                                                                |                                                        |                                                        |
| Cyst volume                | Cyst length × width × depth                                                                   | Cyst measurements:                                    | Cyst measurements (from inner cortex to inner cortex): |
|                            | Increase in volume is indicative of recurrence and potential fracture risk.24,31            | 1) longest length (AP)                                | 1) longest length (from level of top to level of bottom of cyst (AP) |
|                            |                                                                                               | 2) widest width (AP)                                  | 2) widest width (AP)                                   |
|                            |                                                                                               | 3) widest depth (AP).                                 | 3) widest depth (Lat).                                 |
| Tubulation40               | Lateral growth and remodelling of the metaphysis into a tubular shape.40 Widening of the bone beyond normal diameter.47 | Defined as widening of the bone beyond normal diameter: | Defined as widening of the bone beyond normal diameter: |
|                            | Wide metaphysis: indicative of non-healing.40                                                 | 1) absent                                             | 1) absent                                             |
|                            |                                                                                               | 2) present                                            | 2) present                                            |
|                            |                                                                                               | 3) unable to measure.                                 | 3) unable to measure due to fracture.                  |
|                            | Only if tubulation present:                                                                  |                                                        |                                                        |
|                            | 4) widest width in area of tubulation                                                        |                                                        |                                                        |
|                            | 5) expected bone width in area of tabulation                                                 |                                                        |                                                        |

4,40,47
Table 1 Cont.

| Parameter | Definition and interpretation | Initial radiographic criteria | Revised radiographic criteria |
|-----------|------------------------------|------------------------------|-----------------------------|
| Scalloping* | Curves along the edges of the cyst border due to thinning of the wall.4 | Defined as thinning of cortex, measure based on two views: 1) absent 2) present 3) percentage of overall cyst border that is scalloped. | Defined as thinning of the cortex, measure based on two views: 1) absent 2) present 3) percentage of total circumference of the cyst. |
| Cyst index13 | Index is an assessment of the mechanical resistance of the cyst wall, with lower cyst indices associated with fracture and recurrence.3,10,12,13,29 | 1) Cyst length 2) cyst width 3) cyst depth 4) the narrowest diameter of the diaphysis on AP film (from outer cortex to outer cortex). | Cyst index equations:  
Cyst index\textsuperscript{13} = \frac{\text{cyst area}}{\text{diameter of diaphysis}^2}  
Cyst index\textsuperscript{13} = \frac{\text{cyst area}}{\text{diameter of diaphysis}^2}  
Cyst index = \frac{\text{cyst length} \cdot \text{cyst width} \cdot \text{cyst depth}}{\text{diameter of diaphysis}^2} |
| Outcome parameters | | | |
| Cyst grade\textsuperscript{14} | Cyst healing. Grade III and Grade IV cysts are considered to be satisfactorily healed.3 | Cole modification of the Neer Classification:1 1) Grade I – clearly visible 2) Grade II – visible but multilocular and opaque 3) Grade III – sclerosis around or within the partially visible cyst 4) Grade IV – complete healing with obliteration of the cyst. | A product scale with an image representative of each grade was provided for visual reference during assessment (Fig. 1). |
| Fracture | A break in the continuity of the bone as seen on the radiograph. Indicative of cyst recurrence. | 1) Absent 2) present. |

*parameters that were illustrated on the visual schematic created for the second reading

Bone measurements consistently demonstrated excellent reliability with the exception of the width of the nearest physis in the first reading.

Overall, the second reading using the revised criteria and schematic displayed either consistent (difference < 0.10) or higher reliability values compared with the first reading for all continuous parameters and for all categorical parameters with the exception of the location in the metaphysis and mid-diaphysis and cortical rim for percentage agreement. Using the schematic and revised criteria, loculations, cyst activity, widest width of tubulation, expected bone width in tubulated area, percentage of scalloping around the cortical rim and bone measurements all demonstrated improved reliability between the first and second reading.

Comparing single and serial-chronological methods of reading

Between the single and serial chronological series methods, inter-rater reliability values did not indicate an obviously superior method. Parameters for which kappa values and percentage agreement had improved by more than 0.2 (20%) were cyst location in the metaphysis and mid-diaphysis. Number of loculations had a kappa and percentage agreement value that decreased by more than 0.2 (20%). Continuous parameters had ICC values that remained stable across methods.

Discussion

Reliable radiographic assessment is essential for the diagnosis, prognosis and outcome of benign bone lesions. Although parameters such as the cyst grade, index and activity had explicitly defined criteria in the literature,\textsuperscript{3,13,14,29} the remainder of SBC parameters did not. This study initially evaluated the inter-rater reliability of parameters based on descriptions that were available, and after a schematic and precise criteria were added, reliability generally improved. Thus, the translation of these descriptive characteristics into specific binary categories, ordinal scales and quantitative measurements depicted in the
Fig. 2 Visual schematic developed to illustrate criteria for radiologists in the second and third readings.

revised criteria should guide clinicians when evaluating benign bone lesions such as SBCs. Many of the SBC parameters that were evaluated in this study are also relevant to other bone lesions.\textsuperscript{30-36} In contrast to prior studies, one notable trend is that the kappa values for the categorical parameters in our study displayed greater variance across readings than those reported in prior literature.\textsuperscript{30-33,35} The likely explanation is that whereas prior studies almost exclusively used binary scales (i.e. absent or present), this study used ordinal scales. Multiple categories provide clinicians with more specific information but may affect reliability.
| Parameter | First reading |  |  | Second reading |  |  | Third reading |  |  |
|-----------|---------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|           | n | % | κ | ICC | 95% CI | n | % | κ | ICC | 95% CI | n | % | κ | ICC | 95% CI |
| Metaphysis | 32 | 0.83 | 0.66 | 0.38 to 0.95 | 32 | 0.59 | 0.19 | -0.14 to 0.53 | 32 | 0.91 | 0.81 | 0.61 to 1.00 |
| Epiphysis | 32 | 0.97 | 0.01 | -0.99 to 1.0 | 32 | 0.88 | -0.067 | -0.13 to -0.0015 | 32 | 0.97 | 0.00 | 0.0 to 0.0 |
| Cyst location | 32 | 0.75 | 0.21 | -0.11 to 0.53 | 32 | 0.84 | -0.053 | -0.14 to 0.033 | 32 | 0.94 | 0.63 | 0.16 to 1.00 |
| Prox. diaphysis | 32 | 0.85 | 0.70 | 0.41 to 0.98 | 32 | 0.63 | 0.26 | -0.061 to 0.58 | 32 | 0.88 | 0.75 | 0.53 to 0.97 |
| Mid diaphysis | 32 | 0.97 | 0.01 | -0.99 to 1.0 | 32 | 0.97 | 0.00 | 0.0 to 0.0 | 32 | 0.97 | 0.00 | 0.0 to 0.0 |
| Dist. diaphysis | 32 | 0.97 | 0.01 | -0.99 to 1.0 | 32 | 0.97 | 0.00 | 0.0 to 0.0 | 32 | 0.97 | 0.00 | 0.0 to 0.0 |
| Loculation | 31 | 0.48 | 0.29 | 0.16 to 0.42 | 32 | 0.66 | 0.65 | 0.45 to 0.84 | 32 | 0.34 | 0.41 | 0.22 to 0.59 |
| Number of loculations | 30 | 0.30 | 0.16 | 0.06 to 0.26 | 32 | 0.094 | 0.005 | -0.005 to 0.016 | 32 | 0.00 | -0.008 | -0.064 to 0.049 |
| Cortical rim | 32 | 0.92 | 0.66 | 0.45 to 0.87 | 32 | 0.84 | 0.45 | 0.047 to 0.86 | 32 | 0.66 | 0.068 | -0.14 to 0.28 |
| % continuous cortical rim | 32 | 0.92 | 0.66 | 0.45 to 0.87 | 32 | 0.84 | 0.45 | 0.047 to 0.86 | 32 | 0.66 | 0.068 | -0.14 to 0.28 |
| Periosteal reaction | 32 | 0.92 | 0.66 | 0.45 to 0.87 | 32 | 0.84 | 0.45 | 0.047 to 0.86 | 32 | 0.66 | 0.068 | -0.14 to 0.28 |
| Cyst activity | 29 | - | - | 0.73 | 0.56 to 0.85 | 31 | - | - | 0.89 | 0.78 to 0.95 | 28 | - | - | 0.95 | 0.90 to 0.98 |
| Distance from physis to cyst top | 31 | - | - | 0.91 | 0.84 to 0.95 | 30 | - | - | 0.83 | 0.68 to 0.92 | 30 | - | - | 0.83 | 0.67 to 0.91 |
| Longest length | 31 | - | - | 0.83 | 0.72 to 0.91 | 31 | - | - | 0.95 | 0.91 to 0.98 | 32 | - | - | 0.95 | 0.90 to 0.98 |
| Widest width | 31 | - | - | 0.92 | 0.83 to 0.96 | 31 | - | - | 0.85 | 0.71 to 0.92 | 32 | - | - | 0.92 | 0.83 to 0.96 |
| Cyst volume | 30 | - | - | 0.95 | 0.89 to 0.97 | 30 | - | - | 0.92 | 0.84 to 0.96 | 30 | - | - | 0.90 | 0.80 to 0.95 |
| Greatest depth | 31 | - | - | 0.91 | 0.84 to 0.95 | 30 | - | - | 0.83 | 0.68 to 0.92 | 30 | - | - | 0.83 | 0.67 to 0.91 |
| Cyst volume | 30 | - | - | 0.92 | 0.83 to 0.96 | 31 | - | - | 0.91 | 0.84 to 0.96 | 32 | - | - | 0.90 | 0.80 to 0.95 |
| Change in volume | 32 | 0.89 | 0.63 | 0.42 to 0.83 | 30 | 0.90 | 0.00 | -0.00 to 0.0 | 32 | 0.88 | 0.46 | 0.04 to 0.88 |
| Present/absent | 32 | 0.92 | 0.66 | 0.45 to 0.87 | 32 | 0.84 | 0.45 | 0.047 to 0.86 | 32 | 0.66 | 0.068 | -0.14 to 0.28 |
| Bone expansion | 32 | 0.92 | 0.66 | 0.45 to 0.87 | 32 | 0.84 | 0.45 | 0.047 to 0.86 | 32 | 0.66 | 0.068 | -0.14 to 0.28 |
| Widest width of tub. | 22 | - | - | 0.97 | 0.93 to 0.99 | 27 | - | - | 0.99 | 0.97 to 0.99 | 26 | - | - | 0.98 | 0.96 to 0.99 |
| Expected bone width | 22 | - | - | 0.85 | 0.73 to 0.93 | 27 | - | - | 0.97 | 0.93 to 0.98 | 26 | - | - | 0.94 | 0.87 to 0.97 |
| Present/absent | 32 | 0.92 | 0.66 | 0.45 to 0.87 | 32 | 0.84 | 0.45 | 0.047 to 0.86 | 32 | 0.66 | 0.068 | -0.14 to 0.28 |
| Scallop | 14 | - | - | 0.37 | 0.037 to 0.70 | 26 | - | - | 0.70 | 0.43 to 0.85 | 22 | - | - | 0.67 | 0.36 to 0.85 |
| Percent of border | 31 | - | - | 0.14 | -0.062 to 0.39 | 30 | - | - | 0.88 | 0.77 to 0.94 | 29 | - | - | 0.85 | 0.70 to 0.92 |
| Cyst index | 30 | - | - | 0.31 | -0.05 to 0.60 | 30 | - | - | 0.86 | 0.72 to 0.93 | 29 | - | - | 0.84 | 0.68 to 0.92 |
| Cyst index A | 32 | 0.81 | 0.37 | -0.21 to 0.53 | 31 | 0.90 | 0.71 | 0.42 to 1.0 | 32 | 0.81 | 0.55 | 0.24 to 0.85 |
| Cyst index B | 32 | 0.47 | 0.27 | -0.15 to 0.39 | 32 | 0.66 | 0.57 | 0.37 to 0.77 | 32 | 0.59 | 0.47 | 0.27 to 0.67 |
| Presence of fracture | 20 | - | 1.00 | 1.0 to 1.0 | 19 | - | - | 1.00 | 1.0 to 1.0 | 18 | - | - | 1.00 | 1.0 to 1.0 |
| Bone measurements | 32 | 0.81 | 0.37 | -0.21 to 0.53 | 31 | 0.90 | 0.71 | 0.42 to 1.0 | 32 | 0.81 | 0.55 | 0.24 to 0.85 |

Single hyphens indicate value not calculated due to data classification as continuous or categorical. Dashed lines are indicative of a parameter that was irrelevant to that particular stage of reading.
Diagnostic factors, such as the cyst’s location, number of loculations, the presence of a cortical rim, and evidence of periosteal reaction have been used to identify SBCs. The inter-rater reliability of the number of loculations and periosteal reaction in this study were similar or better compared with studies on other benign bone lesions such as aneurysmal bone cysts or malignant bone lesions, and are useful parameters. The criteria for the cortical rim demonstrated consistently poor reliability in this study due to the ambiguity of the cyst walls. Comparable reliability values for cortical rim in literature have varied depending on whether the study evaluated cortical rim thickening or destruction. The lack of a consistent method of assessment and the variability in reliability indicates this parameter as an unreliable factor, therefore, we recommend against its use to assess SBCs.

Prognostic information such as scalloping of the cyst wall and tubulation (widening) have been associated with non-healing and cyst recurrence. Additionally, active cysts or an increase in cyst volume have been associated with fracture. Prior studies of lesions such as enchondroma and chondrosarcoma have demonstrated high percent agreement with slight to fair kappa values. The reliability of tubulation, lesion activity and volume have not been previously investigated but demonstrated excellent values in this study and thus are useful parameters. The cyst index as proposed by Kaelin and McEwen is another prognostic tool that has been used to assess SBCs. Vasconcellos et al. reported poor inter-observer reliability most likely due to the use of trapezoidal shapes to approximate the cyst area for the equation. This can create unreliability due to the variable configuration of SBCs. To address this, one study used tracing software to determine the area and reported substantial reliability. The current study created two modified cyst indices (A and B) based on a rectangular approximation. Both demonstrated excellent reliability in the second and third readings. However, as Cyst index B was more reliable in the first reading, more specific criteria and marking for measurements and assisted in cases where images within a series varied in quality and thus should be the preferred method.

In conclusion, the majority of the radiographic criteria developed for the diagnosis, prognosis and outcome of SBC were found to be reliable and improved with explicit criteria and a visual schematic. Inter-rater reliability did not differ between single and serial-chronological readings.

Received 09 August 2018; accepted after revision 14 January 2019.

COMPLIANCE WITH ETHICAL STANDARDS

FUNDING STATEMENT

This study was funded by the Canadian Institutes of Health Research (CIHR, FRN 130272).

OA LICENCE TEXT

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0) licence (https://creativecommons.org/licenses/by-nc/4.0/)
INTER-RATER RELIABILITY OF SBC RADIOGRAPH CRITERIA

Reproducibility
The equivalence of weighted kappa and the intraclass correlation coefficient was assessed to determine reproducibility (ranging from 0.80 to 1.00).

Pathological fractures secondary to unicameral bone cysts.
Behavior and interpretation of the kappa statistic: a review of the literature.

Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:37-81.

Fleiss JL. The design and analysis of clinical experiments. First ed. New York, NY: John Wiley and Sons, 1986.

Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1971;31:447-451.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-174.

Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 1990;43:543-549.

Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: resolution of the two paradoxes. J Clin Epidemiol 1996;49:431-434.

Warrens MJ. A formal proof of a paradox associated with Cohen’s kappa. J Classif 2010;27:322-332.

Shankar V, Bangdiwala SI. Observer agreement paradoxes in 2x2 tables: comparison of agreement measures. BMJ Med Res Methodol 2014;14:100.

Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-268.

Nelson JC, Pepe MS. Statistical description of interrater variability in ordinal ratings. Stat Methods Med Res 2000;9:475-496.

Oppenheim WL, Galleno H. Operative treatment versus steroid injection in the management of unicameral bone cysts. J Pediatr Orthop 1984;4:1-7.

Mahnken AH, Nolte-Ernsting CC, Wildberger JE, et al. Aneursymal bone cyst: value of MR imaging and conventional radiography. Eur Radiol 2003;13:1118-1124.

Geirnaerdt MJ, Hermans J, Bloem JL, et al. Usefulness of radiography in differentiating enchondroma from central grade 1 chondrosarcoma. AJR Am J Roentgenol 1997;169:1097-1104.

Goursiotyanni S, Hwang S, Panicek DM, et al. Reproducibility and clinical correlations of post-treatment changes on CT of prostate cancer bone metastases treated with chemotherapy. Br J Radiol 2012;85:1243-1249.
33. **Gajewski DA, Burnette JB, Murphey MD, Temple HT.** Differentiating clinical and radiographic features of enchondroma and secondary chondrosarcoma in the foot. *Foot Ankle Int* 2006;27:240-244.

34. **Vasconcellos DA, Yandow SM, Grace AM, et al.** Cyst index: a nonpredictor of simple bone cyst fracture. *J Pediatr Orthop* 2007;27:307-310.

35. **Cha SM, Shin HD, Kim KC, Kang DH.** Flexible intramedullary nailing in simple bone cysts of the proximal humerus: prospective study for high-risk cases of pathologic fracture. *J Pediatr Orthop B* 2013;22:475-480.

36. **Cho HS, Seo SH, Park SH, et al.** Minimal invasive surgery for unicameral bone cyst using demineralized bone matrix: a case series. *BMC Musculoskelet Disord* 2012;13:33.

37. **Campanacci M, Capanna R, Picci P.** Unicameral and aneurysmal bone cysts. *Clin Orthop Relat Res* 1986;204:25-36.

38. **Chang CH, Stanton RP, Glutting J.** Unicameral bone cysts treated by injection of bone marrow or methylprednisolone. *J Bone Joint Surg [Br]* 2002;84-B: 407-412.

39. **Ragsdale BD, Madewell JE, Sweet DE.** Radiologic and pathologic analysis of solitary bone lesions. Part II: periosteal reactions. *Radiol Clin North Am* 1981;19:749-783.

40. **Manaster BJ, May DA, Disler DG.** Musculoskeletal imaging: The requisites. 4th ed. Philadelphia: Elsevier Health Sciences, 2013.

41. **Glowacki M, Ignys-O’Byrne A, Ignys I, Mankowski P, Melzer P.** Evaluation of volume and solitary bone cyst remodeling using conventional radiological examination. *Skeletal Radiol* 2010;39:251-259.

42. **Ulici A, Balanescu R, Topor L, Barbu M.** The modern treatment of the simple bone cysts. *J Med Life* 2012;5:469-473.

43. **Ferrara R, Priolo F, Cammisa M, et al.** Clinical trials in rheumatoid arthritis: methodological suggestions for assessing radiographs arising from the GRISAR Study. Gruppo Reumatologi Italiani Studio Artrite Reumatoide. *Ann Rheum Dis* 1997;56:608-612.

44. **Wright JG, Feinstein AR.** Improving the reliability of orthopaedic measurements. *J Bone Joint Surg [Br]* 1992;74-B:287-291.