Evaluating the accuracy of cranial indices in ancestry estimation among South African groups

Leandi Liebenberg, MSc[1*], Kyra E. Stull, PhD[2], Ericka N. L'Abbé, PhD, D-ABFA[1], Deona Botha, MSc[1]

[1]Department of Anatomy, University of Pretoria, Private Bag x323, 0007, Arcadia, South Africa

[2]Department of Anthropology, Idaho State University, Pocatello, ID 83204

* Corresponding author.

Leandi Liebenberg

Department of Anatomy, University of Pretoria

Private Bag x323, Arcadia 0007 South Africa

Tel: +27-82 072 0057/12-319-2438

E-mail: leandil@hotmail.co.za

ABSTRACT

Historically, population differences were quantified using cranial indices. Even though the application of indices is associated with numerous statistical and methodological problems, the use of cranial indices to estimate ancestry persists as demonstrated by its inclusion in several recent papers and conference presentations. The purpose of this study was to classify 207 South African crania and compare the results of five standard cranial indices to linear discriminant analysis (LDA). New sectioning points were created to contend with low classification accuracies (40% - 79%) and possible secular trends. Although the accuracies of the new sectioning points increased (66% - 87%), the accuracies associated with the stepwise LDA were higher (84%) and could classify the crania into one of the three South African groups. The results of the study demonstrate that indices cannot compete with multivariate techniques and should not be used in forensic anthropological analyses for ancestry estimation.

Ancestry estimation is an essential component of the biological profile, particularly in heterogeneous countries, such as South Africa and the United States. Various aspects of the skeleton can be used to estimate ancestry; however, both metric and morphoscopic features of the cranium have been identified as the most accurate, with mid-facial features often producing the best results (1-8). While emerging research around the world challenges the historical approaches in biological anthropology, the continual use of cranial indices in South Africa perpetuates typological thinking in the assessment of ancestry and understanding of human variation.

Indices are utilized to explore shape differences independent of size and are frequently used to quantify differences between morphological features (9). Traditionally, skulls were classified into broad categories based on shape variations derived from indices; for instance crania for which the length was greater than the breadth were classified as long and termed dolichocephalic; crania of intermediate length were called mesocephalic; and crania with a greater breadth than length were known as brachycephalic (10). Shape differences among crania were associated with specific ancestral groups such that crania of Africans were predominantly found to be dolichocephalic while the crania of Europeans were found to be brachycephalic (11-14). In past traditional systems of classification, the idea of ancestry as discrete, exclusive "types" inspired notions of the existence of ideal categories. In turn, cranial indices became a tool for presumptions of racial superiority and inferiority (15,16). Historically in South Africa, morphological differences among populations were quantified with cranial indices derived from a suite of cranial dimensions as employed by DeVilliers (11).

Perhaps the greatest appeal of indices is the simplicity of its application. However, the apparent simplicity conceals numerous statistical complications (17,18). Atchley et al. (19) argue that the generation of ratios and indices from continuous variables changes the structure and underlying distribution of data. Specifically, the standard deviations and coefficients of variation are affected (19). Furthermore, the reliance on single number representations of complex biological data risks producing misleading results (20). By reducing multifaceted features, such as cranial morphology, to a single index value, pertinent

information may be masked or, more likely, lost (21,22). For instance, the score calculated with the cranial index can result from a large cranial length, large cranial breadth, or a combination of both. The variation associated with each measurement is lost when translated into an index.

Besides the statistical issues associated with indices, several methodological problems also exist. The dependence on univariate measures for classification ignores the literature that states the inclusion of more variables rather than less variables maximizes differences when categorical response variables exhibit some degree of overlap, such as sex and ancestry (23,24). Second, the use of indices fails to satisfy *Daubert* criteria. Not only does *Daubert* demand that scientists ensure their methodology is accurate, precise, testable and peer-reviewed, it also requires techniques to have associated error rates to be admissible in a court of law (25-27). Third, cranial indices were generally derived from antiquated populations (e.g. 11), and their continued inclusion in analyses ignores the surplus of data revealing secular increases in cranial dimensions and body size (22, 28, 29). Admittedly secular change may potentially influence all variables and the studies that utilize them – not just the indices. However, secular changes may have a larger effect on classification when only looking at two variables (i.e. indices) as opposed to the cranium as a whole (i.e. multivariate) (29).

Despite the recognized problems, numerous – and recent – peer-reviewed publications and conference presentations continue to use cranial indices for forensic purposes, citing "clear racial trends" and emphasizing the importance of indices in evaluating both ancestry and sex differences (30-38). Because it is unknown what methods practitioners depend on during forensic analyses, a recent survey was conducted in South Africa. Among the 30 forensic practitioners (e.g. anthropologists and pathologists) in South Africa that responded to the survey, 30% stated they include cranial indices in the suite of methodologies used to analyze skeletal remains. While 30% does not make up the majority, the number is surprisingly large considering the inherent problems associated with the use of indices and the availability of more robust statistical methods. Furthermore, the participants that use indices ranged from having an honours degree to a doctoral degree and self-identify anywhere as a novice to an expert. Such results demonstrate that the use of indices is prevalent in the classroom and laboratory setting and is not limited to a specific generation.

While the practical application of cranial indices were largely abolished in the 19th and 20th centuries with vocalizations against a typological approach, and despite the fact that cranial indices are not recommended in the Scientific Working Group for Forensic

Anthropology (SWGANTH) best practices, the use of cranial indices persists in South Africa and other countries around the world (e.g. 30-38). Although it was believed that most anthropologists recognize multivariate approaches outperform indices, the continued use is evident without a reference to prove otherwise. Thus, there is a need to critically assess the continued use of cranial indices in forensic anthropology. The purpose of this paper was to test the accuracy of five standard cranial indices, namely the cranial index (CI), upper facial index (UI), orbital index (OI), nasal index (NI) and gnathic index (GI), and evaluate their performance in differentiating South African groups compared to linear discriminant analysis.

Materials and Methods

The sample consisted of 79 crania of black and 78 crania of white modern South Africans from the Pretoria Bone Collection (PBC), housed at the University of Pretoria. Within each modern South African group, the sexes were equally distributed. The ages for the modern sample ranged from 21 to 90 years, with mean ages of 49 years and 69 years for black and white South Africans, respectively. In addition, 50 historic Khoesan crania, of estimated sex and unknown age, were used from the Rudolf Pöch Collection at the University of Vienna in Austria. The name Khoesan merges two indigenous groups from sub-Saharan Africa, namely the Khoikhoi and the San. The combination of these morphologically and culturally distinct groups is primarily based on their similar click-languages (39,40). Dr Pöch and his associates illegally exhumed suspected Khoesan remains from South Africa between 1907 and 1909 and transported them to Austria in order to conduct typological research on race (41). Due to the paucity of historical records and illegitimate manner in which the remains were collected, the population affinity of individuals in the collection has been questioned (42). All crania utilised in the current study were of adult individuals and were without extensive alveolar resorption, post-mortem damage, pathology or healed fractures.

Ten standard measurements, which included the maximum cranial breadth (XCB), maximum cranial length (GOL), bizygomatic breadth (ZYB), upper facial height (NPH), orbital height (OBH), orbital breadth (OBB), nasal height (NLH), nasal breadth (NLB), basion-prosthion length (BPL) and cranial base height (BBH) were taken to the nearest millimeter with standard spreading and sliding calipers. The ten measurements were used to calculate five indices pertaining to cranial dimensions, namely the CI, UI, NI, OI and GI (Table 1). The sectioning points that were used to classify the skull as either South African

Table 1. The formulae used to calculate the indices.

	Formula
Cranial index (CI)	XCB/GOL*100
Upper facial index (UI)	NPH/ZYB*100
Orbital index (OI)	OBH/OBB*100
Nasal index (NI)	NLB/NLH*100
Gnathic index (GI)	BPL/BBH*100

black, South African white or Khoesan were obtained following De Villiers (11) and Steyn (43) (Table 2).

Table 2. The original sectioning points for the cranial dimensions of South African blacks, whites and Khoesan. Refer to Table 1 for abbreviations.

White	Khoesan
>80	75-79.9
>55	<44.9
>85	<75.9
<46.9	>51
<97.9	98-102.9

Percent correct was utilized to gauge the accuracy of the indices and revealed how often the unknown crania were classified within the correct group. A classification accuracy of 75% or greater was considered acceptable as it offers practical applicability. A multitude of statistical tests were applied using R (version 2.15.3) (44) and included an analysis of variance (ANOVA), Tukey's range test, Fisher natural breaks classification, and linear discriminant analysis (LDA). An ANOVA was employed to simultaneously compare the group means per variable without having to adjust alpha for the effect of multiple tests (45).

Tukey's range test was used in conjunction with ANOVA in order to determine which of the means of the three groups were significantly different from each other (46). Simply, where the ANOVA shows the significance of each model when all three populations are compared at once, Tukey's range test demonstrates significant differences within the model.

Where correct classifications were low, new sectioning points were created manually and with Fisher's natural break classification based on the modern samples and the ANOVA results. Manual sectioning points were created by taking the mean of two means, thus a point between two groups. The Fisher natural breaks classification method identifies the optimal arrangement of values into different groups by making use of natural breaks in the data. The method decreases the variance within groups and maximizes the variance between groups while simultaneously preserving the clustering of the data (47, 48). Break points appear where an actual break is in the data rather than equal intervals as is with manual sectioning points. The classIntervals function was used, specifying the Fisher method, from the classInt package in R (49).

Discriminant analysis was employed to discern population differences based on the original ten measurements. Discriminant analysis identifies relationships between qualitative criteria and quantitative predictor variables and ultimately defines boundaries used to distinguish between different groups (50). A linear discriminant analysis (LDA) is used if the requirements of a sufficient sample size, multivariate normal data and equal variance-covariance matrices are met (23,24,51). While these requirements should be satisfied, LDA is capable of performing sufficiently even when the requirements are violated (23). The ten measurements were analyzed with the inclusion of all variables as well as a backward stepwise variable selection to determine which combination of variables achieved the highest overall correctness rate. Stepwise variable selection is a technique that identifies the measurements that best separates the groups and removes the superfluous variables that do not contribute to increased classification (52). All LDA models were conducted with equal, rather than proportional, priors and leave-one-out cross-validation (LOOCV). With LOOCV, one individual is removed from the sample and a discriminant function is calculated using all of the remaining individuals; the function is then used to classify the removed individual into the most similar group (23). A direct comparison was made between the accuracies of the sectioning points and the accuracies of the LDA to determine which approach yields the highest number of correct classifications. The packages MASS and klaR were utilized in R to employ the all measurement and stepwise models, respectively (53,54).

Results

  The percent correct was calculated for each population based on the De Villiers indices (Table 3); the NI performed the best overall (79%), while the CI and UI performed

Table 3. Percent correct for each ancestral group using original sectioning points. Refer to Table 1 for abbreviations.

	Correct Classification %				
	CI	*UI*	*OI*	*NI*	*GI*
Black	80	85	70	89	19
White	14	17	67	59	93
Khoesan	11	10	26	96	33
Overall	40	41	58	79	50

the worst overall (40% and 41%, respectively). With both CI and UI, a large number of individuals misclassified as black and as a consequence, resulted in low accuracy rates for the South African whites (14% and 17% for CI and UI respectively) and Khoesan (11% and 10%, for CI and UI respectively). The GI misclassified a large proportion of the crania as white and thus South African whites presented with the highest correct classifications (93%). In contrast, South African blacks and Khoesan presented with the lowest number of correct classifications for the GI (19% and 33%, respectively). Because De Villiers noted similar values among groups for both the OI and NI, assignment was into two rather than three groups. Thus, South African blacks and whites were pooled for OI and South African blacks and Khoesan were pooled for NI. Although NI had the overall best classification, the correct classification for white South Africans was only 59%, which was in contrast to the high correct classification of black South Africans (89%) and Khoesan (96%).

  The ANOVA resulted in statistically significant differences ($p < 0.05$) for all five indices. When further explored with Tukey's range test, results revealed statistically significant differences between two of the three groups for all five indices. The lack of significant differences ($p > 0.05$) between the group means for the UI, CI, GI, and NI

suggested the Khoesan and South African blacks should be pooled. The OI was the only index to demonstrate significant differences between the Khoesan and South African blacks ($p < 0.001$), and Khoesan and South African whites ($p < 0.001$); suggesting the black and white groups should be pooled for the OI. When the original measurements were visualized with density plots, the overlap between the populations is evident (Figure 1).
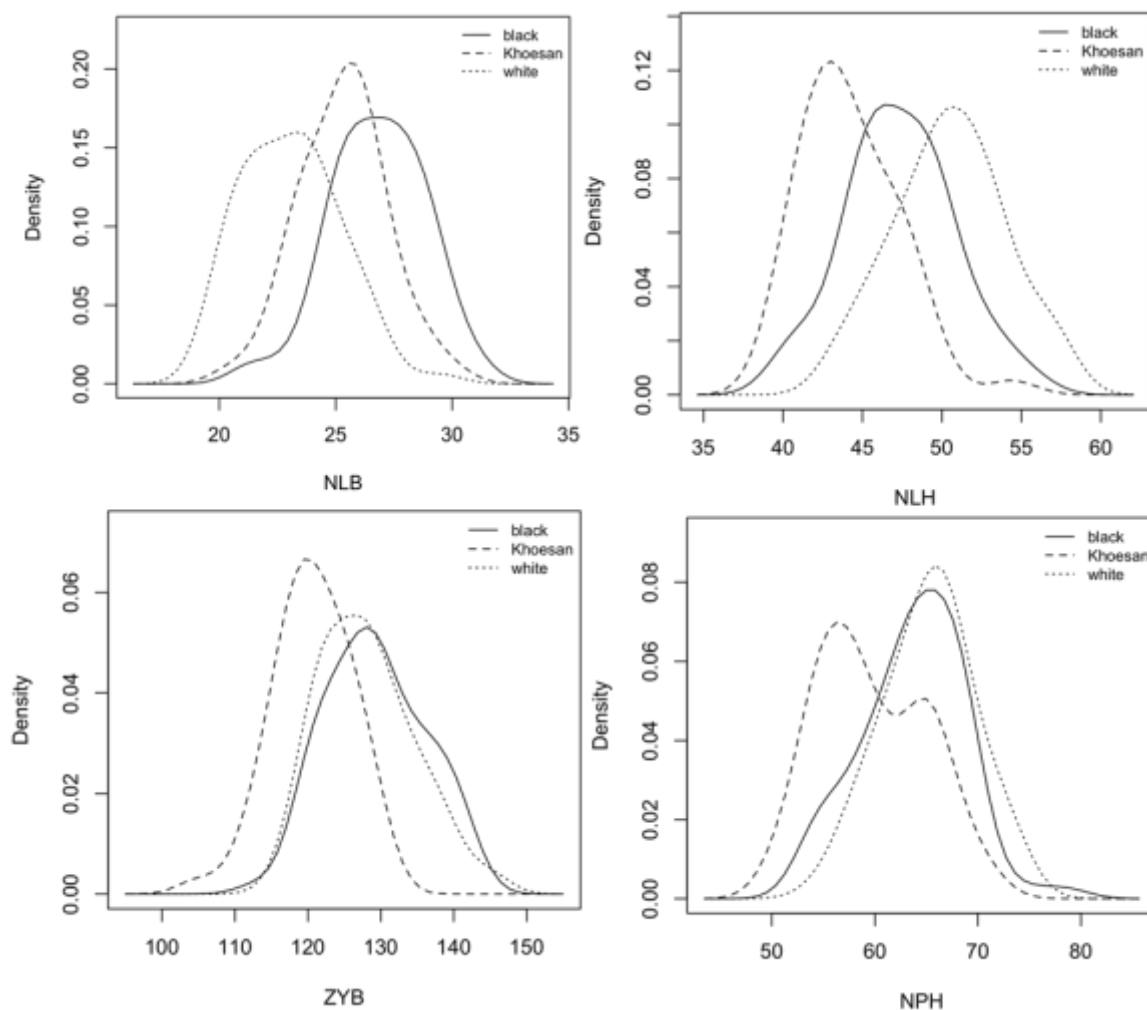
Figure 1. Density plots for each of the variables illustrating the overlap between the population groups. Keeping best (NI, top graphs) and worst (UI, lower graphs) variables.

Based on low percent correct results, new sectioning points were created for all indices except UI. Although the ANOVA noted significant differences between the groups for UI, minimal differences between group means made the creation of sectioning points impractical. For the remaining four indices, sectioning points were created manually and with Fisher natural breaks; percent correct was subsequently recalculated (Table 4). While the

Table 4. New sectioning points created manually and using the Fisher natural breaks method. Refer to Table 1 for abbreviations.

	Manual	Fisher
CI	BK > 74 > W	BK > 74 > W
OI	K > 83 > BW	K > 86 > BW
NI	W > 52 > BK	W > 52 > BK
GI	W > 95 > BK	W > 96 > BK

manual and Fisher natural breaks method resulted with the same sectioning points for CI and NI, the two approaches produced different sectioning points for OI and GI. It is worth emphasizing that all of the newly derived indices classify individuals into two groups rather than three groups because of a lack of statistical differences among the three groups (ANOVA results). The new sectioning points for the indices resulted in increased accuracies (66% - 87%) (Table 5). However, only the NI was able to produce a percent correct of 75% for all three South African groups.

Table 5. Percent correct for new sectioning points. Abbreviations: Mn= Manual; F= Fisher. Refer to Table 1 for abbreviations.

	Correct Classification %					
	CI (Mn, F)	OI (Mn)	OI (F)	NI (Mn, F)	GI (Mn)	GI (F)
Black	73	77	61	85	73	65
White	69	58	88	90	78	83
Khoesan	78	79	59	90	88	84
Overall	72	73	66	87	78	76

The stepwise selected LDA model achieved better results than the LDA model utilizing all of the variables (Table 6). The three-way LDA resulted in an overall correctness rate of 84% with seven stepwise selected variables, namely XCB, GOL, OBH, OBB, NLB,

Table 6. Classification accuracies based on the LDA models.

	Stepwise	All measurements
Black	84%	81%
White	88%	87%
Khoesan	80%	80%
Overall	84%	83%

NLH, BPL and BBH. The NLB and NLH contributed the most to the first linear discriminant variable while OBH contributed the most to the second linear discriminant variable. Evaluation of population-specific classification accuracies revealed that white and black South Africans achieved the highest correct classifications (88% and 84% respectively), and the Khoesan achieved a slightly lower correct classification (80%).

Discussion and Conclusions

The results indicate that multivariate techniques produce higher accuracies than cranial indices in ancestry estimation, which can be attributed to the limitations of two-dimensions and numerous statistical and methodological problems. In sum, the high misclassifications result from indices not capturing the wide range of human variation present within and among the groups. Although the application of cranial indices was once a common technique in the 19th and 20th centuries, the classification accuracies obtained in this study using the DeVilliers (11) and Steyn (43) indices were poor.

Univariate comparisons indicated that the crania of Khoesan are not significantly different from the crania of black South Africans; therefore, distinction between Khoesan and black South Africans is not possible when using a univariate approach. While the historic Khoesan exhibit both genetic and morphological characteristics that distinguish them from black South Africans, there is evidence of gene flow between the historical Bantu-speaking

and indigenous groups early in history (55). Currently, language is one of the major factors of self-identification among black South Africans and nine of the eleven official languages are Bantu-speaking languages (56). Herbert (57) suggests that the past relationship between the Khoesan and Bantu-speaking groups were characterized by frequent and intimate interactions, which included trade and intermarriage and occurred over several centuries. The heterogamy resulted in composite groups existing all over southern Africa (57). Population history indicates that the two groups have shared histories and genetics and the similar genotypes may result in phenotypes that one index cannot differentiate. In contrast to the indices, the LDA could differentiate the two groups and yield high accuracies.

Another possible influential factor in the accuracies that should be considered is the age distribution of the sample. A large disparity between the age distributions may skew the results as previous studies have noted remodeling and increased dimensions in the skeletal structure of middle-aged to old individuals (58,59). While the age distribution of the Khoesan sample used in this study is unknown, the age distribution of the South African blacks and whites are similar and inclusive of individuals throughout the entire adult age spectrum. Furthermore, the age distributions are analogous to the age range used by DeVilliers (11) (17 to 95 years). Albert et al. (58) reported only slight increases (between 1.1mm and 1.6mm) in the craniofacial dimensions of older individuals, with facial height exhibiting the greatest change due to antemortem tooth loss. Thus, while the effects of age-related craniofacial remodeling should be acknowledged, age is not expected to be a major contributing factor to the rate of misclassification observed with the indices.

Although a substantial improvement in correct classifications was noted with the new sectioning points, the improved results are likely due to the pooling of at least two of the three groups for all the indices. As expected, the highest classification accuracies were obtained with the NI and GI as these indices focus on mid-facial features of the cranium. The majority of the measurements chosen for the stepwise model correlated with the indices that produced the highest percent correct, namely the OI, NI and GI. Although the NI produced an accuracy (87%) that appears to be slightly higher than the LDA accuracy (84%), the NI only classified the crania into two groups while the LDA classified the crania into three groups. The highest correct classifications were achieved by white and black South Africans, which suggest the stated groups have the largest between-group variation.

The current study emphasizes the need to exclude cranial indices from forensic anthropological analyses, especially the historically derived indices of DeVilliers (11). As expected, the multivariate approach using the original measurements yielded higher accuracies that differentiated the three South African groups. The findings of this study reiterate the statement by Albrecht (17), "In practice, it may be more troublesome to validate the use of ratios for a particular problem than to proceed directly to alternative, more elaborate, statistical procedures" (p.70).

References

1. Patriquin ML, Steyn M, Loth SR. Metric assessment of race from the pelvis in South Africans. Forensic Sci Int 2002; 127:104-13.

2. Gill GW. Craniofacial Criteria in the Skeletal Attribution of Race. In: Reichs KJ, editor. Forensic Osteology: Advances in the identification of Human Remains. Charles C. Thomas, Springfield, IL; 1998. p.293-317.

3. İşcan MY, Steyn M. Craniometric determination of population affinity in South Africans. Int J Legal Med 1999; 112:91-7.

4. Hefner JT. Cranial Nonmetric Variation and Estimating Ancestry. J Forensic Sci 2009; 54:985-95.

5. L'Abbé EN, van Rooyen C, Nawrocki SP, Becker PJ. An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Sci Int 2011; 209.

6. McDowell JL, L'Abbé EN, Kenyhercz MW. Nasal aperture shape evaluation between black and white South Africans. Forensic Sci Int 2012; 222:397.e1-397.e6.

7. L'Abbé E, Kenyhercz M, Stull K, Ousley S. 2013. Craniometric Assessment of Modern 20th century black, white and coloured South Africans. Proceedings of the 65th annual meeting of the Academy of Forensic Sciences. 2013. Washington, D.C.

8. Hefner JT, Spradley MK, Anderson B. Ancestry assessment using random forest modeling. J Forensic Sci 2014; 59:1-7.

9. Mosimann JE. Size allometry: Size and shape variables with characterizations of the lognormal and gamma distributions. J Am Stat Assoc 1970; 56:930-45.

10. Boas F. The cephalic index. Am Anthropol 1899; 1:448-61.

11. De Villiers H. The skull of the South African Negro: A biometrical and morphological study. Witwatersrand University Press, Johannesburg; 1968b.

12. Crawfurd J. On the classification of the races of man according to the form of the skull. Transactions of the Ethnological Society of London 1868; 6:127-34.

13. Aravanan KP. Notable Negroid elements in Dravidian India. JTS 1980; 17:20-45.

14. Odukuma EI, Akpuaka FC, Igbigbi PS, Otuaga PO, Ejebe D. Patterns of cephalic indexes in three West African populations. Afr J Biotechnol 2010; 9:1658-62.

15. Gravlee CC. How race becomes biology: Embodiment of social inequality. Am J Phys Anthropol 2009;139:47-57.

16. Brace CL. Region does not mean "race": Reality versus convention in Forensic Anthropology. J Forensic Sci 1995; 40:171-5.

17. Albrecht GH. Some comments on the use of ratios. Syst Zool 1978; 27:66-71.

18. Jungers WL, Falsetti AB, Wall CE. Shape, relative size, and size-adjustments in morphometrics. Yearb Phys Anthropol 1995; 38:137-61.

19. Atchley WR, Gaskins CT, Anderson D. Statistical properties of ratios. I. Empirical results. Syst Zool 1976; 25:137-48.

20. Green R, Chapman PM. Editorial: The problem with indices. Mar Pollut Bull 2011; 62:1377-80.

21. Holliday TW, Ruff CB. Relative variation in human proximal and distal limb segment lengths. Am J Phys Anthropol 2001; 116:26-33.

22. Jantz RL. Cranial change in Americans: 1850-1975. J Forensic Sci 2001; 46:784-7.

23. Ousley SD, Jantz RL. FORDISC 3 and Statistical Methods for Estimating Sex and Ancestry. In: Dirkmaat DC, editor. A companion to Forensic Anthropology. Blackwell Publishing LTD; 2012. p.311-29.

24. Tabachnick BG, Fiddel LS. Using multivariate statistics. 4th ed. Allyn and Bacon; 2001.

25. Hummel TJ, Sligo JR. Empirical comparison of univariate and multivariate analysis of variance procedures. Psychol Bull 1971; 76:49-57.

26. Dirkmaat DC, Cabo LL, Ousley SD, Symes SA. New perspectives in forensic anthropology. Yearb Phys Anthropol 2008; 51:33-52.

27. Melnick RL. A *Daubert* motion: A legal strategy to exclude essential scientific evidence in toxic tort litigation. Am J Public Health 2005; 95:s30-s34.

28. Stewart TD. Responses of the human skeleton to changes in the quality of life. J Forensic Sci 1980; 25:912-21.

29. Cameron N, Tobias PV, WJ Frasier, Nagdee M. Search for secular trends in calvarial diameters, cranial base height, indices and capacity in South African negro crania. Am J Hum Biol 1990; 2:53-61.

30. Shah GV, Jadhav HR. The study of cephalic index in students of Gujarat. J Anat Soc India 2004; 53:25-6.

31. Umar MBT, Singh R, Shugaba AI. Cephalometric indices among Nigerians. J Applied Sci 2006; 6:939-42.

32. Oladipo GS, Oluto JE, Suleiman Y. Anthropometric studies of cephalic indices of the Ogonis in Nigeria. Asian J Med Sci 2009; 1:15-17.

33. Oladipo GS, Okoh PD, Isong EE. Anthropometric studies of cephalic length, cephalic breadth and cephalic indices of the Ibibios in Nigeria. Asian J Med Sci 2010; 2:104-106.

34. Oladipo GS, Oyakhire MO, Ugboma-Henry AA. Anthropometric studies of nasal indices of the Ekpeye and Ikwere ethnic groups in Nigeria. Asian J Med Sci 2010; 2:167-69.

35. Ilayperuma I. Evaluation of cephalic indices: A clue for racial and sex diversity. Int J Morphol 2011: 29:112-17.

36. Yagain VK, Pai SR, Kalthur SG, Hemalatha P, Hemalatha I. Study of cephalic index in Indian students. Int J Morphol 2012; 30:125-29.

37. Kumar M, Gopichand PVV. The study of cephalic index in Haryanvi population. Int J Pure App Biosci 2013; 1:1-6.

38. Kanchan T, Gupta A, Krishan K. Categorization of human crania using cranial indices in a South Indian population. Proceedings of the 66th annual meeting of the Academy of Forensic Sciences. 2014. Seattle, WA.

39. Schlebusch C. Issues raised by use of ethnic-group names in genome study. Nature 2010; 464:487.

40. Petersen DC, Libiger O, Tindall EA, Hardie R, Hannick LI, Glashoff RH, et al. Complex patterns of genomic admixture within southern Africa. PLOS Genet 2013; 9:e1003309.

41. Legassick M, Rassool C. Skeletons in the cupboard: South African museums and the trade in human remains 1907-1917. South African Museum; 2000. p. 9-24.

42. Morris AG. Anthropology at the Southern tip of Africa: Carrying European baggage in an African context. Curr Anthropol 2012; 53:s152-s160.

43. Steyn M. Fisiese Antropologie. University of Pretoria: Department of Anatomy; 1993.

44. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

45. Samuels ML, Witmer JA. Statistics for the Life Sciences. Pearson Education, Inc.; 2003. p. 463-76.

46. Fenech AP. Tukey's method of multiple comparison in the randomized blocks model. Am J Statist Assoc 1979; 74:881-84.

47. Jenks GF. The data model concept in statistical mapping. International Yearbook of Cartography 1967; 7:186-90.

48. Murray AT, Shyy TK. Integrating attribute and space characteristics in choropleth display and spatial data mining. Int J Geographical Information Science 2000; 14:649-67.

49. Bivand R. classInt: Choose univariate class intervals. R package version 0.1-21. 2013. http://CRAN.R-project.org/package=classInt.

50. Kachigan SK. Multivariate statistical analysis: A conceptual introduction. Radius Press; 1991. p. 216-35.

51. Stull KE, Godde K. Sex estimation of infants between birth and one year through discriminant analysis of the humerus and femur. J Forensic Sci 2012; 1-8.

52. Ousley SD, Jantz RL, Freid D. Understanding Race and Human Variation: Why Forensic Anthropologists are good at Identifying Race. Am J Phys Anthropol 2009; 139:68-75.

53. Venables W, Ripley B. Modern Applied Statistics with S. 4th ed. New York: Springer; 2002.

54. Roever C, Raabe N, Leubke K, Ligges U, Szepannek G, Zentgraf M. klaR: Classification and visualization. 2013.

55. Stynder DD. Craniometric evidence for South African Later Stone Age herders and hunter-gatherers being a single biological population. J Archaeol Sci 2009; 36:798-806.

56. Ribot I. Differentiation of Modern Sub-Saharan African Populations: Craniometric interpretations in relation to geography and history. Bull Mém Soc Anthropol Paris 2004; 16:143-65.

57. Herbert RK. The sociohistory of clicks in Southern Bantu. Anthropol Linguist 1990; 32:295-315.

58. Albert AM, Ricanek Jr K, Patterson E. A review of the literature on the aging adult skull and face: Implications for forensic science research and applications. Forensic Sci Int 2007; 172:1-9.

59. Vance VL, Steyn M, L'Abbé EN, Becker PJ. A cross-sectional analysis of age related changes in the osteometric dimensions of long bones in modern South Africans of European and African descent. Forensic Sci Int 2010; 199:110-e1.

Additional information and reprint requests:

Leandi Liebenberg, MSc

Department of Anatomy

University of Pretoria

Private Bag x323

Arcadia 0007 South Africa

E-mail: leandil@hotmail.co.za