Katherine A. Murray,[1] M.A. and Tracy Murray,[1] Ph.D.

A Test of the Auricular Surface Aging Technique

REFERENCE: Murray, K. A. and Murray, T., "**A Test of the Auricular Surface Aging Technique,**" *Journal of Forensic Sciences*, JFSCA, Vol. 36, No. 4, July 1991, pp. 1162–1169.

ABSTRACT: A recently developed technique used in the estimation of age at death is based upon the metamorphosis of the auricular surface of the ilium. This technique was designed to be used in seriation to estimate the age distribution of a skeletal sample. However, the technique has also been used in forensic science cases, which must be analyzed on a case-by-case basis.

The present study examines the use of the auricular surface technique as the single aging factor. Two hypotheses are tested using a sample of 189 individuals drawn from the Terry Collection, housed at the Smithsonian Institution. The hypotheses are (1) is the method equally applicable across race and sex, and (2) how well does the method perform as the single aging factor?

The results indicate that the amount of degenerative change in the auricular surface is not dependent upon race or sex in any given age category. However, the rate of degenerative change is too variable to be used as a single criterion for the estimation of age; the range of estimation error is simply too large for forensic science purposes.

KEYWORDS: physical anthropology, auricular surface, human identification, age at death, race, sex

One of the most important questions facing the forensic anthropologist is the estimation of age at death. Although dental and skeletal maturation can provide a fairly accurate estimate for children, age estimation of an adult individual must be based upon macroscopic and microscopic morphologic changes in the skeleton. Traditional macroscopic methods of age estimation have relied on changes in the pubic symphysis, the degree of cranial suture closure, degenerative changes such as vertebral osteoarthritis, and the degree of dental attrition. Newly developed aging techniques include a modified Todd system for aging the pubic symphysis, and morphologic change of the sternal rib end. Microscopic techniques include osteon counting of long-bone segments and dental histology. Resorption of cancellous bone as seen in studies of radiographs may also provide a general indication of age [*1,2*].

A recently developed technique of estimating age is based upon the metamorphosis of the auricular surface (AS) of the ilium [*3*]. The technique was developed using the Libben archaeological population and the Hamann-Todd autopsy collection. More than 250 individuals from the Libben series and approximately 500 from the Hamann-Todd collection were examined in detail. Lovejoy et al. established eight modal age phases on the basis of textural, organizational, and degenerative transformations on the auricular surface. The technique was then tested using two independent samples drawn from the Hamann-Todd collection; none of the individuals in these samples had been used in the

Received for publication 24 March 1990; revised manuscript received 5 Sept. 1990; accepted for publication 7 Sept. 1990.

[1]Forensic osteologist, Department of Anthropology, and distinguished professor, Department of Economics, respectively, University of Arkansas, Fayetteville, AR.

Copyright © 1991 by ASTM International

development of the technique. The system was designed to "serve as age modes in an idealized metamorphosis of the auricular surface" (Ref *3*, p. 26) and to be used in seriation with other population specific auricular surfaces. When used with other aging techniques, this technique provides the researcher with data that are vital for the development of demographic profiles of a population [*3*].

Although developed for application to large population analyses, the AS technique is often used during forensic science analysis. However, in such cases the analysis generally involves only a single individual, and therefore, the age at death must be determined on a case-by-case basis. When presented with a skeleton for analysis, the investigator is often hampered by loss of some of the key components; field recovery techniques may overlook some vital elements, and animal activity may result in the loss of important surfaces such as rib ends and pubic symphysis. Lack of specialized equipment, time, and training may eliminate microscopic analysis as an alternative technique. In some cases, the auricular surface may be the only evidence available from which to estimate age.

The present study examines the AS aging technique as a single aging factor. Two questions are addressed. First, is the AS method equally applicable across race and sex? Although Lovejoy et al. (Ref *3*, p. 27) maintain that the same age modes could be used to estimate age for both male and female, the applicability to individuals of different races was not addressed. Second, how well does the AS method perform when used as the single aging factor? Lovejoy et al. clearly recognize shortcomings in the AS method; they state, "while many AS will clearly conform to one of the above stage descriptions, many will not, the latter containing combinations of features from two or more phase mode descriptions" (Ref *3*, p. 26). These two questions are subjected to statistical tests using a random sample drawn from the Terry Collection, housed at the Smithsonian Institution.

The Terry Collection

The Terry Collection was assembled from 1914 to 1965 and contains individuals who, by and large, were indigent and lived in or near St. Louis. Robert J. Terry, head of the Department of Anatomy at Washington University in St. Louis, began the collection, and Mildred Trotter continued the work after she became head of the Department. The collection consists of over 1600 skeletons, most of which are of known sex, race, and age at death. Racially, the collection is composed mainly of American white and black individuals, but contains a few Mongolians as well.

The data in Table 1 present the age, sex, and race demographics of the Terry Collection, excluding Mongolians. There are approximately 538 black males (BM), 451 white males (WM), 350 black females (BF), and 276 white females (WF). Some of the individuals in the collection are of unknown or estimated age. This is indicated in the morgue record by a question mark or an age range for age at death. The collection reflects the socioeconomic group from which it is derived and is weighted heavier toward males, black, and older age groups.

Methodology

The first question to be addressed is whether the AS method is equally applicable to individuals of different sex and race. The specific null hypothesis to be tested is that the AS method is independent of sex and race. To test this hypothesis, 50 individuals from each of four categories (white male, white female, black male, and black female) were randomly selected from the Terry Collection. Age at death for each of these individuals was known from the Terry Collection documentation; any individual whose age at death was reported as unknown or estimated was rejected from the sample.

TABLE 1—*Demographic profile of the Terry Collection (without Mongolians).*

Age	WM	WF	BM	BF	Total
11–15				1	1
16–20	5		14	8	27
21–25	1		29	26	56
26–30	10	7	51	27	95
31–35	7	5	52	24	88
36–40	26	11	63	36	136
41–45	29	10	50	26	115
46–50	38	13	54	33	138
51–60	103	46	106	44	299
61+	220	172	106	110	608
Unknown age	12	12	13	15	52
Total	451	276	538	350	1615

Each individual was independently examined and aged according to the age modes as described by Lovejoy et al. [*3*]. Both the black-and-white photographs and the published description of each phase were used to determine an age estimation.[2] To assure that the investigator (KM) was truly objective, an assistant collected the right ilium for each individual in the sample. Some innominates were not available, and therefore, the sample was reduced from 200 total observations to 189. For each individual in the sample, the true age, race, and sex were independently recorded for later use in the statistical analysis. An index number was added to the record so that the age estimated by the investigator could be concorded with the true age, sex, and race. This index number was simply the order in which the individual was examined and aged. To summarize, data for each individual in the sample were collected on the following variables: true age at death, estimated age at death, race, sex, and the index number.

Even though the investigator was experienced in using the AS method, the index number was also used to test whether the estimation error (estimated age minus true age) varied as the investigator moved through the sample. The investigator might unconsciously use experience gained through examining individuals to improve the estimation of age for individuals examined later in the sample. If this were to occur, the variance of the estimation error would be larger for groups of individuals examined earlier and smaller for those groups of individuals examined later. A statistical test for such an error bias was conducted; it was strongly rejected.

The primary hypotheses to be tested are that the AS method of estimating age at death is independent of race and sex. These hypotheses were tested using traditional statistical analysis of variance (ANOVA); the dependent variable is the error in estimating age at death and the independent variables are race and sex. The estimation error was defined as the difference between the midpoint of the estimated age range and the true age, that is,

$$\text{Error} = (\text{AGEu} + \text{AGEl})/2 - \text{True Age}$$

where AGEu and AGEl are the upper and lower ages of the estimated age range.

Results

The sample calculations yielded *P* values of less than 0.01 for race and 0.55 for sex, thereby rejecting the independence hypothesis for race but not for sex. This result leads

[2]Recently, Bedford et al. presented a revised version of the method using color photographs [*4*]. Although our study was not conducted using the color photographs, it is possible this method may prove to be more successful.

to the conclusion that the AS method can be applied equally to males and females, as claimed by Lovejoy et al. [3]. However, it is not equally valid for black and white individuals.

A linear regression model was developed to estimate the effect of race on the bias in using the AS method. Since sex tested insignificant in the analysis of variance test, it was not included in the regression model. We were concerned, however, that the magnitude of the bias for race might be related to the true age of the individual. Therefore, the regression model specified the estimation error as the dependent variable and included race and true age as independent variables. The results were unexpected. The estimated bias caused by race was small and statistically insignificant—a result that contradicts the previous analysis of variance test. A second surprise was that the error was positively related to true age (the P value was less than 0.01).

These surprising results led to a careful reexamination of the analysis of variance results. It was observed that, on average, the true age was underestimated by almost 13 years. This latter result was puzzling given that the investigator applying the AS method was experienced in its use; furthermore, the statistical test for observer bias (as discussed in the methodology section) was strongly rejected.

These results raise serious questions regarding the AS method: Is the method more suitable for estimating age at death for some age groups than other age groups? To examine the influence of true age on the results, the total sample was divided into three subsamples by age: ages 20 through 49 (63 observations), ages 50 through 60 (41 observations), and ages 61 and over (85 observations). The analysis of variance procedure was rerun using each of these three samples, testing only for independence of race. Race was significant only for the youngest sample. But, interestingly, the average error was less than one year for this sample. In contrast, the true age was significantly underestimated, on average, for the older two samples—by 10.5 years for the middle-aged sample and by 23.7 for the oldest sample.

This puzzling result raised questions about the age, race, and sex distribution of the Terry Collection sample. A casual inspection of the sample revealed that the average age is over 50, that the older population is predominantly white and the younger population predominantly black. The age, race, and sex distribution of the sample is presented in Table 2, which presents the number of sample observations and the average true age by race, by sex, and by age group are found.

Since the age and race distribution of the sample is not uniform, one questions whether the test of the hypothesis that estimated age using the AS method is independent of race is really a surrogate for an untested hypothesis that the AS method is biased by true age. A valid test of the independence of race would require a sample with similar age distributions for both races. Such a sample could be obtained by removing the younger and older observations from the initial sample. Roughly one third of the youngest and oldest observations fall below age 40 and above age 60; the remaining 69 observations included

TABLE 2—*Age, race, and sex characteristics of the Terry Collection sample.*

	Black				White				
	Male		Female		Male		Female		Total No.
	No.	Age	No.	Age	No.	Age	No.	Age	Observations
Age 20–49	25	35	22	36	11	42	5	38	63
Age 50–60	10	55	8	55	12	56	11	56	41
Age 61 over	10	77	17	80	24	70	34	73	85
Full sample	45	49	47	55	47	60	50	66	189

16 black females, 19 black males, 13 white females, and 21 white males, for a total of 35 blacks and 34 whites.

This subsample (ages 40 through 60) was used to test the hypothesis of independence of race and sex; sample *P* values of 0.21 and 0.57 strongly reject the hypothesis that the AS method is influenced by race or sex. These results lead to the conclusion that the AS method is equally applicable to individuals of both races and sexes in these age groups. However, even for this subsample, the true age exceeded the estimated age by an average of 9.2 years.

This last result implies that while the AS method is unbiased regarding race and sex, it is not very accurate when used as the only instrument for estimating age. On the other hand, the fact that the average error is large may not be telling the entire story. Table 3 presents the distribution of the estimation error by true age categories.

Clearly, the majority of cases in which the true age is underestimated by at least two age categories involves individuals with true ages of 50 and over. Is it possible that the AS method would perform better if the older individuals were excluded? In all, 44% of the individuals were aged correctly or within one age category. Eliminating those individuals with true ages of 50 and over raises the above percentage to 57%. Thus, even without the problem of the older individuals, the AS method yielded incorrect age category estimates (by at least two categories) in more than 40% of the observations in this subsample. This age bias is considered too large for the AS technique to be used as a single and reliable aging method in forensic science cases.

Figure 1 represents the demographic profiles both of the true age at death for the entire Terry Collection and of the true and estimated age at death of the sample. The similar profiles of the sample and of the Terry Collection illustrate the representativeness of the sample. The profile produced by the estimated age clearly demonstrates the tendency to underestimate the true age. The percentage of age estimates falling between the ages 30 to 50 is much larger than those which actually occur in the sample.

The uniform tendency to overage younger individuals and to underage older individuals lead us to believe that the AS method could be adjusted to yield reliable estimates. Further efforts to develop a regression model to correct for the age bias in applying the AS method were disappointing. Regression equations were developed with acceptable

TABLE 3—*Distribution of estimation errors by true age.*[a]

	Frequency Distribution of Estimation Error					
True Age Category	Underestimate Two or More Age Categories	−1	0	+1	Overestimate Two or More Age Categories	Total
20–24	*	*	1	0	5	6
25–29	*	1	1	1	2	5
30–34	0	1	2	4	4	11
35–39	0	4	4	5	0	13
40–44	9	5	2	2	0	18
45–49	7	2	1	0	0	10
50–60	25	8	5	3	*	41
61+	53	17	15	*	*	85

[a]NOTE: The estimation error is reported in "number" of age categories. For example, there were 11 individuals with true age between 30 and 34; one was estimated in category 25–29; two were correctly estimated to fall in age category 30–34; four were estimated to be in age category 35–39; and four were estimated to fall in age categories 40–44 and above.

Since the sample was limited to adult individuals, one cannot underestimate the true age of the youngest age group nor overestimate the true age of the oldest age group. Similar reasoning prohibits large estimating errors for the second youngest and second from the oldest age categories. An asterisk (*) denotes such cases.

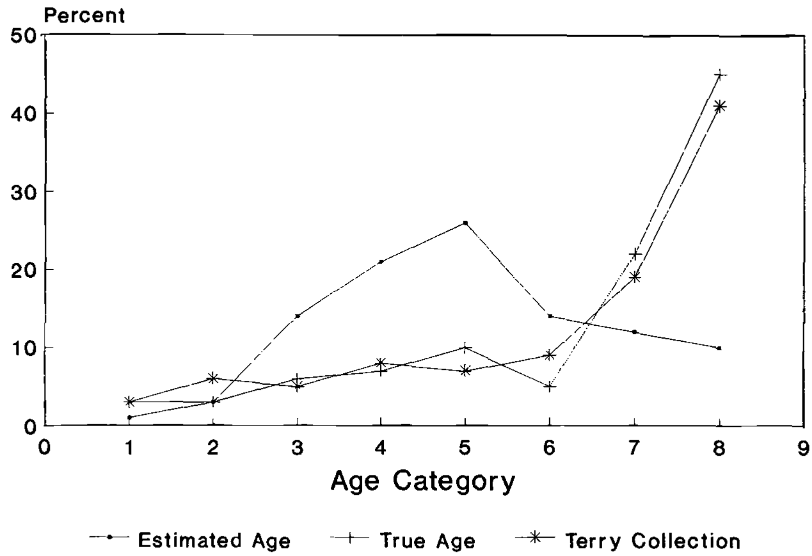

FIG. 1—*Demographic profiles both of the true age at death for the entire Terry Collection and of the true and estimated ages at death of the sample. Age categories for the estimated age and the true age are* (1) *20–24 years,* (2) *25–29,* (3) *30–34,* (4) *35–39,* (5) *40–44,* (6) *45–49,* (7) *50–60, and* (8) *61+. Age categories for the Terry Collection are* (1) *21–25,* (2) *26–30,* (3) *31–35,* (4) *36–40,* (5) *41–45,* (6) *46–50,* (7) *51–60, and* (8) *61+.*

statistical properties; that is, the regression estimates were unbiased. However, the confidence intervals were so large that the resulting estimated age ranges were considered too wide for forensic science purposes.

Comparison with Results of Lovejoy et al.

Lovejoy et al. conclude that "the accuracy and reliability of the present system compares favorably with other systems utilizing the pubic symphyseal face" (Ref *3*, p. 28). In contrast, the results of the statistical tests presented in this paper lead to the conclusion that the AS method is too unreliable to be used as a single aging technique for forensic science purposes. Why is it that our conclusions seem to contradict those of Lovejoy et al.?

There are three obvious sources of contradiction. The first deals with the reliability of the investigators using the method to estimate age at death. The description of the methodology used by Lovejoy at al. (Ref *5*, p. 4) to assure the objectivity of the investigator leaves us with no doubts regarding the credibility of their methods. We also took similar precautions and, therefore, do not consider this a source of contradiction.

Second, the studies are conducted using different samples. If either or both of the samples are unrepresentative of the underlying real population, then differences in results are to be expected. In this case, however, serious questions would exist as to which study, if either, is reliable.

Both studies measure the bias as the difference between the estimated age and the true age; a positive error means that the individual's age was overestimated. Even though the age distributions of the two samples were very different (the median age of the Lovejoy et al. sample is approximately 40, whereas for our sample the median age is 60), the pattern of the estimation errors is very similar. On average, we overestimated the age of individuals whose true age was 34 or under and underestimated those with true age of 40 and up; our estimates were unbiased only for that group of individuals falling in the age category of 35 to 39 (see Table 3). In comparison, Lovejoy et al. (Ref *3*, Table 3, p. 27) overestimated individuals whose true age was 39 or under and underestimated those with true age of 40 and up.

Regarding the magnitude of the bias, for the age category of 40 through 60, our average bias was 9.2 years, whereas they report a bias of 5.9 years for the 40-to-50 age group and a bias of 8.3 years for the 50-to-60 age group. Clearly, the magnitude of our bias is somewhat larger. However, we have a higher percentage of older individuals, which tends to yield larger underestimates; this might explain the difference in the magnitude of the average bias.

From the evidence available, it appears that the two samples are telling the same story. The AS method yields overestimates of true age for younger individuals and underestimates for older individuals, with the crossover age between 35 and 40. We do not believe that the difference between the samples drawn from the Hamann-Todd and Terry Collections is the source of contradiction between the studies.

The third source of possible contradiction deals with statistical methodology, which includes both the objective of the study and the statistical tests being used. Lovejoy et al. *compare* the errors in estimating age at death using the AS method with the errors obtained using other methods. Their concern is whether the AS method compares favorably with alternative methods. They use as criteria the bias (tendency to over- or underestimate the true age), inaccuracy (the magnitude of the error without regard to whether the error is positive or negative), and reliability (whether different observers would achieve similar results). They report summary calculations of the average bias and the average inaccuracy for the alternative methods and find that the AS method compares favorably (Table 2 in Ref *5*, p. 7).

Our objective is quite different. Initially, we set out to *test the formal hypotheses* that the AS method was applicable across race and sex; Lovejoy et al. state that, in general, the AS method was equally applicable to individuals of both sexes, though they made no claims regarding race (Ref *3*, p. 27). Our results showed that indeed the AS method could be applied equally to all individuals without adjustments for race or sex. However, interim statistical results persuaded us to question whether the AS method was suitable for forensic science purposes *as a single aging technique*. We found that it was not. We further suspect that if Lovejoy et al. had conducted our tests on their sample, they would have arrived at the same conclusion. But since we test the AS method as a single aging technique, our results are not suitable for comparing the AS method with other methods. Moreover, our results do not refute their primary conclusion that "when used in conjunction with other age indicators [the AS method] will also improve the reliability of estimates of age at death in forensic anthropology" (Ref *3*, p. 28).

Conclusions

The results of this study indicate that the amount of degenerative change in the auricular surface is not dependent upon race or sex in any given age category. However, degenerative change in the auricular surface is much too variable across individuals to be used as a single criterion for age estimation. Used in conjunction with other aging techniques, the Lovejoy et al. AS aging technique may provide additional information which will help the investigator to estimate an age at death, but its tendency to overage younger individuals (below 30 in our sample) and underage older individuals (above 35 in our sample) makes the method too unreliable for forensic science purposes, and therefore it should not be used as the only criteria for estimating age at death.

Acknowledgment

The authors would like to thank the Department of Anthropology, Smithsonian Institution, for providing access to the Terry Collection. In addition, we would like to thank Dr. Jerome Rose (University of Arkansas-Fayetteville), and Dr. Ted Rathbun (University of South Carolina-Columbia) for their comments and suggestions on earlier drafts. We are also indebted to Scott Murray and Todd Johnston, who provided invaluable research assistance.

References

[1] Ubelaker, D. H., *Human Skeletal Remains Excavation, Analysis, Interpretation*, Taraxacum, Washington, DC, 1989.

[2] Krogman, W. M. and Iscan, M. Y., *The Human Skeleton in Forensic Medicine*, Charles C Thomas, Springfield, IL, 1986.

[3] Lovejoy, C. O., Meindl, R. S., Pryzbeck, T. R., and Mensforth, R. P., "Chronological Metamorphosis of the Auricular Surface of the Ilium: A New Method for the Determination of Age at Death," *American Journal of Physical Anthropology*, Vol. 68, No. 1, Sept. 1985, pp. 15–28.

[4] Bedford, M. E., Russel, K. F., and Lovejoy, C. O., "Auricular Surface Aging Technique," poster presentation, 58th Annual Meeting of the American Association of Physical Anthropologists, 7 April 1989.

[5] Lovejoy, C. O., Meindl, R. S., Mensforth, R. P., and Barton, T. J., "Multifactorial Determination of Skeletal Age at Death: A New Method with Blind Tests of its Accuracy," *American Journal of Physical Anthropology*, Vol. 68. No. 1, Sept. 1985, pp. 1–14.

Address requests for reprints or additional information to
Katherine A. Murray, M.A.
Department of Anthropology
417 Hotz Hall
University of Arkansas
Fayetteville, AR 72701