A112 Age-Cohort Categorization and Multi-Factorial Age Estimation in Machine Learning Environments

Joseph T. Hefner, PhD, Michigan State University, Dept of Anthropology, 355 Baker Hall, East Lansing, MI 48824; and Nicholas V. Passalacqua, PhD, Western Carolina University, 101 McKee, Cullowhee, NC 28723*

After attending this presentation, attendees will better understand the utility of age-cohort categorization using a multi-factorial approach to the estimation of age at death from adult human skeletal remains. As one part of a larger study into multi-factorial age estimation methods, this presentation addresses a straightforward approach to age estimates based on the combined data of three commonly used adult aging methods.

This presentation will impact the forensic science community by addressing the knowledge gap in multi-factorial age estimate strategies. The method proposed has the added benefits of quantification, verification, and robust statistical classification algorithms necessary for empirically based multi-factorial age estimates.

The traditional approach to producing a final age estimate from multiple skeletal indicators relies, in part, on heuristic cut-off points, minimum-to-maximum estimates, or non-statistical intervals developed using experience and expertise. This traditional approach is generally preferred over other, more robust methods because it is easier to apply during casework than current multi-factorial approaches, such as ADBOU, that provide age-at-death probabilities by means of transition analysis. Motivated by this knowledge gap and recent calls for research into multi-factorial age estimation strategies from the National Institute of Standards and Technology Overseas Security Advisory Council (NIST OSAC), this study tested the efficacy of an age-cohort categorization strategy using three commonly incorporated adult age indicators: the pubic symphysis, the sternal end of the fourth rib, and the auricular surface of the innominate. These age estimation methods were selected because: (1) they are commonly collected during forensic anthropological casework; (2) they constitute verified methods of age estimation; and, (3) as this research seeks to provide proof-of-concept results, these data represent current standards in the field.

Age data were collected following the Suchey-Brooks (pubic symphysis), Iscan et al. (rib), and Osborne et al. (auricular surface) methods. Data were collected for 358 (male, *n*=240; female, *n*=118) individuals from the William M. Bass Donated Skeletal Collection in Knoxville, TN, and the Hamann-Todd Collection Cleveland, OH. The sample represents modern American individuals comprising males aged 18 to 97 ($\bar{x}$=55.7, *Standard Deviation (SD)*=16.22) and females aged 20 to 94 ($\bar{x}$=62.4, *SD*=13.1). Age-cohort categorization was achieved using three strategies. First, a broad age class (young adult <39; middle adult 35-59; and old adult >60) was implemented. Next, the data was sub-divided into 10-year (e.g., 18-29, 30-39) and 5-year intervals (e.g., 18-24, 25-29). Male and female data were analyzed both separately and pooled. A two-part data analysis approach measured the ability of three classification algorithms to accurately predict the broad, 10-year, and 5-year age cohorts. To assess how well these data could be combined into a straightforward multi-factorial age estimator, quadratic Discriminant Function Analysis (qDFA) – Leave One Out (LOO) x-validation, Canonical Analysis of Principal coordinates (CAP – 4,999 permutations), and artificial Neural Networks (r-a-NN – hold-out samples for testing and validation) were conducted.

Variable importance measures indicate rib morphology was the single most important variable in every analysis, followed by pubic symphysis, then the auricular surface. This is not surprising given previously documented issues with auricular surfaces as an aging criteria. Classification accuracies for the broad age-cohort categorization were acceptable using qDFA (*F*=64%; *M*=67%; *Pooled*=66%), CAP (*F*=71%; *M*=71%; *Pooled*=68%), and r-a-NN (*F*=77%; *M*=75%; *Pooled*=77%). The qDFA and CAP performed well-below expectation for the 10-year and 5-year interval age-cohort categorizations; however, the regression analysis (with classification) developed using an artificial neural network performed promisingly well, correctly classifying the test samples (hold-out) for the 10-year (73%) and 5-year (74%) age cohorts. To improve network generalization and avoid overfitting, all successful networks were retrained using early stopping and regularization. Retraining generated age-cohort categorization models with mean squared errors lower than the individual performances and, as such, generalized even better to the testing and validation samples while avoiding overfitting.

Developing new multi-factorial approaches to age estimation may require novel applications of data mining and machine learning methods implementing current skeletal age-at-death indicators to predict "reportable" age estimates. The heuristic combination of age-at-death estimates may not properly weigh each indicator, placing emphasis on methods of very little import. In the end, age-cohort categorizations can provide accurate broad, 10-year, and 5-year interval age ranges derived from reliable and verifiable models.

Age Estimation, Machine Learning, Artificial Neural Networks