H69 Statistical Considerations of the Histomorphometric Test Protocol for Determination of Human Origin of Skeletal Remains

MariaTeresa A. Tersigni-Tarrant, PhD, JPAC-CIL, 310 Worchester Ave, Bldg 45, Joint Base Pearl Harbor-Hickam, HI 96853-5530; and John E. Byrd, PhD, 95-033 Hokuiwa St, #51, Mililani, HI 96853-5530*

After attending this presentation, the attendees will have a better understanding of a test protocol using osteon area as the basis for recognizing remains as likely human in origin. They will also receive an exposition of the philosophy underlying the statistics used in the test, and how that philosophy underpins interpretation of test results.

This presentation will impact the forensic community by explaining a new test protocol that is widely applicable and by clarifying the appropriate use of statistics in a forensic test.

Tersigni-Tarrant et al. proposed the use of osteon area as an effective measure for recognizing human bone versus other species.[1,2] It is widely known that humans have exceptionally large osteons on average, and that most species that share similar histological structure have smaller osteons. However, it is also clear that there is variation in individual osteon size within species, within individuals, and even within the same histological section. Any test that will utilize osteon size must have a proper statistical foundation. A test protocol has been developed that provides a significance test as the basis for interpretation of the osteon size. The test protocol requires one to: (1) section the case specimen and prepare a digital image of the section with scale; (2) import the image into the program Image J; (3) randomly select 30 osteons to measure; (4) measure the osteon areas and calculate the arithmetic mean; (5) test the null hypothesis that the remains are human by formal comparison (t-test) of the case specimen mean with the mean and standard error (with N=30) from a reference database consisting of 1,204 osteons from multiple long bones of five known individuals; and, (6) assess the strength of the evidence using Deborah Mayo's concept of severity.[3] A critical component of this protocol is the use of the randomly selected osteons as the basis for the test.

Human reference data has been analyzed and statistically compared to a close living relative, the chimpanzee (*Pan troglodytes*). This comparison is arguably a rigorous way to evaluate the power of the test to exclude non-human bone specimens. This test has also been applied to 37 specimens from 35 different known human individuals. Collectively, these efforts serve as validation of the test protocol. From the reference data (N=1,204 osteons), the mean osteon area was calculated at 37,298.5μm^2, standard deviation at 14.623.7, and noted that the distribution was skewed to the left (non-normal). The chimpanzee data (N=794 osteons from 17 individuals) showed a mean osteon area at 25,324.8μm^2, a standard deviation of 14,346.5, and a distribution skewed to the left. Bootstrapping was used to derive estimates of the means and standard deviations. The original reference data for each species was sampled with replacement 1,000 times, with N=30 each iteration. The mean was calculated for each iteration, which facilitated estimation of the population mean and offered a standard deviation of the estimated means. These values were: human mean at 37,365.3, standard deviation (of means) at 2,727.7; chimpanzee mean at 25,377.6, standard deviation (of means) at 2,692.1. Both sets of means exhibit a normal distribution. The closeness of the bootstrap means to the original estimated means was noted, as was the closeness of the bootstrap standard deviations of the means to the standard errors of the original estimates when basing it on a sample size N=30 (human standard error at 2,669.9; chimpanzee standard error at 2,619.3). Subsequent analyses employ bootstrap estimates since they are slightly more conservative than the original estimated means. The power of the test with the human mean as the null hypothesis and the chimpanzee mean as the alternate is approximately 0.90. The high power suggests that the test will have utility. Severity estimates were utilized to evaluate individual test results. Severity of a test, as defined by Mayo, is based on the probability that the test would reject the null hypothesis if in fact an alternative is true.[3] The more severe the test, the more likely one will reject the null when it is false. High severity justifies greater confidence in the null hypothesis when it is accepted, supporting the interpretation as "human" in many instances.

References:

1. Tersigni MT, Michael A, Byrd JE. Osteon area and circularity: a method for the assessment of human and non-human fragmentary remains. *Proceedings of the American Academy of Forensic Sciences*; 60th Anniversary Scientific Meeting. Washington, DC. 2008:14:375.
2. Tersigni-Tarrant MT, Byrd JE, Manabe J. Test of osteon circularity as a method of human/non-human identification. *Proceedings of the American Academy of Forensic Sciences*; 63rd Annual Scientific Meeting. Chicago, IL; 2011;17:381.
3. Mayo DG. Error and the gowth of experimental knowledge. Chicago (IL): University of Chicago Press, 1996.

Hypothesis Test, Statistical Power, Osteon Area

 ** Presenting Author*