A125 Subadult Age Estimation Using a Mixed Cumulative Probit and Its Application in KidStats

Kyra E. Stull, PhD, University of Nevada, Reno, Reno, NV 89557; Michael H. Price, PhD, Santa Fe Institute, Santa Fe, NM 87501; Louise K. Corron, PhD, University of Nevada, Reno, Reno, NV 89557; Elaine Y. Chu, MA, University of Nevada, Reno, Reno, NV 89557*

Learning Overview: After attending this presentation, attendees will be able to better estimate subadult age both methodologically, with the introduction of the mixed cumulative probit, and practically, with the introduction of a freely available graphical user interface, KidStats, and the R package, yada.

Impact on the Forensic Science Community: This presentation will impact the forensic science community as there is currently no method to build a multivariable, multi-indicator model using both continuous and ordinal data to estimate a continuous variable (i.e., age). Additionally, this presentation will introduce the virtual subadult anthropological database, which contains the largest number of de-identified computed tomography images of individuals ranging in age from birth to 20 years from around the world, housed at the University of Nevada, Reno, and available for research.

To date, positive identification of a subadult is largely dependent on age estimation. There has been a stagnation in the development and validation of age estimation methods, in part due to the lack of available subadult data in skeletal collections. In an effort to advance the forensic anthropological toolkit, as well as better understand growth and development of modern children, a virtual multiregional skeletal collection of approximately 2,500 subadults has been amassed and is housed at the University of Nevada, Reno. The North American sample comes from the University of New Mexico (UNM) Health Sciences Center, Office of the Medical Investigator (OMI) and the Office of the Chief Medical Examiner in Baltimore, MD. The worldwide samples come from Taiwan, France, the Netherlands, South Africa, Angola, Brazil, and Colombia. Supplementing skeletal collections with computed tomography images and full body X-rays has provided a unique opportunity to collect a large number of age indicators—diaphyseal dimensions, epiphyseal fusion, and dental formation—on these individuals to better understand the inherent variability of the age indicators and how it impacts age estimation.

A new algorithm, the mixed cumulative probit, was developed. It is an improved extension of the cumulative probit that handles both ordinal and continuous predictor variables to estimate a single continuous variable, deals with missing data, takes into account correlations among indicators, allows heteroscedasticity in component variables, and provides an estimate of age with 95% confidence intervals. The algorithm is available for forensic anthropologists through the user-friendly graphical user interface KidStats, and also as a separate R package, yada, so it can be employed by other researchers.

This presentation will demonstrate the performance of the mixed cumulative probit using 1,100 North American individuals aged between birth and 20 years, with different variable combinations. Cross-validation was built into model development and was applied to a hold-out sample, offering the most realistic assessment performance. The results indicate narrower confidence intervals than previously provided by any age estimation technique, ranging from 0.1 months to 5 years. The larger confidence intervals are not automatically linked to older individuals, but rather to individuals that had fewer variables input into the algorithm. Results demonstrate that combining all subadult indicators yields a more precise estimate. This study did not see any patterned bias in the age estimates, likely because the model is Bayesian in nature (and therefore has no regression to the mean effect) and because the parametric model used for inference is more flexible than previous models. Comparable performance in the training and test samples suggests there is no over- or under-fitting.

The freely available graphical user interface KidStats allows practitioners to employ this robust method in casework; it yields an age estimate given one of the predictor variables or up to the maximum number of 27 predictor variables. Additionally, the large number of individuals in model development enables the user to choose from a multitude of reference samples to use a global model, a population-specific model, or a model based on any combination of reference samples. If individuals are interested in using this algorithm to develop different estimation methods, it is also available in the yada R package.

The mixed cumulative probit will not only improve subadult age estimation but broaden methodological possibilities for adult age estimation as well. Furthermore, it has a number of applications both within and outside the field, such as paleodemography, economics, and human biology.

Subadult, Age Estimation, Cumulative Probit

*Presenting Author