Accuracy Tests of Tooth Formation Age Estimations for Human Skeletal Remains

SHELLEY SAUNDERS, CAROL DeVITO, ANN HERRING, REBECCA SOUTHERN, AND ROBERT HOPPA
Department of Anthropology, McMaster University, Hamilton, Ontario, Canada L8S 4L9

KEY WORDS Juvenile, Aging, Dental development

ABSTRACT Estimations of age from tooth formation standards for a large (n = 282) sample of subadult skeletal remains from a 19th century historic cemetery sample were analyzed. The standards of Moorrees et al. (1963a,b) for the permanent and deciduous teeth, and Anderson et al. (1976) for the formation of permanent dentition were employed in a variety of combinations to calculate mean dental ages. Tests of accuracy and bias were made on a small sample (n = 17) of personally identified individuals, and age of attainment scores were compared to age of prediction scores for each individual. The resulting dental age distributions for the skeletal sample were compared to documented burial records for the cemetery to determine the representativeness of the skeletal sample. These comparisons showed little difference between age of attainment versus age of prediction methodologies. The standards of Moorrees et al. (1963a,b) were observed to provide the most accurate estimates of age with a standard deviation of one-half year. The standards of Anderson et al. (1976), while easier to use and more extensive, are problematic in that the original reference sample begins at three years of age, while the sample used by Moorrees and colleagues begins at birth. The skeletal age distributions compare well to the overall chronological age distribution for the cemetery. These results affirm that tooth formation age estimates for subadult skeletal remains from archaeological or forensic samples provide accurate assessments of age at both the individual and population level. © 1993 Wiley-Liss, Inc.

The dentition is the best physiological indicator of chronological age in juveniles (Holtz, 1959; Green, 1961; Demirjian, 1986; Smith, 1991b). Further, tooth formation, or more specifically, chronologies of tooth mineralization as visualized radiographically, are now recognized to be the best means of estimating chronological ages of immature human skeletons from archaeological sites and for forensic cases (El-Nofely and Iscan, 1989; Merchant and Ubelaker, 1977; Saunders, 1992; Ubelaker, 1987, 1989).

The formation of tooth crowns and roots has been shown to be much less affected by hormonal influences, local and general environmental factors, and nutritional and social factors than tooth emergence, skeletal development, weight or height (Demirjian, 1986; El-Nofely and Iscan, 1989; Smith, 1991b). Evidence for this assertion includes relatively small variations in the timing of dental development in children with major abnormalities affecting maturation (Garn et al., 1965a; Kuhns et al., 1972; Niswander and Sujaku, 1965), lower coefficients of variation in the stages of tooth formation than skeletal development (Lewis and Garn, 1969), and low correlations between tooth formation and weight, fatness, stature or
bone age (Anderson et al., 1976; Garn et al., 1965b and Gran, 1962, Gyulavari, 1966; Lacey et al., 1973; Prahl-Andersen and Roede, 1979; Steel, 1965). In addition, tooth formation has been found to be highly heritable (Garn et al., 1965b; Moorrees and Kent, 1981).

A further advantage to using tooth formation rather than tooth emergence in the estimation of chronological age is that it provides a continuous series of developmental changes from before birth to about 20 years (Demirjian, 1986). While stages of tooth mineralization in dry bones are directly comparable to radiographic standards from living samples, living population standards for tooth emergence through the gingiva are not the same as observations of tooth emergence through the alveolus on dry bones (Demirjian, 1986; Merchant, 1973; Trodden, 1982). In addition, only a few published studies are suitable for use in age prediction because they include data for a number of teeth, consist of large reference samples and have an applicable methodology for reporting ages (Anderson et al., 1976; Demirjian et al., 1973; Demirjian and Goldstein, 1976; Haavikko, 1970; Kataja et al., 1989; Moorrees et al., 1963a,b; Nystrom et al., 1986).

Other studies that use maturity scales (Demirjian et al., 1973; Demirjian and Goldstein, 1976) in reporting tooth formation chronologies are difficult for osteological researchers to apply. Only the standards of Moorrees et al. (1963a,b) include data for both deciduous and permanent teeth from the same series of children. In addition, subjects in all of the above studies are of American or European derivation with almost no data available for other geographic areas of the world.

While these difficulties are widely known and accepted, skeletal biologists working with archaeological material generally have no recourse to test the accuracy of dental age estimates because the true chronological ages of their samples are rarely known. In fact, the question of age accuracy is laid aside in growth-related studies of past populations by invoking the assumption that the most useful comparisons of any group are between the two physiological indicators of maturation, dental age, which is more stable, and skeletal age, which is more sensitive to environmental insults. This observation has been the basis for a number of growth-related studies over the last twenty years (Hoppa, 1992; Hummert, 1983a,b; Jantz and Owsley, 1984a,b; Lallo, 1973; Lovejoy et al., 1990; Mensforth, 1985; Merchant and Ubelaker, 1977; Owsley and Jantz, 1985; Saunders and Melbye, 1990; Sundick, 1978; Walker, 1969; Wall, 1991). While the logic is justified, this assumption ignores the possibility that the dental age distribution of a skeletal sample may not even be representative of the dental age distribution in the living population because of sampling error and/or a bias in mortality.

Only a few researchers (Bowman et al., 1990) have examined the accuracy of dental age estimates in archaeological or forensic samples. Tests of age prediction using children of known age have been reported from studies of living children (for a review of this literature see Smith, 1991b). In a report on a small sample of 23 children, Crossner and Mansfield (1983) found 70% of tooth formation estimates for permanent mandibular and anterior maxillary teeth fell within ± 3 months of true age and discrepancies of no more than 6 months were found for age estimates when tested against the standards of Liliequist and Lundberg (1971) and Gustafsson and Koch (1974). Another comparison of several dental formation standards using permanent teeth (Haag and Matsson, 1985) found standard deviations of difference between dental and chronological age to be approximately 10% of age. The tooth formation standards of Demirjian et al. (1973) estimated a subject’s chronological age (from 2 to 20 years) within 15–25 months of 85% confidence. Later, tests of these Canadian standards on a sample of French children in Lyons found a mean advancement of nine months in the dental development of the French children (Proy et al., 1981). A further test of the methods developed by Demirjian...
and co-workers on Finnish children found that the models predicted dental age reliably (Kataja et al., 1989).

Smith (1991b) points out that none of the tested systems are particularly suited to age prediction. In fact, she has provided a series of recommended formation values for age prediction based on her determination of the appropriate method for constructing chronologies of growth stages. She argues that for age prediction it is more appropriate to assign an age that is the midpoint between the mean age of attainment of a subject's current stage of formation and the subsequent one, since at the time of observation, the subject is in-between the attainment of one stage and the next. Smith's own test of age prediction accuracy utilized four Canadian children of British origin who were used to test dental age standards in the study by Anderson et al. (1976). The chronological ages of these children were compared to her calculated values for predicting age from the stages of permanent mandibular tooth formation derived from the data of Moorrees et al. (1963a) (Smith, 1991b). The results were "remarkably accurate" (Smith, 1991b), differing by a maximum of 0.2 years when schedules for the correct sex were applied and even when schedules for both sexes were combined and applied. Within-individual inaccuracy2 based on a single tooth yielded a standard deviation of ± 0.56 years while mean values for five or more teeth decreased the standard deviation to ± 0.09 years, suggesting that dental age can be estimated to within 2 months for young children. Smith expresses some surprise that the few earlier empirical tests of dental age prediction show good success despite her enumeration of considerable theoretical and methodological difficulties associated with existing dental development standards. As yet, no one has tested her recommended values for age prediction based on stage of development.

The difficulties of dental age estimation of immature human skeletons would appear to be magnified with archaeological samples. These samples tend to be fragmentary and the number of teeth available for mean dental age estimates varies considerably. Consequently, examinations of both the accuracy of individual dental age estimates of personally identified skeletons (from archaeological or forensic circumstances) and the accuracy of skeletal sample distributions of dental age against known burial age distributions would help address many of the concerns cited above. The present study reports on dental formation age estimates of a small sample (n = 17) of personally identified children's skeletons from a nineteenth century historic cemetery from Upper Canada. In addition, distributions of dental age derived from several standards and methods were compared for the total juvenile sample of dentitions (n = 241) to the expected distribution of chronological ages-at-death based on a complete series of burial registers for the same cemetery. The study considers differences in age derived from combinations of three age estimation systems: the standards of Anderson et al. (1976) for permanent teeth, the standards of Moorrees et al. for the permanent (1963a) and deciduous (1963b) teeth. In addition, the variation between the results for these systems is examined as well as the variation in dental age based on different combinations of teeth and the relative confidence that can be expected for age predictions from archaeological cases.

MATERIALS AND METHODS

The St. Thomas' Church cemetery skeletal sample

In 1989, St. Thomas' Anglican Church, Belleville, Ontario was given legal permission to close and partially disinter human remains from a nineteenth century cemetery located on land adjacent to the church property. The church hired a private archaeological contract firm, Northeastern Archaeological Associates, to carry out the disinterments of skeletal remains and associated artifacts. A total of 579 burial shafts and
some 576 individuals were excavated over four months from three-fourths of a hectare or an area of almost one and three-fourths acres. While the true area of the total cemetery is not known, the excavated area is believed to represent approximately one-third of the original cemetery grounds (McKillop et al., 1989; Saunders et al., 1991). Tombstones or grave markers had long since been disturbed by historical events (i.e., two church fires, one in each century) and only a few remained standing on the property.

The excavation area was cleared of soil overburden and individual grave shafts were identified by shovel shining and then carefully excavated by hand. Additional test pitting was used to check for graves not identified by soil changes. The well-drained sandy soil promoted easy and rapid removal of the remains as well as excellent bone preservation. The skeletons were all studied initially at the church and then in the Department of Anthropology, McMaster University, so that full scale collection of data was carried out for one year prior to reburial on church property.

The cemetery was known from parish registers and church vestry minutes to have been used from August 30, 1821 to April 14, 1874 whereafter all subsequent town burials took place at the municipal cemetery (Bellestedt, 1969; Diocese of Ontario Archives [DOA] Series I, part 35A). Based on comparisons to the parish registers, the skeletal sample represents 37% of the total number of 1,564 interments during that period (Saunders et al., 1991). While a subsample of individuals (n = 80) is personally identified from legible coffin plates or coffin fastenings, no cemetery plan or map survives, probably because of its loss during the earlier church fires.

The source of documentation: quality of the records

Records of burials in St. Thomas' cemetery, as well as marriages and baptisms in the church, were kept by the church's ministers and are available for the entire period during which the cemetery was in use. Data include name, age, death date, name of the registrar, burial date, and occasional notes on family relationships as well as cause of death. All records were transcribed to a database management program (Borland International, 1990) and checked twice by different individuals for transcription errors.

A total of 1,564 individuals were recorded as having been interred in the cemetery. An additional 17 individuals were identified as having been buried in other locations besides St. Thomas's cemetery, all of which indicate the attention to detail adhered to by the recorders. Consequently, a series of historical demographic tests were applied to evaluate the quality of the parish registers (Drake, 1974). These involve determining the adequacy of sample size (a minimum of 100 registry entries per year); whether there are obvious and serious gaps in the registers; evidence of under registration; that persons buried in other churchyards are so indicated in the records; and that recorded sexes and ages are not themselves estimates. The analysis of St. Thomas' parish registers (Rogers, 1991) showed that some individuals were buried without record but that these are rare, isolated cases. Reductions in the number of parish events occurred between 1831 and 1835, but the number of burials was unaffected. No significant gaps in the registers were detected and parishioners buried elsewhere were so noted in the register. Individuals of unknown sex represent a nonsignificant proportion of the total sample (less than 2%) and those of unrecorded age appear to be randomly distributed throughout the age range. Consequently, it was concluded that the register data can be confidently treated as a reliable source for comparison to skeletally derived sex and age profiles.

Of the total burials registered, 712 individuals were listed as having died under the age of 15. The ages of 637 of these children (89%) were clearly identified in the burial register, leaving 75 cases that could not be included in the sample. However, by cross-checking the latter against the recorded baptisms in the parish registers, it was possible to identify the specific ages of 45 additional individuals, increasing the registry sample of subadults up to 15 years to 682, or 96%. This is a good example of how record linkage techniques improve the quality of parish record data.
Nineteenth century settlers in the Upper Canada town of Belleville were comprised mainly of immigrants from the British Isles, Ireland and western Europe, as well as descendants of United Empire Loyalists from the United States (Boyce, 1967, 1991; Mika and Mika, 1986). It is also possible to identify the country of origin for many of the buried individuals by reference to baptism or marriage records. In addition, early censuses and assessment records for the town summarize the geographic origins for Belleville's inhabitants (Canada (Province) Board of Registration and Statistics, 1853; Canada Bureau of Agriculture and Statistics, 1878). While it might be expected that all burials in St. Thomas's cemetery would be Anglican or Church of England, in fact, the church sold burial plots to some other denominations such as Methodists and Presbyterians throughout the century, thereby increasing some of the religious and ethnic variability of the burial sample (Herring et al., 1992). One Mohawk Indian and two “persons of colour” were identified in the burial register as having been buried in the cemetery. However, these three individuals were all adults at the time of their deaths.

Age estimation of subadults

Subadult skeletons were first identified on the basis of active tooth eruption and skeleton epiphyseal development and fusion. A total of 282 juvenile or nonmature individuals were initially identified in the skeletal sample. Of these, it was possible to evaluate tooth formation for 241 cases (86%) up to 15 years using X-rays (both panoramic and lateral) and macroscopic observations. Seventeen of these individuals (7%) had been personally identified on a separate occasion by reference to coffin plates (Saunders et al., 1992a).

The 14 developmental stages proposed by Moorrees, Fanning and Hunt (MFH) (1963a) were determined for all present and observable teeth. A single observer was trained to recognize the stages of crown and root formation and double determinations were made from the X-rays. Interpolation charts for the mean dental age of three deciduous teeth (mandibular canine, first and second molars) (MFHd) and ten permanent teeth (MFHp) were prepared from the norm charts published by Moorrees et al. (1963a,b). These values were then calculated by cross comparisons using a database management program to the tooth formation stages recorded for the sample. In addition, mean dental ages for 16 permanent teeth were taken from the tables provided by Anderson, Thompson and Popovich (ATP) (1976) and the same individual tooth mean ages calculated by cross comparisons.

Although Moorrees et al. (1963a) divide permanent mandibular root formation stages into mesial and distal roots, a separate test of the Belleville data found no significant differences between the two roots for degrees of calcification. Consequently, these two root stages were averaged after data collection. Also, since it is not possible to determine the elapsed period since an observed tooth has attained stage 14 (root apex complete), these means were not included in the overall age assessment unless the tooth was deemed transitional between stages 13 and 14, after which a midpoint between the two means was taken.

Overall individual mean age estimates (values for the two sexes pooled) were calculated from the sums of individual mean tooth ages based on the published values for each of the three standards. These values represent “mean age of attainment” estimates. Next, three combinations of these standards were also calculated by averaging the values for each method: (1) ATP (permanent) and MFH (permanent); (2) ATP (permanent) and MFH (permanent) and MFH (deciduous), and (3) MFH (permanent) and MFH (deciduous). First, case-by-case comparisons for the personally identified “knowns” in the skeletal sample were made using the three individual standards and the three combinations of age estimation techniques and compared to documented age at death. In addition, the absolute and mean differences between these standards and documented age were calculated. Simple standard deviations for the overall mean dental age estimates for each individual in the entire sample were also calculated as a gauge of relative between-tooth agreement. Next, within-individual coefficients of variation based on the variation contributed by
different teeth to the overall mean age estimate were calculated for all individuals aged by more than one tooth. Finally, the sample age distributions calculated from the dental remains were compared to the sample juvenile age distribution for the total cemetery, derived from the parish registers.

A second set of age estimates were calculated using "age of prediction" tables derived from the MFH permanent tooth standards interpolated for this study and compared to the "mean age of attainment" estimates as well as to the known chronological ages for the personally identified sample. In all, calculations of mean dental ages were weighted by the number of teeth used in each method for any one individual. The absolute difference between the estimated age of attainment or prediction and documented age can be used as a gauge of accuracy (Smith, 1991b) or corresponding inaccuracy (Lovejoy et al., 1985) of the dental age estimates. The sign of the difference will indicate bias of direction in the method of age estimation. Mean levels and standard deviations of accuracy can also be calculated depending on the number of teeth used to derive an estimate within any one individual.

RESULTS

Table 1 lists the dental age estimates for the three sets of standards (ATP, MFHp and MFHd) alone as well as the three combinations of standards as described in the Materials and Methods for the seventeen individuals personally identified from coffin plates and burial records. The absolute differences between these six estimates for each case and the documented age of each individual as well as the mean differences and standard deviations for the overall sample are listed in Table 2. While considerable variation exists between the estimates for any one individual, the two combination methods of ATP + MFHp + MFHd and MFHp + MFHd produce the best overall accuracy in age estimates as well as the lowest standard deviations. The average difference in age estimation for these two combination methods is approximately one-half year.

Individuals dying under one year of age are, not surprisingly, more closely aged by the MFH deciduous standards but overall sample estimates favour the two combination methods identified above. In addition, very young infants dying at less than one month often cannot be aged by any of these methods because of the lack of sufficiently formed deciduous teeth. In the case of the St. Thomas skeletal sample, a number (n = 25) of such individuals required assessment based on a combination of studies of early deciduous development (Deutsch et al., 1985; Kraus and Jordan, 1965; Lunt and Law, 1974; Prahl-Andersen and van der Linden, 1972).

The proportions of individuals representing annual age cohorts for each of the three

---

**Table 1. Estimated age at death from tooth formation standards (pooled sexes) versus documented age of death for 17 personally identified individuals from St. Thomas' Cemetery skeletal sample**

<table>
<thead>
<tr>
<th>Burial no.</th>
<th>Sex</th>
<th>Documented</th>
<th>MFHp</th>
<th>MFHd</th>
<th>ATP</th>
<th>ATP/MFHp</th>
<th>ATP/MFHd</th>
</tr>
</thead>
<tbody>
<tr>
<td>230</td>
<td>M</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>346</td>
<td>F</td>
<td>0.17</td>
<td>0.39</td>
<td>0.29</td>
<td>0.39</td>
<td>0.29</td>
<td>0.39</td>
</tr>
<tr>
<td>88</td>
<td>M</td>
<td>0.69</td>
<td>0.86</td>
<td>1.29</td>
<td>0.71</td>
<td>3.50</td>
<td>2.40</td>
</tr>
<tr>
<td>93</td>
<td>M</td>
<td>0.75</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>91</td>
<td>M</td>
<td>0.83</td>
<td>1.06</td>
<td>1.00</td>
<td>1.08</td>
<td>3.40</td>
<td>2.20</td>
</tr>
<tr>
<td>62</td>
<td>M</td>
<td>0.86</td>
<td>0.87</td>
<td>1.00</td>
<td>0.83</td>
<td>3.60</td>
<td>2.73</td>
</tr>
<tr>
<td>352</td>
<td>M</td>
<td>1.50</td>
<td>1.52</td>
<td>2.17</td>
<td>1.39</td>
<td>3.70</td>
<td>2.94</td>
</tr>
<tr>
<td>438 B</td>
<td>M</td>
<td>2.00</td>
<td>2.55</td>
<td>3.73</td>
<td>1.57</td>
<td>3.71</td>
<td>3.72</td>
</tr>
<tr>
<td>113</td>
<td>M</td>
<td>2.26</td>
<td>2.41</td>
<td>3.38</td>
<td>1.55</td>
<td>3.81</td>
<td>3.66</td>
</tr>
<tr>
<td>147</td>
<td>F</td>
<td>2.33</td>
<td>1.51</td>
<td>2.05</td>
<td>1.15</td>
<td>3.71</td>
<td>3.34</td>
</tr>
<tr>
<td>463</td>
<td>F</td>
<td>2.33</td>
<td>1.90</td>
<td>2.72</td>
<td>1.28</td>
<td>3.79</td>
<td>3.47</td>
</tr>
<tr>
<td>318</td>
<td>F</td>
<td>2.75</td>
<td>2.83</td>
<td>3.65</td>
<td>1.83</td>
<td>4.09</td>
<td>3.91</td>
</tr>
<tr>
<td>288</td>
<td>F</td>
<td>3.85</td>
<td>3.19</td>
<td>4.11</td>
<td>2.26</td>
<td>4.37</td>
<td>4.30</td>
</tr>
<tr>
<td>398</td>
<td>F</td>
<td>3.94</td>
<td>3.10</td>
<td>3.72</td>
<td>1.88</td>
<td>3.75</td>
<td>3.74</td>
</tr>
<tr>
<td>109</td>
<td>M</td>
<td>4.13</td>
<td>3.28</td>
<td>4.07</td>
<td>2.02</td>
<td>3.94</td>
<td>3.99</td>
</tr>
<tr>
<td>174</td>
<td>F</td>
<td>8.00</td>
<td>7.22</td>
<td>7.22</td>
<td>7.25</td>
<td>7.60</td>
<td>7.46</td>
</tr>
</tbody>
</table>

1ATP, Anderson, Thompson and Popovich (1976); MFH, Moorrees, Fanning and Hunt (deciduous (1963a), and (permanent (1963b)).
TABLE 2. Absolute differences between estimated age at death (years) from tooth formation standards (pooled sexes) versus documented age at death for 17 personally identified individuals from St. Thomas' Cemetery skeletal sample

<table>
<thead>
<tr>
<th>Burial no.</th>
<th>Documented</th>
<th>MFHpM</th>
<th>MFHd</th>
<th>ATP</th>
<th>ATP/MFHp</th>
<th>ATP/MFHpd</th>
</tr>
</thead>
<tbody>
<tr>
<td>230</td>
<td>0.01</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>346</td>
<td>0.17</td>
<td>0.220</td>
<td>0.170</td>
<td>0.220</td>
<td>0.170</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>0.69</td>
<td>0.165</td>
<td>0.600</td>
<td>0.020</td>
<td>2.810</td>
<td>1.705</td>
</tr>
<tr>
<td>93</td>
<td>0.75</td>
<td>0.027</td>
<td>0.750</td>
<td>0.027</td>
<td>0.750</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>0.83</td>
<td>0.233</td>
<td>0.170</td>
<td>0.253</td>
<td>2.570</td>
<td>1.370</td>
</tr>
<tr>
<td>62</td>
<td>0.86</td>
<td>0.006</td>
<td>0.140</td>
<td>0.028</td>
<td>2.740</td>
<td>1.873</td>
</tr>
<tr>
<td>352</td>
<td>1.50</td>
<td>0.020</td>
<td>0.670</td>
<td>0.110</td>
<td>2.290</td>
<td>1.435</td>
</tr>
<tr>
<td>428 B</td>
<td>2.00</td>
<td>0.552</td>
<td>1.734</td>
<td>0.433</td>
<td>1.710</td>
<td>1.723</td>
</tr>
<tr>
<td>113</td>
<td>2.26</td>
<td>0.152</td>
<td>1.017</td>
<td>0.713</td>
<td>1.550</td>
<td>1.405</td>
</tr>
<tr>
<td>147</td>
<td>2.33</td>
<td>0.820</td>
<td>0.285</td>
<td>1.177</td>
<td>1.380</td>
<td>1.010</td>
</tr>
<tr>
<td>463</td>
<td>2.33</td>
<td>0.434</td>
<td>0.393</td>
<td>1.055</td>
<td>1.460</td>
<td>1.140</td>
</tr>
<tr>
<td>318</td>
<td>2.75</td>
<td>0.061</td>
<td>0.900</td>
<td>0.918</td>
<td>1.340</td>
<td>1.159</td>
</tr>
<tr>
<td>298</td>
<td>3.85</td>
<td>0.662</td>
<td>0.363</td>
<td>1.587</td>
<td>0.520</td>
<td>0.450</td>
</tr>
<tr>
<td>398</td>
<td>3.94</td>
<td>0.638</td>
<td>0.225</td>
<td>2.065</td>
<td>0.190</td>
<td>0.203</td>
</tr>
<tr>
<td>109</td>
<td>4.13</td>
<td>0.849</td>
<td>0.060</td>
<td>2.112</td>
<td>0.190</td>
<td>0.143</td>
</tr>
<tr>
<td>174</td>
<td>8.00</td>
<td>0.779</td>
<td>0.782</td>
<td>0.750</td>
<td>0.400</td>
<td>0.542</td>
</tr>
<tr>
<td>148</td>
<td>8.42</td>
<td>2.139</td>
<td>2.139</td>
<td>8.420</td>
<td>2.200</td>
<td>2.178</td>
</tr>
<tr>
<td>Absolute Mean Difference</td>
<td>0.47</td>
<td>0.64</td>
<td>1.17</td>
<td>1.39</td>
<td>1.17</td>
<td>0.58</td>
</tr>
<tr>
<td>s.d.</td>
<td>0.52</td>
<td>0.57</td>
<td>1.93</td>
<td>0.91</td>
<td>0.61</td>
<td>0.49</td>
</tr>
</tbody>
</table>

1ATP, Anderson, Thompson and Popovich (1976); MFH, Moorrees, Fanning and Hunt (deciduous (1963a), and (permanent (1963b)).

individual standards alone: (1) ATP (stippled); (2) MFH permanent (grey), and (3) MFH deciduous (black) are shown in Figure 1. This figure illustrates the differences in the distributions which reflect the sampling biases and truncations of the reference samples from which the original standards were derived. Particularly noteworthy are: the inability of deciduous standards to age beyond nine years, and the omission of cases under three years in the ATP for permanent teeth. The original growth sample used by Anderson et al. (1976) does not include individuals younger than three years of age, whereas the Moorrees et al. (1963a,b) sample begins at birth. As a result, the distribution of age estimates based on the Anderson et al. (1976) standards gradually approaches a minimum of three years but does not extend below this age. This trend is more clearly shown in Figure 2. When individual age estimates using the ATP standards are plotted against corresponding estimates using the MFH permanent standards, the effect of the truncation of the original ATP sample is evident by the over-aging of children below five years.

Variation in age estimations also increases with increasing age of the subjects when within-individual standard deviations are plotted against mean ages. The calculated correlation between the two variables is significant (see Fig. 3) \( r = 0.694, P < 0.001 \). However, a certain proportion of individuals falling in the 2-to-5-year-old age categories seem to have high standard deviations beyond what would be expected if a regression line was calculated. In fact, when these cases were examined, it was determined that the permanent MFH maxillary incisor age estimates were high compared to estimates for other teeth. Since the published formation charts for these teeth in the Moorrees, Fanning and Hunt publication begin at a relatively late stage of development—crown complete—it might be expected that these values would have a biasing effect on age estimations of young children much like the Anderson, Thompson and Popovich standards. This difference is attributed to the Moorrees et al. (1963a) values for the anterior and lateral teeth being drawn from two separate samples.

Calculated within-individual coefficients of variation of age estimates ranged from less than 1 to 52, a broad range of variation. These coefficients were plotted against the number of teeth contributing to a mean dental age (Fig. 4). There is an inverse relationship between variation in age and the number of teeth contributing to a mean dental age. This effect does not occur until at least six teeth contribute to the age estimate but
most of the St. Thomas’ cases preserve five or fewer teeth.

Consequently, the combination method of the permanent and deciduous MFH standards, which had been identified as the best combination method for age estimation, were compared to the 682 burials of individuals up to 15 years listed in the burial records, with the permanent maxillary incisors included and excluded in the estimates (Fig. 5). The comparisons of the proportions of individuals in each age category by year for the dental age estimates are compared to the records’ proportions. It can be seen that both dental age methods are quite close to the records’ values but the method which includes only permanent mandibular teeth from the MFH standards fares slightly better.

To test this observation, documented ages for the cemetery sample were treated as expected values and chi square tests of goodness of fit were run using the two methods (Table 3). The MFHpd (mandibular and maxillary) age estimates were compared to the burial records and returned a likelihood ratio chi square value of 19.10 at a significance level of $P = 0.024$. The calculation of standardized residuals\(^3\) showed that the 6-to-7-year-old frequencies were significantly overrepresented (standardized residual = 3.21) in the skeletal sample and significantly underrepresented in the documents sample (standardized residual = −1.91). When the likelihood ratio chi square was run for the comparison of the MFHpd (mandibular teeth only) dental age estimates to the burial records, a value of 15.77 resulted with a probability level of $P = 0.072$. The calculated standardized residuals again

---

\(^3\)Standardized residuals in excess of ± 1.64 (the 95th percentile of the standard normal distribution) indicate a significant deviation in the observed cell values from those expected. (Reynolds, 1977).
yielded a significant value for the 6-to-7-year age category in both the skeletal sample (standardized residual = 2.99) and the burial records sample (standardized residual = -1.78). This finding, again, indicated an overrepresentation of 6-to-7-year-olds in the skeletal sample. Since cell sizes in the 9 years or over cohorts of the skeletal sample are small, it was necessary to combine the 9–14 year cohorts in the chi square calculation. However, a further assessment of the ratios of 10-to-14-year-olds versus all other cohorts showed no significant differences between the skeletal and documents sample.

Finally, for the ten personally identified individuals with observable developing permanent teeth, determinations were made of the absolute differences between estimates using the MFH permanent mandibular tooth standards, the ATP permanent tooth standards and finally, ages of prediction based on the correct sex and incorrect sex (Table 4). Prediction values were derived following the method described by Smith (1991b) from our originally interpolated MFH permanent mandibular standards.

These results, presented in Table 4 show, first of all, that age estimates using the correct sex standards only, are not necessarily closer to the documented ages than are the standards used from the incorrect sex. In addition, the estimates based on age of prediction standards for Moorrees et al. (1963a,b) are not substantially different from age of attainment standards. While it can be seen that for most individual cases, the age of prediction values move upwards towards an older age estimate, as indicated by the numerical sign of the reported differences, the absolute mean differences (i.e., inaccuracy) for the two types of standards are almost identical.

**DISCUSSION**

Comparisons of mean dental age estimates using several combinations of different dental formation standards, permanent teeth, deciduous teeth, and two different source samples, show very clearly that the combination of permanent and deciduous standards based on the sample of children published by Moorrees et al. (1963a,b) is the best possible method to use in estimating mean dental age from skeletal samples of juvenile skeletons, at least when general population background of the samples is comparable. In the present study, it was assumed that family origins based in the British Isles, Ireland, and western Europe are similar enough to those of the American white children examined by these researchers to have comparable dental development schedules. While slight differences in dental maturity have been found between modern North American and European samples of children (Demirjian, 1986; Demirjian and Lamarch, 1979; Prahl-Andersen et al., 1979; Proy et al., 1981), the absence of maturity schedules reported in the same way as the MFH standards precludes our examination of such differences. The use of, or addition of, the standards published by Anderson et al. (1976) is not recommended because of the truncation of their reference sample at three years of age. This course is recommended even though the ATP standards are easier to apply (the availability of published means and standard deviation values in tabular form for age of prediction) than the MFHpdp.
standards which require interpolation from graphs of the means and variation of dental formation. Here then, is a troublesome methodological problem, since each researcher seems to interpolate the MFH ages each time they are required or create their own charts of interpolation.4

The need for the inclusion of deciduous tooth standards is made more evident by the fact that most skeletal samples contain large proportions of infants and young children. However, there is still a distinct shortage of detailed standards for the early formation of deciduous crowns that can be applied to fetal and neonatal skeletons (Skinner and Goodman, 1992) which cannot be included in MFHp mean age estimates.

Within-individual calculated coefficients of variation (CV) for the total sample of permanent dentitions span a wide range (1 to 52) with an average value of 20. This is in contrast to the low range of CVs (2–16) reported by Smith (1991a) for a small sample of humans. The average CV in her study is 10. In fact, the total range of coefficients of variation that Smith reports for humans along with several Pan and fossil hominid specimens reaches a maximum of 36. There are probably several explanations to explain the wider range in the St. Thomas' sample. These include the fact that most of the St. Thomas' sample of dental ages are based on five teeth or less, whereas Smith's sample of humans all have six or more teeth included in their estimates. Maxillary teeth tend to have higher CVs than mandibular teeth (Demirjian, 1986; Smith, 1991a). Our examination of the St. Thomas cases with CVs over 40 suggests that variability in age estimates is often increased because of large discrepancies between the maxillary incisors (aging upward) and mandibular teeth (aging downward). In addition, some of the highest

---

4The original mean and standard deviation values from the study are not available (Moorrees, personal communication).
standard deviations of within-individual age estimates are found among individuals under five years of age. These factors are produced by an additional sample bias effect from the use of the permanent MFH maxillary incisor standards which only include values from approximately 4 years and beyond, and which are derived from a different reference group that the other values. These observations of coefficients of variation should serve to emphasize that studies of small samples of fossils and source cases may not necessarily present the full picture of population and sample variability (Cope and Lacy, 1992; Plavcan, 1989).

When both deciduous and permanent maxillary and mandibular tooth formation stages are included in mean dental age estimates for this group of young children, the total sample of dental ages using the MF-Hpd combination is significantly different from the chronological age-at-death distribution provided by the parish registers. However, exclusion of the MFH permanent maxillary incisor standards from estimates improves the goodness of fit between the skeletons and records as shown by a decrease in the chi square value even though the 6-to-7-year-old cohort of skeletons still appears overrepresented. Investigation of the absolute numbers of individuals represented at the 6-to-7-year age level in the two samples showed that these ages must be reasonable accurate. It is unlikely that estimations of age would be any more problematic in one age cohort versus all others when the tooth development standards are derived from a longitudinally studied reference sample. In addition, further examination of the mean dental ages of the 15 individuals placed in this cohort finds them normally distributed over the age range.

Rather, the authors speculate that the excavation of St. Thomas' cemetery produced a biased selection of burials of children at this age, and suspect that there was selective burial in the excavated portion of the cemetery resulting from several children dy-
ing from a common cause, possibly drowning. While the recorded causes of death in the parish registers are not numerous, those burials which do have a cause of death listed show an unusually high occurrence of drowning deaths, a problem for nineteenth century Belleville where the fast flowing Moira River which runs through the town experienced dramatic spring runoffs (Moodie, 1853; Herring et al., 1992). Consequently, it appears that any apparent discrepancies between the skeletal sample and the documented burials are due to sampling problems and not to problems of age estimation. In fact, when a further sample of 64 subadult skeletons not ageable by the Moorrees, Fanning and Hunt tooth formation criteria are added into the skeletal sample, the proportion of infants under one year increases dramatically and differs significantly from the documented data. This dif-

**TABLE 3. Standardized residuals for likelihood ratio chi square calculations comparing skeletal sample (n = 241) dental age estimates to documented burials (n = 682)**

<table>
<thead>
<tr>
<th>Age cohort</th>
<th>Skeletal¹ Documents</th>
<th>Skeletal² Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.71</td>
<td>-.42</td>
</tr>
<tr>
<td>2</td>
<td>-.35</td>
<td>.80</td>
</tr>
<tr>
<td>3</td>
<td>-.47</td>
<td>.28</td>
</tr>
<tr>
<td>4</td>
<td>.84</td>
<td>-.50</td>
</tr>
<tr>
<td>5</td>
<td>-.63</td>
<td>.37</td>
</tr>
<tr>
<td>6</td>
<td>-.11</td>
<td>.06</td>
</tr>
<tr>
<td>7</td>
<td>3.21¹</td>
<td>-1.91³</td>
</tr>
<tr>
<td>8</td>
<td>.02</td>
<td>-.01</td>
</tr>
<tr>
<td>9</td>
<td>-.60</td>
<td>.36</td>
</tr>
<tr>
<td>9–15</td>
<td>-1.04</td>
<td>.62</td>
</tr>
</tbody>
</table>

¹MFH permanent and deciduous tooth age estimates with maxillary standards included, compared to documented burials from parish registers.
²MFH permanent and deciduous tooth age estimates with maxillary standards excluded, compared to documented burials parish registers.
³Standardized residuals in excess of ± 1.64 indicating significant deviations in the observed cell values.

---

Two out of the fifteen 6-to-7-year-old children listed in the parish registers have circumstances of death described and both of these drowned in the Moira River.
TOOTH FORMATION AGE ESTIMATION

185

TABLE 4. Differences and absolute overall mean differences between estimated and documented age for age of attainment versus age of prediction schedules for the permanent dentition only

<table>
<thead>
<tr>
<th>Method</th>
<th>MFH, correct sex</th>
<th>MFH, incorrect sex</th>
<th>ATP, correct sex</th>
<th>ATP, incorrect sex</th>
<th>MFH Prediction, correct sex</th>
<th>MFH Prediction, incorrect sex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B109</td>
<td>B113</td>
<td>B148</td>
<td>B352</td>
<td>B147</td>
<td>B174</td>
</tr>
<tr>
<td></td>
<td>-0.74</td>
<td>-0.05</td>
<td>-2.62</td>
<td>0.67</td>
<td>0.50</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>-0.71</td>
<td>-0.22</td>
<td>-2.74</td>
<td>0.67</td>
<td>0.09</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>-0.01</td>
<td>1.52</td>
<td>-1.99</td>
<td>2.20</td>
<td>1.77</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>-0.37</td>
<td>1.37</td>
<td>-2.60</td>
<td>1.65</td>
<td>1.27</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>0.31</td>
<td>0.39</td>
<td>-2.08</td>
<td>0.92</td>
<td>0.15</td>
<td>-0.67</td>
</tr>
<tr>
<td></td>
<td>0.29</td>
<td>2.35</td>
<td>0.96</td>
<td>0.46</td>
<td>0.11</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Mean s.d.       | -0.29            | 0.26              | 0.26             | 0.79              | 0.26                        |
| Mean s.d.       | 0.35             | -0.38             | 0.17             | 0.68              | 0.68                        |
|                   | 1.13             | 1.25              | 1.25             | 0.65              | 0.65                        |

The use of age prediction tables versus age of attainment tables for estimating age does not seem to make much of a difference in the average level of accuracy of age estimation for the Belleville sample of personally identified children, even though it has been argued that age prediction standards should be closer to true age (Smith, 1991b). The calculation of prediction values for the Moorrees et al. (1963a) standards requires two stages, first to interpolate the attainment values from the graphs, and second to calculate the prediction values from these interpolated values. Given the problem of interobserver error in interpolation from these graphs, any errors will be further translated into the prediction values. While it is clear that the differences between estimated and documented age for the MFH age of documented age for the MFH age of prediction values are generally positive, indicating higher estimated ages as expected, this is not always the case so that the absolute mean value of the difference is almost identical to that for the age of attainment estimates. The fact that attainment values do not prove to be any less accurate than secondarily calculated prediction values for the St. Thomas' Cemetery sample would suggest that the added step provides no added benefit. Consequently, the use of age of prediction standards is not necessarily likely to improve dental age estimation in archaeological and forensic cases as indicated by this study. Other methodological problems as identified above are probably more pressing.

The lack of a clear difference in the accuracy of correct sex versus incorrect sex estimates seems puzzling. However, this is likely a result of the fact that most of this sample of ten individuals is under five years of age when there is greater similarity between girls and boys in development timing. In fact, Demirjian and Levesque (1980) found no difference in the timing of dental development between girls and boys up to five years of age. The number of teeth used in age estimation does make a difference as Smith (1991b) showed for the four cases she tested. Dental age estimates in the St. Thomas' sample based on a single tooth have a standard deviation of ± 0.94, while the average standard deviation when all possible teeth are used is ± 0.38 years. However, estimates that include all available teeth, deciduous as well as permanent, are better indicators of overall accuracy so that the standard deviation of error for the MFH pd method, excluding maxillary incisors, as reported in Table 2 which is ± 0.53 years, should be taken as the gauge of how closely one might expect to estimate the dental age of an archaeological or forensic case.

In summary, it is recommended that osteological researchers use the tooth formation standards of Moorrees, Fanning and Hunt for permanent and deciduous teeth as has been recommended by previous researchers (El-Nofely and Iscan, 1989; Merchant and Ubelaker, 1977; Ubelaker, 1987,1989). The standards of Anderson, Thompson and Popovich should not be used for individuals less than five years of age to avoid problems of sample truncation. This study shows that there is no reason to add the ATP estimates to MFH estimates in older children since dental age estimates for these latter standards alone closely approximate chronological age distributions. However, values for permanent maxillary incisors published by Moorrees et al. (1963a) should be excluded from age estimates because of their biasing effect especially on in-
individually in the 2-to-5-year-old range. There does not seem to be a problem with using the MFH "age of attainment" schedules for age prediction although this observation may change with larger samples of tested cases. Researchers can be optimistic about their dental age estimates for archaeological and forensic cases but only as long as skeletal and dental preservation is good, as was the case for the St. Thomas' sample, and observation error is controlled for as is the ever present problem of population variability.

ACKNOWLEDGMENTS

The authors thank St. Thomas' Anglican Church, Belleville, for giving its permission to study the skeletal remains and to transcribe parish records. We also thank the following individuals for their assistance in the completion of this project: Gerry Boyce, Heather McKillop, Loralyn Geise, Andrea Lagan, Laurie Dailey, and Warren Pollett. The following institutions provided facilities and technical assistance: Chedoke Hospital, Hamilton, McMaster University Medical Centre Radiology Department and McMaster University Faculty of Social Sciences. The authors also thank four anonymous reviewers for their comments on an earlier draft of this manuscript. This research was funded in part by the Ontario Heritage Foundation, McMaster University Faculty of Social Sciences and the Social Sciences and Humanities Research Council of Canada (Grant 410-92-1493).

LITERATURE CITED

Canada (Province) Board of Registration and Statistics (1853) Census of the Canadas 1851–52. Presented to both houses of Parliament by command of His Excellency. Quebec: John Lovell, Volume 1.


