JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, March 2011, Vol. 56, No. 2
doi: 10.1111/j.1556-4029.2011.01700.x
Available online at: onlinelibrary.wiley.com

PAPER

ANTHROPOLOGY

Stephanie E. Calce,[1] *M.Sc. and Tracy L. Rogers,*[2] *Ph.D.*

Evaluation of Age Estimation Technique: Testing Traits of the Acetabulum To Estimate Age at Death in Adult Males*

ABSTRACT: This study evaluates the accuracy and precision of a skeletal age estimation method, using the acetabulum of 100 male ossa coxae from the Grant Collection (GRO) at the University of Toronto, Canada. Age at death was obtained using Bayesian inference and a computational application (IDADE2) that requires a reference population, close in geographic and temporal distribution to the target case, to calibrate age ranges from scores generated by the technique. The inaccuracy of this method is 8 years. The direction of bias indicates the acetabulum technique tends to underestimate age. The categories 46–65 and 76–90 years exhibit the smallest inaccuracy (0.2), suggesting that this method may be appropriate for individuals over 40 years. Eighty-three percent of age estimates were ±12 years of known age; 79% were ±10 years of known age; and 62% were ±5 years of known age. Identifying a suitable reference population is the most significant limitation of this technique for forensic applications.

KEYWORDS: forensic science, forensic anthropology, skeletal age estimation, acetabulum, adult male, os coxae

The ability to estimate age at death of human remains is an important tool in medico-legal investigations. Recently, it has been suggested that the acetabulum provides age-related changes that may aid in accurate age estimation of unknown individuals (1–4). Studies by Rissech et al. (1,4), describe seven age changes in the fused acetabulum and propose a method for estimating age with 89–100% accuracy based on these characteristics (Fig. 1). Rissech et al. (1) conclude that acetabular observations enable accurate age at death estimates of adults over 40 years of age.

The results of Rissech et al. (1) are significant for a number of reasons: (i) standard techniques for age estimation using the pubic symphysis and auricular surface exhibit inaccuracies for individuals over 40 years of age (5–11), (ii) most techniques lump individuals over 60 years of age into one age class (5–16), (iii) the acetabulum is more likely to survive postdepositional processes than fragile areas of the skeleton, such as the pubic symphysis (17), and (iv) few age estimation techniques report such high accuracy for individuals over 40 years of age. The purpose of this research is to (i) test the precision of Rissech et al.'s (1) method of scoring each trait, (ii) evaluate the accuracy of age at death estimates for individuals over 40 years, and (iii) compare the results achieved by using different reference populations to determine the impact of choosing an inappropriate reference sample.

[1]Department of Anthropology, University of Victoria, P.O. Box 3050, STN CSC, Victoria, British Columbia, Canada, V8W 3P5.
[2]Department of Anthropology, University of Toronto at Mississauga, 3359 Mississauga Rd. N., Mississauga, Ontario, Canada L5L 1C6.
*Presented in part at the Canadian Association for Physical Anthropology (CAPA) Annual Meeting, November 17, 2007, University of Calgary Alberta, Banff, Alberta, Canada.

Received 27 July 2009; and in revised form 19 Oct. 2009; accepted 10 Dec. 2009.

Methods and Materials

The Grant Collection (GRO), housed at the University of Toronto, is composed of mostly male subjects (n = 147) over the age of 40 (n = 133). Age at death is known for each individual (8). It is therefore an appropriate collection to test the method developed by Rissech et al. (1,4). A sample of 100 male individuals was randomly selected from the GRO and, as in the original publications, the left os coxae of each individual was examined. The nondestructive method of evaluation proposed by Rissech et al. (1,4), comprises close morphological examination of seven traits of the acetabular region: (i) acetabular groove, (ii) acetabular rim shape, (iii) acetabular rim porosity, (iv) apex activity, (v) activity of acetabular fossa, (vi) activity of outer edge of acetabular fossa, and (vii) porosities of the acetabular fossa (Figs 2–8). Each variable was scored based on a series of states within each observed morphological condition (1,4), for example acetabular groove can be scored as no groove [0]; groove [1]; pronounced groove [2]; or very pronounced groove [3]. See Table 1 for a description of each trait, variable states, and characteristics as published by Rissech et al. (1). Individuals with noninflammatory osteoarthritis or diffuse idiopathic skeletal hyperostosis were not excluded, as was the case in the original study. Linear regression was used to associate the degree of trait expression with age at death. Standard error of the estimate versus known age was conducted using a paired t-test.

To test for intra-observer error, every third individual (for a total of 34 individuals) was selected from the same sample and subjected to re-examination by the author (Calce). Intra-observer error was calculated from the difference in years between the first and second age estimates (n = 34). A comparison of percentage wrong estimates greater than ±5 years and p-value <0.05 from paired t-test was considered significant. Standard error of the age estimate, that

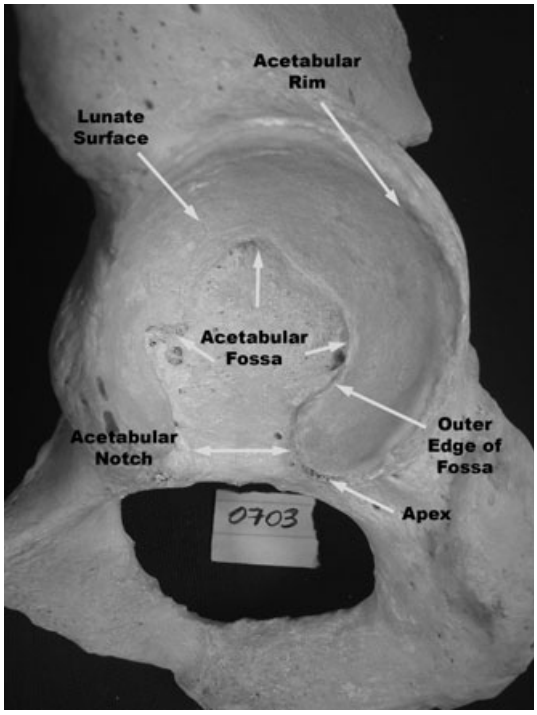

FIG. 1—*Features of the acetabulum affected by the degenerative aging process. Denoted by arrows, seven variables for examinations are based on the changes in the lunate surface, acetabular rim, acetabular fossa, acetabular notch, and apex.*

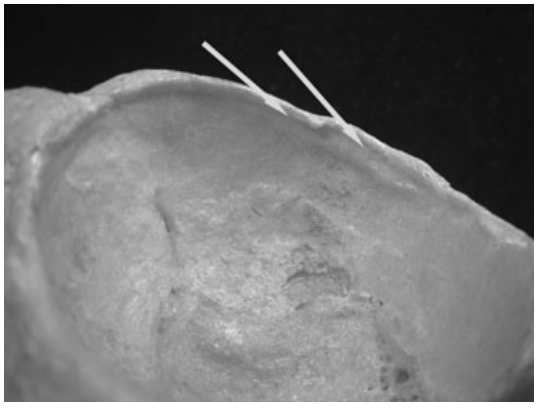

FIG. 2—*Acetabular groove. Found along the rim the groove becomes more pronounced and covers a larger surface area between the lunate surface and the acetabular rim with increased age. Each state can be coded 0–3 with no groove below the acetabular rim to extreme osteophytic development.*

is, $\sqrt{|\Sigma\ (y_1\text{-}y_2)/n|}$, was calculated to measure the accuracy of predictions.

Known ages for each individual were collected after each specimen was examined to eliminate observer bias. Specimen identification number, known ages for each specimen, and scored states for each of the seven variables were entered into an excel spreadsheet and saved as a comma-separated delineated version for easy file export to the statistical software, IDADE2, designed by the original authors (1,4, http://www-personal.umich.edu/~gfe/ [accessed July 17, 2009]).

As Bayes' theorem is used to compute posterior probabilities given an observation, IDADE2 calculates age at death by

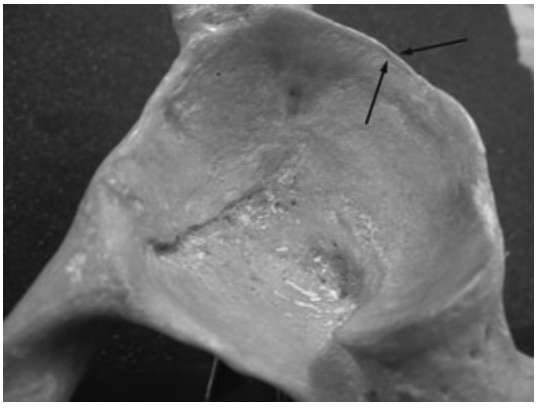

FIG. 3—*Acetabular rim shape. In younger individuals, the rim appears rounded, smooth, and dense. As age increases, osteophytic development causes narrowing of the rim, sharp crest formation, and spongy-like appearance of bone. In the area denoted by arrows, rim shape is coded 0 through 6 states.*

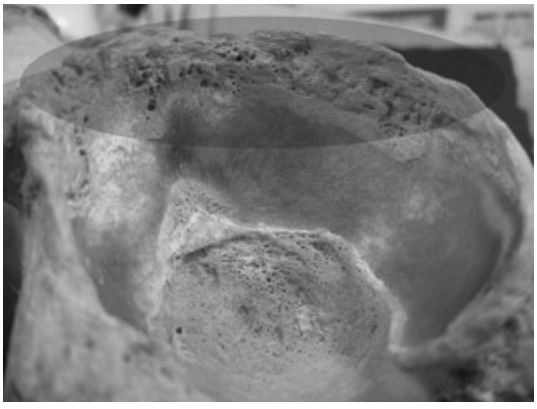

FIG. 4—*Acetabular rim porosity. Microporosity of bone indicative of younger individuals begins on the anterior inferior iliac spine and travels along the acetabular rim. Coded 0–5, macroporosity and large perforations (>1 mm) invade the superior area of the lunate surface as age increases.*

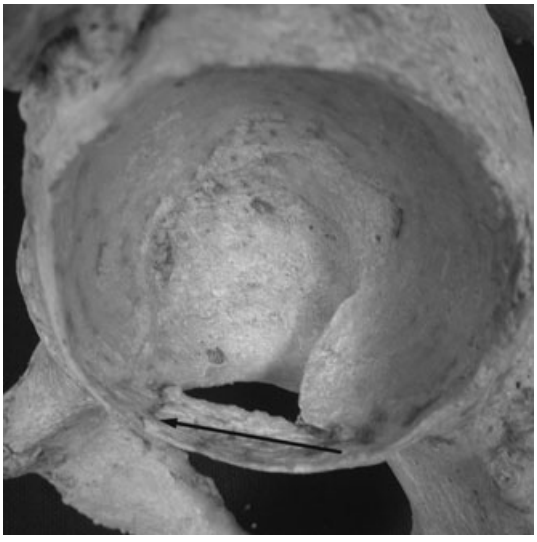

FIG. 5—*Apex activity. The posterior horn of the lunate surface primarily appears round and smooth. With age, a sharp spicule forms and osteophytic development may enter the acetabular notch to make contact with the anterior horn of the lunate surface. Each state is coded 0–4.*

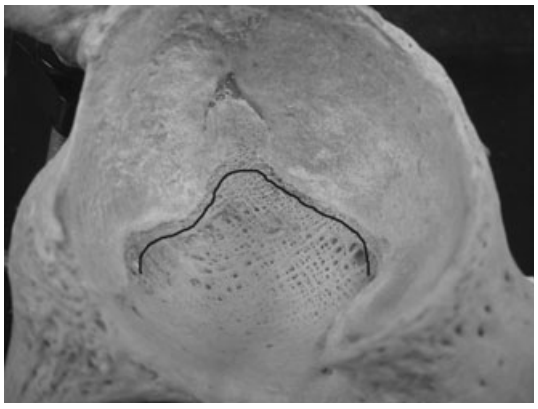

FIG. 6—*Activity of the outer edge of acetabular fossa. In older individuals, a visible crest forms on the outer edge of the fossa adjacent to the lunate surface and in some cases may cover the acetabular fossa. Coded 0–5, this ridge is often not visible in younger specimens but can be felt by moving the thumb from the superior region of the acetabular fossa to the inferior lunate surface.*

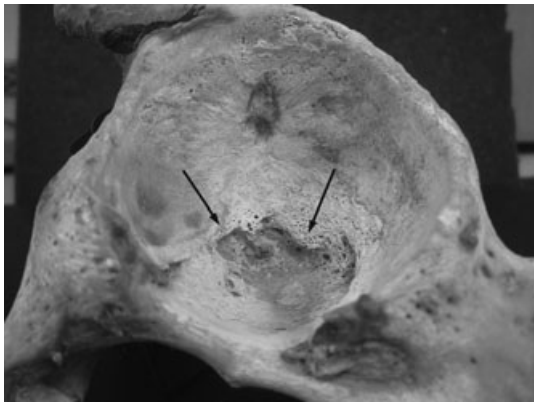

FIG. 7—*Activity of the acetabular fossa. The lunate surface and acetabular fossa are initially level, dense, and smooth. With age, the fossa descends into a more internal position appearing deeper than the lunate surface. Coded 0–5, growth of spongy bone from the inferior line of the lunate surface toward the fossa may appear porotic and eventually obliterate the fossa entirely.*

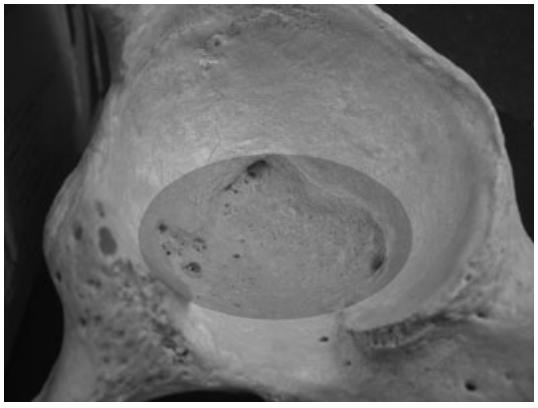

FIG. 8—*Porosities of the acetabular fossa. Through the aging process, fossa becomes porous and turns into trabecular bone. The evolution of micro- and macroporosity is coded 0–6 with fossa obliteration in the latest stage of development.*

multiplying prior probability distribution (scored states) by the likelihood function (center of the age class), divided by the normalizing constant (age class). IDADE2 uses the known ages at death of the reference population to estimate a likelihood distribution for age at death estimates for all other observed unknown specimens ([1,4,18,19]; http://www-personal.umich.edu/~gfe/ [accessed July 17, 2009]). Reference populations are used to generate age ranges for the scored states of each variable and for measures of fit. For example, using the reference population, the statistical analysis might reveal that all acetabular grooves that score as "2" correspond to an age range of "41–55." The same variables and states used to describe specimens in the reference collection must also be used in testing for an unknown individual.

The technique was tested on the GRO using three test reference populations to determine the importance of selecting an appropriate comparative population to generate age ranges. The test reference collections include the following: (i) the GRO (n = 100), a contemporary collection comprised of human skeletal remains from the early to mid-20th century, housed at the University of Toronto, Toronto, Ontario, Canada, (ii) *Esqueletos Identificados* of Coimbra (n = 242), comprised of Portuguese nationals interred between the late 19th and early 20th centuries, housed in the Anthropology Museum of the University of Coimbra in Portugal, Coimbra, Portugal, and (iii) an Iberian Collection (4) comprised of human skeletal remains from four separate collections in Western Europe (n = 394). The Iberian Collection includes *Esqueletos Identificados* of Coimbra, Anthropology Museum of the University of Coimbra in Portugal, Coimbra, Portugal (n = 242); the Lisbon Collection (19th and 20th centuries) housed at the Bocage Museum, Lisbon, Portugal (n = 57); the UAB Collection (20th century) housed at the University Autononoma Barcelona, Barcelona, Spain (n = 18); and the St. Bride Collection (18th and 19th centuries) housed in St. Bride's Church, London, U.K. (n = 77). The author (Calce) collected data for the GRO independently; the authors of the original articles provided state frequency data for the Comibra and Iberian Collections, which were then utilized by the author (Calce) in this research.

IDADE2 was used to find (i) differences between the estimated distribution and the known age at death, (ii) youngest and oldest age of age classes in central 95% of the estimating distribution, and (iii) estimated age at death for each individual. IDADE2 also allows the user to choose preferred age classes based on the youngest age and yearly intervals, which determine how state frequencies are presented in the results file. In this experiment, data were presented based on 5-year age classes. State frequencies (numbers of individuals in each age class for each state of the seven traits) were calculated to identify anomalies or errors in scoring. Bias, that is, Σ (EAD-KA/n), and inaccuracy, that is, Σ |EAD-KA|/n, were calculated to identify whether the application of this method resulted in over- or underestimation of age and find the average absolute error of age estimation (10,15).

Results

Results are expressed as the difference between estimated and known age at death for each specimen in three analyses. Figures 9–11 demonstrate the differences between the known and estimated ages for each individual, using each of the three reference collections.

The GRO as a Reference Population

The age range of specimens used in the GRO was 17–89 years (Table 2). Across all age classes, 83% of age estimates were within

TABLE 1—*Morphological description of the seven acetabular variables and their states.*

Variable	Description of the Variable	States of the Variable	Characteristics of the States	Code
(1) Acetabular groove	This groove appears below and surrounds the internal margin of the acetabular rim. With age, the acetabular groove can become more or less pronounced either along the entire acetabular rim or along only a part of it.	No groove (fig. 1)	There is no groove below the acetabular rim. There is no anatomic interruption between the lunate surface and the acetabular rim.	0
		Groove (fig. 2)	An anatomic interruption is observed between the lunate surface and the acetabular rim. Although it might be short or shallow, it surrounds some or much of the acetabular rim.	1
		Pronounced groove (fig. 3)	A deeper groove surrounds a large part of the acetabular rim.	2
		Very pronounced groove (fig. 4)	An extremely pronounced groove surrounds nearly all the acetabular rim. In some specimens, extreme growth of the rim has obscured the groove so that only a tissue discontinuity between the lunate surface and the acetabular rim can be observed.	3
(2) Acetabular rim shape	With age, the acetabular rim loses its round and smooth form as a consequence of the progressive development of osteophytes, which can become a crest.	Rounded acetabular rim (fig. 5)	The acetabular rim is dense, round, and smooth, typical of young specimens.	0
		Partially narrow acetabular rim (fig. 6)	The acetabular rim keeps its round and smooth form in some areas but in others is narrower. There are two possibilities: (i) the iliac part of the acetabular rim narrows but not the ischial part or (ii) the external part of the acetabular rim retains its rounded form but its internal part has an upright form. In all of these cases, the acetabular rim is smooth to the touch.	1
		Narrow or rough acetabular rim (fig. 7)	There are two possibilities: (i) whole acetabular rim is narrow or (ii) some part of the acetabular rim might be rough to the touch because of the presence of little grooves. In both possibilities, there is no osteophytic construction.	2
		Partially crested rim (fig. 8)	Osteophytic constructions form a small chain (*c.* 1 mm in height) on some small part of rim; a bigger osteophyte linked or not to the chain might be observed.	3
		Crested rim (fig. 9)	An osteophytic formation makes either (i) a low crest (*c.* 1 mm in height) along the entire acetabular rim or (ii) a high crest (2–4 mm in height) along only part of it. This crest appears dense.	4
		Very high crested rim (fig. 10)	A very high crest (44 mm in height) has developed as a consequence of bone construction and destruction. This crest is thin and sharp or rounded with a spongy appearance.	5
		Destructured rim (fig. 11)	An extremely high crest (48 mm in height) has developed. It may be either thin and sharp and leaning slightly toward the lunate surface, or rounded, spongy and fragile with swollen and hollow bone.	6
(3) Acetabular rim porosity	With aging, porosity appears on the acetabular rim and on the adjacent ilio-ischiatic area of the acetabulum. Two kinds of porosity can be defined: (i) microporosity, which refers to a fine, just optically visible perforation (< 1 mm) and (ii) macroporosity, which refers to an oval or round perforation larger than 1 mm.	Normal porosity (fig. 12)	Acetabular rim is smooth without porosities and roughness. The area adjacent to the acetabular rim also has normal porosity.	0
		External porosity (fig. 13)	On the area around the acetabulum, microporosity is lightly increased on the anterior inferior iliac spine, on the posterior wall of the acetabulum, and on the area below the two extremities of the lunate surface. There is no porosity on the acetabular rim, which is dense and smooth.	1
		Rim porosities (fig. 14)	Some microporosities on the acetabular rim may be large (=1 mm) but the acetabular rim always has a round and dense appearance. There is no bone destruction.	2
		Rough rim (fig. 15)	The acetabular rim is not smooth to the touch, and there may be some macroporosity on the rim.	3
		Destructured rim (fig. 16)	Newly constructed bone has become very porous with many micro- and macroporosities, or it has suffered subsequent destruction.	4
		Extremely destructured rim (fig. 17)	Macro- and microporosities of the destructured acetabular rim have partially invaded the lunate surface. Usually, this invasion occurs on the superior area of the lunate surface below the anterior inferior iliac spine.	5
(4) Apex activity	Apex activity refers to the bone activity observed on the apex of the posterior horn of the lunate surface. With aging, this apex loses its rounded form, gradually becoming sharper and finally developing a spicule, which can become quite large.	No activity (fig. 18)	The apex is round and smooth to the touch. There is no spicule.	0
		Apex activity (fig. 19)	The apex has become longer and is sharp to the touch, or a small spicule can be felt.	1
		Osteophytic activity (>1 mm; fig. 20)	A developed and conspicuous osteophyte larger than 1 mm can be seen with the naked eye.	2
		Much osteophytic activity (>3 mm; fig. 21)	The apex has an osteophyte larger than 3 mm, which may cover the entire horn of the lunate surface.	3
		Very much osteophytic activity (>5 mm; fig. 22)	An osteophyte is so large (> 45 mm) that it enters the acetabular notch and may completely cross it, in which case the anterior horn of the lunate surface also has activity.	4

Continued.

TABLE 1—*Continued.*

Variable	Description of the Variable	States of the Variable	Characteristics of the States	Code
(5) Activity on the outer edge of the acetabular fossa	This activity refers to an osteophytic formation that grows as a minicrest from the outer edge of the acetabular fossa toward the lunate surface. Usually, it can be felt but not seen. When it is present, the edge is rough to the touch and can be detected by repeatedly moving the finger along the outer edge of the acetabular fossa toward the acetabular fossa surface (fig. 23). Sometimes, this osteophytic formation becomes visible and extensive enough to cover the acetabular fossa.	No activity on the outer edge	The outer edge feels smooth, or at least not rough, and the finger moves smoothly over it toward the fossa.	0
		Slight activity (<1/4) on the outer edge	A minicrest can be felt (but not seen) on less than one-quarter of the outer edge of the acetabular fossa. It is usually found on only one of the two horns of the lunate surface, near the apex.	1
		Medium activity (<1/2) on the outer edge	Bone growth can be felt (but not seen) on between a quarter and half of the outer edge of the acetabular fossa. Usually, this bone growth is not continuous; therefore, all the active parts must be considered to estimate the proportion.	2
		Much activity (<3/4) on the outer edge	Bone growth can be felt on between one-half and three-quarters of the outer edge of the acetabular fossa.	3
		Extreme activity (>3/4) on the outer edge	Bone growth can be felt, and sometimes, it can be seen on more than three-quarters of the outer edge.	4
		Destructured outer edge (fig. 24)	There is so much visible bone growth on the outer edge toward the fossa that it partially covers the fossa parallel to the outer edge.	5
(6) Activity of the acetabular fossa	The young acetabular fossa appears dense and smooth and is almost level with the lunate surface. With aging, the acetabular fossa moves to a more internal position and clearly appears deeper than the lunate surface. Also the activity, expressed as relief, porosities, and bone production, is present on the fossa. When this activity is extreme, the acetabular fossa may be obliterated.	No activity (fig. 25)	The lunate surface is level with the acetabular fossa, which appears dense and smooth.	0
		Slight activity (fig. 26)	The lunate surface is clearly no longer level with the acetabular fossa, which still appears dense and smooth.	1
		Peripheral activity (fig. 27)	The acetabular fossa shows activity between one-quarter and half of its surface. This activity is usually located on the posterior area of the fossa or sometimes on peripheral areas, but never on the center. This activity results in relief, porosities, and spongy bone, which grow toward the lunate surface from small parts of the external border of the fossa. Areas of the acetabular fossa without activity appear dense and smooth.	2
		Central activity (fig. 28)	There is activity on about half of the fossa. It is usually found on the posterior half and always extends to the center. Activity on the center of the acetabular fossa usually produces a relief that is similar to the trabeculae. Peripheral activity is usually expressed by porosities. There may be some growth of spongy bone toward the lunate surface.	3
		Major activity (fig. 29)	Activity is observed on more than three-quarters of the fossa. This activity produces relief and porosities, but the fossa does not lose its consistency and density.	4
		Generalized activity	The entire fossa, or nearly all of it, is covered by extensive formation. There are two possibilities: (i) the fossa is not consistent nor dense (fig. 31a) and (ii) the fossa is partially or totally obliterated (fig. 30b).	5
(7) Porosities of the acetabular fossa	Through the aging process, microporosities first become macroporosities, then trabecular bone, and finally destruction invades the entire fossa. There are two types of macroporosities. (1) Smaller (≤1.5 mm) macroporosities occur as a transition of microporosities into trabecular bone; these have a blunt perimeter and will be called smaller macroporosities. (2) Larger (>1.5 mm) macroporosities have a sharp perimeter because of destruction; these are conspicuous, larger and either round or less regular and will be called macroporosities with destruction.	Dense acetabular fossa (fig. 31)	The acetabular fossa is dense and smooth, but it may have a few normal peripheral microporosities.	0
		Acetabular fossa with microporosities (fig. 32)	The acetabular fossa appears dense but there are small areas with some microporosities. These areas look like ''orange skin,'' usually on the superior lobe of the fossa, but sometimes elsewhere.	1
		Macroporosities or peripheral trabecular bone	Part of the fossa is covered with microporosities and smaller macroporosities. These porosities occur on about one half of the fossa, which can include the center, but not on all three lobes (fig. 33a). Some trabecular bone may occur on the peripheral area of the fossa (fig. 33b).	2
		Macroporosities on the three lobes (fig. 34)	Porosities occur on about three quarters of the fossa. The three lobes and the center of the fossa are covered with smaller macroporosities and microporosities, but not the area of the acetabular notch. Trabecular bone may occur on the peripheral area of the fossa.	3
		Macroporosities with destruction (fig. 35)	Macroporosities with destruction occur on a base of microporosities and smaller macroporosities. This may be observed over most of the fossa or only over a restricted area.	4
		Bone destruction on most of the fossa (fig. 36)	Most of the fossa is covered with trabecular bone. There are no microporosities. There is much destruction evidenced by large irregular macroporosities with destruction. The bone of the fossa is swollen and has lost consistency as a result of bone destruction.	5
		Bone proliferation (fig. 37)	Bone proliferation on the acetabular fossa obliterates the fossa.	6

Source: Rissech et al.[1]

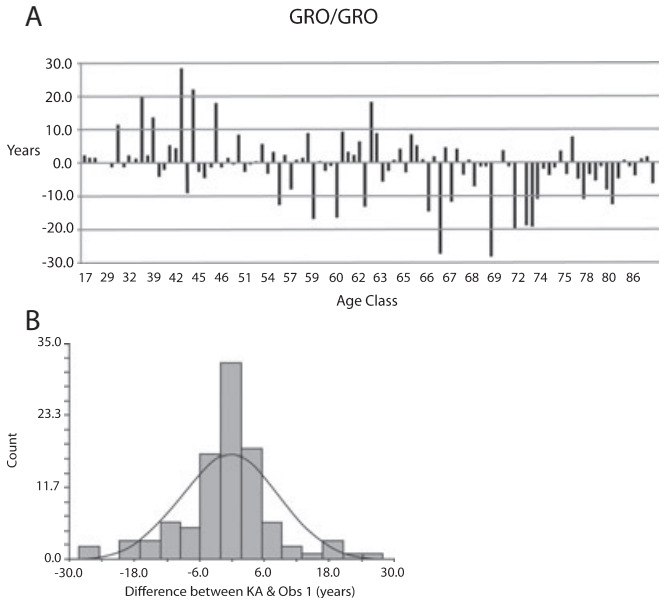

FIG. 9—*(A) Difference (in years) between the known ages and estimated age at death for each specimen in the GRO Collection, when the same collection is used as reference. (B) Age at death was estimated correctly for a large proportion of individuals; 83% of age estimates were ±12 years from the known age.*

TABLE 2—*Sample age distribution for the GRO Collection of individuals.*

Age Class	*n*
16–20	2
21–25	1
26–30	4
31–35	3
36–40	4
41–45	10
46–50	3
51–55	8
56–60	11
61–65	11
66–70	16
71–75	12
76–80	9
81–85	2
86–90	4
Total	100

GRO, Grant Collection.

±12 years of known age; 79% within ±10 years of known age; and 62% within ±5 years of known age. Results are summarized and presented in Fig. 9 and Table 3. Across all age classes, the inaccuracy in using this method is an average of 8-year difference between the estimated and known age. The bias (mean error of age estimates incorporating direction of error [±]) indicates that overall (across all age classes 16–90) this technique tends to underestimate the age of the individual. As expected, for age classes 16–55, the bias shows that this technique *over*estimates age at death; while for age classes 56–90, the technique *under*estimates age at death (Table 4). This trend is prevalent in most age-estimation techniques, so the result is not considered unusual (5–16). Considering specific age categories, 41–45 and 66–75 showed the largest inaccuracies (0.6 and −0.7, respectively); notably, the age categories 46–65 and 76–90 showed the smallest inaccuracy (−0.2). These results indicate that individuals between 41–45 and 66–75 years of age are most difficult to estimate

TABLE 3—*Number of specimens with fit* (f), *within specified amounts, tabulated within age classes, and over all specimens. Fit is the expected value of absolute difference between the known age and the nearest age in any age class, except the one in which the known age falls. Of GRO individuals, 62% had an expected difference between known age at death and the estimating distribution of <5 years; 37% received values of fit over 10 years.*

Age	*f* 1	*f* 2	*f* 3	*f* 4	*f* 5	*f* 10	*f* > 11	Total
16–20	2	0	0	0	0	0	0	2
21–25	1	0	0	0	0	0	0	1
26–30	3	0	0	0	0	0	1	4
31–35	3	0	0	0	0	0	0	3
36–40	2	1	0	0	0	0	1	4
41–45	2	1	2	0	1	0	4	10
46–50	0	2	1	0	0	0	0	3
51–55	0	0	2	1	0	4	1	8
56–60	4	1	0	0	0	4	2	11
61–65	1	0	1	1	1	3	4	11
66–70	1	1	1	5	3	2	3	16
71–75	1	3	1	2	1	1	3	12
76–80	0	1	2	2	1	1	2	9
81–85	1	1	0	0	0	0	0	2
86–90	3	0	0	0	0	1	0	4
Total	24	11	10	11	7	16	21	100

GRO, Grant Collection.

TABLE 4—*Results of bias and inaccuracy tests, that is, Σ (EAD-KA/*n), *based on scores obtained using the GRO.*

Age Class	*n*	Bias/Inaccuracy
16–20	2	0.0
21–25	1	0.0
26–30	4	0.1
31–35	3	0.0
36–40	4	0.3
41–45	10	0.6
46–50	3	0.0
51–55	8	0.0
56–60	11	−0.2
61–65	11	−0.2
66–70	16	−0.7
71–75	12	−0.7
76–80	9	−0.4
81–85	2	0.0
86–90	4	−0.062

GRO, Grant Collection.

accurately for chronological age. The authors consider three explanations for this result. First, age-related changes occurring in individuals 41–45 and 66–75 mark transitional periods between young-to-middle and middle-to-older age adults, respectively, highlighting variability and imprecise differentiation between neighboring age classes at these intermediary phases. Second, large inaccuracies are merely a statistical artifact of the sample, because the largest proportion of GRO individuals (38%) belongs to these age classes. Third, age-related morphological changes are highly variable for individuals 41–45 and 66–75 years of age, where further research into specific degenerative modifications targeting middle-aged persons is required. Results of inaccuracy testing imply that very little error exists when estimating age for the elderly, >76 years.

A further comparison of the estimates generated by the GRO to the actual ages of the GRO reveals that individuals between 40 and 90 years of age were assigned accurate age estimates in 79% of all cases, suggesting that this method would be appropriate to estimate age for individuals over 40 years. Results of the bias/inaccuracy test across all age classes are presented in Table 4.

The standard error of the measurement between estimated and known age is 0.83 years (95% confidence intervals [CI] 2.61–0.96; Fig. 12). Using linear regression, a positive association between trait expression and known chronological age is occurring within the data (r^2 = 0.7176, correlation = 0.8471; Fig. 13). States are correlated with age in the way that as the scored value increases so does estimated age at death. Significant differences of intra-observer error testing were found. Results of intra-observer tests are discussed independently herein.

The Iberian and Coimbra Collections as Reference Populations

The age range of specimens used in the Iberian and Coimbra Collections was 15–96 years. The Iberian and Coimbra Collections, although tested separately, produced identical results. When using the Iberian and Coimbra Collections as reference populations, ID-ADE2 was unable to estimate age at death for 39 individuals in the GRO. The combination of scores was so rare in the context of these collections that an age estimate was not obtainable for over one-third of the sampled population. Of the 61 estimated ages obtained, 41% of age estimates were within ±12 years of known age; 36.1% were within ±10 years of known age; and 19.7% were within ±5 years of known age (Figs 10 and 11). Results show that the most distant reference collection, in terms of space, time, and ancestry, produced the least accurate estimates.

Scoring Effectiveness and Intra-observer Error

All seven traits of the acetabulum for each individual in the GRO were in good condition for evaluation. Each trait was scored independently. Familiarity with the traits of the acetabulum required some study; traits were scored using only the information given in the descriptions of variables (Table 1) and accompanied photos (1). Thirty-four GRO individuals were randomly selected and re-scored to calculate intra-observer error. The results show that differences in age at death estimates between the two tests fell within ±5 years from the first age estimate in 82% of cases. Figure 14 illustrates results of the intra-observer test showing the difference (in years) between the first and second tests of age at death estimates for these individuals. Results of the intra-observer test also show that 100% of age at death estimates are within ±8 years from the first age at death estimate. Age classes 56–60 and 71–75 showed the largest gap between two estimates, 7.2 and 7.3 years, respectively, suggesting that age-related expressions of traits are more highly variable with increased age. A paired *t*-test for observations 1 and 2 revealed that the average difference in age estimates is 1.26 years

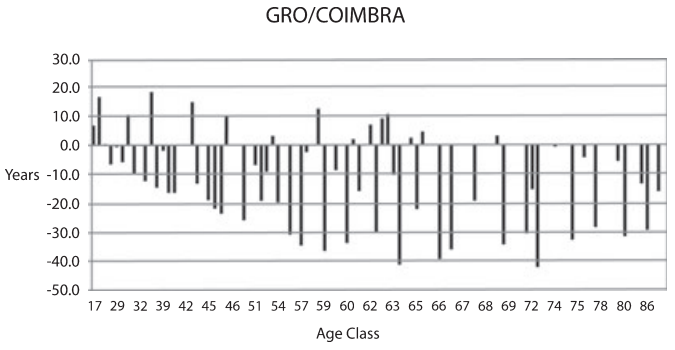

FIG. 10—*Difference (in years) between the known ages and estimated age at death for each specimen in the GRO Collection, when the Coimbra Collection is used as a reference.*

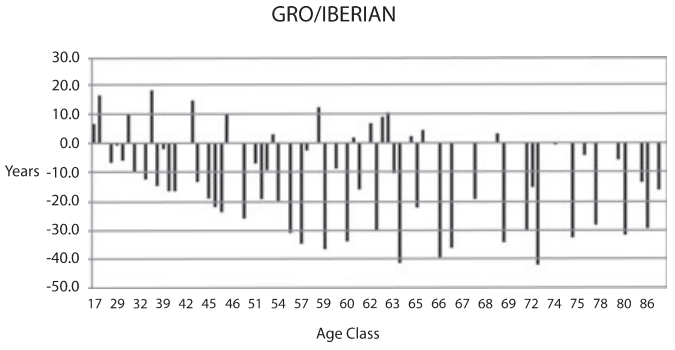

FIG. 11—*Difference (in years) between the known ages and estimated age at death for each specimen in the GRO Collection, when the Iberian Collection is used as a reference.*

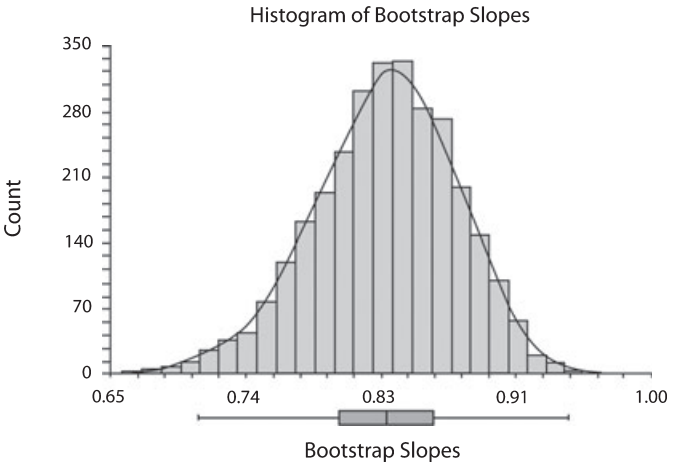

FIG. 12—*Bootstrap confidence intervals are reported for 3000 iterations where gross outliers have been removed and sample is representative of the population. Data show that the average difference in age estimates from chronological age is <1 year (0.83).*

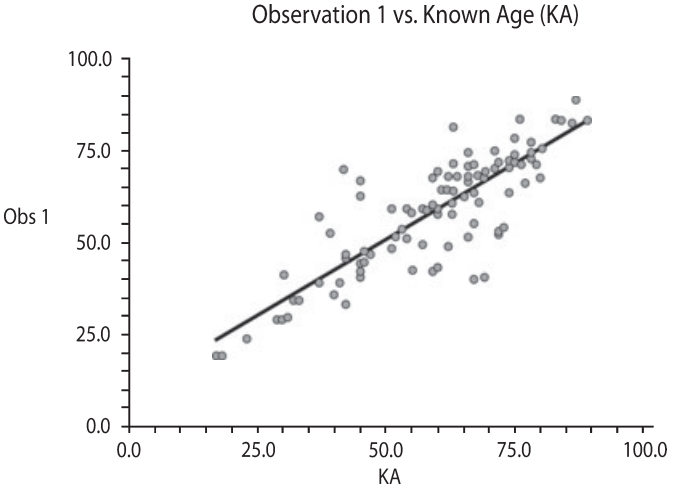

FIG. 13—*The equation of the straight line relating observation 1 (obs_1) and known age (KA) is estimated as: obs_1 = (9.38) + (0.83) KA for* n = *100. The y-intercept, the estimated value of obs_1 when KA is zero, is 9.38 with a standard error of 3.23. The slope, the estimated change in obs_1 per unit change in KA, is 0.83 (95% CI 0.73–0.93) with a standard error of 0.053;* r^2 *= 0.7176, correlation = 0.8471.*

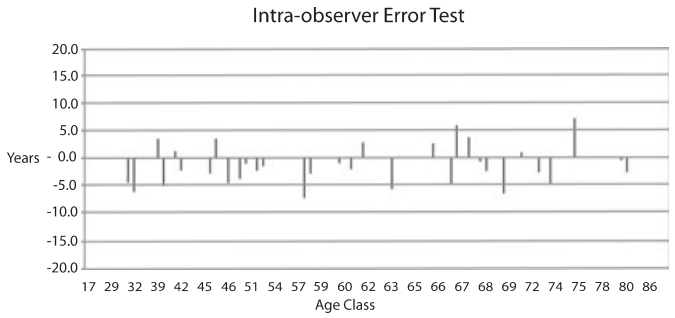

FIG. 14—*Difference (in years) between first and second tests to estimate age at death for 34 individuals in the GRO Collection. The second test (illustrated by graph) shows that 82% of age at death estimates* (n = *34) is within ±5 years from the first age at death estimate. The standard error of the second estimate is −1.1 years.*

(95% CI 2.52–1.11 years). Results of the paired *t*-test suggest a marginally significant difference between age estimates of observations 1 and 2 ($t = 2.036$, $p = 0.049$).

States for all seven variables, across all specimens ($n = 34$), were scored differently between the first and second tests, contributing to dissimilar results. By comparison of scored states for individuals, 63.8% of all scores for every trait were identical between first and second tests. States of acetabular groove and activity of acetabular fossa were scored alike in 74.1% and 71.1% of cases, respectively. Across all states of all variables, 22.3% were scored within one state of the previous test and 13.9% were scored within two states. The maximum difference in scoring was 2 with the average difference in scoring 0.5. Table 5 illustrates differences in scoring between first and second tests for each trait. High dissimilarity in scoring between observations indicates that the trait is difficult to distinguish, that is, score precisely. Trait two (acetabular rim shape), trait seven (fossa porosity), and age range 61–70 were affected by the most variable results in scoring. This may have resulted from vague and contradictory state definitions provided by Rissech et al.[1] (Table 1). For example, anatomic surfaces of the ischium, iliac, and pubic bone are referred to as "internal" or "external" parts. Precise terminology describing locations of lesions or specific regional changes should be standardized among investigators (Table 1). In addition, state 5 of acetabular rim shape indicates that the visible crest can *either* be "thin and sharp" or be "rounded with a spongy appearance" (Table 1). If states are correlated with age in the way that as the scored value increases so does estimated age at death, the explanation of state observations contradicts with the variable's description, which specifies that "with age, the acetabular rim loses its round and smooth form." The authors suggest redefining terms, such as "swollen," "consistency," "relief," "bone proliferation," and "destruction," to help the investigator understand verbal descriptions as they pertain to macroscopic morphological changes.

Discussion and Conclusions

As the acetabular region is well preserved in the os coxae, it is a valuable skeletal element for research, with great potential for use in skeletal age estimation. In the original articles (1,4), Rissech et al. establishes that the acetabulum method of age estimation cannot be applied if one or more of the seven traits are missing or damaged, as in the case of taphonomic factors. Using the GRO, each trait of the os coxae for all randomly chosen individuals was well preserved and robust for scoring evaluation. No individuals were eliminated because of damaged areas of the acetabulum, nor is evidence of exposure to taphonomic variables present. Surfaces of the bones are smooth with no degradation of cortical bone, which indicates that soft tissue was removed by dissection (following autopsy), as opposed to natural decomposition, that is, results of a burial environment.

Intuitively, in cases of scattered or scavenged remains, disarticulation of the femur from the acetabular joint can damage the rim of the acetabulum; however, Haglund's (1991) study of 53 canid-scavenged human remains showed that ligamentous attachments of joint surfaces prevented scavenger-assisted disarticulation (20). In fact, Haglund (1997) reports that the os coxae was recovered in 60–79% of all cases of scavenged remains for the observed postmortem interval of 4 h–52 months (20). The high frequency of recovery for the os coxae bone is significant because lengthy postmortem intervals negatively impact the process of personal identification using the skull, that is, comparison of antemortem dental records that no longer exist, or the inability to extract DNA which declines over time (21). Where surface finds make up more than 45% of depositional environments (22), methods to estimate age using postcranial skeletal elements are necessary to aid in personal identification, especially in cases where the skull is not recovered with the rest of the body. In addition to this, preservation of the acetabulum in postdepositional scenarios renders the technique more useful than pubic symphyseal aging as this area of the skeleton is often damaged and not available for examination (17). The high preservation of the os coxae in forensic cases makes a technique like Rissech et al.'s (1) acetabular age estimation method a potentially valuable contribution to forensic anthropology. Unfortunately, there are problems estimating the degree of development of features.

Consistency in scoring of each trait is imperative to yield reliable results. Descriptions of variables and accompanied photos provided by the original authors (1,4) are well detailed but results of scoring effectiveness and intra-observer error indicate that revisions or clarifications may be necessary. States for variables 2, 3, and 5–7 were not easily identified through intra-observer inspection, indicating that these descriptions should be better defined, combined, or eliminated for clarity between observers. Specifically, the authors suggest a revision of the description of variables and associated states for variables 6 and 7. Porotic expression and bone production at the site of the acetabular fossa should be explicit and separated from the position of the fossa in relation to the lunate surface, that is, level versus a more

TABLE 5—*Differences between scores of observations 1 and 2 for each variable 1–7 using the GRO.*

Variable	Variable Name	Same Estimate (%)	Within 1 State (%)	Within 2 States (%)	Average Difference in Scoring	Maximum Difference in Scoring
1	Acetabular groove	74.1	23.5	2.4	0.29	2
2	Acetabular rim shape	50	26.5	23.5	0.74	2
3	Acetabular rim porosity	62.1	23.5	14.3	0.53	2
4	Apex activity	62.1	29.4	8.5	0.47	2
5	Outer edge of fossa	71.1	14.7	14.2	0.44	2
6	Acetabular fossa	71.1	14.7	14.2	0.44	2
7	Fossa porosity	58.8	23.5	17.7	0.59	2

GRO, Grant Collection.

internal position. In Rissech et al.'s original work (1), nonsignificant differences in inter- and intra-observer tests are reported but results of the Friedman test are not made explicit in the article. From results of the current study on the GRO, it is clear that trait scoring of the acetabulum is not consistent between observations leading to subjective coding of variable states. Over one-third of secondary assessments were scored differently, highlighting the discrepancy between observations that resulted in significantly different age estimates. Acetabular characteristics frequently appeared to overlap designated states and the observer was obliged to choose one over another, with no consistent means of doing so. Given the variation in skeletal aging, it is extremely difficult to comprise the full range of morphological changes in every age-related expression. For this reason, precise definitions of fewer, more encompassing variable states reduce subjectivity in scoring and increase the utility of the method, particularly for older-aged individuals where joint surfaces are complicated by extrinsic factors, such as weight, infection, and physical activity. Overall, investigator confidence was low estimating trait prevalence for states 2–5 of variables 2 and 5–7. This is reflected by the high dissimilarity in scoring (Table 5) and reinforces the need for familiarity with descriptions and clear explanations of traits.

Existing methods of age estimation that lump individuals over the age of 60 into one age class pose a problem in the positive identification of remains for the aging populace (5–10,23). Because the anatomic features of the acetabulum described by Rissech et al., develop through a longer maturation and aging time course than other features used in age estimation, the technique can be used to distinguish stages within the senior age category (1,4). For men in the GRO, we found similar levels of accuracy for age classes 46–65 and 76–90, suggesting that this method would be appropriate in estimating age at death for much older individuals. As estimating the age of those over 60 years is notoriously difficult for many morphological age estimation techniques, the acetabulum method has a valuable role to play in the identification of the elderly.

Adult age markers confirm that there is a large amount of variability in the aging process both between geographic populations and between individuals themselves, dependent, of course, on a number of factors, such as diet, nutrition, and physical activity (6,12–14). From the results presented here, it is clear that the Iberian and Coimbra Collections are not appropriate as reference populations to test age estimation for individuals in the GRO. Outcomes of this study show that to yield reliable results, the chosen reference population *must* be close in geographic and temporal context to the test specimen to account for variability in the aging process of the skeleton; less accurate age estimates will occur when a more distant reference collection is used. In a bioarchaeological context, biographic data, such as ancestry, can be collected from a number of sources, such as cemetery registers and civil registration records. In a forensic context, where positive identification or ancestry of an individual may not be known to investigators, it may be difficult to select an appropriate reference collection. Even if an appropriate population is known (e.g., based on ancestry assessments), such a reference collection may not exist, as most collections in North America consist largely of European-derived (White) or African-derived (Black) peoples.

The accuracy for the acetabulum method increases as the number of specimens increases. In a forensic context, this may also be problematic because the minimum number of individuals for examination may be one, in which case, the reliability of this technique is unclear. Looking closer at accuracy and precision in using this technique, the authors of the original research fail to highlight that reporting 89% accuracy in 10-year intervals actually signifies an age range of 20 years. In fact, our reported result of 83% accuracy within 12 years of known age results in an age range of 24 years, which is fairly large in a forensic context. On the other hand, in comparison with the pubic symphyseal and auricular surface aging methods, the acetabulum method seems to perform better at reducing broad categories of age for estimates of unknown individuals (5,24). Osteologists are divided on which skeletal age estimation techniques produce the most accurate and precise estimation of age, resulting in perpetual testing and refinement of methods (15,17), as many available criteria and methods for assessing skeletal age at death should be applied to account for variability in the aging process. Even if the reliability between methods differs based on varying age indicators, the sequential addition of other age markers should improve the accuracy of determination in positively identifying human remains (17). Standardized methodologies and objective determination principles using empirical testing are the foundation of admissibility rules of evidence and central to the role of the forensic anthropologist (25). Forensic anthropologists frequently assume the role of an expert witness; therefore, understanding and abiding by the rules of scientific evidence admissibility is vital to constructing new techniques and providing applicable testimony (25–27). Continuous testing of applicable methods of age estimation is essential so that flaws or errors potentially inhibiting the technique or its application can be detected (25–27). To satisfy acceptable measures of scientific evidence through "reliability" and "relevance" (26,27), it is necessary to conduct further testing of Rissech et al.'s (1,4) acetabulum technique on more geographically specific samples.

That said, we encourage other researchers to test this method on available collections. Perhaps an appropriate forum for this technique would be FORDISC® (28), an interactive discriminant functions program that classifies unknown adult skeletal elements based on known samples. One suggestion is to add Rissech et al.'s (1) age estimation technique to the FORDISC® database where researchers from around the globe can submit state frequency data, making it available to every forensic anthropologist in the field. Finally, the investigators suggest that this method be used in conjunction with other age estimation techniques to increase the accuracy and precision of positively identifying human remains.

Acknowledgments

The authors acknowledge the University of Toronto for allowing access to the Grant Collection for research purposes. The authors thank Dr. Carme Rissech and Dr. George Estabrook for lending state frequency data and providing instruction on use of this method. Dr. Michael Schillaci is acknowledged for his assistance with statistical software and preparation of manuscript.

References

1. Rissech C, Estabrook GF, Cunha E, Malgosa A. Using the acetabulum to estimate age at death of adult males. J Forensic Sci 2006;51(2):213–29.
2. Rougé-Maillart CL, Telmon N, Rissech C, Malgosa A, Rougé D. The determination of male adult age by central and posterior coxal analysis—a preliminary study. J Forensic Sci 2004;49(2):208–14.
3. Rissech C, Sañudo JR, Malagosa A. The acetabular point: a morphological and ontogenic study. J Anat 2001;198(6):743–8.
4. Rissech C, Estabrook GF, Cunha E, Malgosa A. Estimation of age at death for adult males using the acetabulum, applied to four western European populations. J Forensic Sci 2007;52(4):774–8.
5. Buckberry JL, Chamberlain AT. Age estimation from the auricular surface of the ilium: a revised method. Am J Phys Anthropol 2002;119(3):231–9.
6. İşcan MY. Age markers in the human skeleton. Springfield, IL: Charles C. Thomas, 1989.

7. Boldsen JL, Milner GR, Konigsberg LW, Wood JW. Transition analysis: a new method for estimating age from skeletons. In: Hoppa RD, Vaupel JW, editors. Paleodemography: age distributions from skeletal samples. Cambridge, UK: Cambridge University Press, 2002; 73–106.
8. Bedford ME, Russell KF, Lovejoy CO, Meindl RS, Simpson SW, Stuart-Macadam PL. Test of the multifactorial aging method using skeletons with known ages at death from the Grant collection. Am J Phys Anthropol 1993;91(3):281–97.
9. Murray KA, Murray T. A test of the auricular surface aging technique. J Forensic Sci 1991;6(4):1162–9.
10. Falys CG, Schutkowski H, Weston DA. Auricular surface aging: worse than expected? A test of the revised method on a documented historic skeletal assemblage. Am J Phys Anthropol 2006;130(4):508–13.
11. Berg GE. Pubic bone age estimation in adult women. J Forensic Sci 2008;53(3):569–77.
12. Saunders SR, Fitzgerald C, Rogers T, Dudar C, McKillop H. A test of several methods of skeletal age estimation using a documented archaeological sample. Canadian Soc Foren Sci 1992;25(2):97–118.
13. 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(6): 1203–9.
14. Kimmerle EH, Konigsberg LW, Jantz RL, Baraybar JP. Analysis of age at death estimation through the use of pubic symphyseal data. J Forensic Sci 2008;53(3):558–68.
15. Hens SM, Rastelli E, Belcastro G. Age estimation from the human os coxa: a test on a documented Italian collection. J Forensic Sci 2008;53(5):1040–3.
16. Konigsberg LW, Herrmann NP, Wescott DJ, Kimmerle EH. Estimation and evidence in forensic anthropology: age at death. J Forensic Sci 2008;53(3):541–57.
17. White TD, Folkens PA. Human osteology, 2nd edn. San Diego, CA: Academic Press, 2000;340–61.
18. Lucy D, Aykroyd RG, Pollard AM, Solheim T. A Bayesian approach to adult human age estimation from dental observations by Johanson's age changes. J Forensic Sci 1996;41(2):189–94.
19. Prince DA, Kimmerle EH, Konigsberg LW. A Bayesian approach to estimate skeletal age at death utilizing dental wear. J Forensic Sci 2008;53(3):588–93.
20. Haglund WD. Dogs and coyotes: postmortem involvement with human remains. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press LLC, 1997;367–81.
21. Komar DA, Potter WE. Percentage of body recovered and its effect on identification rates and cause and manner of death determination. J Forensic Sci 2007;52(3):528–31.
22. Komar DA. Twenty-seven years of forensic anthropology casework in New Mexico. J Forensic Sci 2003;48(3):521–4.
23. Laurent M. Comparison of four skeletal methods for the estimation of age at death on white and black adults. J Forensic Sci 2007;52(2):302–7.
24. Suchey JM, Katz D. Applications of pubic age determination in a forensic setting. In: Reichs KJ, editor. Forensic osteology: advances in the identification of human remains, 2nd edn. Springfield, IL: Charles C. Thomas, 1998;204–36.
25. Christensen AM. The impact of *Daubert*: implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification). J Forensic Sci 2004;49(3):427–30.
26. *Federal Rules of Evidence* (1975; 2000).
27. *Daubert v. Merrell Dow Pharmaceuticals Inc.*, 509 U.S. 579 (1993).
28. Ousley SD, Jantz RL. FORDISC 3.0: Personal Computer Forensic Discriminant Functions, 3.0 ed. Knoxville, TN: University of Tennessee, 2005.

Additional information and reprint requests:
Stephanie E. Calce, M.Sc.
Department of Anthropology
University of Victoria
P.O. Box 3050, STN CSC
Victoria, BC, V8W 3P5
Canada
E-mail: scalce@uvic.ca