J Forensic Sci, Sept. 2004, Vol. 49, No. 5
Paper ID JFS2003348
Available online at: www.astm.org

Daniel L. Osborne,[1] *M.A.; Tal L. Simmons,*[2] *Ph.D.; and Stephen P. Nawrocki,*[3] *Ph.D.*

Reconsidering the Auricular Surface as an Indicator of Age at Death*

ABSTRACT: Using standards established by Lovejoy et al. (1) to estimate age at death from auricular surface morphology, 266 individuals of documented age, sex, and ancestry from the Terry and Bass Donated Collections were scored. Analysis of covariance (ANCOVA) indicates that for the factors that could be controlled, age is the sole influence on auricular surface morphology. Ancestry and sex had no significant effect on auricular phase expression. No evidence of secular changes was detected when comparing the Terry Collection (early 20th century) to the Bass Collection (later 20th century). Pearson correlations reveal that several of the subcomponents of the auricular surface (superior and inferior demifaces, left and right sides, transverse organization, texture) correspond with age equally well, although a combined scoring of all features performs slightly better than any one indicator taken alone. Not surprisingly, only 33% of the sample was correctly aged when using the 5-year age ranges provided by Lovejoy et al. (1), suggesting that the published ranges are much too narrow to be used in forensic contexts. To assess the variation in age per phase, standard descriptive statistics and error ranges were calculated and can be employed by forensic anthropologists when estimating the age of an unidentified decedent. Because the mean ages of some of the eight phases did not differ significantly from one another, a modified six-phase system is presented. The auricular surface performs as well as any other single skeletal indicator of adult age. This research suggests that a statistically-informed approach should be taken in order to fully understand the drawbacks and limitations of any aging method.

KEYWORDS: forensic sciences, forensic anthropology, skeletal age estimation, auricular surface, os coxae

Perhaps the greatest challenge to those investigating human skeletal remains is the problem of accurate age estimation. Although standards are currently available for the estimation of adult age at death from several skeletal indicators (1–9), many fall short of their desired levels of accuracy (7–12). Confusing the issue is that the age ranges provided by some authors for their methods (e.g., auricular surface and sternal rib end techniques) do not adequately describe the full range of variation in age that exists per phase or stage (7–13).

Lovejoy et al. (1) developed a method by which age at death can be estimated by examining morphological features of the ilium. In this study the authors selected specimens from the Hamann-Todd Collection whose known ages fell within specific five-year increments. A phase system was created based on the modal morphological condition for each age cohort. The most immediate application of this method is in paleodemographic research, whereby the relative ages of individuals in a population are determined by seriation (1,14). However, forensic anthropologists have employed the method to estimate specific ages for unidentified individuals. Unfortunately, actual error ranges for each phase have never been provided. This study documents the application of the auricular surface technique of age estimation on a case-by-case basis to two U.S. population skeletal samples of documented age at death in an attempt to increase the method's applicability in forensic settings.

Literature Review

Lovejoy et al. (1) conducted the original study on age-related changes of the auricular surface. Materials used in this analysis include the Libben Collection housed at Kent State University ($n = 250$), the Hamann-Todd Collection curated at the Cleveland Museum of Natural History ($n = 500$), and forensic cases from the Cuyahoga County Coroner's Office ($n = 14$). The authors noticed a correlation between age estimated through several regions of the skeleton and auricular surface morphology (1,14). After recognizing the regularity with which auricular surface morphology appears to have changed with age, the authors analyzed how these changes in morphology correlated with age.

Their study included a new method of age estimation based on metamorphosis of the auricular surface, the application of a case-by-case seriation and a systematic multifactorial method of age determination (15). The results of the study defined eight stages of metamorphosis divided into five- and ten-year increments, spanning a range of 20–60+ years. Seriation was applied to minimize research time and reduce the chance of inter-observer error (15). The age ranges were created and the modal features recorded for each phase in the technique. Sample photographs of each stage of metamorphosis were provided in the original study and have since been updated (16). Subsequent tests of accuracy were conducted using specimens from the Hamann-Todd Collection. Two blind tests were run with sample sizes of 100 and 110, respectively, which were drawn randomly from the Hamann-Todd Collection. Results showed that the use of the auricular surface aging technique as a single indicator of age-at-death is comparable to or better than any other adult aging technique. The importance of the auricular surface technique in a multifactorial determination of skeletal age was

[1] Department of Anthropology, Western Michigan University, Kalamazoo, MI.
[2] Department of Forensic and Investigative Sciences, University of Central Lancashire, Preston, Lancashire PR1 2HE UK.
[3] Department of Biology, University of Indianapolis, Indianapolis, IN.
* This research was partially funded by the Western Michigan University Graduate Research Grant, and was presented at the 2001 American Academy of Forensic Sciences meeting. A version of this paper won the J. Lawrence Angel Student Prize.
Received 18 Oct. 2003; and in revised form 1 Feb. 2004 and 7 April 2004; accepted 9 April 2004; published 4 Aug. 2004.

also demonstrated by Lovejoy et al. (15). The authors presented a multifactorial method that uses a principle components weighting of five indicators of age. These indicators included the pubic symphyseal face, auricular surface, radiographs of the proximal femur, dental wear, and suture closure.

Lovejoy et al. (16) used the comparative anatomy of primates to demonstrate the efficacy of the pubic symphysis and the auricular surface in the estimation of age-at-death. In this study, the authors argue that the pubic symphysis is valuable for age estimation up to and immediately following the formation of the ventral rampart (typically in the third or fourth decade of life; phases I-V in the Todd system) and that changes in this region following the fusion of the ventral rampart "offer only minimal special ability to systematically chronicle advancing age" (16, p. 33). Auricular surface morphology, however, changes in a regular way throughout life. Therefore, while the pubic symphysis is a helpful indicator of age into the third and fourth decades of life, the auricular surface is a more reliable indicator of age beyond the fourth decade of life. The authors also argue that any attempt to estimate age from skeletal remains should consider all regions available and when faced with large demographic samples, seriation should be employed in order to avoid or reduce inter-observer error.

Meindl and Lovejoy (14) reviewed their original study in *Age Markers in the Human Skeleton* (17), stressing the importance of the auricular surface aging technique in the study of paleodemography. In this article the authors argued that auricular surface morphology is a valuable indicator of age because: (1) it is more durable than other regions through which age can be estimated (i.e., pubic symphyses and sternal ends of ribs), and thus more likely to be present upon examination of skeletal material recovered from an archaeological context; and (2) the results produced in tests of the auricular surface method indicate a higher frequency of correctly aging individuals past the fifth decade of life, thus giving a more accurate paleodemographic profile. The authors also argued that the most accurate means of determining age at death in the human skeleton is accomplished via the multifactorial method. The use of the auricular surface technique in the context of forensic anthropology was not discussed.

A test of the accuracy of the auricular surface aging technique on a case-by-case basis was conducted by Murray and Murray (13). In a blind study of 189 autopsied individuals of known age at death from the Terry Collection, the auricular surface aging technique was employed to determine the accuracy of the technique across ancestry and sex. In their conclusions the authors suggest that degenerative change is not dependent upon either ancestry or sex. The investigators discovered, however, that the auricular surface technique had a tendency to underage specimens by almost 13 years. This suggested that the rate of degenerative change was unreliable as a single indicator of age and that errors made in estimating age with auricular surface morphology occurred too frequently for the method to be useful to forensic anthropology. This also suggested that while the age ranges defined by Lovejoy et al. (1) may seriate a large sample, they do not reflect individual variability of the auricular surface morphology related to age.

Buckberry and Chamberlain (7) offered a revision of Lovejoy et al.'s (1) method using a components scoring system. In this study the authors created a scoring system in which morphological features are scored independently and then summed to get a composite score. The composite scores were divided into seven stages, which correspond with age ranges derived using methods similar to Brooks and Suchey (2) and Katz and Suchey (5). Initial development of this technique utilized a medieval cemetery from Blackgate, Newcastle, UK. Subsequent testing of the technique took place utilizing the Spitalfields Collection, a known-age sample from London, UK. Buckberry and Chamberlain (7) found no sex-related differences in the method and reported a 0.63 correlation with age.

Osborne (8) tested and refined the Lovejoy et al. (1) method using the Terry and Bass Collections of known-age individuals. The goals of this analysis were to create more discrete phase categories based on exclusive morphological features and to determine what statistical method would provide the most accurate age range per phase. In doing this, standard descriptive statistics and probit analysis were compared using inaccuracy and bias statistics. Results indicate that use of standard descriptive statistics provides the most accurate representation of age-related variation in auricular surface morphology. The author used Lovejoy et al.'s (1) definitions of features, with the exception of microporosity as it has been deemed difficult to differentiate from weathering. The results of this analysis included the collapsing of the eight-phase system to a six-phase system that utilized mean ages and 95% prediction intervals to create age ranges.

Schmitt et al. (9) provide a multifactorial method of age estimation using various features of the auricular surface and pubic symphysis. In this study the authors discuss the application of a separate scoring system for individual indicators of age using a geographically heterogeneous sample, but they do not elaborate on the details (i.e., actual application) of this method. Schmitt et al. (9) do report greater repeatability with this technique, but find no evidence that multifactorial analysis of age estimation is any better than using a single indicator alone. The authors suggest that the variable sex does not influence age estimation, but that there is variation at the population level in rates of senescence.

Materials and Methods

The two skeletal collections used are the Bass Donated Collection ($n = 72$), housed at the University of Tennessee, Knoxville, and the Terry Collection ($n = 194$), housed at the United States National Museum of Natural History at the Smithsonian Institution in Washington, DC (Table 1). The Terry Collection was collected between 1914 and 1965 by Terry and Trotter (18). The Bass Donated Collection, collected since 1981, represents a contemporary sample of the U.S. population. During the examination, the left and right superior and inferior demifaces were scored independently of each other using standards set forth by Lovejoy et al. (1). The superior and inferior demifaces were designated using an arbitrary sectioning point extending posteriorly from the arcuate line. Surface features were recorded without the knowledge of the individual's true age. The senior author collected all data (for more details see Osborne (8)).

TABLE 1—*Sample distribution by sex, ancestry, collection, and decade.*

	Black Males	Black Females	White Males	White Females	Total
16–20	2	7	1	0	10
20–29	9	8	11	3	31
30–39	8	9	19	11	47
40–49	12	10	18	9	49
50–59	12	8	16	11	47
60–69	9	10	12	12	43
70–79	6	9	9	9	33
80–89	2	0	3	1	6
Total Terry	49	58	45	42	194
Total Bass	11	2	44	15	72
Combined	60	60	89	57	266

TABLE 2—*ANCOVA model.*

Dependent Variable	Independent Variables
Phase =	Sex + Ancestry + Collection + (Sex * Ancestry) + (Sex * Collection) + (Ancestry * Collection) + (Sex * Ancestry * Collection) + Age

The data were statistically analyzed in four ways: (1) Pearson correlations, (2) analysis of covariance (ANCOVA), (3) calculation and comparison of inaccuracy and bias statistics, and (4) calculation and comparison of means and 95% prediction intervals. SPSS v. 9.0 (19) and SYSTAT v. 5.2 (20) were used in this analysis.

In order to determine whether auricular surface morphology is influenced by factors other than age, an analysis of covariance (ANCOVA) was run on the combined dataset. ANCOVA is designed to test whether treatments (e.g., sex, ancestry, and collection) affect the dependent variable phase while holding age constant. The general model for the ANCOVA used here is provided in Table 2 (21). Variables in parentheses are interactions. Age is a continuous covariate and therefore does not enter into the interactions.

Although analysis of (co)variance is traditionally used on continuous dependent variables (such as osteometrics), we believe that the use of phase as the dependent variable is justified. First, auricular surface morphology changes in a continuous fashion with increasing age as the joint is subjected to repeated microtrauma, and discrete phases are only assigned for ease of scoring during osteological analysis. Second, there are a relatively large number of phase categories (eight), which helps to more closely mimic a truly continuous variable. Third, while the differences in morphology between each adjacent phase may not necessarily be exactly equivalent (proportional) across the entire range, the method of scoring was originally designed to track modal morphological changes that occur in subsequent five-year age brackets (1), which imparts a degree of regularity to the system. In short, the auricular surface phases contain more information than traditional ordinal data, and the use of nonparametric tests designed for nominal or ordinal variables would probably result in a reduction in sensitivity in hypothesis testing. Instead, a parametric test that can control for a number of variables simultaneously, such as ANCOVA, would seem to be more appropriate for age-degenerative biological data.

Analysis of this dataset indicates that for phases with more than a few individuals, age is normally distributed or nearly so within each phase. Levene's test on the residuals of the full-model ANCOVA is insignificant for each main effect and each interaction, indicating homogeneity of variances. Furthermore, the covariate age was entered into the ANCOVA model with the main effects (sex, ancestry, and collection) to test for heterogeneity of slopes. The results of these tests were insignificant ($p < 0.42$), suggesting that the data meet the assumption of homogeneity of slopes. Since the underlying requirements for ANCOVA have been met, and given the recognized robusticity of ANCOVA in minor departures from normality (22) in addition to the quasi-continuous nature of phase, the use of ANCOVA is justified and, indeed, should provide a very powerful test for the effects of other variables on auricular surface morphology.

Results

Pearson correlations were used to assess the strength of the relationship between age and phase (Table 3). Correlations were calculated separately for transverse organization, texture, the left and right superior and inferior demifaces, and all indicators combined.

TABLE 3—*Pearson correlations between age and phase by surface indicator.*

Indicator	*r* Value	*p* Value
All indicators	0.589	<0.001
Transverse organization	0.538	<0.001
Texture	0.543	<0.001
Left superior demiface	0.538	<0.001
Left inferior demiface	0.544	<0.001
Right superior demiface	0.565	<0.001
Right inferior demiface	0.536	<0.001

TABLE 4—*Results of ANCOVA with phase as the dependent variable (n = 266).*

Factor	Sums of Squares	F-Ratio	*p* Value
Sex	0.722	0.284	0.595
Ancestry	0.587	0.230	0.632
Collection	0.012	0.005	0.945
Sex * Ancestry	4.382	1.722	0.191
Sex * Collection	0.243	0.095	0.758
Ancestry * Collection	9.831	3.652	0.064
Sex * Ancestry * Collection	1.222	0.480	0.489
Age	356.722	140.183	<0.001*

The results indicate that while age is most highly correlated with an assessment of phase when using all indicators, all values are very similar and there is good concordance between different sections of the auricular surface. Given these results, the phase estimate data derived from examining all morphological indicators were used in all of the following analyses.

The results of the ANCOVA are presented in Table 4. Age was the only significant influence ($p < 0.05$) on auricular surface morphology. None of the main effects (sex, ancestry, or collection) are significant, and so it is not necessary to create different sex- or group-specific standards for age estimation from the auricular surface. The negative results for collection seem to indicate that secular trends have not affected the relationship between phase and age in the past century in U.S. populations. None of the four interactions are significant either, although the ancestry * collection interaction, while not quite reaching significance ($p = 0.064$), is suggestive. Detailed examination of the data indicates that one of the four subgroups (blacks in the Bass Collection) has a higher adjusted least-squares mean for phase and yet does not substantially differ from the other three groups in mean age. This subgroup is much smaller ($n = 13$) than any of the other subgroups and may represent a case of sampling error. In summary, because there are no clear or substantive effects of the independent variables on phase other than age, the entire sample has been combined for subsequent statistical analyses.

The r^2 for the ANCOVA model yields a value of 0.363, indicating that most of the variation in auricular surface morphology cannot be explained by age, sex, ancestry or collection. Some other unknown variable or variables are therefore implicated. The adjusted r^2 for age (0.343) indicates that 34% of the variation in auricular surface morphology in the population as a whole is attributable to age.

Inaccuracy and bias statistics were utilized to determine the accuracy of the standards created in the Lovejoy et al. (1) study. Inaccuracy refers to the average error in years regardless of over- or under-estimation of age, while bias refers to the average error in years taking into consideration the direction of the deviation (15). In calculating the inaccuracy and bias statistics using the five-year

TABLE 5—*Inaccuracy and bias for the combined sample using Lovejoy et al.'s aging standards.*

Phase	Age Range	*n*	Inaccuracy*	Inaccuracy Range*	Bias*	Bias Range*
1	20–24	5	5.3	2.5–8.5	4.3	−2.5–8.5
2	25–29	10	7.0	2.5–11.5	7.0	2.5–11.5
3	30–34	13	7.2	0.5–13.5	3.3	−12.5–13.5
4	35–39	37	11.2	0.5–38.5	−4.9	−38.5–13.5
5	40–44	52	11.6	0.5–36.5	−4.8	−36.5–19.5
6	45–49	30	11.5	1.5–32.5	−1.2	−32.5–18.5
7	50–60	17	10.1	0.0–32.0	1.9	−32.0–24.0
8	>60	102	13.0	0.0–39.0	6.1	−22.0–39.0
All phases		266	11.4	0.0–39.0	1.2	−38.5–39.0

* Inaccuracy and bias are reported in years.

TABLE 6—*Percentage of individuals correctly aged using Lovejoy et al.'s aging standards and the inclusion of adjacent phase intervals.*

Phases	*n*	Intervals	% Individuals Correctly Aged	Expanded Intervals	% Individuals Correctly Aged
1	5	20–24	80	20–29	100
2	10	25–29	20	20–34	100
3	13	30–34	15	25–39	54
4	37	35–39	16	30–44	41
5	52	40–44	15	35–49	42
6	30	45–49	10	40–60	43
7	17	50–60	53	45–60	71
8	102	>60	52	≥50	72
All Phases	266	N/A	33	N/A	59

TABLE 7—*Mean ages and 95% prediction intervals by phase for the combined sample.*

Phase	*n*	Mean*	S.D.*	95% P.I*	Accuracy	Inaccuracy*	Inaccuracy Range*	Bias*	Bias Range*
1	5	18.2	4.09	5.8–30.6	100%	2.7	0.2–6.8	0.00	−6.8–4.2
2	10	20.5	3.10	13.1–27.8	100%	2.5	0.5–4.5	0.00	−4.5–4.5
3	13	29.2	7.91	11.3–47.1	100%	6.1	0.2–15.8	−0.03	−15.8–10.2
4	37	42.4	13.67	14.4–70.4	97%	11.2	0.6–33.6	0.00	−33.6–18.4
5	52	47.3	14.20	18.6–76.0	96%	11.4	0.3–31.7	0.00	−31.7–24.3
6	30	48.7	13.70	20.1–77.3	97%	11.3	0.3–31.3	0.00	−31.3–19.7
7	17	53.1	14.14	22.3–83.9	94%	10.3	0.9–33.9	−0.02	−33.9–22.1
8	102	58.9	15.24	28.4–89.4	98%	12.9	0.1–32.9	0.00	−28.1–32.9

* Figures are in years; "S.D." refers to standard deviation; "P.I." stands for prediction interval; "accuracy" refers to the percentage of individuals from the original sample correctly aged using the 95% prediction interval; "Inaccuracy" and "Bias" are reported.

age ranges associated with the Lovejoy et al. (1) standards, the mid-range of each interval was used as the target age. The results (Table 5) indicate that the method becomes increasingly inaccurate with the progression to higher phases, which is to be expected with any aging system (6). Furthermore, as indicated by the bias values, individuals estimated in Phases 1–3 tend to be over-aged, those in Phases 4–6 are under-aged, and individuals scored in Phases 7 and 8 are over-aged (Table 5).

If the original age ranges provided in the Lovejoy et al. (1) study for each phase are used to estimate the age of the individuals used in this study, it becomes readily apparent that their five-year intervals do not reflect the true variation inherent in this aging system. Table 6 lists the percentages of individuals correctly aged using the original five-year ranges for each of the eight phases and for the total sample. The observed ranges for each phase in the study sample greatly exceed the five-year age ranges. Even if the age ranges provided in the original study are expanded to include the phase preceding and following the phase in which an individual is scored (e.g., providing 15+ year ranges), the results are still disappointing, with only 42% of individuals correctly aged in the middle three phases.

The mean ages and the 95% prediction intervals for each phase were calculated to examine age variation by phase (Table 7). While the results are not discrete (meaning that there are large age distributions and significant overlap of ages between phases), they are a good representation of the amount of variation in age present in each phase of development that would exist in the parent population as a whole. In some cases (e.g., between Phases 1 and 2; and Phases 5 and 6) there does not appear to be a significant difference between the mean ages of each phase. This may warrant the collapsing of phases with similar age distributions into one phase, similar to other refined aging systems (2,5).

Once the 95% prediction intervals were determined (11,23), the data were examined to establish how many of the individuals used in the study would be aged correctly using the new standards. Table 7 details the percentages of individuals correctly aged using the 95% prediction intervals derived from this study. The overall trend shows that age estimation becomes less accurate as individuals progress into more advanced phases, as is common in all studies of skeletal aging (2,5,6). The inaccuracy and bias results for the entire sample as a whole are 11.1 and 0.01 years, respectively. The near zero bias values are a byproduct of the statistical method employed (6). In general, inaccuracy values are only slightly lower per phase than those obtained using the Lovejoy et al. (1) modal ages. Excluding Phase 1, the mean ages for the sample used in this study actually fall outside of the modal five-year ranges provided by Lovejoy et al. (1) for Phases 2, 3, 4, and 5.

To determine whether or not the mean ages of Phases 1 and 2, and Phases 5 and 6 were significantly different, *t*-tests were performed. The results indicate that the mean ages for Phases 1 and 2 are not significantly different ($p > 0.05$), nor are the mean ages for Phases 5 and 6 ($p > 0.05$), thereby justifying the combination of these phases. After collapsing the phases, all individuals under the age of 18 were removed from the sample and the mean ages and 95% prediction intervals for each of the six new phases were recalculated. Table 8 presents the new mean ages, suggested age ranges, and the inaccuracy and bias values for each new phase. The suggested age ranges are modified prediction intervals for each phase. They are rounded to the nearest year, and when the lower age limit falls below 18 the range is listed as less than or equal to the upper limit of the age range (e.g., Phase 1 is ≤27 years). Table 8 lists the percentages of individuals from this study correctly aged using these standards. The morphological features defining each of these revised stages are presented in

TABLE 8—*Mean ages and suggested age ranges in years for the collapsed phases for the combined sample.*

Phase	*n*	Mean*	S.D.*	Suggested Age Range	Inaccuracy*	Inaccuracy Range*	Bias*	Bias Range*	Percent Correctly Aged
1	11	21.1	2.98	≤27	2.7	1.1–3.9	0.0	−3.9–3.1	100
2	13	29.5	8.20	≤46	6.2	0.5–15.5	0.3	−15.5–10.5	100
3	37	42.0	13.74	≤69	11.2	1.0–34.0	−0.4	−34.0–18.0	97
4	82	47.8	13.95	20–75	11.5	0.2–32.2	0.0	−32.2–24.8	98
5	17	53.1	14.14	24–82	10.3	0.9–33.9	0.0	−33.9–22.1	100
6	102	58.9	15.24	29–89	12.9	0.1–32.9	0.1	−28.1–32.9	94
All	262				11.3		−0.0		

* Figures are in years; "S.D." refers to standard deviation; "Inaccuracy" and "Bias" are reported; "All" refers to all phases combined in which "*n*" is totaled and "Inaccuracy" and "Bias" represent the data for the entire sample as a whole.

TABLE 9—*Refined auricular surface phase descriptions.*

Phase	Morphological Features
1	Billowing with possible striae; mostly fine granularity with some coarse granularity possible
2	Striae; coarse granularity with residual fine granularity; retroauricular activity may be present
3	Decreased striae with transverse organization; coarse granularity; retroauricular activity present beginnings of apical change
4	Remnants of transverse organization; coarse granularity becoming replaced by densification; retroauricular activity present; apical change; macroporosity is present
5	Surface becomes irregular; surface texture is largely dense; moderate retroauricular activity; moderate apical change; macroporosity
6	Irregular surface; densification accompanied by subchondral destruction; severe retroauricular activity; severe apical change; macroporosity

Table 9. Individuals under the age of 18 are excluded from the age ranges because it is assumed that the ages of those individuals will be estimated with greater accuracy using developmental indicators. The inaccuracy and bias values of the six-phase system are roughly equivalent to those derived from the collapsed eight-phase system.

In order for a direct comparison to be made between Lovejoy et al.'s (1) recommended mean ages by phase and those presented here, inaccuracy and bias statistics were computed by decade (Table 10). For all three methods, inaccuracy is lowest in the middle decades and highest in the oldest decade. Also, bias is positive (overestimated) for the lower decades and negative (underestimated) for the higher decades. Each of these results is expected given the statistical methodology used in the analysis (6). The Osborne six-phase method is no different in terms of accuracy than the eight-phase method, suggesting that the use of the eight-phase method is no better than the use of six phases (i.e., there is no statistical justification for splitting auricular surface degeneration into eight phases).

Discussion

The results of the ANCOVA test are perhaps the most interesting and important results derived from this study. ANCOVA indicates that for factors that may influence change in auricular surface morphology (age, sex, ancestry, and collection), age is the only factor that influences such change. Given these findings there is no need to create population-specific standards for blacks and whites or males and females for the use of the auricular surface as an indicator of age at death. Caution is warranted, however, when using the auricular surface as an age indicator on a population not represented in the Terry and Tennessee collections (e.g., non-U.S. populations), as well as populations derived from archaeological contexts (9,24,25).

Auricular surface morphology has potential as an indicator of age in skeletal remains of a forensic nature, particularly in the United States. The reasons for this are two-fold: (1) the auricular surface is highly durable with regards to taphonomic processes and, thus, extremely useful in the presence of fragmentary remains; and (2) since auricular surface morphology is not affected by sex or ancestry, an assessment of age at death can be made without knowledge of the individual in question's sex or ancestral classification, the latter being problematic on many levels (26).

While age does account for differences in auricular surface morphology, it is not the sole contributor to such differences. As indicated by the adjusted r^2 value (0.34), age only accounts for a small amount of the observed variation in auricular surface morphology. Since the categorical independent variables and their interactions do not contribute significantly to auricular surface morphology, there must be other factors for which effects cannot be controlled that influence change in auricular surface morphology.

TABLE 10—*Inaccuracy and bias statistics by decade: a comparison of Lovejoy et al. and Osborne methods.*

		Lovejoy et al.		Osborne 8*		Osborne 6	
Decade	*n*	Inaccuracy	Bias	Inaccuracy	Bias	Inaccuracy	Bias
20–29	31	12.7	12.6	13.5	12.2	13.3	12.4
30–39	47	13.9	13.6	14.6	14.0	14.5	13.9
40–49	49	9.3	4.4	7.0	5.0	7.1	5.0
50–59	47	9.8	−1.3	6.0	−2.4	6.1	−2.6
60–69	43	9.9	−8.5	10.5	−10.5	10.6	−10.6
70–79	33	13.1	−13.1	17.5	−17.5	17.5	−17.5
80–89	6	23.1	−23.1	27.3	−27.3	27.4	−27.4
Total	256	11.6	0.9	11.4	−0.1	11.4	−0.1

* Osborne 8 refers to the eight-phase method described in Osborne (8); Osborne 6 refers to the six-phase method described in Osborne (8).

Such factors may include individual differences in joint cartilage thickness, occupational stresses, life history variables, and the size and shape of the joint surface itself.

The results of testing the inaccuracy of the method show that the average estimation error increases with the progression into higher phases. This seems to mirror the overall trend in age estimation, as the accuracy of age estimation decreases as the age of the individual increases. This is likely a result of individual differences in anatomy and life history following completion of growth.

The percentage of individuals correctly aged using the original standards presented by Lovejoy et al. (1) demonstrates how inadequate the age ranges for this system are if misused as error ranges. Some authors argue that an aging system should *not* be tested on the sample from which it was developed, as a different sample may be more variable than the sample used to create the standards (27). Thus, such a different sample would truly test an aging method's applicability in a broader sense. Since neither the Terry nor Bass Donated Collections were used in the original study (1), they should provide an excellent test of this aging system. It is clear that the original data as presented (1) do not adequately reflect the true range of variation in auricular surface morphology per phase, and that the uncritical application of the five-year intervals is problematic.

If, however, Lovejoy et al.'s (1) age ranges are expanded to include the ranges adjacent to the estimated phase, there is some improvement in the percentage of individuals correctly aged. This may seem appealing to the forensic investigator whose primary goal in estimating age at death is to provide a broad enough age range so that a potential positive identification is not excluded, but a range that is also narrow enough to facilitate the identification process. Another benefit to this practice is the decreased chance of incorrect phase assignment. Typically, the morphological features of adjacent phases differ only slightly, so by including their age ranges into a more robust age range the investigator would reduce the chances of incorrect phase assignment due to slight, but incorrect, assessments of morphological features. Therefore, the age range of an individual estimated as being in Phase 3 would incorporate the age ranges for Phases 2, 3, and 4. In so doing, the original five-year age range is expanded to 15 years. While this results in a greater percentage of individuals correctly aged, particularly for Phases 1 and 2, the results still do not describe the full range of variation present in auricular surface morphology and the practice will clearly result in the incorrect age assessment of a significant number of forensic cases.

The 95% prediction intervals calculated for each phase indicate that change in auricular surface morphology with age is highly variable. The 95% prediction intervals project, for each phase, the ages that 95% of the population can be expected to fall between. If these prediction intervals were used as age ranges in an aging system, they would be as robust as other aging systems that are based on rates of degenerative change, e.g., the pubic symphysis (1,4).

Conclusion

Given the findings presented herein, we suggest a modified six-phase system for age estimation using auricular surface morphology. The revised method provides more robust phase categories and a more realistic view of the variation associated with auricular surface morphology and age. Although the method is somewhat inaccurate, this is more of a reflection on adult age estimation in general rather than a problem specific to the auricular surface. These results should be considered typical for adult age estimation via skeletal analysis.

We propose that ANCOVA is a useful tool in examining the effects of multiple controlled factors on the aging process. Furthermore, we suggest the continued use of inaccuracy and bias statistics to gage the precision of an aging system. Standard descriptive statistics and the use of 95% prediction intervals appear to provide the most accurate representation of age per phase of development.

While it is apparent that the original five-year intervals are insufficient for the needs of forensic anthropology, only recently has this issue been examined in a statistically meaningful way (7–9). Here we have argued that in creating standards for the estimation of age at death the full range of variation inherent in the aging system should be presented so that the reader is afforded an understanding of how imprecise such endeavors may be. Forensic anthropologists have the unique ability to utilize multiple regions in the skeleton to create a target age range for the subsequent identification of the decedent. This ability, however, may be undermined without knowledge of the limitations of the specific aging systems in question. Maples (10) described skeletal age estimation as art rather than science. While this is a good description, the process might best be described as more of a gestalt, with our intuitive hunches being moderated by an informed understanding of the underlying statistical realities and limitations of our methods.

Acknowledgments

This research partially fulfilled the requirements for the Masters degree in Anthropology at Western Michigan University, Kalamazoo, MI. Thanks to Robert Anemone, Della Cook, and the anonymous reviewers for comments on various drafts of this paper. The senior author would like to thank C. Owen Lovejoy and Heather York from Kent State University for taking time out of their busy schedules to instruct me in the application of the auricular surface method. Without their help this research would not have been possible. The senior author would also like to thank Murray Marks and Lee Meadows Jantz from the University of Tennessee, Knoxville, for assistance and access to the Bass Donated Collection and David Hunt of the Smithsonian Institute, Washington, DC, for help with the Terry Collection. Finally, the senior author thanks Megan Lagrou for her assistance during this research.

References

1. Lovejoy CO, Meindl R, Pryzbeck TR, Mensforth RP. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am J Phys Anthropol 1985;68:15–28. [PubMed]
2. Brooks S, Suchey JM. Skeletal age determination based on the os pubis: a comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. Human Evolution 1990;5:227–38.
3. Iscan MY, Loth S, Wright RK. Age estimation from the rib by phase analysis: white males. J Forensic Sci 1984;29:1094–1104. [PubMed]
4. Iscan MY, Loth S, Wright RK. Age estimation from the rib by phase analysis: white females. J Forensic Sci 1984;30:853–63.
5. Katz D, Suchey J. Age determination of the male os pubis. Am J Phys Anthropol 1986;69:427–35. [PubMed]
6. Nawrocki SP Regression formulae for estimating age at death from cranial suture closure. In: Reichs KJ, editor. Forensic osteology: advances in the identification of human remains. New York: Wiley-Liss, 1998; 276–92.
7. Buckberry J, Chamberlain AT. Age estimation from the auricular surface of the ilium: a revised method. Am J Phys Anthropol 2002;119:231–9. [PubMed]
8. Osborne D. Reconsidering the auricular surface as an indicator of age at death [masters thesis in anthropology]. Kalamazoo (MI): Western Michigan University, 2000.

9. Schmitt A, Murail P, Cunha E, Rougé D. Variability of the pattern of aging on the human skeleton: evidence from bone indicators and implications on age at death estimation. J Forensic Sci 2002;47: 1–7.
10. Maples WR. The practical application of age-estimation techniques. In: Iscan MY, editor. Age markers in the human skeleton. Springfield: Charles C. Thomas, 1989;319–24.
11. Nawrocki SP, Osborne DL. Serious problems in the estimation of age from human skeletal remains. Proceedings of the 7th Midwest Bioarcheology & Forensic Anthropology Association, Columbia, MO 2000.
12. Simmons T. Revising age estimation standards for a Bosnian forensic population: clavicle, rib and pubic symphysis. Proceedings of the American Association of Physical Anthropology. Columbus, OH 1999.
13. Murray KA, Murray T. A test of the auricular surface aging technique. [PubMed] J Forensic Sci 1991;36:1162–9.
14. Meindl RS, Lovejoy CO. Age changes in the pelvis: implications for paleodemography. In: Iscan MY, editor. Age markers in the human skeleton. Springfield: Charles C. Thomas, 1989; 137–67.
15. Lovejoy CO, Meindl R, Mensforth RP, Barton TJ. Multifactorial determination of skeletal age at death: a method and blind tests of its accuracy. [PubMed] Am J Phys Anthropol 1985;68:1–14.
16. Lovejoy CO, Meindl R, Tague RG, Latimer B. The senescent biology of the hominoid pelvis: its bearing on the pubic symphysis and auricular surface as age-at-death indicators in the human skeleton. Rivista di Antropologia (Roma) 1995;73:31–49.
17. Iscan MY. Age markers in the human skeleton. Springfield: Charles C. Thomas, 1989.
18. Thompson D. Forensic anthropology. In: Spencer F, editor. A history of American physical anthropology: 1930–1980. New York: Academic Press, 1982;357–69.
19. SPSS for Windows. Release 9.0.5. Chicago: SPSS Inc. 1998.
20. Wilkinson L. SYSTAT for the Macintosh. Version 5.2. Evanston: Systat Inc. 1992.
21. Daniel WW. Biostatistics: a foundation for analysis in the health sciences. New York: John Wiley & Sons, Inc., 1999.
22. Keppel G. Design and analysis: a researcher's handbook. New Jersey: Prentice Hall, 1991.
23. Hahn GJ, Meeker WQ. Statistical intervals: a guide for practitioners. New York: Wiley, 1991.
24. Lagrou M. A comparison of multiple age indicators in US and Bosnian skeletal samples [masters thesis in anthropology]. Kalamazoo (MI): Western Michigan University, 2000.
25. Lagrou MK, Simmons T. Age indicators are population specific-caution! Proceedings of the American Academy of Forensic Sciences. Seattle, WA 2001.
26. Marks J. Human biodiversity: genes, race, and history. New York: Aldine de Guyter, 1995.
27. Smith BH. Standards of human tooth formation and dental age assessment. In: Kelley M, Larsen C, editors. Advances in dental anthropology. New York: Wiley- Liss Inc., 1991;143–68.

Additional information and reprint requests:
Daniel L. Osborne, M.A.
Department of Anthropology
Indiana University
130 Student Building
Bloomington, IN 47405
E-mail: dosborne@indiana.edu