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. ISSN 0301-0449

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

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:	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. Objective To compare the diagnostic yield of FUSS performed for suspected child abuse to double reading initial exams. 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. 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, $p<0.0001$. 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.

Original article

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

Abstract

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.

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

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.

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, $p<0.0001$.

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.

Introduction

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

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

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

Materials and Methods

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

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

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

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

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

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

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

Results

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

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

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

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

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

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

Discussion

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

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

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

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

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

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

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

Conclusion

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

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

References

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

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

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

Legends

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

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

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

Table 1. Six categories reflecting the concordance among single readings of initial skeletal surveys, double readings 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

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

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

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

		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)

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

		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)*

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

Figure 1

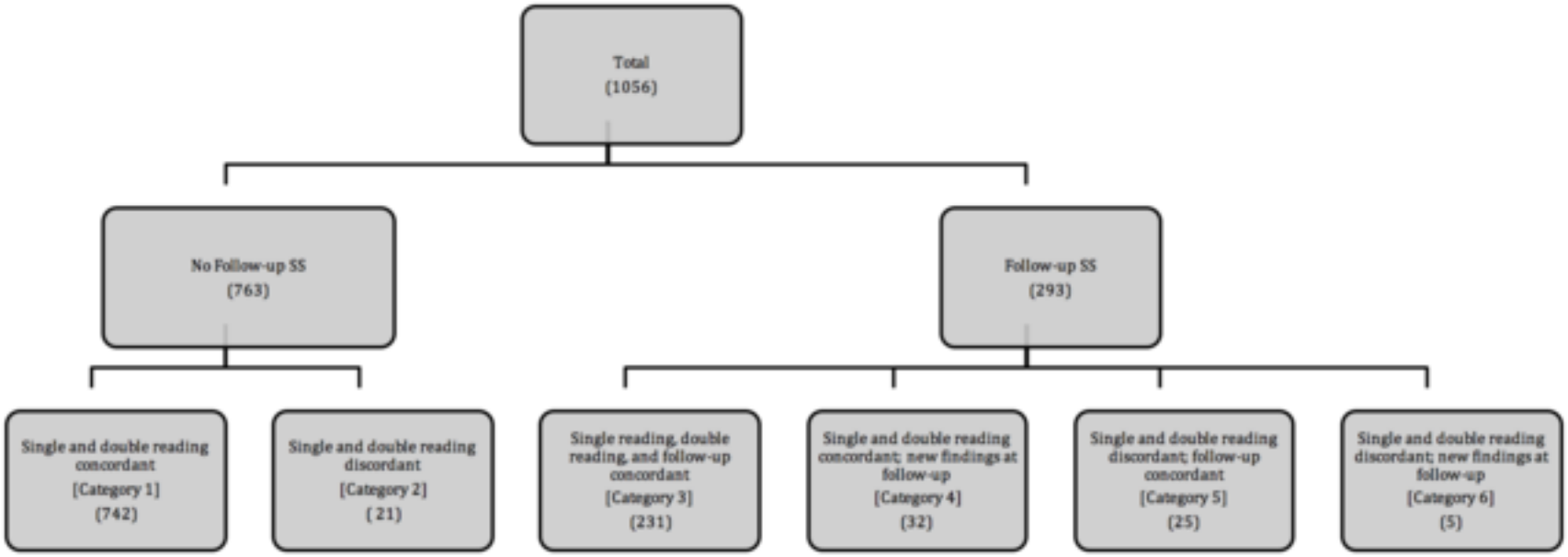

Figure 2

Click here to access/download;Figure;Fig 2.png

Figure 3 a

Click here to access/download;Figure;Fig 3 a.png

Figure 3 b

Click here to access/download;Figure;Fig 3 b.png