A77 Reference Sample Optimization for Juvenile Age Estimation From Diaphyseal Lengths

Valerie Sgheiza, MA, University of Illinois at Urbana-Champaign, Urbana, IL 61801; An-Di Yim, MA*, University of Illinois at Urbana-Champaign, Urbana, IL 61801*

Learning Overview: After attending this presentation, attendees will understand how to evaluate the parameters of a good reference sample in juvenile age estimation from diaphyseal lengths, including prioritizing reference sample size over specificity.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by challenging the assumption that reference sample specificity should be prioritized in juvenile age estimation and demonstrating the research potential of evaluating methods using age intervals rather than point estimates alone.

Population specificity of reference samples for biological profile methods is seeing increased attention in forensic anthropology. The assumption that a more specific reference sample will produce better estimates deserves a close look in juvenile age estimation. Relative to adults, juvenile trait-age correlations are high, and sex and ancestry estimation is challenging. The project objective is to compare the importance of reference sample size and specificity when estimating age from long bone lengths. Two competing hypotheses are tested: first, whether better age estimation results will be obtained by using a more specific reference sample, and second, whether better age estimation results will be obtained by using a larger reference sample.

The data consist of diaphyseal length measurements of the humerus, radius, ulna, femur, tibia, and fibula from 102 South African girls, 203 South African boys, 52 United States girls, and 86 United States boys (age range: 1–13 years). First, model stability was measured under varying conditions of reference sample heterogeneity and size. Model stability was calculated as the pooled standard error of two univariate linear models built using reference samples of equal sizes. Reference sample size was varied from 40%–100% of the original size. Reference sample heterogeneity by country for a single sex and sex for a single country was varied from 0%–50% for each of the six long bones in a full factorial design.

To test model performance, ages were estimated using a Bayesian (length dependent on age) multivariate normal linear model with trait covariances included. With each of the four specific sex/country combinations as a test sample, the model was trained on the full sample, the opposite specific group, and the same specific group. When the same group was in the training and test samples, leave-one-out cross-validation was used so that the same individual was never in both the training and test samples for an age estimate. Success rate was measured as the proportion of individuals whose true age fell within a 95% estimated probability interval for age.

In the model stability testing, standard error increased with decreasing sample size, while heterogeneity of reference sample sex and country had no directional effect. In the performance testing, reference sample specificity did not consistently affect performance of the model. In the age estimation testing, slopes and R^2 values of true age versus maximum likelihood estimates showed minimal differences between training/test sample combinations. R^2 values ranged from 0.84 to 0.91, and slopes were between 0.94 and 1.03, indicating strong correlations with minimal bias. Absolute distance of the success rate from 0.95 was moderately negatively correlated with the size of the training sample (rho = -0.5761, $p = 0.0499$). The width of the estimated age interval (scaled by true age) was not significantly correlated with the size of the training sample. Larger reference samples therefore produced equally precise age estimates that were more likely to match the expected error rate.

In conclusion, reference sample size is more important than homogeneity for producing age estimation models with low standard error. Furthermore, using a reference sample that is specific by sex and nationality does not produce consistently more accurate, more precise, or less biased age estimates than pooled reference samples.

Population Specificity, Skeletal Development, Method Evaluation

Copyright 2021 by the AAFS. Permission to reprint, publish, or otherwise reproduce such material in any form other than photocopying must be obtained by the AAFS.