Eric Baccino,[1] M.D.; Douglas H. Ubelaker,[2] Ph.D.; Lee-Ann C. Hayek,[2] Ph.D.; and A. Zerilli,[3] D.D.S.

Evaluation of Seven Methods of Estimating Age at Death from Mature Human Skeletal Remains

REFERENCE: Baccino E, Ubelaker DH, Hayek L-AC, Zerilli A. Evaluation of seven methods of estimating age at death from mature human skeletal remains. J Forensic Sci 1999;44(5):931–936.

ABSTRACT: Different approaches to the estimation of age at death in mature human skeletal remains were evaluated utilizing samples from 19 recent French autopsy individuals of known age at death. Methods of estimating age at death from single-rooted teeth, the sternal ends of the fourth ribs, the symphyseal face of the pubis and femoral cortical remodeling were evaluated by two independent observers (three observers for the teeth). Comparison included ages estimated from three more comprehensive approaches utilizing data from the application of two or more of the individual methods. The results indicate that the comprehensive approaches are superior to the individual ones and the success of the latter reflects not only the morphological expression of the aging process, but also the technique complexity and the experience of the investigator. Of the individual techniques, the "Lamendin" dental technique was most effective for individuals of ages greater than 25 years.

KEYWORDS: forensic science, skeletal age estimation, comparison

Accurate estimation of age at death of adult human skeletal remains represents an important component in forensic anthropological analysis. Currently, many individual techniques are available to assist professionals in this endeavor (1). These techniques have been developed from a variety of approaches and source materials and vary in their complexity. Such variation complicates decisions on which techniques should be employed in analysis and which are likely to provide the most accurate results. Because many of these techniques were developed from different samples, their relative accuracy is difficult to compare from the published literature alone.

The research reported here examines the success of four different approaches to the estimation of age at death in mature human skeletal remains of recent French origin. The study compares the relative success of these different methods when applied separately to the same individuals of known age at death. These particular techniques were chosen for examination because they can be applied to samples that are easily acquired in autopsy of complete or partially decomposed skeletons. Comprehensive approaches utilizing multiple age indicators and methods as well as interobserver applications were also evaluated. The research was designed to examine the relative value of using a single technique versus a multifactor technique, and the contributions to age estimation of: a) technique complexity; and b) investigator/observer experience and training.

[1] Centre Hospitalier Universitaire, Montpellier, France.
[2] National Museum of Natural History, Smithsonian Institution, Washington, DC.
[3] Centre Hospitalier Universitaire, Brest, France.
Received 10 Sept. 1998; and in revised form 11 Jan. 1999; 29 Jan. 1999; accepted 29 Jan. 1999.

Materials and Methods

The test sample consisted of skeletal material removed at autopsy from 19 adult individuals. The autopsies were conducted in Brest, France between September 1996 and June 1997. The 19 individuals were of known age at death, ranging from 19 to 54 years with a mean of 37.6 years and a standard deviation of 10.0. Four females (ages 21 to 52, mean 40.5, standard deviation 13.5) and 15 males (ages 19 to 54, mean 36.9, standard deviation 9.4) all were of European ancestry.

The specimens removed at autopsy and available for analysis consisted of the symphyseal area of both pubic bones, the sternal ends of the fourth ribs, at least one single rooted anterior tooth, and complete cross-sections of the mid-shaft right femur. Following removal at autopsy, all specimens had been cleaned and labeled with a specimen number. Information on age at death was maintained separately. For the dental technique, a maxillary central incisor was utilized when available.

Four specific techniques were selected for evaluation. The symphyseal face of the pubis was evaluated using the Suchey-Brooks system (2). This system consists of the assignment of the symphyseal face to one of six phases originally defined from study of a large well-documented autopsy sample from Los Angeles, California. Casts of examples of bones in all phases are available for comparison for both males and females. For each of the six phases for males and females separately, mean and standard deviation values are available. Independent testing of various methods of estimating age at death from the mature pubic symphysis suggests the Suchey-Brooks system offers advantages over the others (3).

The sternal end of the fourth rib was evaluated using the system of İşcan et al. (4–6) and İşcan and Loth (7,8), and the rib data in Ubelaker (1). An eight phase classification system is presented for both males and females. Summary statistics are available in this literature for each phase, including mean, standard deviation and "95% confidence interval."

Single-rooted teeth were evaluated using the technique of Lamendin et al. (9). This method was developed from a French sample of 306 modern teeth originating from 208 individuals. Using this method, age at death is estimated from calculations of the extent of periodontosis and root transparency. The authors (9) provide mean error by decade.

Femoral cortical remodeling was evaluated utilizing the Kerley method (10,11). This method calls for the quantification of the percentage of circumferential lamellar bone and the numbers of primary osteons, secondary osteons, and osteon fragments in four circular fields located at the periosteal border of the cortex. Ages were estimated using the Kerley profile chart (12). Microphotographs were prepared for each of the four fields and utilized by both observers in assessing the structures within the required field size (11). To facilitate data collection, the circular field was marked on each photograph. Each observer independently assessed structures within this field.

Two observers independently examined the four data sources for all 19 individuals in an attempt to compare estimates of age at death. Observer 1 is a forensic pathologist having considerable experience with all of the individual techniques except the Kerley method. Observer 2 is a forensic anthropologist with considerable experience with all of the techniques except the Lamendin method. To ensure an independent assessment, each investigator evaluated each type of sample separately, e.g., dental samples separately from ribs, pubic bones, and cortical sections. This was done to eliminate the possibility that knowledge of the stage of development of one structure might influence the assessment of another structure from that same individual. Ages were not calculated until all observations from one sample group (e.g., teeth) were complete. These two investigators gathered their data at separate times and did not discuss their evaluations with each other prior to the calculation of age.

A third investigator who was a dental specialist with extensive experience with the Lamendin method evaluated only the teeth, using the same procedures as outlined above.

In addition to the individual technique assessments, three comprehensive approaches were utilized by Observers 1 and 2: 1. the mathematical average of the four ages derived from the individual techniques; 2. a two-step procedure (13); and a "global" approach in which each investigator evaluated all the data available for each individual and estimated age at death relying on experience to interpret the results. In the "two-step" procedure, the age generated from the Suchey-Brooks system is utilized if the pubic bones are judged to be within the first three phases. If the pubic bones are in the older phases, then the age generated from the Lamendin method is utilized.

For the global approach, each of the two investigators who examined all of the structures, utilized all of their notes, their calculations of age from the structures, and a re-examination of the structures if needed to produce a comprehensive age estimate. This estimate was generated without knowledge of actual age or age estimated by the other investigator. Actual age at death was revealed only after all of the above procedures were completed.

To facilitate comparison and analysis, single year age estimates were generated. For those techniques that produce a range (e.g., Suchey-Brooks), the mid-point of the range was used as the estimate.

For each of the 19 individuals we were able to evaluate each of the variables. Seven sets of estimated ages resulted from the seven methods being considered. Descriptive statistics were calculated on each set of estimated ages and for the true age of each skeleton. Tests on paired samples were performed for observers over variables and for true versus estimated ages. Hartley's F-max test (maximum variance divided by minimum variance) was used to compare scorer variability on observer variables. A t-test performed on data under conditions of paired observations is more powerful than the usual t-test on the means on the same data set. The t-tests that use the paired observer results tested the hypothesis that the mean difference was zero between observers on each of the methods. The paired samples test for each method tested the hypothesis that the mean difference is zero between the true age of the individual and the age estimated from applying one of the methods. Finally, we tested the hypothesis that each set of estimated ages had the same true mean as that of the actual skeletons.

In order to compare interobserver results we calculated both bias and inaccuracy (bias is the signed difference between estimated and actual age; inaccuracy is the absolute difference). Tests on means of these quantities over each method and for observers between methods were performed. In addition, both parametric and non-parametric correlation analyses were performed. Kolmogorov-Smirnov one sample tests of Normality were run on each of the seven sets of estimated ages for each observer and in each case the distribution of the populations could not be distinguished from Normal. Therefore, the results of the two correlation analyses agreed in each case.

Results

Interobserver Agreement

Histological Method—Both the estimated average ages and the standard deviations of those estimated ages were quite similar, regardless of observers' experience.

No significant interobserver differences were detected when using this method (Table 1). The histological method was the sole method for age estimation that yielded no significant correlation between the estimated ages from each observer. Table 2 shows that Observer 1 recorded estimated ages that on average provided for

TABLE 1—*Descriptive statistics, correlations, and tests for interobserver comparisons of estimated ages of* n = *19 skeletons for each method. Paired sample t-test (column 8) tests the hypothesis that there is no difference between the average difference between the ages estimated by each observer. A significance (column 9) of 'ns' indicates that the two estimates, on average, cannot be said to be different. Column 1 lists each test method with a number (1,2) to indicate the observer. For example, the row labeled P1 contains results for Observer 1 using the pubis method. For the sample of 19 French skeletons, the mean age was 37.6 years; standard deviation, 10.0; standard error 2.3; variance, 100.8; ages ranged from 54 years to 19 years, with an overall age range of 35 years.*

Variable Pair	Means (Est. Ages)	St. Devs.	Pearson Correlation	Spearmann Correlation	Mean Difference	St. Dev of Difference	t	Significance
Avg1, Avg2	40.5, 40.9	7.9, 9.7	.912, p = .000	.873, .000	−0.42	4.0732	−.451	ns
Global1, Global2	38.3, 40.6	11.0, 11.7	.833, p = .000	.745, .000	−2.32	6.62	−1.53	ns
Two1, Two2	37.7, 37.7	11.0, 9.6	.978, p = .000	.982, .000	0.00	2.58	0.00	ns
H1, H2	39.5, 41.5	10.6, 10.8	.413, ns	.297, ns	−1.58	11.57	−.595	ns
P1, P2	36.5, 40.3	11.0, 12.8	.776, p = .000	.827, .000	−3.84	8.16	−2.05	.055, ns
R1, R2	45.4, 42.4	15.0, 14.9	.881, p = .000	.901, .000	3.05	7.28	1.83	ns
D1, D2	40.1, 40.1	8.2, 6.6	.967, p = .000	.972, .000	0.00	2.49	0.00	ns

TABLE 2—*Paired t-tests of the difference between each observer's age estimate from each method with the true age. Tests were run at the 0.05 level; observed p-values listed in column 5.*

Paired Variables	Mean Difference	St. Dev. Diff.	t	p
Avg1, true	2.9	4.3	2.913	0.009
Avg2, true	3.3	4.7	3.105	0.006
Global1, true	0.7	6.1	0.486	ns
Global2, true	3.0	3.8	3.4	0.003
Two1, true	0.0	6.3	0.036	ns
Two2, true	0.0	6.1	0.038	ns
H1, true	2.3	12.5	0.791	ns
H2, true	3.8	7.3	2.297	0.034
P1, true	−1.2	8.2	−0.617	ns
P2, true	2.7	7.7	1.5	ns
R1, true	7.8	8.8	3.851	0.001
R2, true	4.7	9.4	2.190	0.042
D1, true	2.5	7.1	1.528	ns
D2, true	2.5	7.4	1.464	ns

the largest standard deviation of the true and estimated age differences. However, because for Observer 1 the standard deviation of this average difference was very large, the t-test of these differences could detect no significant difference between the histological-estimated and the true ages, on average. This was not the case for Observer 2 (Table 2).

Pubis Method—There was a significant correlation between each observer's results (Table 1), but not a significant difference between each set of estimated ages with the set of true ages (Table 2). That is, each set of estimated ages with the set of true ages, and, both sets of estimated ages were, on average, non-significantly different from the true age values. Observer 2's variability, as measured by the standard deviation, was 12.8, and that for Observer 1 was 11.0 (Table 1) while the true age standard deviation for the skeletons was 10.0. For this method, as with the others, there was no difference between the observers' mean age estimates (Table 1) or between the true ages and either of the set of observers' age estimates (Table 2), on average. However, this method had the largest mean difference between the observers (Table 1). The correlations of true age with estimated age for each observer were highly significant. The one-sample hypothesis test that the true mean was 37.63 for the estimated ages for each observer resulted in non-significance.

Rib Method—The mean age estimates for the two observers based upon the rib method show no significant difference at the 0.01 level, as determined by the paired sample t-test (Table 1). The two sets of age estimates were highly positively correlated, both overestimating the true ages on average (Table 2). The standard deviations of estimated ages obtained with this method were the largest of those from any method (Table 1). The test that each observer's estimated ages averaged to the known average value of 37.63 was non-significant for Observer 2 only. The hypothesis was rejected at the 0.05 level for the estimates for Observer 1. The paired t-test of true age with estimated age for each observer rejected the hypothesis of no difference at the 0.05 level (Table 2).

Dental Method—For the dental method, the mean estimated ages for both observers were the same, within rounding error, while Observer 1 had a slightly, but not significantly, higher variance (Table 1). Both sets of age estimates for this method were highly correlated (Table 1). The paired sample test of no difference between the observer estimates could not be rejected (Table 1). The paired sample test of no difference between each observer's mean estimate and the true value was also non-significant (Table 2). Finally the test of the hypothesis that the true mean of each set of estimates was the actual value of 37.63 was non-significant.

Average Method—For this method, Observer 1 estimated ages with lower variability (Table 1) than did Observer 2. The standard deviation of 7.9 for the first observer for considerably lower, while the value of 9.7 for Observer 2 was within rounding error of the average true standard deviation of 10 for all skeletons in the sample. However, the results of Hartley's F-max test, which compares variances, revealed no significant differences although both observers' standard deviations were small as expected. A paired t-test on the hypothesis that there was no difference between the mean difference of the observers' estimated ages could not be rejected (Table 1). That is, on average, both observers' estimated ages were similar. There was high correlation and no significance found between the two data sets (Table 1). However, when each observer's set of estimated ages from this average method was compared to the true skeletal age (Table 2), both observers' results were significantly different from the appropriate true age, on average, at the 0.05 test level (observed probabilities, $p = 0.009$; $p = 0.006$). Only for this method and the rib method were results significant for both observers when testing that there was no difference, on average, between estimated and true age (Table 2).

Global Method—The mean ages for this method were the second highest, after the rib method, but the standard deviations were considerably smaller than for the rib method (Table 1). The mean difference between the two sets of estimated ages was ranked smaller than only the pubis and rib methods. There was a significant correlation between observers' estimates (Table 1) and also a correlation of each observer's results with the true ages (Table 2). When the paired t-test of the difference between each observer's result with the true ages was computed, Observer 2's result was significantly different (Table 2). The t-tests for the true age as the true mean of the estimates were both non-significant.

Two-step Method—Both observers had an average of 37.7 for this method (Table 1), while the true age was 37.6. The variability for both observers was quite close to that calculated for the true ages. There was high correlation as well as no significant difference between the two sets of estimates (Table 1). There was high correlation between the observers' results (Table 1) and the data set of true ages. The paired t results testing the comparability of the estimated ages for each observer and the true ages were both non-significant (Table 2), as were the results of the one-sample test that the mean estimated age was the true value of 37.63.

Bias and Inaccuracy

Histological Method—Observer 1's bias for the histological method ranged from −17 to +27 (Table 3). The bias for Observer 2 ranged from −12 to +14. While the mean bias for both observers was positive, Observer 2 had a higher average value with a smaller standard error indicating more consistency of scoring. A test of the mean interobserver bias was non-significant. The test that bias, on average, equals zero was non-significant for only Observer 1. This

TABLE 3—*Interobserver bias for each method using* n = 9. *Descriptive statistics, correlations, and paired t-tests of the hypothesis that on average there is no mean bias difference between observers. Note the numerator in the Table 2 t-test uses the means in this table.*

Paired Variables	Minimum	Maximum	Mean	St. Error	Mean Diff.	Std. Error	t	p	Pearson Corr., p	Spearman Corr., p
Global1, Global2	−11, −4	15, 8	0.68, 3.00	1.41, 0.88	−2.3158	1.1580	−1.526	ns	.184, ns	.148, ns
Two1, Two2	−10, −10	17, 13	0.00, 0.00	1.45, 1.40	0.0	0.5923	0.0	ns	.914, p = .000	.896, .000
Avg1, Avg2	−4, −6	10, 12	2.89, 3.32	0.99, 1.07	−.4211	0.9344	−.451	ns	.591, p = .008	582, .009
H1, H2	−17, −12	27, 14	2.26, 3.84	2.86, 1.67	−1.5789	2.6549	−.595	ns	.412, ns	.288, ns
P1, P2	−11, −10	17, 14	−1.16, 2.68	1.88, 1.77	−3.8421	1.8720	−2.052	0.055, ns	.475, p = .040	.563, .012
R1, R2	−5, −12	26, 26	7.79, 4.74	2.02, 2.16	3.0526	1.6710	1.827	ns	.683, p = .001	.616, .005
D1, D2	−10, 10	17, 16	2.47, 2.47	1.62, 1.69	0.0	0.5723	0.0	ns	.941, p = .000	.936, .000

TABLE 4—*Interobserver inaccuracy results for each method using* n = 19. *Descriptive statistics, correlation, and paired t-tests of the hypothesis that, on average, there is no mean inaccuracy difference between observers.*

Paired Variables	Minimum	Maximum	Mean	St. Error	Mean Diff.	Std. Error	t	p	Pearson Corr., p	Spearman Corr., p
Global1, Global2	0, 1	15, 8	4.37, 4.26	0.97, 0.52	.1053	1.1468	0.092	ns	−.098, ns	−.024, ns
Two1, Two2	0, 0	17, 13	4.47, 4.79	1.00, 0.82	−.3158	.5018	−.629	ns	.865, p = .000	.746, .000
Avg1, Avg2	0, 0	10, 12	4.16, 4.89	0.70, 0.65	−.7368	.7717	−.955	ns	.349, ns	.470, .000
H1, H2	1, 0	27, 14	9.53, 6.47	1.85, 1.14	3.0526	1.6050	1.902	ns	.510, p = .026	.599, .007
P1, P2	0, 0	17, 14	6.42, 6.68	1.14, 1.03	−.2632	1.6056	−.164	ns	−.091, ns	−.008, ns
R1, R2	0, 0	26, 26	8.84, 8.00	1.77, 1.54	.8421	1.2871	.654	ns	.705, p = .001	.547, .015
D1, D2	1, 1	17, 16	5.74, 6.26	1.06, 1.01	−.5263	.5587	−.942	ns	.856, p = .000	.762, .000

was the only method for which there was no significant correlation between the two sets of bias values.

The t-tests for mean inaccuracy showed that each of the observer's values were significantly different from 0 at approximately the same level of significance. The two observers mean inaccuracies were quite different; that for Observer 1 was higher, but not significantly. The paired sample test comparing inaccuracy over observer was non-significantly different, and there was a significant correlation between the two sets of values (observed $p = 0.026$) (Table 4).

Pubis Method—The range of bias values for Observer 1 for this method was −11 to 17, with a negative mean of −1.16, while Observer 2's range was −10 to 14, with a positive mean bias of 2.68. When the two sets of observer biases were compared with a t-test no significant difference was found. There was a significant correlation between the bias values at the $p = 0.04$ level. Tests that the average bias was zero were both non-significant.

The observers' mean inaccuracies and standard errors were approximately equivalent. A t-test that these means were equal to 0 was highly significant. However, the paired sample t-test of mean difference of observer's values was not significant and there was no correlation detected between these sets of data.

Rib Method—Observer 1 had a higher mean bias than Observer 2, but not significantly so when tested with 18 degrees of freedom. Bias values for Observer 1 ranged from −5 to 26, while for Observer 2, the range was −12 to 26. The test of the mean difference between the two sets of bias values was not significant for this method, however, the tests for each observer that their average bias was zero were both highly significant. There was a highly significant correlation between the two sets of biases.

For the two observers the average or mean difference between the inaccuracy of their age estimates was non-significant, and their values were highly correlated.

Dental Method—For the dental method, the two observers' mean bias values were identical, with approximately equal standard errors. Both the minimum bias values were −10 while Observer 1 had a maximum of 17, compared with Observer 2's maximum of 16. When the paired sample t-test was run to examine the mean difference in observers' bias, the result was as expected non-significant, since the sample bias was actually 0. This was the only method for which there was zero bias. There was a highly significant correlation between the two bias data sets.

Since inaccuracy is the absolute value the results for the means and standard errors are similar to those for bias. Both tests for zero inaccuracy were highly significant. The paired sample t-test for comparing observers' inaccuracy values was non-significant and there was a highly significant correlation between the two sets of values.

Average Method—The mean bias for the observers was approximately the same; the mean bias was −0.42 and this method had one of the smaller standard errors of the difference. The t-test that this mean difference was actually zero was not significant and there was a high correlation of observer bias results. For both observers their bias scores tested significantly different from zero.

The mean inaccuracy for each observer was similar, with similarly sized small standard errors. The test that inaccuracy of each observer was zero was rejected in each case. The paired t-test for comparability of observer inaccuracy averages was non-significant. There was no correlation between the two observers' inaccuracy results, however.

Global Method—For the global method, which involved degree of experience, we found the most disparate mean biases (Observer 1: 0.6842; Observer 2: 3.0000) and standard errors (Observer 1: 1.4081; Observer 2: 0.8819). The test that these observers' mean biases were actually zero was accepted only for Observer 1. Although there was no significant difference detected between these two data sets, there was no correlation, either.

Two-step Method—This method provided the only mean bias results equal to zero within rounding error, of course resulting in non-significance for a test of that hypothesis. The size of the significant correlation between observers was second only to that obtained for the dental method. When each observer's average bias was tested, the hypothesis of zero mean bias could not be rejected.

This method had a small mean bias difference and the smallest standard error of inaccuracy difference. The t-test was not significant for observer difference and there was a highly significant correlation between observer values.

Three-Observer Comparison

Comparison of three observers is possible only for the dental method. For this method, the mean estimated ages for all three observers were the same, within rounding error, while Observer 1 had a slightly, but not significantly, higher variance, and the third observer's variance was the smallest, but not significantly.

For the 19 specimens, Observer 3's estimated ages averaged 40.11 with a standard deviation of 5.50. The true age range was 19 years to 54 years while Observer 3's estimates ranged from 30 to 53 years, showing difficulty of estimation for both the younger and the older specimens. Observer 1's range was 27 to 60 and Observer 2's range was 27 to 56. Observer 3's results correlated more highly with those of Observer 2, but all three interobserver correlations were highly significant. The paired sample t-test of no difference between each observer's mean age estimate and the true age was non-significant for the first two observers, but highly significantly different for the third observer. When linear regressions were run to attempt to predict true age with observers' estimated ages, the lowest value of the R-square (percent of the variation explained) was for Observer 3's results although all regression tests were significant.

The mean bias for Observer 3 was 2.47, which resulted in a non-significant test result with 18 degrees of freedom that this mean was zero. The inaccuracy was 6.89 for this third observer. The test that mean inaccuracy was zero was highly significant, however, the test that this third observer's estimated ages could have the same 37.63 value as the true ages was non-significant.

Summary

Observer Comparability—The ages estimated on the basis of each of the methods for each observer were non-significantly different. The interobserver correlation was the highest for the two-step and the dental methods but each of the other methods had highly significant correlation between observer scores except for the histological method. The correlation of true age with estimated age was also non-significant for the histological method. Both the pubis and rib methods had very high mean bias interobserver differences (Table 3; column 6), with the pubis method having the larger standard error. The histological method had the largest standard error of any method; the two-step method had no interobserver mean difference and the smallest standard error.

Bias—The range of bias values was the largest for the histological and the rib methods. Observer bias scores for the two-step method were the second most highly correlated of those for any other method. The highest correlation of bias scores was with the dental method, which also had a zero mean observer bias difference. All t-tests of interobserver bias difference were non-significant.

Inaccuracy—The histological method had the largest observer differences; the difference between the mean values was the largest of any method. The variability for this method as defined by the standard error was equal to that of the pubis method and both were considerably higher than those for any of the other methods. Although the pubis method has a small mean interobserver inaccuracy, its variability was quite high between observers and there was no correlation between the observers' values. This is not a good characteristic for an observational method since it indicates that observers' values are unrelated in size. This result is reinforced by the mean bias values, which have opposite signs (Table 3; column 4). Although the global method had the smallest mean inaccuracy difference, the standard error of that mean difference was quite large, but produced a non-significant test result. This is to be expected in a method relying upon expert judgement and experience.

In general, observer standard errors were lower for all combined methods than for the single methods. Tests that mean accuracy was zero were highly significant for both observers.

Conclusion

Comparison of the inaccuracy values between Observers 1 and 2 reveals that: 1. the comprehensive approaches to the estimation of age at death fared better than single methods and 2. the dental "Lamendin" method offered superior results in this French sample when compared to the other single techniques. Variation in accuracy between Observers 1 and 2 in the different approaches appears to represent their relative experience and familiarity with the techniques.

For Observer 1, the forensic pathologist, the lowest mean inaccuracy value (most accurate) was obtained for the average technique (4.2), followed by the global (4.4) and two-step procedure (4.5). Of the individual techniques, most accurate results were obtained from the dental approach (5.7), followed by the pubis (6.4), rib (8.8), and femoral cortical remodeling (9.5) method. The considerable experience of this observer with the Lamendin technique and little experience with the complex histological technique are reflected in these results.

Observer 2, the forensic anthropologist, produced the most accurate estimates with the subjective global technique (4.3) followed by the two-step (4.8), and average (4.9) approaches. Of the individual techniques, the dental method was most accurate (6.3) for Observer 2, in spite of his lack of experience with the technique. The dental technique was followed in accuracy by the femoral cortical remodeling method (6.5), pubis (6.7), and ribs (8.0).

Age estimates by Observer 3, the dental specialist, using the Lamendin method produced an inaccuracy value of 6.9. This value is similar to those produced by the other two observers in spite of Observer 3's greater experience with this technique.

Results reported here generally are consistent with those found by others using similar research designs. Lovejoy et al. (14) found that methods using multiple age indicators offered superior results to those obtained using single indicators.

Saunders et al. (15) applied four methods, including a version of the Suchey Brooks pubis system and the sternal end of the rib technique to a series with documentation on age at death. Adults of both sexes ranging in age from 20 to 87 years were examined. These authors emphasize the need for experience and judgement in evaluating evidence for age at death. They recommend consideration of multiple indicators. Their study reported higher inaccuracy values than found in the present study. However, they examined an age range for the pubis and rib techniques of 72 years compared to only 35 years in this study. Their greater range included the range of ages between 54 and 87, for which ages it is known that age estimation involves larger error.

The present study strongly suggests that comprehensive approaches to age estimation that consider multiple age indicators are superior to isolated methods. Of the individual methods, themselves, the dental Lamendin method offered superior results in this study to others, even with variable experience of observers. This technique also has an advantage of simplicity and utilizing accessible and easily extracted single-rooted teeth. Other single techniques, especially the histological "osteon" method appear more vulnerable to the experience of the observer. Note that in this application of the Kerley method, only the femur was utilized, whereas data are available in the published method for the tibia and fibula as well. This study supports the conclusion of Saunders et al. (15) that all techniques have merit but that the most appropriate approach to age estimation should be one that considers all available evidence and recognizes the value of professional training and experience.

References

1. Ubelaker DH. Human skeletal remains. 2nd ed. Washington: Taraxacum, 1989.
2. Brooks S, Suchey JM. Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. Human Evolution 1990;5:227–38.
3. Klepinger LL, Katz D, Micozzi MS, Carroll L. Evaluation of cast methods for estimating age from the os pubis. J Forensic Sci 1992;37:763–70.
4. İşcan MY, Loth SR, Wright RK. Metamorphosis at the sternal rib end: A new method to estimate age at death in white males. Am J Phys Anthropol 1984;65:147–56.
5. İşcan MY, Loth SR, Wright RK. Age estimation from the rib by phase analysis: White males. J Forensic Sci 1984;29:1094–104.
6. İşcan MY, Loth SR, Wright RK. Age estimation from the rib by phase analysis: White females. J Forensic Sci 1985;30:853–63.
7. İşcan MY, Loth SR. Determination of age from the sternal rib in White males: A test of the phase method. J Forensic Sci 1986;31:122–32.
8. İşcan MY, Loth SR. Determination of age from the sternal rib in White females: A test of the phase method. J Forensic Sci 1986;31:990–9.
9. Lamendin H, Baccino E, Humbert JF, Tavernier JC, Nossintchouk RM, Zerilli A. A simple technique for age estimation in adult corpses: The two criteria dental method. J Forensic Sci 1992;37:1373–9.
10. Kerley ER. The microscopic determination of age in human bone. Am J Phys Anthropol 1965;23:149–63.
11. Kerley ER, Ubelaker DH. Revisions in the microscopic method of estimating age at death in human cortical bone. Am J Phys Anthropol 1978;49: 545–6.
12. Kerley ER. Estimation of skeletal age: After about age 30. In: Stewart TD, editor. Personal Identification in Mass Disasters. Washington (D.C.): Smithsonian Institution. 1970;57–70.
13. Baccino E, Zerilli A. The two step strategy (TSS) or the right way to combine a dental (Lamendin) and an anthropological (Suchey-Brooks System) method for age determination (abstract). Proceedings of American Academy of Forensic Sciences, 1997:150.
14. Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ. Multifactorial determination of skeletal age at death: A method and blind tests of its accuracy. Am J Phys Anthropol 1985;68:1–14.
15. Saunders SR, Fitzgerald C, Rogers T, Dudar C, McKillop H. A test of several methods of skeletal age estimation using a documented archaeological sample. Can Soc Forensic Sci J 1992;25:97–118.

Additional information and reprint requests:
Eric Baccino
Hopital Lapeyronie
371, av. Du Doyen Gaston Giraud
34295 Montpellier Cedex 5
France