Population-Specific Variation in the Accuracy of Rogers' Method of Sex Estimation

Honours Thesis

Rachel Simpson

ABSTRACT

Rogers' (1) method of sex estimation is a visual technique that evaluates morphological variation in four traits of the distal posterior humerus. This method has the potential for widespread application in biological anthropology, but previous tests have been unable to replicate Rogers' initial accuracy rate of 92%. Additionally, the role of populations in the accuracy of the method has not been sufficiently explored, as only one study (2) has controlled for it. Wanek (2) found differences in the accuracy of Rogers' method correlated with different populations but concluded the method could be used on all human populations, regardless. This study tests Wanek's (2) conclusion through a blind test of Rogers' (1) original method, though it differs methodologically from previous studies (1–7) by seriating humeri according to trait expression, and by using logistic regression for analysis of results. In conducting a blind test on a sample of American black and white individuals from The Hamann-Todd Osteological Collection, I found that the method was 67% accurate overall, and that odds for a correct classification were 2.03 more likely for a white individual than for a black individual. Prior to applying this method in the future, bioarchaeologists and forensic anthropologists should consider these results within the context of their study.

Introduction

Rogers' (1) method of skeletal sex estimation from the distal humerus is valuable in estimating sex from fragmented, commingled, or subadult human remains. However, subsequent tests of Rogers' method (2–7) were unable to replicate her initial accuracy rate of 92%. Additionally, with the exception of Wanek (2), scholarly tests of the method have failed to control for population. Wanek found substantial differences in the accuracy of the method correlated with different populations; nevertheless she concluded that the method could still be applied on all human populations. This study set out to test Wanek's conclusion by exploring whether Rogers' method is equally effective when applied to white and non-white individuals. The methodology of this study differs from previous works in its use of a seriation technique in testing Rogers' method, and in its use of logistic regression analysis of the results.

The term 'population' generally refers to a group of people from a given area during a given time period and their interactions with social, economic and environmental stressors (8). Within the context of the present study, the populations investigated are American black and American white dating from approximately 1910–1940 and located in Cleveland OH (9). It should be noted that these categories are inherently biocultural, and that they, along with related concepts such as biological ancestry, biological affinity, or race, are highly problematic. An extended discussion of the limitations of these concepts is found in the discussion below.

Background

Accurate sex estimation of individuals is critical in bioarchaeology because it forms the basis of many paleodemographic and societal inferences (10, 11). In forensic contexts, sex estimation is valuable in identifying individuals (12). Biological anthropologists most often rely upon macroscopic methods of sex estimation (13–15) that revolve around assumptions of sexual dimorphism in skull robusticity (16) and pelvic adaptations for pregnancy and parturition in females (17, 18). Reliance upon these methods is problematic, however. Visual morphological methods tend to be subjective and, therefore, substantial rates of intra- and inter-observer error have been established (19–21). Second, use of the skull and pelvis is not always possible when remains are fragmented, incomplete, or commingled (22, 23). Furthermore, sexual dimorphism of the skull and pelvis tends to emerge only post-adolescence, making these techniques less useful for the evaluation of sub-adults. Advances in metric (e.g., 24–26) and genetic (27) sex estimation methods have partly overcome these limitations; however, most metric methods are population-specific, and there are often problems of cost, preservation, amplification, and contamination in the application of genetic methods (28).

Rogers (1) developed a visual method of sex estimation for humans using the distal humerus that she found to be 92% accurate. As a visual method, it is quick and inexpensive to apply, and because the method uses the distal humerus as opposed to the skull or pelvis, it overcomes many of the limitations caused by the overreliance upon dominant cranial and pelvic methods outlined above. Additionally, Rogers (4) reported that the method is also useful in estimating sex for adolescent individuals, which as noted above, is less reliable with the skull or pelvis (28). Due to these clear advantages, the method has been widely applied in various bioarchaeological and forensic contexts (30–39).

Rogers' method relies on sexual dimorphism of the carrying angle in the elbow, which is the angle formed by the lateral divergence of the forearm from the upper arm at the elbow joint during supination (40). Studies have often found there are significant differences in carrying angle between the sexes (41, 42). The average carrying angle for males is 11.6º ± 3.2 while for females it is 16.7º ± 2.6 (41). One hypothesis suggests that carrying angle is a secondary sex characteristic; differences in carrying angle are caused by males having broader shoulders relative to hips and by females having broader hips relative to shoulders (43, 44), although this theory has been rejected by some scholars (40).

Rogers' (1) method evaluates sexual dimorphism within the following four features of the distal posterior humerus: trochlear constriction, trochlear symmetry, olecranon fossa size/shape, and the angle of the medial epicondyle.

Trochlear constriction (Figure 1) refers to the extent that the trochlea appears 'pinched'. In males, the trochlea is less constricted whereas in females, the trochlea is more constricted, resulting in a "spool-shaped" trochlea (1).

FIGURE 1—*Trochlear constriction, from Rogers* (1).

Trochlear symmetry (Figure 2) refers to the degree to which the medial edge of the trochlea extends distally relative to the lateral edge. The male trochlea is typically assymetrical while the female trochlea is typically symmetrical (1, 5).

The olecranon fossa (Figure 2) differs between males and females in shape and depth, which males typically having a relatively shallower and more triangular fossa and females typically having a relatively deeper and more oval fossa.

FIGURE 2—*Trochlear symmetry and olecranon fossa size and shape, from Rogers* (1).

Lastly, the *angle of the medial epicondyle* (Figure 3) refers to the extent in which the medial epicondyle is raised relative to the trochlea (1). This trait is best assessed when viewed at the distal end while placed on a flat tabletop, anterior side down. The medial epicondyle of males is more or less parallel to the tabletop, while the medial epicondyle of females is angled upward from the tabletop.

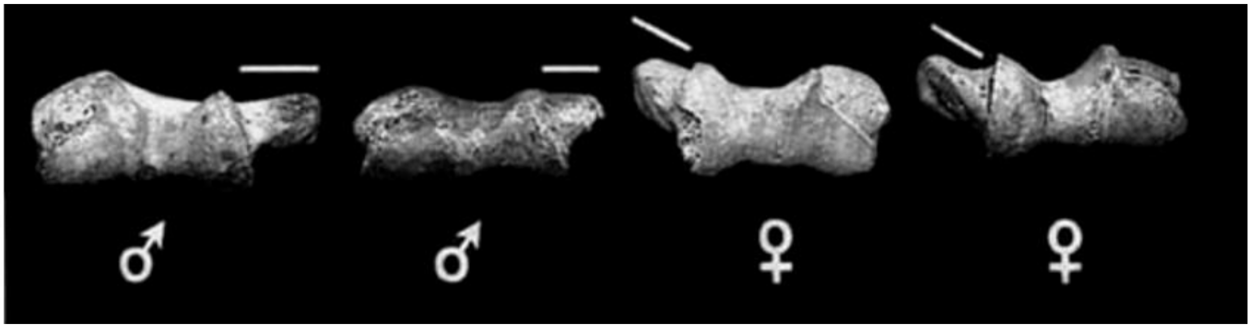

FIGURE 3—*Angle of the medial epicondyle, from Rogers* (1).

Rogers (1) developed her method by evaluating skeletal remains in the Grant Documented Skeletal Collection (University Toronto), the Maxwell Documented Skeletal Collection (University of New Mexico), and the William M. Bass Donated Skeletal Collection (University of Tennessee) in which all evaluated individuals were identified as white (i.e. considered of European biological ancestry). Rogers (1) acknowledged the lack of testing on individuals of non-white populations and suggested that further research be conducted on collections comprised of more diverse populations.

The accuracy of her method has since been tested numerous times (2–7); however, her call for increased diversity has been insufficiently addressed, as only one study (2) has controlled for population. Falys et al. (3), Rogers (4), and Watkinson (6) tested the method solely on European skeletal collections. Vance et al. (5) and Harrison (7) tested the method South African and American skeletal collections respectively, which were comprised of black and white individuals, but neither study evaluated the role of population in the accuracy of the method.

Wanek (2) tested the method on a large sample stemming of the following categories: American whites, American blacks, Native Americans, Alaskan Eskimos, Chinese, Alaskan Aleutians, Japanese, and other (Egyptian, Australian Aborigine, Hawaiian, Arikara Indian, Jamaican, Puerto Rican, Patagonian, Chilean). She argued that while differences exist in the accuracy of this method across different populations (Table 1), the method can be applied on all humans, regardless of population. Wanek's test, however, is not an exact test of Rogers' (1) method because she used two additional features of medial epicondylar symmetry and distal spool curvature in her analysis, and the overall accuracy rate of 84% reflects removal of the trochlear constriction feature from analysis and exclusion of ambiguous cases from analysis. Additionally, the sample sizes of different populations range to a considerable degree (e.g. only five individuals comprise the Japanese subsample).

TABLE 1—*Accuracy of Rogers' method by biological affinity, from Wanek* (2).

	Overall Accuracy (%)	**Accuracy by Trait (%)**			
		Trochlear Constriction	**Trochlear Symmetry**	**Olecranon Fossa Shape**	**Angle of the Medial Epicondyle**
American whites (n = 208)	85	70	74	82	74
American blacks (n = 234)	78	62	71	77	70
Native Americans (n = 54)	80	70	59	70	76
Alaskan Eskimos (n = 44)	83	69	73	77	66
Chinese (n = 41)	94	57	60	60	98
Alaskan Aleutians (n = 40)	94	65	73	85	85
Japanese (n = 5)	80	80	80	80	80
Total (n = 649)	**83**	**65**	**70**	**77**	**74**

In my study, I evaluated Rogers' (1) four original features, restricting analysis to a sample comprised of American black (n= 85) and American white (n=114) individuals. The overall accuracy rate of this study reflects the accuracy of the four features, and any ambiguous classifications were considered incorrect.

Materials and Methods

Data Collection

I tested the method on a sample of 199 humeri from the Hamann-Todd Osteological Collection at the Cleveland Museum of Natural History. Often called “the unclaimed dead of Cleveland” (2, 45, 46), this documented skeletal collection of more than three thousand individuals was amassed by T. Wingate Todd in the early twentieth century (47). Like many other American documented skeletal collections, the Hamann-Todd Osteological Collection is mainly comprised of cadavers not claimed from the morgue (47).

Humeri of individuals classified in the collection as American black and white males and females were selected by a third party to allow for a blind test. No additional steps were required to ensure a blind test as no information beyond an individual identification number was written on the skeletal elements. Similarly to Wanek (2) and Vance et al. (5), I evaluated left humeri to maintain consistency.

Unlike previous applications of this method, rather than evaluating all four traits for each humerus simultaneously, I instead assessed each trait independently by repeatedly seriating the entire sample collection and assigning sex on a three-point scale for each individual trait (male, ambiguous, or female). Between seriations, I randomized the order of the humeri, allowing me to assess each trait without bias from my assessment of the previous trait. When compiling data from all four traits, I assigned an overall sex to each humerus on a five point scale (male, probable male, ambiguous, probable female, or female) based on the following criteria (Table 1).

TABLE 2—*Criteria for overall sex assignment.*

Assigned Sex	Criteria
Male	All four traits consistent with the male sex
Probable Male	Three out of four traits consistent with the male sex
Ambiguous	Two traits consistent with each sex

Probable Female	Three out of four traits consistent with the female sex
Female	All four traits consistent with the female sex

As proposed by Rogers (1), olecranon fossa shape (if unambiguous itself) was given extra weight for humeri classified as ambiguous by the above criteria (Table 1), allowing some 'ambiguous' humeri to be classified as probable males or probable females.

After completing the blind test of the method, I collected data on the actual sex, actual or estimated age, and black or white designation of each individual, based on records. Potential problems with these categories will be examined in the discussion.

Data Analysis

For each humerus belonging to a female individual, a classification was considered correct if the humerus was assigned to the female or probable female category, and it was considered incorrect if assigned to the male, probable male, or ambiguous category. Similarly, each male humerus was considered correct if assigned to the male or probable male categories, and it was considered incorrect if assigned to the female, probable female, or ambiguous category.

Previous tests of Rogers' method have used Fisher's Exact Probability Test (1), which is beneficial for establishing statistical significance within small sample sizes, or Chi-Squared tests (2, 3, 5) which assess the statistical significance of differences between expected versus observed results of a single variable. However, I instead used logistic regression to evaluate the relationship between the predictors of population and sex on the accuracy of the method. This allowed me to assess whether the effect of population on the accuracy of this method is different between males and females, or, in other words, whether there was interaction between the two predictors of population and actual sex. The logistic regression was modelled by the following equation: logit(P) = $\beta_0 + \beta_1$ (Population) + β_2 (Sex) + β_3 (Interaction), in which the effect of population, sex, and interaction on the dependent variable of accuracy (β_0) can be evaluated.

When interaction was shown not to be statistically significant, meaning that the effect of population on the accuracy of this method is equally significant for both sexes, it was removed and the logistic regression was accordingly modeled by the equation: logit(P) = $\beta_0 + \beta_1$ (Population) + β_2 (Sex). Results from logistic regression are presented in an odds ratio as opposed to standard probability.

Results

Overall, the method was 67% accurate, ranging from 58% accuracy for black individuals and 73% accuracy for white individuals (Figure 4 and Table 3).

TABLE 3—*Accuracy of Rogers's method on black and white individuals.*

Population	**Sex**	**Overall Accuracy**		
			Incorrect	Correct
Black Individuals	Female (n=35)	Count	13	22
		%	37	63
	Male (n=50)	Count	23	27
		%	46	54
	Total (n=85)	Count	36	49
		%	42	58
White Individuals	Female (n=29)	Count	8	21
		%	28	72
	Male (n=85)	Count	23	62
		%	27	73
	Total (n=114)	Count	31	83
		%	27	73
Total	Female (n=64)	Count	21	43
		%	33	67
	Male (n=135)	Count	46	89
		%	34	66
	Total (n=199)	Count	67	132
		%	33	67

The highest overall accuracy is achieved when the four traits are analyzed in conjunction. However, when examining the traits individually, olecranon fossa shape is the most accurate trait and trochlear symmetry is the least accurate trait. However, this differs marginally when controlling for population (Table 4).

TABLE 4—*Accuracy of Rogers' method by trait.*

Trait	**Accuracy (%)**				
	Black Individuals		**White Individuals**		**Total**
	Females	Males	Females	Males	
Trochlear Constriction	51	54	41	58	53
Trochlear Symmetry	40	38	38	62	49
Olecranon Fossa Size and Shape	60	46	62	66	59
Angle of the Medial Epicondyle	49	62	72	51	56

The logistic regression results without interaction (Table 5) show that population is a statistically significant predictor of accuracy ($p < 0.05$). As demonstrated by the results, the odds for correct sex estimation using Rogers' method are 2.027 times more likely for a white individual than for a black individual.

TABLE 5—*Logistic regression results (with interaction).*

	df	P-value	Odds
Population	1	0.419	1.551
Actual Sex	1	0.417	0.694
Interaction	1	0.552	1.480

TABLE 6—*Logistic regression results (without interaction).*

	df	P-value	Odds
Population	1	0.022	2.027
Actual Sex	1	0.830	0.830

To help understand how population is important, I assessed of the nature of the incorrect classifications of sex (males, probable males, ambiguous, probable females, and females) between the samples of American black and white individuals. As demonstrated by Figure 5, there is a slightly smaller proportion of ambiguous classifications for the white individuals. In addition, for the sample of black individuals, more females were incorrectly classified as males or probable males; whereas for the sample of white individuals, more males were incorrectly classified as females or probable females.

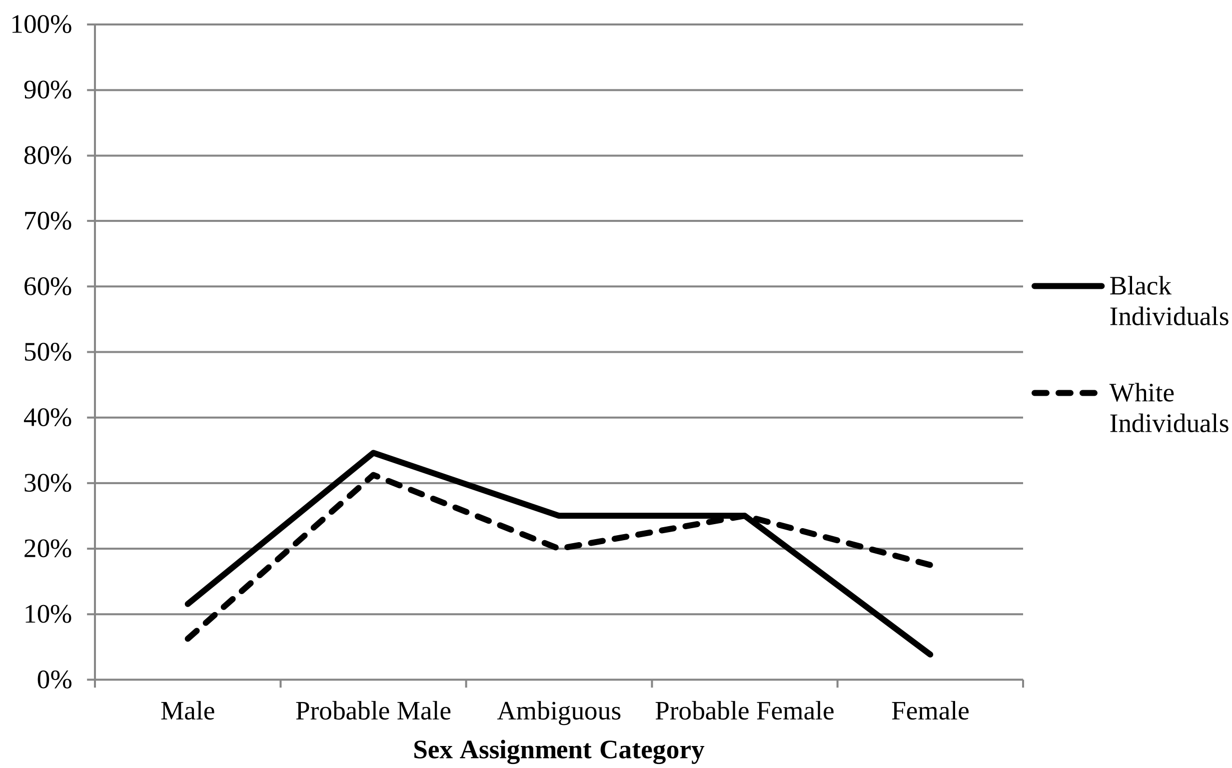

FIGURE 4—*Proportion of incorrect classifications for black (n=36) and white (n= 31) individuals*.

Discussion

Comparison to Previous Studies

In this study, the method was 67% accurate, ranging from 58 to 73% accuracy for black and white individuals respectively. The logistic regression model showed that the odds for a correct sex classification were 2.027 times more likely for a white individual than for a black individual. The effect of population on the accuracy of the method did not differ between males and females. Wanek (2) also found significant differences in the accuracy of this method corresponding to population designations; the method was 78% accurate on black individuals and 85% accurate on white individuals. However, she concluded that although differences exist in the accuracy of the method among individuals of different populations, the method can still be used on all human populations. My results do not support Wanek's conclusion; the odds ratio of 2.027 suggests that the application of the technique to non-white populations may be more problematic than Wanek concluded.

As shown by Table 6, the overall rate of accuracy of 67% is lower than that found in previous tests of Rogers' method. This is to be expected as most of these studies (1, 3, 4, 6) tested the method on only white individuals, and the method has been shown to be more effective on white individuals. Wanek's (2), Vance et al.'s (5), and Harrison's (7) samples consisted of individuals from different populations, though only Wanek (2) evaluated the role of population differences in the accuracy of the method. The results from this study are closest to the accuracy rates from Vance et al.'s (5) and Harrison's (7) studies, particularly in terms of the accuracy of the individual features. However, as neither Vance et al. (5) nor Harrison (7) controlled for population in their studies, it is impossible to draw direct comparisons between their studies and this study on the role of population in the accuracy of Rogers' method. Wanek's (2) study produced considerably higher accuracy, though her rate is inflated by her exclusion of ambiguous classifications, and she tested the method on a wider range of populations.

TABLE 7—*Comparison of results of previous tests of Rogers' method (to nearest whole number)*

Study	**Overall Accuracy (%)**	**Accuracy of Traits (%)**			
		Trochlear Constriction	**Trochlear Symmetry**	**Olecranon Fossa Shape and Size**	**Angle of the Medial Epicondyle**
Simpson (current study)	67	53	49	59	56
Rogers (1999)	92	74	74	91	86
Wanek (2002)	84	65	79	77	74
Falys et al. (2005)	79	69	79	82	75
Rogers (2009)	81	—	—	—	—
Vance et al. (2011)	76	—	45–56	57–61	55–70
Watkinson (2012)	83	74	65	78	76
Harrison (2017)	—	66	60	67	71

In Rogers' (1, 4), Falys et al.'s (3), and Vance et al.'s (5) studies, more females were accurately estimated whereas in Wanek (2) and Watkinson (6), more males were accurately estimated. In this study, differences between males and females were not statistically significant and varied according to the different traits. This means that either there are subjective differences in observer perception of traits or that the expression of these traits as more "male" or "female" varies according to the sample analyzed. Tests of Rogers' method also vary with respect to the accuracy of the different features. For Rogers (1), Falys et al. (3), and this study, olecranon fossa size and shape was the most accurate predictor of sex, but for other studies, trochlear symmetry (2), or angle of the medial epicondyle (5, 7) was most effective.

Limitations

There are some limitations to this study, particularly pertaining to the concept of race as assigned in documented skeletal collections. There are some also limitations to using a documented skeletal collection. The cadaver-based anatomical collections, including the Hamann-Todd Osteological Collection, are biased samples of society often comprised of individuals of lower socioeconomic status (47, 48). However, these collections are often treated in scientific studies as representative cross-sections of populations. To further compound this issue, the designations of ‘black’ and ‘white’ within these anatomical collections are a product of the time in which they were made (47, 49). Todd assembled the Hamann-Todd Osteological Collection in the early twentieth century, and his racial designations of individuals likely reflect the notions of race of this time period—notions which do not align with twenty-first century academic notions of race nor of biological ancestry.

It is outdated and problematic to assume that humans can be neatly bounded within national or continental categories, and further problematic to assume that human variation and diversity directly and unambiguously corresponds to said arbitrary divisions (50–52). In using concepts such as race (or biological ancestry), there is in part a packaging together of various biological traits, such as skin colour, into groupings that oversimplifies the reality of human genetics (50). These typological designations are also often made within a social context in which power differentials between the races partly dictates who is assigned to which group, and these assignments often do not correspond directly with biological ancestry. An example of this would be the ‘one drop’ rule prevalent throughout American history in which individuals were legally considered black if they had any African ancestry (53, 54). Consequently, it is difficult to distinguish whether the variation exhibited by individuals of designated groups actually stems from genetic differences or several other variables inherent in the life experiences of socially constructed groups. In other words, is it possible that the variation between populations is actually the skeletal manifestation of stressors and different life history traits (e.g. average age at death) brought on by a legacy of marginalization toward the black individuals of this sample?

These limitations do not discount the value of this study; research has shown that the accuracy of some sex estimation methods varies population-to-population (55). Researchers cannot assume that such methods are universally accurate for all humans across time and space. This study has shown that there are significant differences in the accuracy of Rogers' method correlated with designations of race. This is not to say that there are differences in distal humerus morphology inherent in all "black" and "white" individuals; it is important to consider the alternative variables that may produce these results but that are masked by the concept of race or biological ancestry. It was not within the bounds of this study to determine the extent to which such variables had a role in the accuracy of Rogers' method but future studies should attempt to tease apart whether these significant differences can be attributed "biological ancestry" or another source of variation.

Conclusions

This study tested Rogers' method on a sample of American black and white individuals from the Hamann-Todd Collection. Similarly to previous tests of Rogers' method (2–7), Rogers' (1) initial accuracy rate of 92% was not replicated. Like Wanek (2), this study found that there are significant differences in the accuracy of this method correlated with different populations; however, the results from this study conflict with Wanek's conclusion that the method can be applied on all human populations regardless of population. As this study evaluated the method on American black and white individuals from the nineteenth and twentieth centuries and the results may be a product of life experience variables, these results should not be generalized to all individuals of African and European ancestry across time and space. However, prior to applying this method in the future, professionals should be cognisant of these findings that Rogers' method is not consistently effective across all populations. Additional studies should be conducted on more populations and attempt to determine the sources of variation for these results.

Acknowledgements

This research was funded by the MacEwan Undergraduate Student Research Initiative (USRI) Grant Fund and was presented at the 2017 Canadian Association of Physical Anthropology Annual Meeting. Special thanks to Dr. Hugh McKenzie for supervision of this project, Dr. Karen Buro for offering her statistical expertise, Pamela Mayne Correia for granting access to the University of Alberta's osteological teaching collection for a pilot study, and Dr. Yohannes Haile-Selassie and Lyman Jellema for granting access to the Hamann-Todd Osteological Collection.

References

1. Rogers TL. A visual method of determining the sex of skeletal remains using the distal humerus. J Forensic Sci 1999; 44: 57–60.
2. Wanek V. A qualitative analysis for sex determination in humans utilizing posterior and medial aspects of the distal humerus [dissertation]. Portland, Oregon: Portland State University, 2002.
3. Falys CG, Schutkowski H, Weston DA. The distal humerus: a blind test of Rogers' sexing technique using a documented skeletal collection. J Forensic Sci 2005; 50(6): 1–5.
4. Rogers TL. Sex determination of adolescent skeletons using the distal humerus. Am J Phys Anthropol 2009; 140: 143–8.
5. Vance VL, Steyn M, L'Abbé EN. Nonmetric sex determination from the distal and posterior humerus in black and white South Africans. J Forensic Sci 2011; 56(3): 710–4.
6. Watkinson L. 'Funny bones': an evaluation of the Rogers visual method of sex estimation, using the posterior distal humerus, on an early medieval population from Bamburgh [dissertation]. Durham: University of Durham, 2012.
7. Harrison DL. An evaluation of the methods used in the estimation of sex [dissertation]. London: University College London, 2017.
8. Cincotta RP, Gorenflo LJ. Preface. In: Cincotta RP, Gorenflo LJ, editors. Human population: its influences on biological diversity. Berlin: Springer, 2011; vii–ix
9. Mensforth RP, Latimer BM. Hamann-Todd Collection aging studies: osteoporosis fracture syndrome. Am J Phys Anthropol 1989; 80: 461–479.

10. Meindl RS, Lovejoy CO, Mensforth RP, Don Carlos L. Accuracy and direction of error in the sexing of the skeleton: implications for paleodemography. Am J Phys Anthropol 1985; 68: 79–85.
11. Schutkowski H. What you are makes you eat different things: interrelations of diet, status, and sex in the early medieval population of Kirchheim unter Teck, FGR. Hum Evol 1995; 10(2): 119–30.
12. Konigsberg LW, Algee-Hewitt BFB, Wolfe Steadman D. Estimation and evidence from forensic anthropology: sex and race. Am J Phys Anthropol 2009; 139(1): 77–90.
13. Phenice TW. A newly developed visual method of sexing the os pubis. Am J Phys Anthropol 1969; 30(2): 297–301.
14. Buikstra JE, Ubelaker DH, editors. Standards for data collection from human skeletal remains: proceedings of a seminar at the field museum of natural history. Fayettville, AR: Arkansas Archaeological Survey, 1994.
15. Bruzek J. A method for visual determination of sex, using the hip bone. Am J Phys Anthropol 2002; 117 (2): 157–68.
16. Walrath DE, Turner P, Bruzek J. Reliability test of the visual assessment of cranial traits for sex determination. Am J Phys Anthropol 2004; 12: 132–7.
17. LeVelle M. Natural selection and developmental sexual variation in the human pelvis. Am J Phys Anthropol 1995; 98(1): 59–72.
18. Tague RG. Costal process of the first sacral vertebrae: sexual dimorphism and obstetrical adaptation. Am J Phys Anthropol 2007; 132(3): 395–405.
19. Rogers T, Sanders S. Accuracy of sex determination using morphological traits of the human pelvis. J Forensic Sci 1994; 39(4): 1047–56.

20. Walker PL. Greater sciatic notch morphology: sex, age, and population differences. Am J Phys Anthropol 2005; 127: 385–91.

21. Blanchard BK. A study of the accuracy and reliability of sex estimation methods of the human pelvis [dissertation]. Chico: California State University, 2010.

22. Kelley MA. Sex determination with fragmented skeletal remains. J Forensic Sci 1979; 24: 154–8.

23. Panakhyo M, Jacobi K Limited circumstances: creating a better understanding of prehistoric peoples through the reanalysis of collections of commingled human remains. In Osterholtz AJ, editor. Theoretical approaches to analysis and interpretation of commingled human remains. New York, NY: Springer, 2016; 75–96.

24. Vodanović M, Demo Ž, Njemirovskij V, Keros J, Brkić H. Odontometrics: A useful method for sex determination in an archaeological skeletal population? J Archaeol Sci 2007; 34: 905–13.

25. Spradley MK, Jantz RL. Sex estimation in forensic anthropology: skull versus postcranial elements. J Forensic Sci 2011; 56(2): 289–96.

26. Kubicka AM, Piontek J. Sex estimation from measurements of the first rib in a contemporary Polish population. Int J Leg Med 2016; 130: 265–72.

27. Stone AC, Milner GR, Pääbo S, Stoneking M. Sex determination of ancient human skeletons using DNA. Am J Phys Anthropol 1996; 99: 231–8.

28. Capelli C, Tschentscher F, Pascali VL. “Ancient” protocols for the crime scene?: similarities and differences between forensic genetics and ancient DNA analysis. Forensic Sci Int 2003; 131(1): 59–64.

29. Cunningham, C., L. Scheuer & S. Black. 2016. *Developmental Juvenile Osteology* (2nd ed.). Academic Press.
30. Milbank D. The Bull Inn, High Street, Sonning, Berkshire: an archaeological watching brief and salvage excavation for Fuller, Smith and Turner plc. Thames Valley Archaeological Services Ltd, 2007.
31. Rossouw LZ. A forensic anthropological investigation of skeletal remains recovered from a 1000-year-old archaeological site in north western Namibia [dissertation]. Cape Town: University of Cape Town, 2010.
32. Schats R. Fysich Antropologische Analyse van de Skeletten can de Paardenmarkt [dissertation]. Leiden, Netherlands: Leiden University, 2012.
33. Shapland F, Lewis M. Brief communication: a proposed osteological method for the estimation of pubertal stage in human skeletal remains. Am J Phys Anthropol 2013; 151: 302–10.
34. Shapland F, Lewis M. Brief communication: a proposed method for the assessment of pubertal stage in human skeletal remains using cervical vertebrae maturation. Am J Phys Anthropol 2014; 158: 144–153.
35. Swanepoel E, Steyn M. Preliminary report on the skeletal remains of the inhabitants of 17th century Dambarare, Zimbabwe. South African Archaeological Society Goodwin Series 2013; 11: 68–77.
36. Pike K. Bearing identity: a biocultural analysis of human remains from Old Mission Point (C1Dq-1), New Brunswick [dissertation]. St. John’s, NL: Memorial University, 2014.

37. Karapetien MK. Costal facet variations on the eighth, ninth, and tenth thoracic vertebrae: association with sex and shifts in the cranio-caudal pattern of the human axial skeleton. European Journal of Anatomy 2015; 19(2): 179–88.

38. Arthur N, Gowland RL, Redfern, RC. Coming of age in Roman Britain: osteological evidence for pubertal timing. Am J Phys Anthropol 2016; 159(4): 698–713.

39. Pandolfi L, Mannino MA, Talamo S, Salari L, Sansò P, Sublimi Saponetti S et al. A reassessment of the infills and faunal assemblages of karst cavities known as ventarole in Salento (Apulia, Southern Italy): a multidisciplinary investigation on Cava Donno (Corigliano D'Otranto, Lecce). Alpine and Mediterranean Quarternary 2017; 30(1): 25–40.

40. Purkait R, Chandra H. An anthropometric investigation into the probable cause of formation of 'carrying angle': a sex indicator. Journal of Indian Academy of Forensic Medicine 2004; 26(1): 14–9.

41. Paraskevas G, Papadopoulos A, Papaziogas B, Spanidou S, Argiriadou H, Gigis J. Study of the carrying angle of the human elbow joint in full extension: a morphometric analysis. Surg Radiol Anat 2004; 26: 19–23.

42. Van Roy P, Baeyens JP, Fauvart D, Lanssiers R, Clarijs JP. Artho-kinematics of the elbow: study of the carrying angle. Ergonomics 2005; 48: 1645–1656.

43. Atkinson WB, Elftman H. The carrying angle of the human arm as a secondary sex character. Anat Rec 1945; 91: 49–52.

44. Hooton EA. Up from the ape. New York, NY: Macmillian, 1946.

45. Krebs V, Incavo SJ, Shields WH. The anatomy of the acetabulum: what is normal? Clin Orthop Rel Res 2009; 467(4): 868–75.

46. Morris WZ, Gebhart JJ, Goldberg VM, Wera GD. Implant size availability affects reproduction of distal femoral anatomy. The Journal of Knee Surgery 2016; 29(5): 409–13.
47. Hunt DR, Albanese J. History and demographic composition of the Robert J. Terry Anatomical Collection. Am J Phys Anthropol 2005; 127: 406–17.
48. Kunos CA, Simpson SW, Russell KF, Hershkovitz I. First rib metamorphosis: its possible utility for human age-at-death estimation. Am J Phys Anthropol 1999; 110: 303–23.
49. Albanese J, Saunders SR. Is it possible to escape racial typology in forensic identification? In: Schmitt A, Cunha E, Pinheiro J, editors. Forensic anthropology and medicine: complementary sciences from recovery to cause of death. Totowa, NJ: Humana Press Inc, 2006; 281–316.
50. Keita SOY, Kittles RA. The persistence of racial thinking and the myth of racial divergence. Am Anthropol; 1997; 99(3): 534–44.
51. Armelagos GJ, van Gerven DP. A century of skeletal biology and paleopathology: contrasts, contradictions, and conflicts. Am Anthropol 2003; 105(1): 53–64.
52. Albanese J, Tuck A, Gomes J, Cardoso HFV. An alternative approach for estimating stature from long bones that is not population- or group-specific. Forensic Sci Int 2016; 259: 59–68.
53. Hollinger DA. The one drop rule and the one hate rule. Daedalus 2005; 134(1): 18–28.
54. Rockquemore KA, Brunsma DL. Socially embedded identities: theories, typologies, and processes of racial identity among black/white biracials. Sociol Q 2002; 43(3): 335–356.

55. Wright LE, Yoder CJ. Recent progress in bioarchaeology: approaches to the osteological paradox. J Archaeol Res 2003; 11(1): 43–70.