Determining when a fracture occurred
Drury, Anne; Cunningham, Craig

Published in:
Journal of Forensic and Legal Medicine

DOI:
10.1016/j.jflm.2017.11.004

Publication date:
2018

Document Version
Peer reviewed version

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. Aug. 2019
Title: Determining when a fracture occurred: Does the method matter?
Analysis of the similarity of three different methods for estimating time since fracture of juvenile long bones

Authors: Anne Drury, Craig Cunningham

Affiliation: Centre for Anatomy and Human Identification, University of Dundee, Dow Street, Dundee, DD15EH, United Kingdom

Publication: Journal of Forensic and Legal Medicine, Volume 53, January 2018, Pages 97–105

ABSTRACT

Radiographic fracture date estimation is a critical component of skeletal trauma analysis in the living. Several timetables have been proposed for how the appearance of radiographic features can be interpreted to provide a likely time frame for fracture occurrence. This study compares three such timetables, by Islam et al. (2000), Malone et al. (2011), and Prosser et al. (2012), in order to determine whether the fracture date ranges produced by using these methods are in agreement with one another. Fracture date ranges were estimated for 112 long bone fractures in 96 children aged 1-17 years, using the three different timetables. The extent of similarity of the intervals was tested by statistically comparing the amount of intersection between the ranges. Results showed that none of the methods were in perfect agreement with one another, but there was greater similarity between the ranges produced by Malone et al. (2011) and the other two studies than there was between Islam et al. (2000) and Prosser et al. (2012), with the greatest similarity existing between Malone et al. (2011) and Islam et al. (2000). The differences between fracture date estimates given by timetables currently existing in the literature indicates that caution should be exercised when estimating the timing of a juvenile fracture. Future research should be undertaken to compare these methods on a population of known fracture timing, and to better understand the relationship between age of the individual, skeletal health, fracture healing rates, and radiographic characteristics of fracture healing.

KEYWORDS: fracture healing, fracture timing, fracture date, healing timelines, juvenile fractures, methodology comparison
INTRODUCTION

Estimation of time since injury is a critical component of trauma analysis in forensic casework. Radiographic skeletal survey is a primary technique used for assessing inflicted skeletal trauma, and can be used to estimate how and when an injury may have been inflicted. This is vital information in cases such as those potentially involving child abuse. It is necessary for medical and forensic experts, such as radiologists, pathologists, and anthropologists, to have a reliable method for establishing a time frame for when a pediatric fracture occurred.

While fracture-dating timelines have been published for adults, these timelines are not appropriate for children, due to the more rapid healing rates of juvenile bone.
We're sorry... but your computer or network may be sending automated queries. To protect our users, we can't process your request right now.

See &lt;a href="https://support.google.com/websearch/answer/86640"&gt;Google Help&lt;/a&gt; for more information.

&amp;copy; 2013 Google - &lt;a href="https://www.google.com"&gt;Google Home&lt;/a&gt;&lt;/div&gt;&lt;/body&gt;&lt;/html&gt;

See &lt;a href="https://support.google.com/websearch/answer/86640"&gt;Google Help&lt;/a&gt; for more information.

© 2013 Google
<table>
<thead>
<tr>
<th>Face</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 g</td>
<td>#0039b6</td>
</tr>
<tr>
<td>30 a72f</td>
<td></td>
</tr>
<tr>
<td>41200</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Padding</th>
<th>Bottom Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 px</td>
<td>50%</td>
</tr>
</tbody>
</table>

Sorry... &lt;/div&gt;&lt;/td&gt;&lt;/tr&gt;&lt;/table&gt;&lt;/div&gt;&lt;/div&gt;&lt;div style="margin-left: 4em;"&gt;&lt;h1&gt;We're sorry...&lt;/h1&gt;&lt;p&gt;... but your computer or network may be sending automated queries. To protect our users, we can't process your request right now.&lt;/p&gt;&lt;/div&gt;&lt;/div&gt;&lt;div style="margin-left: 4em;"&gt;See &lt;a href="https://support.google.com/websearch/answer/86640">Google Help</a> for more information.&lt;br/&gt;&lt;br/&gt;&lt;/div&gt;&lt;/div&gt;&lt;div style="text-align: center; border-top: 1px solid #dfdfdf;"&gt;&amp;amp;copy; 2013 Google &amp;lt;a href="https://www.google.com"&gt;Google Home&lt;/a&gt;&lt;/div&gt;&lt;/body&gt;&lt;/html&gt;&lt;/location&gt;&lt;title&gt;Skeletal Injury in the Child&lt;/title&gt;&lt;uuid&gt;D6779C3B-991F-42CA-B6EE-9FC07EE447A9&lt;/uuid&gt;&lt;subtype&gt;0&lt;/subtype&gt;&lt;publisher&gt;Springer-Verlag&lt;/publisher&gt;&lt;/version&gt;3rd edition&lt;/version&gt;&lt;/place&gt;New York&lt;/place&gt;&lt;/type&gt;0&lt;/type&gt;&lt;/publication_date&gt;9920000000120000000200000&lt;/publication_date&gt;&lt;/authors&gt;&lt;author&gt;&lt;firstName&gt;John&lt;/firstName&gt;&lt;middleNames&gt;A&lt;/middleNames&gt;&lt;lastName&gt;Ogden&lt;/lastName&gt;&lt;/author&gt;&lt;/authors&gt;&lt;/publication&gt;&lt;/citations&gt;&lt;/citation&gt;}. The hematoma may dissect more extensively along the bone shaft due to how relatively easily the periosteum lifts away from underlying immature bone, which results in greater degree of subsequent new periosteal bone formation in juveniles { ADDIN PAPERS2_CITATIONS &lt;citation&gt;&lt;uuid&gt;95A752A2-F155-4C5D-97F2-1FA58C88FDDC&lt;/uuid&gt;&lt;/citation&gt;}. As the periosteum is relatively thicker, it is more likely to stay intact in juvenile fractures, allowing tissue to remain continuous over the fracture site { ADDIN PAPERS2_CITATIONS &lt;citation&gt;&lt;uuid&gt;39A2AEF0-4E4F-4400-978F-
Sub-adult bone constantly undergoes routine growth and remodelling, and the thicker periosteum results in enhanced osteogenic development to resorb and remodel the callus to form regularly aligned trabecular and cortical bone.
In spite of the important applications of pediatric bone healing, existing published scientific research into creating a reliable timeline of fracture healing in children is relatively limited.
The first researchers to examine a large population of children and assess multiple features of fracture healing with the aim of creating a timeline for fracture dating were Islam et al. (2000), who examined forearm fractures in children aged 1-17 years. Researchers examined six radiographically visible features of fracture healing: fracture margins, fracture gap, periosteal reaction, callus, bridging, and remodelling. The authors determined when the features first appeared after injury and the duration for which they were visible.

A similar model was explored by Malone et al. in 2011, with their development of a radiographic assessment method for time since injury.
A Radiographic Assessment of Pediatric Fracture Healing and Time Since Injury

The six features used by these authors were lack of healing, granulation, callus, bridging, clinical union, and completion, and assessed forearm and leg fractures in children aged 0 to 5 years. Prosser et al. (2012) also proposed a six-feature method for estimating fracture age in long bones of children aged 0-5 years, with features listed as soft-tissue swelling, periosteal reaction, soft callus, hard callus, bridging, and remodelling. Research presented in three other articles in the past five years completes all research from this field since the 2000 publication by Islam et al. 

ADDIN PAPERS2_CITATIONS

Dating fractures in infants

Retrospective evaluation and dating of non-accidental rib fractures in infants

The six features used by these authors were lack of healing, granulation, callus, bridging, clinical union, and completion, and assessed forearm and leg fractures in children aged 0 to 5 years. Prosser et al. (2012) also proposed a six-feature method for estimating fracture age in long bones of children aged 0-5 years, with features listed as soft-tissue swelling, periosteal reaction, soft callus, hard callus, bridging, and remodelling. Research presented in three other articles in the past five years completes all research from this field since the 2000 publication by Islam et al.
Healing patterns of clavicular birth injuries as a guide to fracture dating in cases of possible infant abuse

Older research is limited to three primary research articles in addition to timelines proposed in textbooks that were based entirely on clinical experience.
Given the information discussed above regarding the difference between adult and juvenile bone healing rates, it is reasonable to question whether juveniles of different ages will have different bone healing rates. A few studies demonstrate evidence of healing time variation in different ages of children, which further complicates the development and evaluation of studies in this area. ADDIN PAPERS2_CITATIONS <citation><uuid>7DC18C3F-645F-43A5-AC60-7962611A9027</uuid><priority>1</priority><publications><publication><volume>56</volume><publication_date>99201106031200000000222000</publication_date><number>5</number><doi>10.1111/j.1556-4029.2011.01820.x</doi><startpage>1123</startpage><title>A Radiographic Assessment of Pediatric Fracture Healing and Time Since Injury</title><uuid>75AB111B-04F3-4831-B196-60D3AC8DA694</uuid><endpage>1130</endpage><type>400</type><url>http://doi.wiley.com/10.1111/j.1556-4029.2011.01820.x</url></publication><publication><volume>59</volume><number>6</number><subtitle>Log-normal correlation with age of speed of healing</subtitle><startpage>704</startpage><endpage>707</endpage><type>400</type><url>http://doi.wiley.com/10.1111/j.1556-4029.2011.01820.x</url></publication></publications><author><firstName>Christina</firstName><middleNames>A</middleNames><lastName>Malone</lastName></author><author><firstName>Norman</firstName><middleNames>J</middleNames><lastName>Sauer</lastName></author><author><firstName>Todd</firstName><middleNames>W</middleNames><lastName>Fenton</lastName></author></citation>
Skak and Jensen proposed that healing times in juveniles (age 1-17) increase with age, following a log-normal pattern. Yeo and Reed subsequently demonstrated a trend toward increased time for callus formation with increasing age in children birth – 14 years, but results were limited by small sample size and not statistically significant. Malone et al. provide evidence supporting a correlation between age and healing duration in certain stages of healing but not others.

Yeo and Reed subsequently demonstrated a trend toward increased time for callus formation with increasing age in children birth – 14 years, but results were limited by small sample size and not statistically significant.

Yeo and Reed subsequently demonstrated a trend toward increased time for callus formation with increasing age in children birth – 14 years, but results were limited by small sample size and not statistically significant. Malone et al. provide evidence supporting a correlation between age and healing duration in certain stages of healing but not others.
Current research focuses on children 5 years of age or younger, stating that the age demographics of child abuse victims necessitate more research on infants and toddlers.
The three main studies that focus on creating timelines for pediatric fracture dating in long bones, specifically Islam et al. 2000, Malone et al. 2011, and Prosser et al. 2012, have all been proposed to help medical and forensic experts to assess date of injury. The relative efficacy and accuracy of the different fracture dating timelines has not been practically tested by an external researcher. If a single radiograph was evaluated using these different methods, it is unknown to what extent the estimated times since injury based on each article would overlap. Time ranges given are influenced by composition of each study sample, and no studies have been undertaken to determine whether these results apply equally to other sample groups. The present study will use the results described in those three articles to assess a sample of children with long bone fractures, and apply each method to estimate when the fracture may have occurred. The objective of this study is to compare these methods to determine their congruency, and the extent to which they produce similar time frames for the date of injury. Thus, this study also aims to better understand the possible limitations of juvenile fracture dating using published timelines of radiographic features of fracture healing.

MATERIALS AND METHODS

This study was carried out on a collection of digitised radiographs held by the Centre for Anatomy and Human Identification at the University of Dundee. All relevant ethical permissions were obtained. Radiographs from the collection were selected that showed evidence of fracture to a long bone, comprised of the humerus, radius, ulna, femur, tibia, fibula, and metapodials and proximal phalanges. The sample was composed of isolated radiographs, with no follow-up radiographs for a single individual. Any radiographs showing that the bone was treated with internal fixation, such as rods or screws, were excluded. This is consistent with protocol in existing research, based on the concern that internal fixation affects healing times to an unknown extent. Epiphyseal fractures were similarly excluded, as were any radiographs exhibiting visible signs of bone disorders or diseases. It is unknown whether these fractures were accidental, or suspected to be the result of non-accidental injury. Medical records were unavailable, and the only information known to the researcher was the age of the child at time the image was taken.

The final sample was comprised of 112 fractures in 96 children aged 1-17 years, with 35 females and 61 males, with distribution as shown in Figure 1. Average age was 9.05 years. Fracture location was distributed as follows: 40 humeral, 13 ulnar, 10 radial, 11 metacarpal, 4 femoral, 5 tibial, 1 fibular, 23 metatarsal, and 5 proximal phalanges.

So that the results of one method would not bias the application of another, each assessment was performed without knowledge of those previously undertaken on that fracture. Testing of intraobserver error was undertaken to determine whether the methods were equally reproducible by a single observer. There were two additional trials of a random sample of ten of the fractures, with methods consistent to those described above. These additional trials

Figure 1 – Frequency of fractures at each age group in sample
were carried out several days after the initial assessment and several days apart from each other in order to minimize memory bias. To select the fractures that would be reassessed, ten were chosen at random from the categories of elbow, hand, knee, and foot, to have a balanced representation of different fracture locations.

In accordance with Islam et al. (2000) to obtain a fracture date estimate using their timetable of healing, the fracture was assessed for the following features: blunting of the fracture margins signs of sclerosis at the margins of the fracture line, periosteal reaction, the incorporation of periosteal reaction into the bone cortex, density of the callus compared to density of the adjacent cortex, partial or complete loss of fracture margins at the site due to bridging, and remodelling of the fracture as indicated by a loss of the cortical bump over the fracture site, shown by a less acute angle of the new bone over the site. Images in the article were used to help assess the presence or absence of each of these features for every fracture. Islam et al. (2000) also recommended assessment of the widening of the gap between fracture fragments, based on comparison of two sequential radiographs of the same fracture. It was not possible to assess this feature in the present study because there were no sequential radiographs. Once presence or absence of each feature was determined, a time range for the fracture was determined using the figures and discussion in the Islam et al. (2000) article. As per the article, emphasis was placed on the importance of using the presence and absence of multiple radiographic signs, in order to make an informed assessment.

For fracture evaluation based on the Malone et al. (2011) study, each fracture was assigned to one of the six stages outlined in their methods section. To define the six stages, Malone et al. (2011) rely on sharpness of the fracture margins, initial callus and bone formation, mature callus presence, and visibility of the fracture line for their dating method.
When selecting which stage was the best fit, all of these features were examined, and the radiograph was compared to the images selected by the authors to represent each stage. Further description of the stages in the results section of the paper was also used for clarification. The stage that best represented the presence and absence of each feature was selected. Once a stage was assigned to the fracture, a range was assigned using summary table information. Range was used as opposed to simply assigning the mean healing time because Malone et al. (2011) advocate caution in their discussion of the results, and the patient sample in the present study included children older than included in the Malone et al. (2011) study.

To estimate the fracture age using the timetable outlined by Prosser et al. (2012), six radiographic features were assessed for each fracture, as outlined in their study methodology. These features were soft-tissue swelling, periosteal reaction, soft callus, hard callus, bridging, and remodelling. In their materials and methods section, the authors define the characteristics of each of these features and how they appear radiographically. These definitions were used in the present study in order to decide whether or not a feature was visible for this particular method. The suite of characters apparent on a particular radiograph was then compared to the descriptions Prosser et al. (2012) give in their results and discussion section, and an estimate of fracture age was assigned accordingly.

Skeletal element injured, fracture location (diaphyseal, diametaphyseal, or metaphyseal), and fracture type were assessed and recorded. After time ranges were established for each fracture using each of the three methods, the minimum estimated fracture date and maximum
estimated fracture date were separated into two columns in the excel spreadsheet. This spreadsheet with file number, age, the variables listed above, and the minimum and maximum columns was then imported into the program R. RStudio was the primary software used to undertake statistical analysis of the data.

The type of statistical and mathematical analysis possible was limited by the structure of the data. Ranges of numbers needed to be compared in order to determine how similar the fracture dates estimated by using each method were to each other, as opposed to just single numbers. Furthermore, the ranges of numbers were not normally distributed, because the values were dictated by the timetables given in each article, which further excluded which tests were feasible. It was determined that the best way to test the similarity between the ranges would be examining the degree of intersection between the ranges produced by assessing each fracture with each method.

To address the primary goal of the project of determining the extent of similarity between each method, tests of intersection were carried out between the date ranges estimated by each method for each fracture. For this type of testing, one range is compared to another to determine the extent to which the ranges intersect, and then the ranges are compared inversely. For example, Range A of 0-7 versus Range B of 0-14 produces an intersection of 1.0, whereas Range B versus Range A would produce an intersection of 0.5. An intersection of 1.0 means that Range A is completely intersected by Range B. An intersection of 0.5 means that only half of Range B is intersected by Range A. To summarise the extent of similarity between two intervals, tests of conformity were undertaken, which account for one of the intervals being larger than and thus containing the other interval. A binomial test was then used to assess whether the differences in conformity between methods were statistically significant. To determine whether the degree of intersection varied based on age of the individuals, it was compared across three age ranges – 0-5 years, 6-10 years, and 11-17 years. These groups were selected to enable comparison of younger, to middle aged, to older children, and because both Prosser et al. 2012 and Malone et al. 2011 based their dating timelines on children of 5 years or younger, while Islam included individuals up to 17.

RESULTS

The 112 fractures examined were distributed among diaphyseal (41 fractures), diaphyseal (17) and metaphyseal (54) locations, as well as type of fracture as summarised in Figure 2.

Prior to comparing fracture dates produced by each method, the relationship between distribution of times since injury and ages of the individual was examined. Based on boxplots dividing children into age groups 1-
5, 6-10, and 11-17, and assessment of median and interquartile range, there was no correlation in any method between time since injury and age of the child. This finding supports an absence of bias in the sample with regard to time since fracture and individual’s age.

The primary aim of this research, to compare the similarity between fracture date intervals estimated by each method for each incidence are summarised in Figures 3.1, 3.2, and 3.3. To simplify the numerous plots, M1 is the label for Islam et al. (2000), M2 is the label for Malone et al. (2011), and M3 is the label for Prosser et al. (2012), based on alphabetical order of the methods. Figure 3.1 shows the relationship between Islam et al. (2000) and Malone et al. (2011). Based on these plots, the two methods produce extremely similar intervals. In 91.1 percent of individuals, the fracture date range estimated using the Islam et al. (2000) method is completely contained within the date range estimated using the Malone et al. (2011) method. The range of dates estimated by using Malone et al. (2011) is completely contained within the Islam et al. (2000) range in 66.1 percent of cases. These results also indicate less greater, less precise date ranges obtained using Malone et al. (2011).

Figure 3.1 – Intersection between fracture date intervals estimated using data from Islam et al. (2000) (M1), and Malone et al. (2011) (M2). Proportion of individuals with each degree of similarity.

Figure 3.2 – Intersection between fracture date intervals estimated using data from Malone et al. (2011) (M2), and Prosser et al. (2012) (M3). Proportion of individuals with each degree of similarity.
Figure 3.2 shows the relationship between Malone et al. (2011) and Prosser et al. (2012) by comparing the intersection between estimated time-since-fracture intervals produced using the two studies. Fracture estimates produced using Prosser et al. (2012) were always entirely contained within the date range produced by using Malone et al. (2011). According to the inverse comparison, in 66.1 percent of fractures, the date range estimated by the Malone et al. (2011) method shared 50 percent of the interval estimated by Prosser et al. (2012). Again, this data supports that the Malone et al. (2011) method estimates less precise fracture date intervals.

Figure 3.3 outlines the relationship between the ranges produced using the Islam et al. (2000) versus the Prosser et al. (2012) methods. For 23 percent of fractures examined, there was zero similarity between the date ranges estimated by these two methods. Zero degree of similarity indicates that one date range estimate ended, the other started. Negative similarity indicates that for some fractures, there was a gap of time between the date ranges estimated by using each timetable.

Figure 3.7 – Intersection between fracture date intervals estimated using data from Islam et al. (2000) (M1), and Prosser et al. (2012) (M3). Proportion of individuals with each degree of similarity.

After considering these basic interval comparisons, similarity between intervals produced was statistically evaluated using a binomial test of the number of fracture date ranges produced with a high degree of similarity (80 to 100 percent) between one method and each of the other two: M1 vs M2, M1 vs M3, M2 vs M1, M2 vs M3, and M3 vs M1 and M3 vs M2. Eighty to 100 percent was selected as a “high” degree of similarity to accommodate the
fact that each article gave timelines of slightly different scale, with Islam et al. (2000) in particular rounding their results to weeks. According to the binomial tests, the difference in similarity between M1 vs M2 and M1 vs M3 was statistically significant. M1 intervals intersect to a high degree with M2 intervals more often than with M3 intervals (p-value < 2.2e-16). The difference in similarity between M2 vs. M1 and M2 vs. M3 was also statistically significant. M2 intervals intersect to a high degree with M1 intervals more often than with M3 intervals (p-value < 2.2e-16). Finally, according to the binomial test the difference in similarity between M3 vs. M1 and M3 vs. M2 is statistically significant as well. M3 intervals intersect to a high degree with M2 intervals more often than with M1 intervals (p-value < 2.2e-16).

Based on these six comparisons of intersection, the fracture date ranges produced using Islam et al. (2000) and those produced using Malone et al. (2011) seem to have more in common, as do the ranges produced by Prosser et al. (2012) and Malone et al. (2011), with the ranges produced by Islam et al. (2000) and Prosser et al. (2012) having the least in common. To simplify the comparisons made above, it is possible to complete the intersection tests again using conformity, which accounts for one of the intervals being larger than the other by choosing the comparison (either 1 vs. 2 or 2 vs. 1, for example) with greater agreement. This allows two intervals to be compared at the same time, as opposed to comparing first one manner of intersection and then the inverse. It highlights similarity between fracture date ranges produced by using two different methods, but disguises when one method consistently produces wider ranges, which is why the initial interval comparison data is also necessary. Based on the tests of conformity, Figure 4 shows the number of fractures where there was high agreement between each pair of methods in the date ranges produced.

Three binomial tests were used to statistically evaluate differences in agreement. Results confirm that the similarity between the fracture date intervals produced by Islam et al. (2000) and Malone et al. (2011) is greater than the similarity between Islam et al. (2000) and Prosser et al. (2012) (p-value = 0.0001851). The similarity between the intervals produced by Malone et al. (2011) and Prosser et al. (2012) is also greater than the similarity between Islam et al. (2000) and Prosser et al. (2012) (p-value = 9.514e-07). In spite of the slight difference seen on the graph, there is no statistical difference between the M1/M2 agreement and the
M2/M3 agreement. The final result is that the fracture date estimates produced by using the timetable in the Malone et al. (2011) article are equally similar in range intersection to the estimates produced using Islam et al. (2000) and the estimates produced using Prosser et al. (2012), but that the latter two articles produce fracture date intervals that are statistically less similar to each other based on overlap of estimated ranges.

To then assess whether one method consistently aged a fracture as more or less recent than other methods, comparisons between methods were performed of the minimum and maximum dates of each range. Based on box plots and assessment of interquartile ranges, it was found that using the Malone et al. (2011) timetable often produces a lower minimum time since fracture than using the Islam et al. (2000) timetable, in a couple cases giving a date of 25 days earlier, with a mean of a fracture date estimate of 3.90 days earlier. Using the Islam et al. (2000) timetable sometimes produces a larger minimum date since fracture, and sometimes a smaller minimum date, than using the Prosser et al. (2012) timetable, though the median is zero. It is more common for the M3 (Prosser) value to be smaller than the M1 (Islam) value, with the mean number of days earlier that is estimated by using the Prosser et al. (2012) timetable being 2.21 days. Using the Malone et al. (2011) timetable produces a fracture date estimate mean of 1.69 days earlier than the M3 (Prosser) minimum. Using the Malone et al. (2011) timetable therefore produces a fracture date estimate with a lower minimum number of days since fracture than using either of the other methods.

A similar comparison can be done on the maximum values of the three date ranges estimated for each fracture. Differences in methodology, however, give several extreme outliers, because the Malone et al. (2011) and Prosser et al. (2012) timetables include extremely high possible maximum age dates based on their samples, while Islam et al. (2000) cap their timetable at fourteen weeks. These outliers make the mean number of days difference between methods a less indicative comparison. In most cases, Malone et al. (2011) and Islam et al. (2000) provide the same maximum days since injury for a fracture (72/112 cases), which was reflected by a median of zero on the box and whisker comparison, but otherwise the maximum values produced by using Malone et al. (2011) are more often greater than the values by Islam et al. (2000).

Based on similar box and whisker plot comparison of Islam et al. (2000) to Prosser et al. (2012), using the former method provides a higher maximum value by a margin of 7 days in 94 out of 112 cases. It was also found that using the Malone et al. (2011) timetable more often provides a higher maximum date since injury than the Prosser et al. (2012) timetable, most frequently by a margin of 7 days in 74 of 112 cases. Use of the Malone et al. (2011) method most often produced the lowest minimum number of days since fracture, and the highest maximum days since fracture. The most precise method was Prosser et al. (2012), which gave a range of only 7 days in 79 out of 112 cases.

When percentage of fractures examined where identical range sizes were given by two methods was examined, similarity was highlighted between results of the Islam et al. (2000) and the Malone et al. (2011) methods. In 64.29 percent of cases, the size of the fracture date range was identical for M1 and M2, as opposed to in only 8.93 percent of cases for both M1 and M3, and M2 and M3. Therefore, the former two methods are most likely to estimate a fracture range of identical duration.

Based on intraobserver error testing, for eight out of the ten fractures, the fracture date estimates were identical for all three methods across all three trials. Two fractures had
differing estimates for at least one trial. The descriptions given of the features visible on each radiograph for the two cases where error is seen provide more information about the possible error. For FLE140001 Method 1, a feature was identified in the first trial that was not identified in subsequent trials, thus decreasing the range. For Method 2, the same features were identified and the same age range given. For Method 3, the same features were identified, but a different fracture date range was assigned. Possible inconsistencies in the application of Method 3 will be elaborated upon in the discussion. For MLE67, periosteal reaction was not identified in trial 1 but was identified when applying each of the three methods in the two later trials, resulting in an increase in time since fracture in all timetables.

DISCUSSION

This study presents evidence of how different proposed timetables for the dating of fractures in juvenile bone compare to each other when used to assess a sample of fractures of unknown timing. Differences in estimate existed in every analysed case. In 74/112 fractures analysed, the fracture was deemed to have no evidence of healing as per any of the three methods. There were no fractures that were assessed as having signs of healing according to one method and not to another. This meant that the earliest possible category from each of the articles’ timetables was used. According to the Islam et al. (2000) suite of fractures analysed and the timetable given in the article, a fracture that exhibited neither periosteal reaction presence nor callus presence was evaluated as being less than two weeks old. According to the Malone et al. (2011) six stages of healing, a fracture with sharp lines and no callus formation or bridging, clarified as “no presence of bone healing” (pp. 1125) was deemed to be at Stage 1, which had a range of 0-14 days healing time. According to Prosser et al. (2012), an injury demonstrating soft-tissue swelling around the fracture site and no other signs of healing was deemed to be acute, or < 1 week old. This was interpreted to mean 0-7 days old, as the authors categorise their next stage of “recent” as 8-35 days.

It stands to reason, therefore, that the two methods that both define the earliest visible signs of fracture healing as 0-14 days, while the third defines it as 0-7 days, will have a better intersection, especially when 66.07 percent (74/112) of the fractures analysed fell in this category. There were also strong differences between the Prosser et al. (2012) assessment of when certain radiographic features appeared and the assessment particularly by Islam et al. (2011). According to Prosser et al. (2012), the presence of periosteal reaction as the sole sign of healing meant that a fracture was likely 5-14 days old. Meanwhile, Islam et al. (2000) show in their timetable that periosteal reaction is unlikely to appear before 14 days since fracture, and expand in their discussion that an injury between 2 and 3 weeks old is likely to show periosteal reaction without sclerosis of the fracture margins. Prosser et al. (2012) did not include marginal sclerosis in their assessment, citing the rarity of its appearance in an initial pilot study, though Islam et al. (2012) claim that in their own study it was visible in 85 percent of fractures between 4 and 6 weeks of healing. Malone et al. (2011) did not use either of these radiographic characteristics in defining their stages of healing.

Based on the literature providing evidence that healing rates are more rapid in younger individuals { ADDIN PAPERS2_CITATIONS <citation><uuid>49FDD950-0F1C-43A2-8C8E-FE40FFB18544</uuid><priority>19</priority><publications><publication><location>&lt;html&gt;&lt;head&gt;&lt;meta http-equiv="content-type" content="text/html; charset=utf-
Sorry...
Jensen}, the explanation for why Islam et al. (2000) cite periosteal reaction as developing at 14 days and Prosser et al. (2012) cite the same feature as developing at 5-14 days could lie in the differing ages of the children in their sample groups. Both articles agree, however, that a fracture showing signs of periosteal reaction and soft callus with no other features should be aged at 2-3 weeks, in spite of the difference in ages of the sample population. According to Malone et al. (2011), soft callus (or as they say, “fluffy callus”) formation without other features places the fracture in Stage 2, which gives a fracture date estimate of 4-50 days. This estimate completely intersects with that of the other two articles for the appearance of soft callus, but provides a less precise range.

Existing literature discussing differing rates of healing depending on the age of the child focuses on determining to what extent the healing process is quicker at a younger age { ADDIN PAPERS2_CITATIONS <citation><uuid>CBB07927-100E-499C-A6D7-203DDFBF21F5</uuid><priority>20</priority><publications><publication><volume>56</volume><publication_date>99201106031200000000222000</publication_date><number>5</number><doi>10.1111/j.1556-4029.2011.01820.x</doi><startpage>1123</startpage><title>A Radiographic Assessment of Pediatric Fracture Healing and Time Since Injury</title><uuid>75AB111B-04F3-4831-B196-60D3AC8DA694</uuid><subtype>400</subtype><endpage>1130</endpage></publication> http://doi.wiley.com/10.1111/j.1556-4029.2011.01820.x <bundle><publication><title>Journal of Forensic Sciences</title><type>-100</type><subtype>-100</subtype><uuid>44C694B8-DCED-4215-A936-DAC3FDC054DF</uuid></publication></bundle><authors><author><firstName>Christina</firstName><middleNames>A</middleNames><lastName>Malone</lastName></author><author><firstName>Norman</firstName><middleNames>J</middleNames><lastName>Sauer</lastName></author><author><firstName>Todd</firstName><middleNames>W</middleNames><lastName>Fenton</lastName></author></authors></publication></publications><citation><uuid>AE1D8B60-1071-4EBB-8353-3F793FC776FF</uuid><priority>400</priority><publications><publication><volume>59</volume><number>6</number><subtitle>Log-normal correlation with age of speed of healing</subtitle><startpage>704</startpage><title>Femoral shaft fracture in 265 children</title><uuid>1C863999-9FA5-40AB-847A-A0F92A4E8227</uuid></publication></publications></citation> Malone et al. (2011) suggest based on their results that greater age causes a delay in the start of the healing process as opposed to a greater length of time spent healing over all, which could account for the later onset of the first sign of healing in older children but a similar appearance time for later signs when comparing the Islam et al. (2000) and Prosser et al. (2012) timetables. This Malone et al. (2011) hypothesis is in contrast, however, to the evidence provided by Skak and Jensen (1988) for a log-normal correlation between the age of a child and the mean amount of time required for healing. The relationship between age, timing of radiographic feature appearance, and rate of healing is still far from understood.
Because both Malone et al. (2011) and Prosser et al. (2012) both developed their timetables for children aged less than 5 years, it would be expected that the results of using their analyses to date fractures would be more similar than that of a timetable developed on children of ages 1-17 like Islam et al. (2000). The sample from the current study included children of ages 1-17. It is possible that the different features evaluated in the Malone et al. (2011) timetable were more adaptable to a sample of children of other ages than the features of the Prosser et al. (2012) timetable, and less likely to conflict with a timetable like that by Islam et al. (2000) developed using older children. Once again, this would draw attention to the question of whether the appearance times for certain radiographic features of healing are more strongly correlated with the age of the individual. It is also simply possible that the larger, less precise ranges given by Malone et al. (2011) meant that the estimates made using that method were more likely to include the ranges developed by Islam et al. (2000). The group of fractures analysed in this study were distributed such that there was no relationship between the age of the fracture and the age of the individual. This indicates that the study would not have been biased by a certain age group having only recent or only old fractures.

Repeatability of this study could be affected by the manner in which each timetable is applied to a fracture, after it has been determined which radiographic features are present. As described in the methodology above, in order to use each of the three articles to assess when a fracture occurred, the timetable presented in each article was examined in association with the related discussion of how the features should be interpreted to provide a fracture date. In interpreting possible causes for the intraobserver error noted, or potential interobserver error, it is necessary to consider whether even if the exact same radiographic features are identified by two observers, it would be possible for them to assign different fracture ranges. For the intraobserver error trials of MLE67, it was noted in the results that for Method 3, the same features were identified in all three trials, but different date ranges were produced. This suggests problems with how the Prosser et al. (2012) timetable was interpreted in each trial. In examining the original article, Prosser et al. (2012) state that it is possible to assign fractures in young children to one of three date ranges: acute (<7 days), recent (8-35 days), and old (≥36 days), based on an amalgamation of features present and absent. There is some discrepancy, however, between these categories and the timeline proposed in the discussion, and it is often unclear in which category the fracture should be placed. For example, if a fracture has periosteal reaction, hard callus, and some signs of bridging, it could be interpreted as either recent or old, because the authors state that a fracture showing periosteal reaction and soft callus is likely 2 to 3 weeks old, and a fracture with hard callus or bridging is 3 weeks old or older. In general, the timeline the authors summarise does not align with their proposed categories of 0-7 days, 8-35 days, and 36 or more days. Using this article as a guide for fracture dating could therefore result in a variety of fracture age ranges depending on which part of the tables, figures, and discussions were emphasized.

After the initial recent fracture period, Islam et al. (2000) identify fracture gap widening as a feature for determining a fracture aged 3-7 weeks. Unfortunately, this feature could not be applied in the current study because this study examined single instances of an injury as opposed to tracking a single fracture across multiple radiographs. It would be useful in the later stages of healing to make more precise fracture date estimates in the 2 months post-fracture period. Prosser et al. (2012) point out that a timetable including features such as blunting of the fracture margins or widening of the fracture gap over time is not conducive to real world scenarios where radiologists need to give a professional opinion on the age of a fracture based on a single radiograph. This study supports the idea that a timetable designed
to date fractures from a single radiograph should concentrate on features that do not need to be observed over time.

One important aspect of fracture dating that was not addressed by this study was interobserver error. Consistency between different individuals analysing a case is necessary for a reliable dating method because it supports the broader application of that method by individuals outside the initial study group. Prosser et al. (2012) claim high levels of interobserver agreement for all radiographs in the dataset in their study, but the issue was not addressed in the other two articles used here. Other literature in the field raises doubts about the quality of interobserver agreement in radiographic features of healing. According to Halliday et al. (2011), the pediatric radiology specialists in their study had poor agreement in the definition of hard callus versus soft callus and in identification of soft-tissue swelling, as well as other features not used by any of the primary three articles used here. Halliday et al. (2011) emphasize the importance of describing exactly how the radiographic feature appears for each level of classification in order to achieve the highest possible level of agreement. Otherwise, the authors state, it is very difficult to assess fracture date with any degree of certainty. They argue that many of the features currently used to estimate fracture age in juvenile long bones, aside from subperiosteal new bone formation (equated to periosteal reaction), are not reproducible and are very unreliable.

The current study draws attention to the differences in radiographic features used to make fracture dating timelines and the differences between the dates given for appearance of certain features, and the degree to which these differences result in contrasting fracture age estimates. When accompanied by doubts of reproducibility and reliability between observers, the caution advocated in fracture dating by several authors is given much more weight, and appears to be a cause for concern. 

---

title: Dating fractures in infants
author: K Halliday, N Broderick, J Somers, R Hawk
 publisher: The Royal College of Radiologists
volume: 66
doi: 10.1016/j.crad.2011.06.001
---

---

title: Development and Duration of Radiographic Signs of Bone Healing in Children
author: Omar Islam, Don stName
 publisher: American Journal of Roentgenology
volume: 175
---
There is a need for further testing of the existing healing timetables on larger fracture populations where the age of the fracture is known, so that the accuracy of each method can be calculated, as well as testing by several different observers, and repeat testing by those same physicians. It would be better if radiologists experienced in dating fractures undertook this experimentation.

The current study was limited by a lack of knowledge about the medical conditions of the individual, including whether the patients experienced concurrent head injury (cited as an issue by multiple authors { ADDIN PAPERS2_CITATIONS <citation><uuid>1A86A894-B33B-4BD5-B2A4-C7E8183FB31F</uuid><priority>22</priority><publications><publication><volume>66</volume><publication_date>99201111011200000000222000</publication_date><doi>10.1016/j.crad.2011.06.001</doi><startpage>1049</startpage><title>Dating fractures in infants</title><publisher>The Royal College of Radiologists</publisher><type>400</type><subtype>400</subtype><uuid>3ADD77FE-23D8-4075-AFD6-9A37770E7464</uuid></publication>}, <publication><volume>198</volume><publication_date>99201205001200000000220000</publication_date><doi>10.2214/AJR.11.6734</doi><startpage>1014</startpage><title>A Timetable for the Radiologic Features of Fracture Healing in Young Children</title><publisher>The Royal College of Radiologists</publisher><type>400</type><subtype>400</subtype><uuid>CC698E5A-514D-4DD9-AD18-0DFA22AD6A6B</uuid></publication>}).
whether the fracture was the result of non-accidental injury, and the health status of the individual overall. Although some radiographs with obvious signs of pathology were excluded from the study, it is possible for some diseases causing increased bone fragility to not be immediately apparent on radiographs, especially to an inexperienced practitioner. Brittle bones can be caused by a variety of disorders and result in the increased susceptibility of bones to fracture, even when the force is not great, leading to concerns where non-accidental injury is in question.
The original purpose for the creation of the collection of radiographs from which this sample was drawn was also not fracture date estimation, which should be cited as a weakness of the study. There was also only one radiograph, with a single angle of view available for some fractures, whereas for others there were multiple angles available. This inconsistency could have affected data collection, as it is possible that having only one viewpoint could have obscured certain radiographic features. Another consideration is the different populations of children used to create these three timetables, and then the test population; Islam et al. (2000) examined Canadian children, Malone et al. (2011) American children, and Prosser et al. (2012) U.K. children, and the current study examined children seen at the Ninewells hospital in Scotland. This also brings up a vast number of other variables commonly considered in anthropology, like ethnicity and social status and background, and the possibility that these variables in the patients in each sample could play a role in the relationship between the fracture estimates provided by using each method.

One assumption that has not been addressed by any of the literature discussed in this paper is the use of the chronological ages of the children whose fractures are being assessed. In the current literature published creating timetables or investigating the timing of appearance of radiographic features of fracture healing, age of the patients is determined chronologically. One consideration is whether fractures in children with differing skeletal and chronological ages would heal at a rate more in agreement with the skeletal age, which would possibly have an effect on the timetable created and its applicability to children from other populations. According to one study, skeletal ages of current adolescents aged 12-15, both male and female, in the U.S.A. are more advanced than their chronological ages { ADDIN PAPERS2_CITATIONS <citation><uuid>B1115F90-C3CF-481C-94FB-BB9E16E6ECE</uuid><priority>24</priority><publications><publication><volume>4</volume><publication_date>9920100902120000000022200</publication_date><number>5</number><doi>10.1007/s11832-010-0289-z</doi><startpage>467</startpage><title>Skeletal and chronological ages in American adolescents: current findings in skeletal maturation</title><authors><author><firstNames>Ryan</firstNames><middleNames>P</middleNames><lastName>Calfee</lastName></author><author><firstNames>Melanie</firstNames><lastName>Sutter</lastName></author><author><firstNames>Jennifer</firstNames><middleNames>A</middleNames><lastName>Steffen</lastName></author><author><firstNames>Charles</firstNames><middleNames>A</middleNames><lastName>Goldfarb</lastName></author></authors></publication></publications><cites></cites></citation>}. It is unclear whether this could then indicate a rate of fracture healing in
agreement with an older individual, or if the stronger rates of skeletal growth would mean faster rates of healing. If skeletal age affects the rate of fracture healing differently to chronological age, a wealth of concerns arise regarding interpopulation variation in determining skeletal age, and how to apply this knowledge to existing fracture dating timetables when using them for children of different populations. No studies could be found that connect chronological age, bone age, skeletal health and fracture healing rates.

CONCLUSION

The primary goal of this study was to analyse existing timetables for estimating the fracture date of injuries to juvenile long bones by comparing the intervals produced by applying three different methods to a single fracture. Data collected showed differences between all three methods in the date ranges produced, though there were greater similarities between the timetable of Malone et al. (2011) and the other two timetables than there were between Islam et al. (2000) and Prosser et al. (2012), with the greatest similarity existing between Malone et al. (2011) and Islam et al. (2000). The lack of agreement between methods indicates that the existing methods of dating fractures in long bones are not equally applicable in all cases. These results signify that great caution should be used when applying existing timetables to date a fracture of unknown age. When applied in criminal cases of possible non-accidental injury, the results of these fracture date determinations have the power to affect the lives of children and their families. It is thus imperative that all possible variables are considered, and an estimate be made with the greatest understanding of the current literature available.

A substantial amount of research into the relationship between timing of radiographic feature appearance, age of the individual, health of the individual’s bone, and other aspects of medical history is yet to be undertaken. In trying to create dating timetables, the simplest solution for the concerns described in the discussion above would be to exclude fractures in any children with indicators of less than optimal health from studies used to establish fracture dating timetables, and thus remove any possibly confounding variables. This would be counterproductive in the long term though, if the goal of the studies is to be able to age fractures in children who have suffered non-accidental injury. Malnutrition, as another aspect of child abuse, can be associated with fractures from non-accidental injury
A timetable developed on completely healthy children may not therefore be entirely accurate when applied in cases of child abuse, depending on the interactions between bone age, health of the individual, and rates of healing. More research is necessary to better understand the interactions between these variables, but it is made extremely difficult by the incorrectly reported injury times that can occur in these cases { ADDIN PAPERS2_CITATIONS citation=<uuid>D61C16C5-C4E9-47C7-A4D3-ADCE19865C22</uuid>}.
At the present time, it appears that a vast amount of experience in interpreting pediatric fractures, paired with a considered application of all relevant existing research in fracture date estimation, and matched by a great degree of caution may be the best route to establishing a fracture date estimate that is as accurate as possible.
References

{ ADDIN PAPERS2_CITATIONS <papers2_bibliography/> }