Int J Legal Med (2011) 125:469–471
DOI 10.1007/s00414-010-0476-y

LETTER TO THE EDITOR

Comments on “Estimation of stature from cranial sutures in a South Indian male population” by P. P. J. Rao et al.

James Christopher Reed · Bridget F. B. Algee-Hewitt

Received: 29 April 2010 / Accepted: 26 May 2010 / Published online: 11 June 2010

Dear Sir,
We are writing to comment on the article “Estimation of stature from cranial sutures in a South Indian male population” printed last year in Volume 123 of this journal [1]. We note that this article is but one in a series of similar publications that attempt to correlate various craniometric variables with the living stature of an individual [2, 3]. These multiple publications not only attest to a growing interest in this line of research but also contribute to the body of acceptable methodological literature for use in medico-legal contexts. As the *International Journal of Legal Medicine* [4] aims “to improve the scientific armamentarium used in the elucidation of crime and related forensic applications at a high level of evidential proof,” we feel that it is important to formally address our concerns with the logic of the study design and the authors' approach to data collection, analysis, and interpretation of their findings. The authors rightly state that, if measurements taken from the cranium can be used to estimate living height, it would be useful in estimating stature for an unknown individual in forensic cases when only the cranium is present. We argue here, however, that the results do not support the author’s claim that this method “could prove useful” to making an identification in medico-legal contexts. After a careful review of this publication, we conclude that both the method and theory are not sound and that this focus of inquiry in stature estimation should not be pursued.

We first take issue with their procedures for data collection and their choice of cranial variables. Measurements of suture lengths taken with a string do not conform to any convention of osteological data collection and, although the authors provide the manufacturing specifications, “string” is not a standard measuring device. We question why the authors chose not to follow the well-known, documented, and tested protocol available in the seminal manuals for anthropological/craniometric research [5–8]. The parietal chord or the interlandmark distance between left and right sphenion (SPHL-SPHR), for example, would have served as reasonable surrogates for sagittal and coronal suture lengths respectively. As these traditional interlandmark distances are often among the set of measurements taken during craniometric data collection, their use in developing this method would have increased its applicability to previously obtained data and, through their familiarity, reduced potential error in measurement collection. Furthermore, intra- and interobserver error must be assessed to test the consistency of the measurements collected and the reliability of the calculated stature estimates for any methods that are to be accepted as tools for stature estimation [9]. We also argue that the authors' belief that in using a single observer they are reducing measurement error (272) only serves to hide such error by obfuscating the issue of repeatability within and among observers and sample specimens. We wonder if these atypical choices in data collection methods reflect a larger unfamiliarity with skeletometric data and, in turn, a lack of understanding of their statistical properties and conventional treatment. For example, the relationship figures provided show the measurements in centimeters as discrete variables and the data points in the graphs do not indicate any values other than whole numbers. This presentation is clearly problematic; stature is a continuous

J. C. Reed (✉) · B. F. B. Algee-Hewitt
Biomedical Sciences Department, Grand Valley State University,
1 Campus Dr.,
Allendale, MI 49401, USA
e-mail: reedjam@gvsu.edu

variable, not categorical. Of even greater concern, however, is the fact that the data, as graphed, suggest little, if any, relationship between suture length and cadaveric stature. Our conclusions based on visual inspection of the graphs are corroborated by the very low correlation coefficients reported: 0.363 for the coronal suture and 0.090 for the sagittal suture.

The authors also report *P*-values as measures of significance for both tests of relationships and, from these, conclude that the correlation between adult height and the sagittal suture is not significant ($P=0.48$) but that the correlation of height with the coronal suture is ($P=0.001$). We take issue with this tacit acceptance of the "significant" *P*-value as conclusive evidence of the existence of a real or biologically meaningful relationship between coronal suture length and measured cadaveric stature. We, like many others concerned with biostatistical data [10, 11], are cautious in equating "statistical significance" with significance in its larger contextual usage as "meaningful" or synonymous usage as "important" [12]. We believe in this case that the test of "significance" is meaningless because the correlation coefficient for the coronal suture is too low to be of practical use. Moreover, as we can easily prove mathematically that it takes very little change in a sample structure to produce a significant *P*-value and that significance tests are sensitive to research design and sample size, we ask: What if the correlation coefficient of the sagittal suture (0.090) had also been found to be "statistically significant"? Should we blindly accept a similar linear regression equation using the sagittal suture despite all logical evidence suggesting otherwise? We do acknowledge that the authors do provide sample statistics, including the 95% confidence interval, but we are concerned with how these values relate to each other given the very small confidence interval reported. Therefore, as an improvement in method, we strongly advise that estimates of sample statistics, effect sizes, and their confidence intervals be included either alongside or in place of *P*-values. Like Fisher himself, we believe strongly in the merit of replication by extension research and recommend the comparison of population estimates using overlapping confidence intervals.

Statistical analyses are only as good as the framework upon which they are built. The biology of the human cranium belies any positive correlative result one would get in the comparison of cranial measurements and stature. The weak and insignificant results of the statistical analysis in Rao et al. [1] should not be surprising. The application of any such correlative analysis should only be attempted if it is strongly supported by the understood biological and morphological character of the human skeletal system. A good understanding of the development, growth, and senescence of the human skeleton has produced a number of excellent techniques for [13, 14] and on-point caveats about [15–17] the estimation of stature from the skeleton. Long bones, specifically the tibia and femur, have been found to have the most predictive power of stature [17]. Given the importance of these bones in their contribution to stature, it is logical that these bones are highly correlated with stature. If the measurements of the cranium, or features therein (suture length), are to have some predicative value for stature, the nature of the growth and development of the cranium must have some relationship with that of the other bones that contribute to stature. The growth of the human cranium, however, has a different ontological foundation and a much different trajectory from that of the postcranium.

The only interlandmark distance measured on the cranium that has a direct relationship with stature is basion–bregma height, which is generally only used when the whole skeleton is available [18]. Without the entire skeleton present for the application of the Fully method for stature estimation, the use of basion–bregma height or any other measurement of the cranium, including suture length, are highly questionable because the growth trajectory of the cranium is significantly different than the postcranium. Maximum cranial size is reached at an earlier age than maximum postcranial height (what we would consider adult stature) and growth trajectories are also visibly different within the postcranium itself, with the long bones completing their longitudinal growth before the rest of the postcranial skeleton [19]. The difference in growth trajectories is logical as the growth and development of the cranium and postcranium and subsequently that of the cranial sutures are known to be guided by different developmental controls: being derived from different embryonic layers and following very different patterns and trajectories of growth. While genetic and epigenetic components affect the growth and development of both the cranium and the postcranium, the epigenetic factors dominate and direct most of the size and shape changes of the cranium [20–23]. Even at the earliest stages of measurable cranial growth, and following Moss' Functional Matrix Hypothesis, it is clear that the brain and other soft tissue structures directly affect the trajectory of bone growth [19, 20, 22]. Furthermore, other hard tissues and muscle mechanics can also affect the growth of the cranium: the teeth and corresponding alveolus and the muscles of mastication can notably affect the shape of the skull [19, 20]. We posit that if the assumption behind testing a correlative value between cranial measures and overall stature is that the genetics of growth and development are affecting the cranium and postcranium similarly and in equal measure, then, in light of our knowledge of human ontogeny, there lacks any sound biological basis for conducting such studies.

Like growth, normal bone degeneration with senescence also affects the stature of an individual, introducing two types of bias that would confound any correlation between measurements of cranial features, including the growth of the cranium, and living height. Self-reported living stature recorded on legal documents is often based on the reporter's maximum height in young adult years [9, 15], and so documented height may not be the actual height at death. Error, therefore, can be introduced in two ways: (1) assuming that the correlative relationship with reported living stature is accurate, a larger error range must be assumed for older individuals and (2) any correlation calculated must consider the age of the individuals of the sample. As a result, we must develop age-based population specific equations: if the sample consists of younger individuals, the application of any correlative methods of stature estimation would not work well for older individuals, with the converse being also true. We argue that, as our stature gradually diminishes with age, the likelihood of producing accurate results using any correlation of craniometrics with living stature will diminish as well. From growth through senescence, age clearly confounds efforts to estimate living stature from the skeleton, and this is particularly important when there is an attempt to correlate cranial measurements with living stature.

We conclude our response with a word of caution to both authors and readers alike. Given the statistical weaknesses and the biological contradictions plaguing the research design of this study and the results of the proposed method, we strongly discourage the use of this or similar methods that rely upon a correlation between stature and any craniometric values without extensive multi-group studies and repeatability tests. We disagree with Rao and colleagues that their presented method "could prove useful" for the estimation of stature in medico-legal investigations or any situation when only the cranium is available for forensic anthropological analysis. Moreover, at the present time, when expectations for expert testimony, demands on evidentiary relevance and definitions of acceptable "science", are changing in response to, for example, developing protocol by international tribunals and rulings using Daubert standards, we caution that poor methodologies disseminated into the realms outside of academia can have real-world impact and ramifications. We appeal to researchers and practitioners to be mindful of the quality of science produced and accepted.

References

1. Rao JPP, Sowmya J, Yoganarasimha K, Menezes RG, Kanchan T, Aswinidutt R (2009) Estimation of stature from cranial sutures in a South Indian male population. Int J Leg Med 123:271–276
2. Ryan I, Bidmos MA (2007) Skeletal height reconstruction from measurements of the skull in indigenous South Africans. Forensic Sci Int 167:16–21
3. Chiba M, Terazawa K (1998) Estimation of stature from somatometry of skull. Forensic Sci Int 97:87–92
4. International Journal of Legal Medicine (2010) Aims and scope. Available from: http://www.springer.com/medicine/forensic/journal/414. Accessed April 16, 2010
5. Martin R (1956) Lehrbuch der anthropolgie. Revised third edition, vol 4. Fischer, Stuttgart
6. Howells WW (1973) Cranial variation in man. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 67, Cambridge
7. Martin R (1956) Lehrbuch der anthropolgie. Revised third edition, vol 3. Fischer, Stuttgart
8. Moore-Jansen PM, Ousley SD, Jantz RL (1994) Data collection procedures for forensics skeletal material: report of investigations, no. 48. The University of Tennessee Department of Anthropology, Knoxville
9. Giles E, Hutchinson DL (1991) Stature- and age-related bias in self-reported stature. J Forensic Sci 36:765–780
10. Gelman A, Stern H (2006) The difference between 'significant' and 'not significant' is not itself statistically significant. Am Stat 60:328–331
11. Hubbard R, Lindsey RM (2008) Why *P* values are not a useful measure of evidence in statistical significance testing. Theory Psychol 18:69–88
12. Oxford English Dictionary Oxford [electronic resource] (2000–) Oxford University Press. Accessed April 16, 2010
13. Trotter M, Gleser GC (1952) Estimation of stature from long bones of American Whites and Negroes. Am J Phys Anthropol 10:463–514
14. Trotter M (1970) Estimation of stature from intact long limb bones. In: Stewart TD (ed) Personal identification in mass disasters. Smithsonian Institution, Washington, pp 71–83
15. Ousley S (1995) Should we estimate biological or forensic stature? J Forensic Sci 40:768–773
16. Meadows Jantz L, Jantz RL (1999) Secular change in long bone length and proportion in the United States. Am J Phys Anthropol 110:57–67
17. Wilson RJ, Herrmann NP, Meadows Jantz L (2010) Evaluation of stature estimation from the database for forensic anthropology. J Forensic Sci 55:684–689
18. Raxter MH, Auerbach BM, Ruff CB (2006) Revision of the Fully technique for estimating statures. Am J Phys Anthropol 130:374–384
19. Humphrey LT (1998) Growth patterns in the modern human skeleton. Am J Phys Anthropol 105:57–72
20. Sardi ML, Ramírez Rozzi FV (2005) A cross sectional study of human craniofacial growth. Ann Hum Biol 32:390–396
21. Moss ML (1997) The functional matrix hypothesis revisited. 1. The role of mechanotransduction. Am J Orthod Dentofacial Orthop 112:8–11
22. Moss ML (ed) (1973) A functional cranial analysis of primate craniofacial growth. Fourth International Congress of Primatologists, Karger, Basel
23. Moss ML (1997) The functional matrix hypothesis revisited. 4. The epigenetic antithesis and the resolving synthesis. Am J Orthod Dentofac 112:410–417