A29 Examining Inter-Observer Reliability of Metric and Morphoscopic Characteristics of the Mandible

Jennifer F. Byrnes, PhD, University of Hawaii - West O'ahu, 91-1001 Farrington Highway, Kapolei, HI 96707; Michael W. Kenyhercz, PhD, University of Tennessee, 250 S Stadium Hall, Knoxville, TN 37996; Samantha C. Torres, BA, University of Hawai'i - West O'ahu, Div of Social Sciences, 91-1001 Farrington Highway, Kapolei, HI 96707; and Gregory E. Berg, PhD, DPAA Identification Laboratory, 310 Worchester Avenue, Joint Base Pearl Harbor-Hickam, HI 96853-5530*

After attending this presentation, attendees will understand the reliability of metric and morphoscopic characteristics of the mandible for sex and ancestry estimation.

This presentation will impact the forensic science community by providing inter-observer tests of reliability using observers of varied experience levels for sex and ancestry estimation methods focused on the mandible.

To date, there has been no large study of the reliability of metric and morphoscopic traits of the mandible. Berg provided the forensic anthropology community with a method for the estimation of sex and ancestry using 6 morphoscopic and 11 metric characteristics of the mandible from various world populations.[1,2] A combination of morphoscopic and metric variables were shown to discriminate sex and ancestry best. Further, intra-observer error was shown to be low.

An inter-observer agreement study of the mandibular variables was conducted by four researchers with varied experience using the standard descriptions of metrics, as well as new metric descriptions and scoring criteria described by Berg.[1,2] The standard metric variables included GNI, HML, TML, GOG, CDL, WRB, XRH, MLT, MAN, and two new measurements, mandibular body breadth at the M2/M3 junction (TML23) and dental arcade width at the third molar (XDA). The morphoscopic traits included Chin Shape (CS), Lower Border of the Mandible (LBM), Ascending Ramus Shape (ARS), Gonial Angle Flare (GAF), Mandibular Torus (MT), and the Posterior Ramus Edge Inversion (PREI). The sample data were derived from the William M. Bass Donated Skeletal Collection at the University of Tennessee, Knoxville. In total, this study examined 183 mandibles (white females=66; White males=96; Black females=2; and Black males=19) from known individuals and an additional 189 mandibles (White females=91; White males=88; Black females=4; Black males=6) were also examined.

To test the agreement among observers, the Intra-Class Coefficient (ICC) was used. The ICC quantified the proportion of variance that was ascribed to observations. For reliability, both complete agreement and consistency were evaluated using a two-way, random model ICC with a 95% confidence interval. Additionally, for the metric variables, the Technical Error of Measurement (TEM) was calculated. Both ICC and TEM were calculated with four observers for one test and three experienced observers in a second test in order to examine the effect of experience on agreement.

The ICC for the morphological variables for four observers ranged from 0.37 (PREI) to 0.75 (MT). The ICC values for the three experienced observers ranged from 0.43 (LBM) to 0.74 (MT). The majority of morphoscopic ICC values were between 0.67 and 0.74 (three observers). Each of the ICC values was significant at $p < 0.001$. The ICC values for the metric variables for four observers ranged from 0.73 (TML23) to 0.98 (CDL). Experienced observers ranged from 0.80 (TML23) to 0.99 (WRB). The majority of metric ICC values were between 0.91 and 0.99, and all were significant at $p < 0.001$. The TEM of the metric variables ranged from 1.16mm or 1.00% TEM (CDL) to 1.62mm or 10.53% TEM (TML). The TEM for MAN was 2.19 degrees or 1.74% TEM.

The results show that each of the variables has significant correlation among observers, though the metric variables were more accurately replicated than the morphological traits. Experience plays a role for scoring and measuring the mandible. The most consistent error found in the metric data was measuring the mandibular angle, where the least experienced observer had 11 instances of being exactly 10 degrees different, indicating measurement reading errors.

The morphoscopic and metric variables are reliable and valid, though recognition of PREI is the most problematic. The morphoscopic traits had moderate agreement between observers, which is related to observer experience. The most variable metric trait between observers was TML (10.53% TEM), which may relate to problems when teeth obstruct a superiorly derived measurement. Overall, the metric measurements among observers have high agreement. In using the methodology presented by Berg, it is suggested that practitioners become sufficiently familiar with the trait definitions and scoring attributes as well as the range of variability in mandibular morphology.

*Presenting Author

Reference(s):

1. Berg G.E. *Biological affinity and sex determination using morphometric and morphoscopic variables from the human mandible* (thesis). Knoxville, TN: Univ. of Tennessee, 2008.
2. Berg G.E. Biological Affinity and Sex from the Mandible Utilizing Multiple World Populations. In Berg G.E., Ta'ala SC, editors. *Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black and White*. Boca Raton: CRC Press, 2015:43-81.

Anthropology, Reliability, Error of Measurement

*Presenting Author