

A49 Using Biological Data to Inform a DNA Sequencing Strategy

Carrie A. Brown, PhD*, Defense POW/MIA Accounting Agency Laboratory, 106 Peacekeeper Drive, Bldg 301, Offutt Air Force Base, NE 68113; and Jeffrey James Lynch, MSc, 501 Park Avenue, Apt 126, Omaha, NE 68105

After attending this presentation, attendees will be familiar with a successful approach using antemortem and postmortem biological data to prioritize DNA sequencing for a large commingled assemblage.

This presentation will impact the forensic science community by outlining a strategy to prioritize DNA sequencing of elements that fall >1 standard deviation from sample means for stature and age estimates, which can then be used to identify individuals prior to completion of all DNA analyses, reducing the missing person pool and the time families wait for notification.

Large commingled assemblages present unique challenges for identification. Unlike individual cases, the segregation of remains into discrete individuals can be difficult and time-consuming, even when DNA is incorporated. Using antemortem data and postmortem estimates from skeletal remains, this study outlines an effective and biologically informed strategy to prioritize DNA sequencing in a large commingled assemblage.

The assemblage is from the USS *Oklahoma* identification project and is comprised of nearly 13,000 skeletal elements; nearly 5,000 were sampled for DNA analyses. The remains are extensively commingled, with the first casket exhumed containing elements representing approximately 25% of the USS *Oklahoma* casualties unidentified following the incident ($n=394$). Analyses of the assemblage, including DNA sequencing, are concurrent with identifications being made.

All elements were inventoried and, as applicable, measured and epiphyseal fusion and pubic symphysis development recorded following McKern and Stewart.¹ The calculation of stature point estimates by element was automated for the entire assemblage using OsteoSort and linear regression equations from Trotter's Black and White male combined dataset in FORDISC® 3.² Only elements sampled for DNA analyses with a stature point estimate or age data were included. Antemortem data was drawn from military records of the unidentified USS *Oklahoma* casualties. Frequency distributions and descriptive statistics for antemortem data and postmortem estimates were produced using R, with the exception of epiphyseal fusion since mean ages per stage are not provided by McKern and Stewart.¹

The frequency distributions for antemortem stature and postmortem stature point estimates are normal and nearly identical — in inches (mean, standard deviation, n): antemortem (68.56, 2.31, 374); humerus (68.26, 1.86, 507); radius (68.44, 1.47, 325); ulna (68.35, 1.49, 328); femur (68.17, 1.97, 642); tibia (67.08, 1.88, 573); and fibula (68.45, 1.97, 360). Antemortem age and pubic symphysis postmortem point estimates are skewed to the right, but also similar — in years: antemortem (24.49, 6.40, 394); and pubic symphysis point estimate (24.31, 5.21, 405). For casualty individuals, 160 have stature and/or age that is >1 standard deviation from the antemortem means.

Because of the demonstrated similarity between antemortem data and postmortem estimates, prioritization was defined in four tiers based on standard deviations of the point estimates from the means: Level 1 — stature estimate >2 standard deviations ($n=89$); Level 2 — age estimate ≤ 20 years and ≥ 30 years ($n=247$); Level 3 — Stage 3 fusion (~ 21 years, $n=106$); Level 4 — stature estimate 1-2 standard deviations from the mean ($n=501$); and total on priority list, $n=943$. Due to the antemortem age distribution, epiphyses that fuse in the mid- to late-20s were not prioritized (e.g., medial clavicle). When elements were found in multiple priority levels, the element was given the higher of the ranked priorities. It should be noted that all cranial and dental samples were sequenced prior to other elements in order to aid with Minimum Number of Individuals (MNI) estimation and dental identifications.

This prioritization strategy has thus far resulted in segregation from the large commingled assemblage of 20 of the 160 individuals whose biological data are >1 standard deviation from antemortem age and/or stature means. These individuals will be identified prior to the completion of DNA sequencing for the entire project, which would not be feasible if relying solely on processing samples in the order they were received. The ability to make identifications can be important for family members, especially when a large period of time exists from incident to identification, as well as the forward progress of the project, as each identification further reduces the pool of missing persons. This also can be applied when sampling elements for DNA. One limitation to this approach is the lack of or incomplete antemortem data, which will affect the ability to develop informed priors.

Reference(s):

1. McKern, Thomas W., and T.D. Stewart. *Skeletal Age Changes in Young American Males: Analyzed from the Standpoint of Age Identification*. Natick, MA: Quartermaster Research and Development Command Technical Report EP-45, 1957.
2. <http://osteosort.net>.

Commingled Skeletal Remains, DNA Sequencing Strategy, Biological Data