Application of an ancestrality estimate method through morphoscopic analysis of Brazilian skulls

Joyce Campos de Miranda[1],
Thais Torralbo Lopez-Capp[1],
Luiz Airton Saavedra de
Paiva[2], Edgard Michel-
Crosato[1], Maria Gabriela
Haye Biazevic[1]

[1] Community Dentistry Department,
School of Dentistry, Universidade de
São Paulo (FOUSP), São Paulo, São
Paulo, Brazil.
[2] Institute of Teaching and Research in
Forensic Sciences, São Paulo, São Paulo,
Brazil

Corresponding author:
biazevic@usp.br

The authors declare that they have no conflict of interest.

KEYWORDS

ABSTRACT
The objective of the investigation was to test a technique originating in the United States to verify the accuracy of the ancestry estimate in a sample of Brazilian skulls. The sample consisted of 120 skulls of Brazilian adults from the collection of the Institute of Education and Research in Forensic Sciences (IEPCF), Guarulhos-SP, Brazil, with ancestry determined by routine examination. In the skull, the following structures were analyzed: anterior nasal spine (ANS), inferior nasal aperture (INA), interorbital breadth (IOB) nasal aperture width (NAW), nasal bone structure (NBS), and post bregmatic depression (PBD). The Hefner and Ousley technique was applied to classify ancestry. Data were registered and correctness percentage was obtained. To improve ancestry prediction, Machine Learning techniques were used. Regarding the error of the studied method and the correctness of the complete sample, the study presented a percentage above 50%, with values such as 56.33% for whites and 53.07% for non-whites. The most frequent ANS parameter was marked (37.5%), INA straight (31.7%), IOB wide (38.3%), NAW medium (32.5%) NBS triangular (29.2%), and PBD present (78.3%). The application of the method in the sample of the study showed a limitation to estimate ancestry, suggesting a need for adaptation for its applicability in the Brazilian population.

INTRODUCTION
The existence of sexual, ancestral, and age dimorphism in human skeletons, and its importance in medical legal investigations is well established. According to contemporary literature, the biological profile is outlined through the analysis of bone structures, in turn, this analysis can be performed through numerous methods that are classified into two main groups: qualitative also known as morphological methodologies, and quantitative also known as morphometric methodologies. Qualitative or morphological methods have as main objective to obtain biological parameters through subjective analysis of anatomical characters. On the other hand, quantitative or morphometric methods recommend the use of linear measures, projections, angles, and indices to outline the demographic characteristics of the remains under analysis [1].
In human identification, one of the main characteristics in the design of the biological profile is the estimation of sex, age,

height, and ancestry. In the latter case, the geographic origin is of fundamental importance in establishing the identification of a forensic case.[2]

In addition, the determination of ancestry has been little explored, though skulls can provide substantial information in this regard. After the pelvis, the skull remains one of the parts of the human skeleton that has the most information on this subject.[3] These characters can be estimated through a visual comparative analysis by the correspondents of known origin called the qualitative method or through measurements of the cranial anatomical structures called the quantitative method. The quantitative methodology applied to the skull is called craniometry, which can be defined as being a technique, or conventional system, which determines the measurement of the skull in a universally systematized way, which allows the comparative evaluation between studies carried out by different researchers.[4]

Within forensic anthropology, there are several discussions regarding the terminologies used in the studies, such as the concept of ethnicity, ancestry, and race.[5] On one hand, the word ethnicity, derived from the Greek ethnos, has a slight differentiation with the term ancestry. Ethnicity means a multi-purpose concept, which builds the identity of an individual summarized in kinship, religion, language, shared territory, and nationality, in addition to physical appearance, while the term ancestry encompasses similar genetic characteristics within a given population, but distinct when comparing two groups. These genetic patterns have a great influence on physical traits and are constant throughout human life.[6] Considering the miscegenation processes of the human population of the 21st century, in the present study we adopted the term ancestry, considering the geographical origin [7] as a parameter that has the potential to assist in the identification of a skeleton in the area of forensic anthropology.

Klales & Kenyhercz[7] applied the Hefner technique [8] to an American collection of whites and blacks (Hamann-Todd Collection) and concluded that the technique could be applied among this population group, since the success rate found was high, and the agreement between the exams was adequate. Considering that techniques based on a reference population may not be suitable for application in other groups, the study's objective was to test a technique originating in the United States to verify the accuracy of the ancestry estimate in a sample of Brazilian skulls.

MATERIAL AND METHODS

This study followed the recommendations of the Declaration of Helsinki for research involving human subjects and was approved by the Ethics and Research Committee of the Faculty of Dentistry of the University of São Paulo (FOUSP) (reference number 71476817.4.0000.0075).

The sample consisted of 120 skulls of Brazilian adults (supplementary table) from the collection of the Institute of Education and Research in Forensic Sciences (IEPCF), Guarulhos-SP, Brazil, with ancestry determined by routine examination.

The Hefner and Ousley technique[9] was applied to classify ancestry. The result of the analysis was compared with the examination carried out by an identification specialist, who determined ancestry in the service routine. The rate of correct answers has been described.

In the skull, the following structures were analyzed [8]: anterior nasal spine (ANS), inferior nasal aperture (INA), interorbital breadth (IOB) nasal aperture width (NAW), nasal bone structure (NBS), and post bregmatic depression (PBD). These structures, analyzed according to the author, were then automatically reclassified according to the index called Optimized Summed Scoring Attributes (OSSA), which considers the 6 structures mentioned in a dichotomous way and integrates them in a decision tree, to estimate ancentrality [9] (Chart 1).

The classification of each skull was made according to the parameters described above, and the results were included in a program provided by the author of the technique, which automatically classifies ancestry. These results were compared to the routine examination (prior registration), given to each skull by traditional methods, by the Service's routine.

To carry out the research, an examiner (JCM) was properly trained by an experienced professional (TTLC), who was the gold standard for calibration. Both performed a re-analysis in 10% of the sample.

To improve ancestry prediction, Machine Learning techniques were used. Figure 1 shows the flowchart of the analysis of this technique. The data were worked in the R language and

Python. The steps were: check the best weight of the studied variables and check the accuracy of the algorithms.

All measurements were performed without any prior knowledge of the ancestors of the skulls by the examiners. The statistical test to be performed was the linear regression model. The STATA 13.0 program was used to enter and process the data.

Chart 1. Hefner's classification (8) dichotomized by Hefner & Ousley (9)

	Hefner's classif. (8)		Hefner & Ousley's classification (9)
ANS - Anterior nasal spine	1 2 3 Missing	Slight Intermediate Marked -	0 1 1 Missing
INA - Inferior nasal apertura	1 2 3 4 5 Missing	Pronounced slope Moderate slope Straight Partial Sill Sill -	0 0 0 1 1 Missing
IOB - Interorbital breadth	1 2 3 Missing	Narrow Intermediate Wide -	1 1 0 Missing
NAW - Nasal aperture width	1 2 3 Missing	Narrow Medium Broad -	1 1 0 Missing
NBS - Nasal bone structure	0 1 2 3 4 Missing	Low/Round Oval Marked Plateau Narrow PLateaw Triangular -	0 0 1 1 1 Missing
PBD - Postbregmatic depression	0 1 Missing	Absent Present -	1 0 Missing

Figure 1. Flowchart of the machine learning technique used to predict ancestry
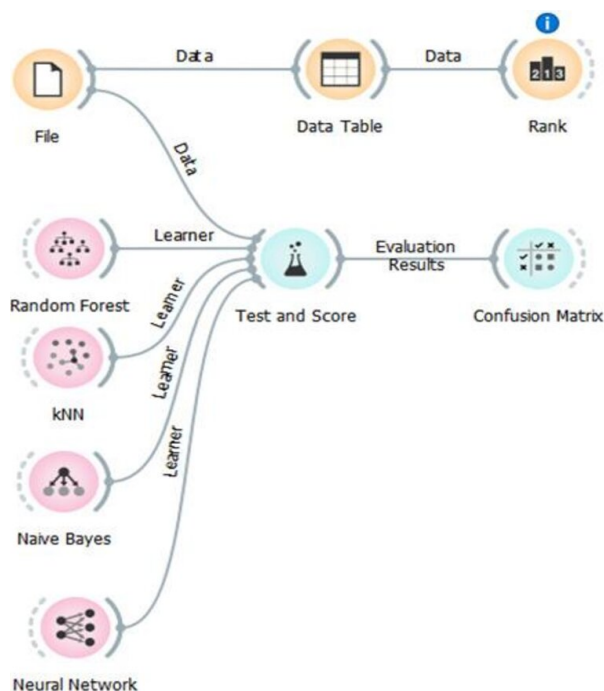
RESULTS

To assess the reproducibility of the ancestry estimation method, the intra-examiner, and inter-examiner concordance tests were performed and the results indicate that the method is reproducible, as they presented an agreement that varied between moderate to perfect in both inter (0.467-1.0) and intra (0.529-1.0) analyses. (Table 1).

When checking the results, it was possible to understand that both in the inter-examiner and intra-examiner analysis, the variable that showed the best performance was the anterior nasal spine (ANS) showing 1.0 and 0.874, respectively, and the one that had the least agreement was the width of the nasal opening (NAW) showing 0.467 and 0.529.

Regarding the error of the studied method and the correctness of the complete sample, the study presented a percentage above 50%, with values such as 56.33% for whites and 53.07% for non-whites (Table 2).

Table 1. Results regarding intra-examiner and inter-examiner agreement for the ancestry estimation method applied in the study

Variable	N(10%)	Intraobserver		Interobserver	
		Kappa	p-valor	Kappa	p-value
ANS	12	0.874	<0.001	1.000	<0.001
INA	12	0.579	<0.001	0.657	<0.001
IOB	12	0.609	<0.002	0.625	<0.002
NAW	12	0.529	<0.002	0.467	<0.024
NBS	12	0.529	<0.002	0.750	<0.001
PBD	12	1.000	<0.001	0.833	<0.001
Ancestry	12	1.000	<0.001	1.000	<0.001

Hefner's method (2009) recommends the qualitative analysis of some structures of the skull that were analyzed and classified as the one proposed, thus generating a number called frequency that expresses the number of skulls that were classified within each specific characteristic. Since some skulls were not eligible for classification of some of the structures proposed by the technique, the method occurred in a small number of the sample without altering the continuity and proposals of the study (Table 3). The most frequent ANS parameter was marked (37.5%), INA straight (31.7%), IOB wide (38.3%), NAW medium (32.5%) NBS triangular (29.2%), and PBD present (78.3%).

Table 2. Results regarding intra-examiner and inter-examiner agreement for the ancestry estimation method applied in the study

	% correctness	**% error**
White	56.33	43.67
Black	53.07	46.93

Table 3. Results for each analyzed sample structure in frequency and percentage for the ancestry estimation method applied in the study

	N	**Freq**	**%**
ANS			
1	120	30	25.0
2	120	39	32.5
3	120	45	37.5
INA			
1	120	12	10
2	120	22	18.3
3	120	38	31.7
4	120	24	20.0
5	120	21	17.5
IOB			
1	120	32	26.7
2	120	41	34.2
3	120	46	38.3
NAW			
1	120	54	45.0
2	120	39	32.5
3	120	25	20.8
NBS			
0	120	13	10.8
1	120	14	11.7
2	120	14	11.7
3	120	33	27.5
4	120	35	29.2
PBD			
0	120	24	20.0
1	120	94	78.3

DISCUSSION

New methods of data analysis are emerging with the increasing complexity of statistical analysis. Techniques for estimating ancestry tend to perform better, when the subjects under analysis comes from a population that already has a cataloged database.[10] In the case of the Brazilian sample, there is no international database that has samples of this population, so it is to be expected that the performance of techniques created in other groups will not be fully satisfactory.

Navega et al. [11] created a database (AncesTrees) where 23 craniometric variables can be inserted and compared with the original database, which has data from some population groups: African, Austro-Melanesian, East Asian, European, Native American, and Polynesian. Thus, when trying to solve a specific case, the expert can select some geographic origins that are more likely or not to select any, and let the program estimate the possibilities of geographic origin.[11]

One possibility to remedy this limitation could be to adopt parameters established for other population groups from Latin America. However, this strategy is not effective either, as it does not historically consider the cultural and historical processes of each group that inhabits the region.[12] When using the support vector machine as a predictor model to differentiate American White, American Black, Guatemalan, and southwest Hispanic (Amerindian ancestry) skulls, Hefner et al.[13] observed that the rate of success was 72% for Guatemalans and 94% for American Black; when associating the Guatemalan and southwest Hispanic groups, an increase in this rate was observed, with results of 96% for American Black, 77% American White and 91% for Hispanics.[13] A similar algorithm correctly classified 60% of a Colombian sample.[14]

When analyzing skulls from the historical collection of Hispanics in the United States, and comparing them with a recent sample of migrants from Mexico, significant differences were found in the morphology of these skulls; thus, the authors state that if the term << Hispanic >> is used within a research context, it should be clearly defined regarding the origin and the time period of the group".[15]

It is known that there is human variation, and therefore, analyses of the skulls and jaws should be performed and the results presented in a probabilistic way. [9] Concerning machine learning, there are many possibilities for analysis [16]; thus, before the data analysis, we performed a model to verify the performance of the various possibilities, allowing an accuracy increase.[16]

In a study with a Brazilian sample using a semi-automatic tool, Jurda & Urbanová [17] analyzed sex and ancestry using mesh to mesh processing (distances of reference anatomical landmarks). In the skull, there is a need to manually establish the reference points, which could hinder the use of fully automated models.[17] Also, examiners must be thoroughly trained to perform the technique.

When using the 3D-ID software to verify the correct classification of sex and ancestry in a sample from Greece, it was observed that the program presented moderate reliability, with a correct classification at 70.9% when comparing individuals to the European population.[18] In a study to verify the intra-examiner error in the analysis of macromorphoscopic characteristics in the identification of skeletal remains, the observer's experience was analyzed, new technologies and inherent error of the method were also analyzed; it was concluded that the possible analysis errors can be predicted and that they do not limit the use of techniques in forensic anthropology.[19] Other studies have shown that, depending on the anatomical characteristic under analysis, there may be wide variations.[20]

CONCLUSIONS

The application of the method described in the sample of the present study showed a limitation to estimate ancestry, suggesting a need for adaptation for its applicability in the Brazilian population.

ACKNOWLEDGEMENT

REFERENCES

1. Sinhorini PA, Costa IAP, Lopez-Capp TT, Biazevic MGH, de Paiva LAS. Comparative analysis of four morphometric methods for sex estimation: A study conducted on human skulls. *Leg Med (Tokyo)* 2019; 39:29-34. doi: 10.1016/j.legalmed.2019.06.001.

2. Kranioti EF, García-Donas JG, Can IO, Ekizoglu O. Ancestry estimation of three Mediterranean populations based on cranial metrics. *Forensic Sci Int* 2018;286:265.e1-265.e8. doi: 10.1016/j.forsciint.2018.02.014.

3. Kranioti EF, Iscan MY, Michalodimitrakis M. Craniometric analysis of the modern Cretan population. *Forensic Sci Int* 2008;180:110 e1-5. doi: 10.1016/j.forsciint.2008.06.018.

4. Burns KR. Race and cranial measurements. In: *Forensic Anthropology Training Manual*. Pearson, 2013. p.222-238.Cunha E, Ubelaker DH. Evaluation of ancestry from human skeletal remains: a concise review. *Forensic Sci Res* 2019;5:89-97. doi: 10.1080/20961790.2019.1697060.

5. Enoch MA, Shen PH, Xu K, Hodgkinson C, Goldman D. Using ancestry-informative markers to define populations and detect population stratification. *Journal of Psychopharmacology (Oxford)*. 2006;20(4 Suppl):19-26. doi: 10.1177/1359786806066041.

6. Klales AR, Kenyhercz MW. Morphological assessment of ancestry using cranial macromorphoscopics. *J Forensic Sci* 2015;60:13-20. doi: 10.1111/1556-4029.12563.

7. Hefner JT. Cranial nonmetric variation and estimating ancestry. *J Forensic Sci* 2009;54:985-95. doi: 10.1111/j.1556-4029.2009.01118.x.

8. Hefner JT, Ousley SD. Statistical classification methods for estimating ancestry using morphoscopic traits. *J Forensic Sci* 2014;59:883-90. doi: 10.1111/1556-4029.12421.

9. Go MC, Hefner JT. Morphoscopic ancestry estimates in Filipino crania using multivariate probit regression models. *Am J Phys Anthropol* 2020;172:386-401. doi: 10.1002/ajpa.24008.

10. Navega D, Coelho C, Vicente R, Ferreira MT, Wasterlain S, Cunha E. AncesTrees: ancestry estimation with randomized decision trees. *Int J Legal Med* 2015;129:1145-53. doi: 10.1007/s00414-014-1050-9.

11. Herrera MD, Tallman SD. Craniometric variation and ancestry estimation in two contemporary Caribbean populations. *Forensic Sci Int* 2019;305:110013. doi: 10.1016/j.forsciint.2019.110013.

12. Hefner JT, Pilloud MA, Black CJ, Anderson BE. Morphoscopic Trait Expression in "Hispanic" Populations. *J Forensic Sci.* 2015;60:1135-9. doi: 10.1111/1556-4029.12826.

13. Monsalve T, Hefner JT. Macromorphoscopic trait expression in a cranial sample from Medellín, Colombia. *Forensic Sci Int* 2016;266:574.e1-574.e8. doi: 10.1016/j.forsciint.2016.07.014.

14. Spradley K, Stull KE, Hefner JT. Craniofacial Secular Change in Recent Mexican Migrants. *Hum Biol* 2016;88:15-29. doi: 10.13110/humanbiology.88.1.0015.

15. Nikita E, Nikitas P. On the use of machine learning algorithms in forensic anthropology. *Leg Med (Tokyo)* 2020;47:101771. doi: 10.1016/j.legalmed.2020.101771.

16. Jurda M, Urbanová P. Sex and ancestry assessment of Brazilian crania using semi-automatic mesh processing tools. *Leg Med (Tokyo)* 2016;23:34-43. doi: 10.1016/j.legalmed.2016.09.004.

17. Bertsatos A, Christaki A, Chovalopoulou ME. Testing the reliability of 3D-ID software in sex and ancestry estimation with a modern Greek sample. *Forensic Sci Int* 2019;297:132-137. doi: 10.1016/j.forsciint.2019.02.004.

18. Kamnikar KR, Plemons AM, Hefner JT. Intraobserver Error in Macromorphoscopic Trait Data. *J Forensic Sci* 2018;63:361-370. doi: 10.1111/1556-4029.13564.

19. Lee M, Gerdau-Radonic K. Variation within physical and digital craniometrics. *Forensic Sci Int* 2020;306:110092. doi: 10.1016/j.forsciint.2019.110092.

20. Lee M, Gerdau-Radonic K. Variation within physical and digital craniometrics. *Forensic Sci Int* 2020;306:110092. doi: 10.1016/j.forsciint.2019.110092.

Supplementary Table. Sample composition

Skull	Sex	Age	Ancestry
1	M	43	B
2	M	37	W
3	F	38	W
4	M	32	W
5	M	43	W
6	M	41	B
7	M	40	W
8	F	70	B
9	F	57	W
10	M	76	W
11	M	61	W
12	F	70	B
13	F	90	W
14	M	56	W
15	M	76	W
16	F	104	W
17	F	64	W
18	M	46	W
19	F	77	W
20	F	56	W
21	M	64	B
22	M	61	W
23	F	44	W
24	F	44	W
25	M	35	B
26	F	56	W
27	M	97	W
28	F	69	B
29	M	72	W

30	M	70	W
31	M	72	W
32	M	46	B
33	M	34	W
34	F	82	W
35	M	57	W
36	F	61	B
37	M	79	W
38	F	37	B
39	F	43	W
40	M	67	W
41	F	74	W
42	M	43	B
43	M	18	W
44	F	78	W
45	F	82	B
46	M	30	B
47	F	77	W
48	M	81	B
49	F	34	W
50	M	58	B
51	F	56	B
52	F	84	W
53	M	36	W
54	F	80	B
55	M	66	W
56	M	61	W
57	F	71	W
58	M	64	W
59	M	65	W
60	M	81	W
61	F	71	W
62	F	79	W
63	F	60	W
64	F	49	B
65	F	79	B
66	F	43	B
67	M	36	B
68	F	20	W
69	F	38	B
70	F	24	B
71	M	46	W
72	F	52	B
73	M	32	B
74	M	48	B
75	M	40	B
76	F	48	W

77	F	36	B
78	M	40	W
79	M	80	W
80	M	25	W
81	M	56	W
82	M	40	W
83	M	68	W
84	M	61	W
85	F	62	W
86	F	68	W
87	F	38	B
88	F	50	B
89	M	60	B
90	M	89	W
91	F	68	W
92	F	61	B
93	M	23	W
94	M	41	W
95	F	40	W
96	F	94	W
97	F	45	W
98	M	87	B
99	F	48	W
100	M	62	B
101	M	40	B
102	M	28	B
103	F	80	B
104	M	28	B
105	M	31	W
106	M	34	W
107	M	28	W
108	M	40	B
109	F	64	W
110	M	65	B
111	M	35	B
112	M	38	B
113	M	21	B
114	M	24	W
115	M	25	B
116	M	23	B
117	M	17	W
118	F	14	W
119	M	30	W
120	M	22	B