H50 Stature Estimation: Are There Any Advantages to Using Principal Component Analysis?

Kalan S. Lynn, BSc, Mercyhurst College, 501 East 38th Street, Erie, PA 16546; and Stephen D. Ousley, PhD, Mercyhurst College, Department of Applied Forensic, Anthropology, 501 East 38th Street, Erie, PA 16546*

After attending this presentation, attendees will develop a basic comprehension of principal component analysis (PCA) and methods used in stature estimation. Some disadvantages that exist in current stature estimation methods and how the current method remedies this situation will be discussed. Attendees will also learn whether PCA is advantageous to increase the precision of stature estimates. They will also learn whether using a greater number of osteometric measurements in a summation method used in Fordisc will result in more precise stature estimates.

This presentatrion will impact the forensic science community by presenting a new method with which stature estimation equations may be created. It will discuss methods with which the most precise stature estimates may be obtained and how this would improve a biological profile and narrow down the list of missing persons in identification efforts.

Stature estimation is an important component of the biological profile in forensic cases. Many methods for estimating stature have been published with various levels of accuracy and precision. Although many stature estimation methods are reasonably accurate, more precise methods would improve identification efforts by narrowing the list of missing persons who may be the deceased. When trying to use multiple correlated long bone measurements to construct stature estimate regression equations, at least one measurement is often found to be *not* statistically significant due to the effect of multicollinearity. Fordisc 3 (Jantz and Ousley, 2005), as in previous methods, circumvents this problem by summing the measurements of several bones, and then regresses these against forensic stature (FSTAT). Fordisc provided the first method to add combinations of three measurements to estimate stature, which enabled narrower prediction intervals. Another solution to the problem of multicollinearity would be to analyze the principal components of several bone measurements.

The present study investigated the use of principal component analysis (PCA) to estimate stature using a sample of 130 White males from the Forensic Data Bank. Ten data subsets were created and each included FSTAT as well as two to five of the following osteometric measurements: maximum length of the femur, condylo-malleolar length of the tibia, maximum length of the humerus, maximum length of the radius, and basion-bregma height. R (R Development Core Team, 2008) was used for all statistical analyses. Principal components were extracted from each data subset and regressed against FSTAT. In each situation, either the first principal component only, or the first and last principal components were found to be statistically significant ($p < 0.05$). The raw measurements were also summed (creating a cumulative variable) and regressed against FSTAT. The precision of the estimates were determined by first using the 95% prediction interval at the mean. Because prediction intervals are narrowest at the mean, the prediction interval two standard deviations above the mean was also evaluated. The precision of the stature estimates obtained using PCA were compared with those obtained by using the summed measurements. Results were interpreted as practically significant if the decreases in the prediction intervals were great enough to affect the stature estimates, which are most often rounded to the nearest inch. Another factor taken into consideration was the great variability that exists in reported statures (Willey and Falsetti, 1991). Thus, the prediction intervals needed to decrease by at least one-half inch to be considered practically significant.

In comparing summed measurements and PCA for stature estimation, PCA provided some small statistically significant prediction interval improvements that were not practically significant. This study also investigated whether there was greater stature estimate precision when using more osteometric measurements to create the cumulative variables. When considering mean values, using four measurements resulted in a prediction interval that was only 0.21 cm (0.1") smaller than using three measurements (the Fordisc maximum) and this difference is not practically significant. When using values that were two standard deviations above the mean, an increase in precision of only 0.40 cm (0.2") was seen using summed variables created from three and four measurements. These decreases in prediction intervals are not practically significant.

While no practical differences in FSTATs were found when using PCA and summed variables, additional studies using cadaver lengths and measured statures are necessary to further investigate the relationship between physical stature and osteological measurements. Areas of further research may include investigating different populations (e.g., White females and Black males and females), both separately and combined, in which the methods above may be more useful.

Stature Estimation, Biological Profile, Principal Component Analysis

 ** Presenting Author*