



This is a repository copy of *Inter- and intra-observer reliability of contrast-enhanced magnetic resonance imaging parameters in children with suspected juvenile idiopathic arthritis of the hip.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/138211/>

Version: Accepted Version

Article:

Porter-Young, F.M. orcid.org/0000-0003-1217-0958, Offiah, A.C. orcid.org/0000-0001-8991-5036, Broadley, P. et al. (4 more authors) (2018) Inter- and intra-observer reliability of contrast-enhanced magnetic resonance imaging parameters in children with suspected juvenile idiopathic arthritis of the hip. *Pediatric Radiology*, 48 (13). pp. 1891-1900. ISSN 0301-0449

<https://doi.org/10.1007/s00247-018-4216-7>

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1007/s00247-018-4216-7>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Pediatric Radiology

Diagnostic yield of double reading initial skeletal surveys versus follow-up skeletal surveys for suspected child physical abuse

--Manuscript Draft--

Manuscript Number:	PRAD-D-18-00527
Full Title:	Diagnostic yield of double reading initial skeletal surveys versus follow-up skeletal surveys for suspected child physical abuse
Article Type:	Original Article
Abstract:	<p>Background Follow-up skeletal surveys (FUSS) are performed in cases of suspected child physical abuse. However, the yield of FUSS compared to double-reading initial skeletal surveys (SS) is not known.</p> <p>Objective To compare the diagnostic yield of FUSS performed for suspected child abuse to double reading initial exams.</p> <p>Materials and methods All initial SS performed between 2/2/2013-3/23/2015 for suspected physical abuse were double-read. McNemar test was applied to compare: 1) yield of FUSS after only single reading versus double reading initial SS; 2) yield of FUSS after only single reading initial SS versus the yield of FUSS after double reading initial SS.</p> <p>Results During the study period, 1056 initial SS were performed (M:F=617:439; age range=2 days-9 years; IQR=4-18 months). Of 293/1056 (28%) cases with FUSS, double reading initial SS showed 30/293 (10%) additional findings. FUSS showed additional findings in 32/263 (12%) not identified by double reading. The difference between the diagnostic yield of FUSS (62/293, 21%) compared to double reading (30/293, 10%) was significant, $p < 0.0001$. Similarly, the difference between the diagnostic yield of FUSS after only single reading initial SS (62/293, 21%) compared to the diagnostic yield of FUSS after double reading (37/293, 13%) was also significant, $p < 0.0001$.</p> <p>Conclusion FUSS identified significantly more new findings than double reading initial SS. However, double reading before performing FUSS significantly decreased the yield of the FUSS. These results show the benefit of double reading and the need for further research to determine if double reading initial SS can obviate follow-up exams in select cases.</p>

1 Original article

2 **Diagnostic yield of double reading initial skeletal surveys versus follow-up**
3 **skeletal surveys for suspected child physical abuse**

4

5 **Abstract**

6 **Background** Follow-up skeletal surveys (FUSS) are performed in cases of suspected child
7 physical abuse. However, the yield of FUSS compared to double-reading initial skeletal surveys
8 (SS) is not known.

9 **Objective** To compare the diagnostic yield of FUSS performed for suspected child abuse to
10 double reading initial exams.

11 **Materials and methods** All initial SS performed between 2/2/2013-3/23/2015 for suspected
12 physical abuse were double-read. McNemar test was applied to compare: 1) yield of FUSS after
13 only single reading versus double reading initial SS; 2) yield of FUSS after only single reading
14 initial SS versus the yield of FUSS after double reading initial SS.

15 **Results** During the study period, 1056 initial SS were performed (M:F=617:439; age range=2
16 days-9 years; IQR=4-18 months). Of 293/1056 (28%) cases with FUSS, double reading initial SS
17 showed 30/293 (10%) additional findings. FUSS showed additional findings in 32/263 (12%) not
18 identified by double reading. The difference between the diagnostic yield of FUSS (62/293, 21%)
19 compared to double reading (30/293, 10%) was significant, $p<0.0001$. Similarly, the difference
20 between the diagnostic yield of FUSS after only single reading initial SS (62/293, 21%)
21 compared to the diagnostic yield of FUSS after double reading (37/293, 13%) was also
22 significant, $p<0.0001$.

23 **Conclusion** FUSS identified significantly more new findings than double reading initial SS.
24 However, double reading before performing FUSS significantly decreased the yield of the FUSS.
25 These results show the benefit of double reading and the need for further research to determine if
26 double reading initial SS can obviate follow-up exams in select cases.

27

28 **Introduction**

29 Child physical abuse is a leading cause of traumatic injury in United States children. In
30 2016 alone, 123,032 U.S. children were documented victims of physical abuse and 639 died as a
31 consequence [1]. Survivors of physical child abuse often suffer long-term consequences. Studies
32 have shown that physically abused children are more prone to physical and mental health
33 conditions as adults including depression and drug abuse, as well as premature mortality [2-5].
34 Because of the magnitude of this problem and its life altering consequences, early recognition
35 and correct diagnosis of physical abuse is essential. Failure to do so can result in a child
36 returning to a hostile environment, putting him or her at risk for repeat injury [6, 7].

37 Accurate diagnosis of physical abuse is not straightforward. More than 80% of children
38 affected are under the age of 18 months and therefore cannot express themselves verbally [8]. In
39 addition, victims often do not have external signs of trauma [9, 10]. One third to half of
40 radiologically detected fractures are clinically occult [11, 12]. Consequently, the American
41 College of Radiology recommends that clinicians order skeletal surveys for any child less than
42 24 months of age in whom physical abuse is suspected, as well as a follow-up skeletal survey
43 two weeks later if the initial exam is abnormal to evaluate for signs of healing at known and
44 previously undetected fracture sites [13].

45 Multiple studies have shown that follow-up skeletal surveys can improve diagnostic yield
46 by identifying additional fractures and adding other helpful information such as fracture age [10,
47 14-16] . We hypothesize that in many cases, additional findings on follow-up skeletal surveys are
48 not new but were in fact present and detectable on the initial skeletal survey. The purpose of this
49 study is to compare the diagnostic yield of follow-up skeletal surveys performed for suspected
50 physical abuse to double reading initial exams.

51

52 **Materials and Methods**

53 Our institutional review board approved this HIPAA-compliant study and the need for
54 written informed consent was waived. All children who underwent a skeletal survey at our
55 tertiary care academic pediatric hospital and affiliate facilities between February 2, 2013 and
56 March 23, 2015 were eligible for inclusion. Patients whose skeletal surveys were performed for
57 indications other than suspected physical abuse were excluded.

58 Skeletal surveys (initial and follow-up) were performed per our institutional protocol and
59 included the following views: anteroposterior (AP) and lateral skull, AP chest and abdomen,
60 lateral spine, right and left oblique ribs; bilateral AP humerus, AP forearm, PA hand, AP femur,
61 AP tibia and fibula; and AP foot. Initial and follow-up skeletal surveys were primarily
62 interpreted by one of 19 pediatric radiologists (2-41 years post-fellowship experience) as part of
63 the patient's standard clinical care. Within 72 hours of single reading, initial skeletal surveys
64 were double read by one of three non-blinded pediatric radiologists (41, 19, and 6 years post-
65 fellowship experience). Discordant results among single readings of initial skeletal surveys,
66 double readings of initial skeletal surveys, and follow-up skeletal surveys were adjudicated by
67 one of three pediatric radiologists (41, 19, and 6 years post-fellowship experience). During the

68 study period, double-reading skeletal surveys was not part of our routine clinical practice, and
69 therefore an official report of the double reading was not generated. However, discrepancies
70 between single and double reading were communicated to both the clinical team and the initial
71 interpreting radiologist; and radiology reports were addended as needed. Final interpretations
72 were communicated to the ordering physician or the institutional Child Abuse Pediatrics team.

73 Patient medical records were reviewed and patient gender and age at the time of the
74 initial skeletal survey recorded. Findings detected by single initial skeletal surveys, double
75 reading initial skeletal surveys, and follow-up skeletal surveys were recorded and compared for
76 each patient. Based on concordance among the interpretations, patients were designated into one
77 of six categories as shown in Table 1. For categories 5 and 6, follow-up skeletal survey
78 interpretations were compared to adjudicated single and double reading interpretations of the
79 initial skeletal surveys.

80 Contingency tables were produced to address two comparisons. The first is the diagnostic
81 yield of follow-up skeletal surveys (without double reading) compared to the diagnostic yield of
82 double reading initial skeletal surveys (Table 2). The second is the diagnostic yield of follow-up
83 skeletal surveys after only single reading initial skeletal surveys compared to the diagnostic yield
84 of follow-up skeletal surveys after double reading initial skeletal surveys (Table 3). McNemar
85 test was applied to assess statistical significance for each of these two comparisons.

86 A member of our institutional Child Abuse Pediatrics team (<BLINDED>) reviewed the
87 electronic medical record for all patients who had findings identified on follow-up skeletal
88 survey not identified on single reading initial skeletal surveys and subjectively determined
89 whether or not the diagnosis of the new findings altered patient management. New findings on
90 follow-up skeletal surveys were considered to have altered management if deemed to have

91 changed the presumed mechanism of injury, refuted an allegation of child abuse, or lead to the
92 diagnosis of a condition that predisposed the patient to pathologic fractures (e.g. osteogenesis
93 imperfecta).

94

95 **Results**

96 During the 26-month study period, 1056 children underwent an initial skeletal survey for
97 suspected physical abuse. All 1056 children were included in this study. The mean age of this
98 population at the time of the initial skeletal survey was 13.2 +/- 13.9 months; age range: 2 days
99 to 9 years; median age = 9 months; interquartile range 4 - 18 months); 42% (439/1056) females.
100 Of the 1056 children enrolled, 293 (28%) had follow-up skeletal surveys.

101 Figure 1 shows the number of patients assigned to each of the six categories from Table 1.
102 For cases without follow-up skeletal surveys, double reading identified additional findings in
103 21/763 (2.8%). Third readers confirmed all 21 new findings. For cases with follow-up exams,
104 single and double readings of initial skeletal surveys were concordant in 263/293 (90%) with
105 follow-up skeletal surveys showing no additional findings in 231/263 (88%) (Category 3) and
106 additional findings in 32/263 (12%) (Category 4). A third reader discarded a 33rd additional
107 fracture diagnosed on follow-up concluding the findings were not definitive. In 30/293 (10%)
108 initial skeletal surveys, additional findings were identified by double reading that were not
109 identified by single reading (Fig. 2 and 3). Third readers confirmed all 30 new findings. For
110 these 30 cases, follow-up skeletal surveys showed no additional findings in 25/30 (83%)
111 (Category 5) and additional findings in 5/30 (17%) (Category 6) not identified by either single or
112 double reading initial skeletal surveys. Third readers confirmed all 5 new findings.

113 Table 4 shows the break down of additional findings identified on double reading and

114 follow-up exams. Correction of false positive interpretations on single reading constituted 18%
115 (16/89) of the additional findings on double reading versus only 2% (2/89) of the additional
116 findings on follow-up exams. Missed fractures constituted 63% (56/89) of the new findings on
117 double reading and 94.3% (84/89) of the new findings on follow-up exams. Most of the new
118 fractures identified both on double reading and follow-up exams were missed rib fractures, 30.4%
119 and 41.2 % respectively.

120 The difference between the diagnostic yield of follow-up skeletal surveys after only
121 single reading initial skeletal surveys (62/293, 21%) compared to the diagnostic yield of double
122 reading initial skeletal surveys (30/293, 10%) was statistically significant - $p < 0.0001$ (Table 2).
123 Similarly, the difference between the diagnostic yield of follow-up skeletal surveys after only
124 single reading initial skeletal surveys (62/293, 21%) compared to the diagnostic yield of follow-
125 up skeletal surveys after double reading initial skeletal surveys (37/293, 13%) was also
126 statistically significant - $p < 0.0001$).

127 After review of the electronic medical record, it was determined that new findings
128 identified on follow-up skeletal surveys altered management in 19% of cases (12/62) by
129 changing the presumed mechanism of injury (9/12), refuting the allegation of child abuse (2/12),
130 or leading to the diagnosis of osteogenesis imperfecta (1/12). Of these, 8 deemed to have altered
131 management were not identified on either single or double reading of initial skeletal surveys
132

133 **Discussion**

134 In cases of suspected physical abuse where follow-up skeletal surveys were performed,
135 we found that double reading initial skeletal surveys showed additional findings in 10% of cases
136 which is approximately double the percentage found by Karmazyn et al (4.5%) [17]. In

137 comparison, follow-up skeletal surveys showed additional fractures in 21% of cases. These
138 results indicate that if choosing between double-reading initial skeletal surveys and performing
139 follow-up skeletal surveys, follow-up skeletal surveys will maximize additional fracture
140 identification in this patient population. However, we also found that double reading initial
141 skeletal surveys before performing follow-up skeletal surveys significantly decreased the number
142 of cases in which additional fractures were identified at follow-up. Therefore, these results also
143 indicate that for a certain subset of this patient population, double reading without follow-up may
144 be sufficient to maximize additional fracture identification.

145 The American college of Pediatrics along with the American College of Radiology
146 recommend a follow-up skeletal survey after 10-21 days if an initial skeletal survey has
147 abnormal findings. This recommendation is based in part on two papers authored by Zimmerman
148 et al. and Kleinman et al. which showed that follow-up skeletal surveys added information and
149 improved diagnostic accuracy for child abuse in 46% and 61% of cases respectively [10, 14] .
150 Authors advocating follow-up skeletal surveys argue that these exams detect fractures not visible
151 initially, differentiate suspected fractures from normal variants, and clarify the age of previously
152 described fractures [14, 18] . It is also argued that follow-up skeletal surveys offer a more
153 thorough assessment when it comes to acute rib and metaphyseal fractures, which have high
154 specificity for abuse [10] .

155 While our study confirmed the results of prior authors that follow-up skeletal surveys
156 detect additional fractures not identified on initial skeletal surveys, the number of additional
157 fractures identified was significantly reduced by double reading initial exams. The time between
158 initial and follow-up skeletal surveys can delay final disposition and create uncertainty and

159 anxiety in cases where physical abuse did not occur or potentially allow a child to remain in an
160 unsafe environment when abusive trauma goes undiagnosed [3].

161 This study has several limitations. Of 1056 children enrolled, only 293 (28%) had follow-
162 up skeletal surveys. A follow-up skeletal survey is ordered for all children under 36 month of age
163 evaluated by our Child Abuse Pediatrics consult service for whom the impression is
164 indeterminate or concerning for abuse. Those children with a history that is plausible as an
165 accidental mechanism of injury do not have clinic follow up.

166 We do not know the reason why the follow-up skeletal surveys were performed in some
167 cases. This introduces a selection bias as the fracture yield of double reading initial skeletal
168 surveys for cases with follow-up skeletal surveys was 10% compared to 2.8% for cases without
169 follow-up skeletal surveys. The effect of double reading and follow-up skeletal surveys on
170 reader confidence was not assessed. It may be that one or both of these techniques not only
171 increased the number of fractures detected but also increased the diagnostic confidence of the
172 interpreting radiologist. Finally, we did not assess the effect of follow-up skeletal surveys on
173 determining fracture age.

174

175 **Conclusion**

176 Follow-up skeletal surveys were significantly better at identifying new fractures compared to
177 double reading initial skeletal surveys in cases of suspected physical abuse. However, double-
178 reading initial skeletal surveys before performing follow-up exams significantly decreased the
179 yield of the follow-up skeletal surveys indicating that these techniques may be complementary.
180 These results show the value of double-reading initial skeletal surveys and indicate the need for

181 further investigation to determine if double reading initial skeletal surveys can obviate the need
182 for follow-up exams in select cases.

183

184 **References**

- 185 1. U.S. Department of Health & Human Services, Administration for Children and Families,
186 Administration on Children, Youth and Families, Children's Bureau. (2018). Child
187 maltreatment 2016. Available from [https://www.acf.hhs.gov/cb/research-data-](https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/child-maltreatment)
188 [technology/statistics-research/child-maltreatment](https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/child-maltreatment).
- 189 2. Brown, D.W., et al., Adverse childhood experiences and the risk of premature mortality.
190 *Am J Prev Med*, 2009. 37(5): p. 389-96.
- 191 3. Christian, C.W., A. Committee on Child, and A.A.o.P. Neglect, The evaluation of
192 suspected child physical abuse. *Pediatrics*, 2015. 135(5): p. e1337-54.
- 193 4. Dube, S.R., et al., Childhood abuse, neglect, and household dysfunction and the risk of
194 illicit drug use: the adverse childhood experiences study. *Pediatrics*, 2003. 111(3): p. 564-
195 72.
- 196 5. Chapman, D.P., et al., Adverse childhood experiences and the risk of depressive disorders
197 in adulthood. *J Affect Disord*, 2004. 82(2): p. 217-25.
- 198 6. Sheets, L.K., et al., Sentinel injuries in infants evaluated for child physical abuse.
199 *Pediatrics*, 2013. 131(4): p. 701-7.
- 200 7. Singh RB, S.J., Fromkin JB, Berger RP, Assessing the use of follow-up skeletal surveys
201 in children with suspected physical abuse. *J of Trauma and Acute Care Surgery*, 2012.
202 73(4): p. 972-976.

- 203 8. Kemp AM, B.A., Morris S, Mann M, Kemp KW, Rolfe K, Sibert JR, Maguire S, Which
204 radiological investigations should be performed to identify fractures in suspected child
205 abuse? *Clin Radiol* 2006 61(9): p. 723-736.
- 206 9. Peters, M.L., et al., The presence of bruising associated with fractures. *Arch Pediatr*
207 *Adolesc Med*, 2008. 162(9): p. 877-81.
- 208 10. Zimmerman s, M.K., Care M, Thomas A, Shapiro R, Utility of follow-up skeletal surveys
209 in suspected child physical abuse evaluations. *Child Abuse Negl* 2005 29(10): p. 1075-
210 1083.
- 211 11. Sonik a, S.-W.R., Rogers KK, Coulter KP, Wootton-Gorges SL, Follow-up skeletal
212 surveys for suspected non-accidental trauma: can a more limited survey be performed
213 without compromising diagnostic information? *Child Abuse Negl* 2010 34(10): p. 804-
214 806.
- 215 12. Wood JN, C.C., Adams, CM, Rubin DM, Skeletal surveys in infants with isolated skull
216 fractures. *Pediatrics* 2009(123): p. 247 – 252.
- 217 13. American College of Radiology. ACR practice guideline for skeletal surveys in children
218 (Res. 47, 17, 35). In: American College of Radiology. *ACR Standards*. Reston, VA:
219 American College of Radiology; 2006:203– 207.
- 220 14. Kleinman PK, N.K., Spevak MR, Rayder SM, Madansky DL, Shelton YA, Patterson MM,
221 Follow-up Skeletal Surveys in Suspected Child Abuse. *AJR*, 1996(167): p. 893-896.
- 222 15. Melville, J.D., et al., Use of Imaging in Children With Witnessed Physical Abuse. *Pediatr*
223 *Emerg Care*, 2017.
- 224 16. Harper, N.S., et al., Follow-up skeletal survey use by child abuse pediatricians. *Child*
225 *Abuse Negl*, 2016. 51: p. 336-42.

- 226 17. Karmazyn, B., et al., Double-read of skeletal surveys in suspected non-accidental trauma:
227 what we learned. *Pediatr Radiol*, 2017. 47(5): p. 584-589.
- 228 18. Bennett BL, C.M., Care M, Kachelmeyer A, Mahabee-Gittens M, Retrospective review
229 to determine the utility of follow-up skeletal surveys in child abuse evaluations when the
230 initial skeletal survey is normal. *BMC Research Notes* 2011(4): p. 354-347.

231

232

233 **Legends**

234 **Fig. 1** Distribution of the patients among the six concordance categories

235 **Fig. 2** 2-year-old boy with suspected child physical abuse. Frontal radiograph of the left tibia
236 and fibula shows periosteal reaction at the mid tibia (black arrow) and nondisplaced distal tibial
237 fracture (white arrow) that were missed on single reading but identified at double reading

238 **Fig. 3** 2-months-old girl with suspected child physical abuse. AP left tibia/fibula radiographs (A:
239 initial exam and B: follow up exam) show a missed subphyseal metaphysis fracture of the
240 proximal tibia (white arrow) that was missed on the initial skeletal survey after single and double
241 reading but identified on the follow up exam

242

243 **Table 1.** Six categories reflecting the concordance among single readings of initial skeletal surveys, double readings
 244 of initial skeletal surveys, and follow up skeletal surveys

Category	Single and double readings of initial skeletal survey	Initial and follow-up skeletal surveys
1	Concordant	Not Performed
2	Discordant	Not Performed
3	Concordant	Concordant
4	Concordant	New findings at follow-up
5	Discordant	Concordant
6	Discordant	New findings at follow-up

245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268

Table 2. Contingency table for comparing the diagnostic yield of follow-up skeletal surveys versus the diagnostic yield of double reading initial skeletal surveys

		Additional findings on double reading	
		No	Yes
Additional findings on follow-up (without double-reading)	No	231 (Category 3)	0

269
270
271
272
273
274

	Yes	32 (Category 4)	30 (Category 5-6)
--	-----	--------------------	----------------------

275 **Table 3.** Contingency table for comparing the diagnostic yield of follow-up skeletal surveys after only single reading
 276 initial skeletal surveys versus the diagnostic yield of follow-up skeletal surveys after both double reading initial
 277 skeletal surveys
 278

		Additional findings on follow-up after double reading	
		No	Yes
Additional findings on follow-up after single reading only	No	231 (Category 3)	0
	Yes	25 (Category 5)	37 (Category 4,6)

279
 280
 281
 282

283 **Table 4.** The number of patients with additional findings identified on double reading and follow up exams per type
 284 of finding.
 285

		Number of patients with additional findings identified on:	
		Double reading initial exams	Follow up exams
Fractures	Ribs	8 (17)*	16 (35)*
	Metacarpals/phalanges	3 (3)*	3 (9)*
	Metatarsals/phalanges	6 (14)*	3 (8)*
	Long bones	6 (12)*	18 (27)*
	Acromion/clavicles	1 (1)*	2 (3)*
	Vertebral bodies	1 (3)*	2 (2)*
	Skull fractures	5 (6)*	0
Osteoporosis		6	2
Wide sutures		4	1
Bulging fontanel		2	0
Acuity		1	0
Failure to thrive		4	0
False positives/overcall		12 (16)*	2
Total		59 (89)*	49 (89)*

286
 287 (*) is the total number of additional findings identified.
 288







