Title: Metric variation of the tibia in the Mediterranean: implications in forensic identification

Authors: Elena F. Kranioti, Julieta G. Garcia-Donas, Mara A. Karell, Liliana Cravo, Oguzhan Ekizoglu, Mario Apostol, Eugénia Cunha

PII: S0379-0738(18)31003-X
DOI: https://doi.org/10.1016/j.forsciint.2019.03.044
Reference: FSI 9739

To appear in: *FSI*

Received date: 31 October 2018
Revised date: 25 March 2019
Accepted date: 27 March 2019

Please cite this article as: Kranioti EF, Garcia-Donas JG, Karell MA, Cravo L, Ekizoglu O, Apostol M, Cunha E, Metric variation of the tibia in the Mediterranean: implications in forensic identification, *Forensic Science International* (2019), https://doi.org/10.1016/j.forsciint.2019.03.044

Metric variation of the tibia in the Mediterranean: implications in forensic identification

Elena F. Kranioti[1,2,*], Julieta G. Garcia-Donas[2,3], Mara A. Karell[2], Liliana Cravo[4], Oguzhan Ekizoglu[5], Mario Apostol[6], Eugénia Cunha[4,7]

[1]Department of Forensic Sciences, Faculty of Medicine, University of Crete, 71110, Heraklion, Crete, Greece.

[2]Edinburgh Unit for Forensic Anthropology, School of History Classics and Archaeology, University of Edinburgh, 4 Teviot place, EH8 9AG, Edinburgh, UK.

[3]School of Anthropology and Conservation, Marlowe Building, the University of Kent, Canterbury, Kent CT2 7NR, Kent, UK

[4]Centre for Functional Ecology, Laboratory of Forensic Sciences, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

[5]Department of Forensic Medicine, Tepecik Training and Research Hospital, Izmir, Turkey

[6]SC Medicina Legale—ASL TO4IvreaItaly

[7]National Institute of Legal Medicine and Forensic Sciences, Lisbon, Portugal

[*]**Author for correspondence and reprint requests:**
Elena F. Kranioti
Department of Forensic Sciences,
Faculty of Medicine,
University of Crete,
71110,
Heraklion, Crete,
Greece.
E-mail: elena.kranioti@ed.ac.uk

Highlights

- Ancestry estimation methods on the postcranial skeleton are limited
- This study explores the anatomical variation of the tibia in a large Mediterranean sample
- The best model can predict ethnic groups with 74% accuracy for the validation sample
- The method is not ideal but seems promising as supplementary method for

ancestry estimation
- More research is needed to expand the sample and verify the results

Abstract

Ancestry estimation from skeletal remains is a challenging task, but essential for the creation of a complete biological profile. As such, the study of human variation between populations is important for the fields of biological and forensic anthropology, as well as medicine. Cranial and dental morphological variation have traditionally been linked to geographic affinity resulting in several methods of ancestry estimation, while the postcranial skeleton has been systematically neglected. The current study explores metric variation of the tibia in six Mediterranean populations and its validity in estimating ancestry in the Mediterranean. The study sample includes 909 individuals (470 males and 439 females) from Cyprus, Greece, Italy, Spain, Portugal and Turkey. The sample was divided in two subsamples: a reference and a validation sample. Multinomial regression models were created based on the reference sample and then applied to validation sample. The first model used three variables and resulted in 57% and 56% accuracy for the two samples respectively, while the second model (6 variables) resulted in 80% and 74% respectively. Classification between groups ranged from 28% to 95% for the reference sample and from 15% to 91% for the validation sample. The highest classification accuracy was noted for the Greek sample (95% and 90% for the reference and validation sample respectively), followed by the Turkish sample (74% and 78% respectively). The Spanish, Portuguese and Italian samples presented greater morphological overlap which resulted in lower classification accuracies. The results indicate that although the tibia presents considerable variation amongst neighbour populations is not suitable as a sole skeletal element to separate all groups successfully. A combination of different skeletal elements may be required in order to achieve the levels of reliability required for forensic applications.

Key-words: Forensic Anthropology, Ancestry, Skeletal Variation, Tibia, Mediterranean

Introduction

Ancestral affinity is defined as the "individual's ancestral geographic region of origin" [1] and its estimation is considered one of the most challenging tasks in creating a biological profile from heavily decomposed human remains. Age, sex and stature tend to be population-specific with the unique exception of the pelvis, which presents a similar pattern of sexual dimorphism across different geographical regions [2]. Thus, the accurate estimation of these features is heavily dependent of the correct allocation of the remains to a given ancestral group. At this point, it is imperative to draw the line between ancestral affinity and "race", a socially structured mechanism of group membership [3], which although it is a non-biological term outdated by all forensic practitioners' textbooks, appears associated with descriptions of missing persons [4]. That race is socially constructed, but real in its societal impact, is also highlighted by the fact that most countries have different census systems with different numbers and types of racial, ancestral and identity-based descriptors[5]. It is thus likely that the phenotypic expression of biological features will not always agree with the ancestry or social identity of the individual, posing severe obstacles in assessing ancestry and establishing positive identification.

Amongst the most popular methods of ancestry estimation of the past are those based on cranial or dental traits that appear predominantly in a given ancestral group. Recent developments in the study of ancestral group discrimination have resulted in sophisticated software such as to FORDISC [6,7] COLIPR [8] and AncesTrees [9] for cranial measurements, 3D-ID [10,11] for cranial shape variation and rASUDAS for dental traits [12]. An obvious disadvantage of these methods is the lack of a reference sample for every population which would inevitably result in erroneous classifications of target individuals not represented in the databases. This is supported by several validation studies [13–15]. Subsequently there is a need for larger databases with each skeletal element available so that the reliability of these methods can be increased.

Traditional or not, the vast majority of studies rely on the cranium while the postcranial skeleton has not been adequately studied. The few postcranial studies conducted have used exclusively American skeletal collections [16], a significant problem for a European context. In an effort to address this issue the current study explores the morphological aspects of the tibia as a possible group indicator in the northeast Mediterranean coast, specifically in Portugal, Spain, Italy, Greece, Turkey and Cyprus. The quantification of the morphometric characteristics of the tibia in each group will also produce

population affinity standards that can be used for ancestry estimation across the Mediterranean, for any situation involving unidentified human remains.

Material and methods

Sample

The sample used in this study is summarised in Table 1. The Greek, Italian and Spanish samples are described in detail in Kranioti & Apostol [17], the Greek-Cypriot sample in Kranioti et al. [18] and the Turkish sample in Ekizoglu et al. [19]. The Portuguese sample derives from the 21st century Identified Skeletal Collection [20] and consists of 62 males and 59 females.

Measurements

Seven measurements were taken on the tibia: Maximum length (ML), Upper epiphyseal breadth (UB), Nutrient foramen anteroposterior diameter (NFap), Nutrient Foramen transverse diameter (NFtrsv), Nutrient foramen circumference (NFCirc), Minimum circumference (MinCirc) and Lower epiphyseal breadth (LB). All measurements were obtained using standard osteometric equipment with the exception of the Turkish sample which used CT scans and 3D reconstructions due to the lack of Turkish osteological collections [19]. According to Stull and colleagues [21] virtual measurements are as accurate as physical measurements and thus the CT scan data used was considered valid for comparison. Virtual measurement of the MinCirc was not possible. In addition, MinCirc, NFap, NFCirc and NFtrsv were not included in the original study by MA, thus were not available for this retrospective study. Measurements for the Greek and Cypriot sample were taken by EK and JG, for the Italians by MA, for Portuguese by LC, for the Turkish by OE and for the Spanish by Del Rio Muñoz [22]. Measurements were taken on the left tibia with the exception of the Italian sample where were taken for both sides and the mean value was used. No bilateral asymmetries were reported in the original study ($p>0.05$) thus we decided to use the mean values. Cases with known or obvious pathological and/or traumatic lesions, or with severe taphonomic changes were excluded from the study sample.

Error estimation

The retrospective nature of the study did not allow for a unified strategy in error quantification. Inter- and Intra-observer errors were calculated for the osteometric [23] and

virtual data [19] using the technical error of measurement (TEM), relative TEM (rTEM), and the coefficient of reliability (R) [24].

Statistical analysis

Variables were tested for normality (Shapiro-Wilk test) and equal variances (Box's M test) between the two groups (males and females) for each sample and between all possible pairs of groups for males and females. Normality was violated in some occasions and the null hypothesis of equal covariance matrices was rejected; thus, non-parametric tests (Kruskal-Wallis test, Mann–Whitney U test) were also used to explore if there were statistically significant differences between sex groups per population and between the populations per sex group.

Multinomial logistic regression (MLR) analysis was used to predict categorical outcomes from categorical and metric data. In this study the population affinity was predicted as a function the dimensions of the tibia (metric data) and sex (categorical data). The sample was divided in two subsamples: a reference and a validation sample. All subcategories and sex groups were proportional to the reference sample (Table 2). The subsamples of each population group for males and females were unequal. For example, the reference sample contained 136 Greeks but only 40 Spanish individuals. To avoid any potential bias, a bootstrap technique was applied, and thus all statistical manipulations were based on the 1000 bootstrapped samples. The Greek sample was set as the reference sample for this analysis. The first regression model (M1) used three metric variables (ML, UB and LB) and one categorical variable (sex) so that all groups could be included. The second regression model (M2) used 6 variables, all but MinCir, and sex which excluded the Italian sample from the analysis.

Data analysis was carried out using the discriminant function subroutines of SPSS 22.0 and Excel 2010.

Results

Error estimation.

For the Greek and Cypriot sample 30 randomly selected tibiae were measured by the same observer within four weeks of the first measurement. The relative TEM was below 5% in all cases, and thus considered valid. The R value was consistently over 0.95, with the exception of TLB which was slightly lower [23]. Similar results were obtained for the Portuguese

sample for all measurements with the exception of TUB, where the R value was 0.89. For the Turkish sample intra-observer error was also low, with rTEMs under 5% and R values over 0.95 [24], while the inter-observer error was relatively higher. Interestingly the variable with the highest error in both cases was TUB, with R=0.73, between the two different observers [19]. For the Portuguese sample, the intra-observer error was calculated and the values for rTEM and R were consistent with the ones reported for the Greek and Turkish sample (See **Supplementary Table 1**).

Sexual dimorphism

Sex differences were evaluated for each sample using non-parametric tests (Mann–Whitney U, Wilcoxon W) and all variables were found to differ significantly ($p<0.001$) between sex in all subsamples. Sexual dimorphism has been explored for most of these subsamples in previous studies [17,19,23] thus was not the focus of this study.

Population differences

The Kruskal-Wallis test for independent samples was applied to males and females and statistically significant differences were found for all variables in both sex groups (Table 3). Pairwise differences were also tested for all possible combinations; the results can be found in **Supplementary Table 2**. The largest differences were noted between the Turkish (TU) and Cypriots (CY) groups, and between the Turkish and Portuguese (PO) group, with the greatest similarity between the Greeks (GR) and Cypriots. Based on these results the Cypriots (CY) and Greeks (GR) were merged in one group called hereafter just the Greeks (GR).

Multinomial logistic regression

Model 1 (M1)

A multinomial logistic regression (MLR) was applied using different combination of variables in order to predict the population group, effectively modelling the relationship between the possible predictors (ML, UB, LB, sex) and the membership in the five groups (GR, TUR, IT, SP and POR)[25]. Statistical significance was determined by a p-value criterion of less than .05. For Model 1, the version with predictors significantly improved the fit between model and data over the intercept-only model, with χ^2 (16, N = 745) = 813.2, Nagelkerke R^2 = .70, p < .001. A Nagelkerke value of 70% is indicative of the proportion of variance of the independent variable explained by the predictors (ML, UB, LB, sex). The X^2 value is the amount by which -2 log likelihood increases when predictor is removed from the

full model [25]. Table 4 shows the parameter estimates contrasting the reference group (GR) versus each of the other groups. The odds ratio (β1) represents the probability of a change in the reference group (GR) versus a change in the comparison group as the independent variable changes. For example, ML-GR vs ML-TUR with β1 = 0.957 means that the chance of being GR versus the chance of being TUR increases with 1:0.957 as the independent variable (ML) increases.

Model 2 (M2)

A second MLR was performed to model the relationship between the predictors (ML, UB, LB, NFap, NFtrsv, NFCirc, sex) and membership in the four population groups (GR, TUR, SP and POR). For Model 2, the version with predictors significantly improved the fit between model and data over the intercept-only model, χ^2 (21, N = 581) = 915.3, Nagelkerke R^2 = .857, $p < .001$. The full description of Model 2 can be found in Table 4.

Classification accuracy

Classification accuracy was calculated for both models (M1 & M2), for both the reference and the validation sample; the results are illustrated in Table 5. M1 resulted in 56% overall classification accuracy, and M2 in 80% overall classification accuracy. Looking at each group separately, the Greeks were accurately classified 80% (M1) and 95% (M2) of the time, while the Turkish were accurately classified 79 % (M1) and 78% (M2) of the time. Only 30% of the Portuguese were correctly classified by M1 but M2 improved the accuracy impressively to 70%. The validation sample, however, resulted in 57% correct group membership for the Portuguese both for M1 and M2. For the Italian reference group, 30% were correctly classified, while the validation sample showed 10% higher accuracy. The least successful group was the Spanish group, with 29% accuracy for M1; this significantly increased when adding the three new variables in M2.

The probability of an unknown case belonging to any given group was 20% for M1 (using five groups and assuming equal prior probabilities) and 25% for M2 (using four groups and assuming equal prior probabilities). Figures 1 and 2 show the probabilities of correct classification for each group using the validation sample.

Discussion

The current study aimed to use metric variables to explore population differences in the tibia between six countries in the Mediterranean: Cyprus, Greece, Turkey, Portugal, Spain

and Italy. Cranial data resulted in high classification accuracy in a sample of Cretans, Greek-Cypriots and the Turkish [26] but no studies on postcranial data currently exist for this part of the world. The main objective of this study was to acquire population-specific data that could be employed in the Mediterranean to aid positive identification of unknown individuals. This is particularly relevant in cases of bodies that are recovered mutilated and/or fragmented, especially as there are recent cases of lower limbs being recovered from both marine and terrestrial environments in Crete (e.g. Figure 3). These cases are particularly challenging due to the frequent migration of people across the Mediterranean, and it is vital to be able to differentiate between locals and foreigners during a forensic investigation.

A previous study on ancestry estimation based on the postcranial skeleton reported higher classification using a combination of postcranial elements[16]. Spradley [16] studied several postcranial elements for ancestry estimation including the tibia and reported only slightly better results for the upper limb bones (63%) as compared to the tibia (61%). According to this study, American white males showed the lowest classification (35%) for the tibia and American Hispanic males the highest (70%). Yet, the three groups were not proportionally represented in the sample (40 Hispanic, 50 black and 185 white) and this may have had an effect on the classification accuracy. The same study reports a 79% overall cross validation accuracy when a combination of 12 measurements are used. However, the sample size for this analysis was even smaller and more skewed population-wise, with no resampling techniques used, and this may also have biased the results. In addition, data for females were not available.

Our Model 1, which uses only three variables, gave an overall 56.3% classification accuracy with very good results for the Greek and the Turkish sample, and very poor results for the Spanish sample. When more variables were added in Model 2 the accuracy improved to 73.8% for the overall validation sample with an impressive 90% accuracy for the Greeks. The overall accuracies obtained for the validation samples were very close to the original classification accuracies in all cases and sometimes higher (e.g. Portuguese and Italians for Model 1). The results obtained here are not ideal for forensic applications but constitute a first step in the creation of a large database of skeletal metric variables for modern Mediterranean and European populations. Future studies should combine more postcranial and cranial elements from each population to create more representative biological signatures for different groups (including a combination of metric and non-metric features) to assist the estimation of ancestry in forensic scenarios.

Ethics

The data were anonymised and treated according to the ethical standards for the use of modern human skeletal collections for research following international and regional guidelines**.**

Competing interests

We have no competing interests

Acknowledgments

EK would like to thank Dr Antonis Papadomanolakis, Head of the Pathology Division of the Greek Ministry of Justice, Transparency and Human Rights in Crete for providing the facilities for the storage of the Cretan collection and the Orthodox Church of Limasol for providing permission for the study of the Cyprus material.

Bibliography

[1] Scientific Working Group on Forensic Anthropology, 2013. Ancestry Assessment Revision 0., 2013.

[2] J. Brůžek, F. Santos, B. Dutailly, P. Murail, E. Cunha, Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology, Am. J. Phys. Anthropol. (2017) 440–449. doi:10.1002/ajpa.23282.

[3] L.W. Konigsberg, B.F.B. Algee-Hewitt, D.W. Steadman, Estimation and evidence in forensic anthropology: Sex and race, Am. J. Phys. Anthropol. 139 (2009) 77–90. doi:10.1002/ajpa.20934.

[4] M. İşcan, M. Steyn, The human skeleton in Forensic Medicine, 3rd ed., CHARLES C THOMAS, Springfield, Illinois, 2013.

[5] A. Morning, Ethnic Classification in Global Perspective: A Cross-National Survey of the 2000 Census Round, Popul. Res. Policy Rev. 27 (2008) 239–272. doi:10.1007/s11113-007-9062-5.

[6] S. Ousley, R. Jantz, FORDISC 3.0: Computerized forensic discriminant functions., (2005).

[7] S. Ousley, R. Jantz, Fordisc 3, Rechtsmedizin. 23 (2013) 97–99. doi:10.1007/s00194-013-0874-9.

[8] M.K. P. Urbanová, COLIPR, (2008).

[9] D. Navega, R. Vicente, D.N. Vieira, A.H. Ross, E. Cunha, AncesTrees: ancestry estimation with randomized decision trees., Int J Leg. Med. 129 (2015) 651.

[10] A.H. Ross, D. Ph, D.E. Slice, D. Ph, S.E. Williams, D. Ph, Geometric Morphometric Tools for the Classification of Human Skulls, (2010).

[11] D.E. Slice, A. Ross, 3D-ID: Geometric Morphometric Classification of Crania for Forensic Scientists, (2010).

[12] G. Scott, D. Joel, J. Coelho, E. Cunha, J. Irish, rASUDAS : A New Method for Estimating Ancestry from Tooth Crown and Root Morphology, in: 85th Annu. Meet. Am. Assoc. Phys. Anthropol., Atlanda, USA, 2016. doi:10.13140/RG.2.1.2792.0241.

[13] P. Urbanová, A.H. Ross, M. Jurda, M. Nogueira, Testing the reliability of software

tools in sex and ancestry estimation in a multi-ancestral Brazilian sample, 16 (2014) 264–273. doi:10.1016/j.legalmed.2014.06.002.

[14] M. Katherine Spradley, R.L. Jantz, Ancestry Estimation in Forensic Anthropology: Geometric Morphometric versus Standard and Nonstandard Interlandmark Distances, J. Forensic Sci. 61 (2016) 892–897. doi:10.1111/1556-4029.13081.

[15] P. Guyomarc'h, J. Brůžek, Accuracy and reliability in sex determination from skulls: A comparison of Fordisc® 3.0 and the discriminant function analysis, Forensic Sci. Int. 208 (2011) 180.e1-180.e6. doi:10.1016/J.FORSCIINT.2011.03.011.

[16] M.K. Spradley, Metric Ancestry Estimation from the Postcranial Skeleton, in: Biol. Affin. Forensic Identif. Hum. Skelet. Remain. Black White, CRC Press, LLC, 2014: pp. 83–94.

[17] E.F. Kranioti, M.A. Apostol, Sexual dimorphism of the tibia in contemporary Greeks , Italians , and Spanish : forensic implications, Int. J. Legal Med. 129 (2015) 357–63. doi:10.1007/s00414-014-1045-6.

[18] E.K. Kranioti, J.G. García-Donas, P.S. Almeida Prado, X.P. Kyriakou, H.C. Langstaff, Sexual dimorphism of the tibia in contemporary Greek-Cypriots and Cretans: Forensic applications, (2016). doi:10.1016/j.forsciint.2016.11.018.

[19] O. Ekizoglu, A. Er, M. Bozdag, M. Akcaoglu, I.O. Can, J.G. García-Donas, E.F. Kranioti, Sex estimation of the tibia in modern Turkish: A computed tomography study, Leg. Med. (Tokyo). 23 (2016) 89–94. doi:10.1016/j.legalmed.2016.10.004.

[20] M.T. Ferreira, R. Vicente, D. Navega, D. Gonçalves, F. Curate, E. Cunha, A new forensic collection housed at the University of Coimbra, Portugal: The 21st century identified skeletal collection, Forensic Sci. Int. 245 (2017) 202.e1-202.e5. doi:10.1016/j.forsciint.2014.09.021.

[21] K.E. Stull, M.L. Tise, Z. Ali, D.R. Fowler, Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images, Forensic Sci. Int. 238 (2014) 133–140. doi:10.1016/j.forsciint.2014.03.005.

[22] P.A. Río Muñoz, Estudio antropológico-forense, antropométrico y morfológico, de la colección de la escuela de medicina legal de Madrid, 2000.

[23] E.F. Kranioti, J.G. García-Donas, P.S. Prado Almeida, X.-P. Kyriakou, H.K. Langstaff, Sexual dimorphism of the tibia in contemporary Greek-Cypriots and Cretans: forensic applications., Forensic Sci. Int. 271 (2017) 29.e1-129.e7. doi:http://dx.doi.org/10.1016/j.forsciint.2016.11.018.

[24] S.J. Ulijaszek, D.A. Kerr, Anthropometric measurement error and the assessment of

nutritional status, Br. J. Nutr. 82 (1999) 165. doi:10.1017/S0007114599001348.

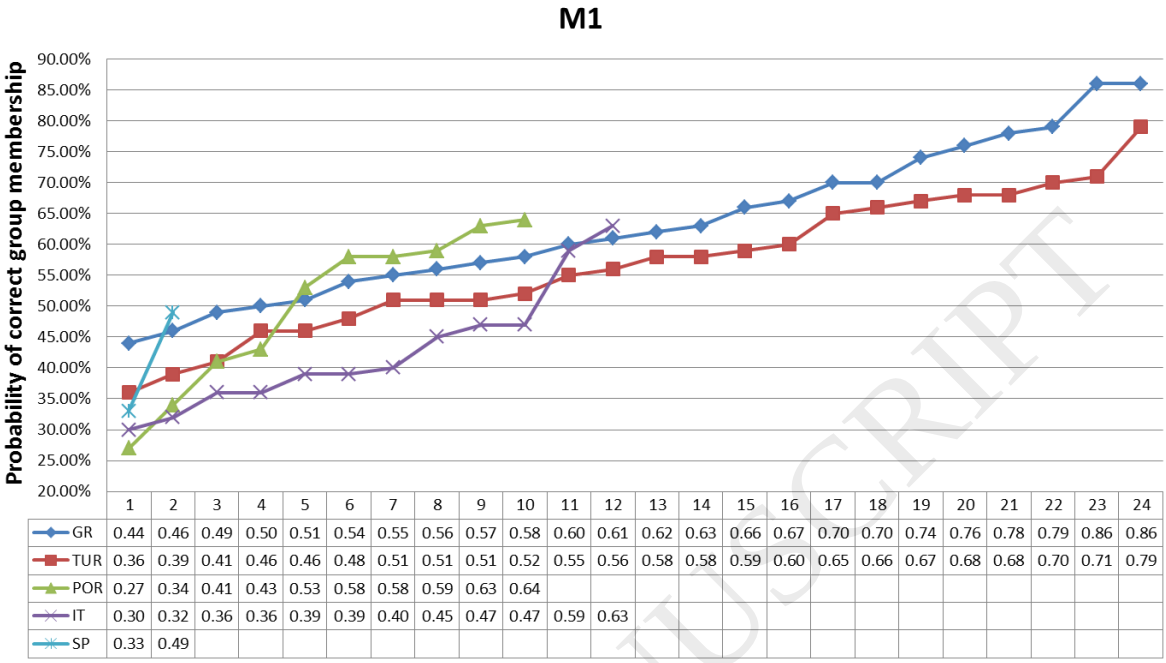
[25] K. Wuensch, Multinomial Logistic Regression with SPSS', (2014) 1–8. http://core.ecu.edu/psyc/wuenschk/MV/multReg/Logistic-Multinomial.pdf .

[26] E.F. Kranioti, J.G. García-Donas, I.O. Can, O. Ekizoglu, Ancestry estimation of three Mediterranean populations based on cranial metrics, Forensic Sci. Int. (2018). doi:10.1016/j.forsciint.2018.02.014.

Figure legends

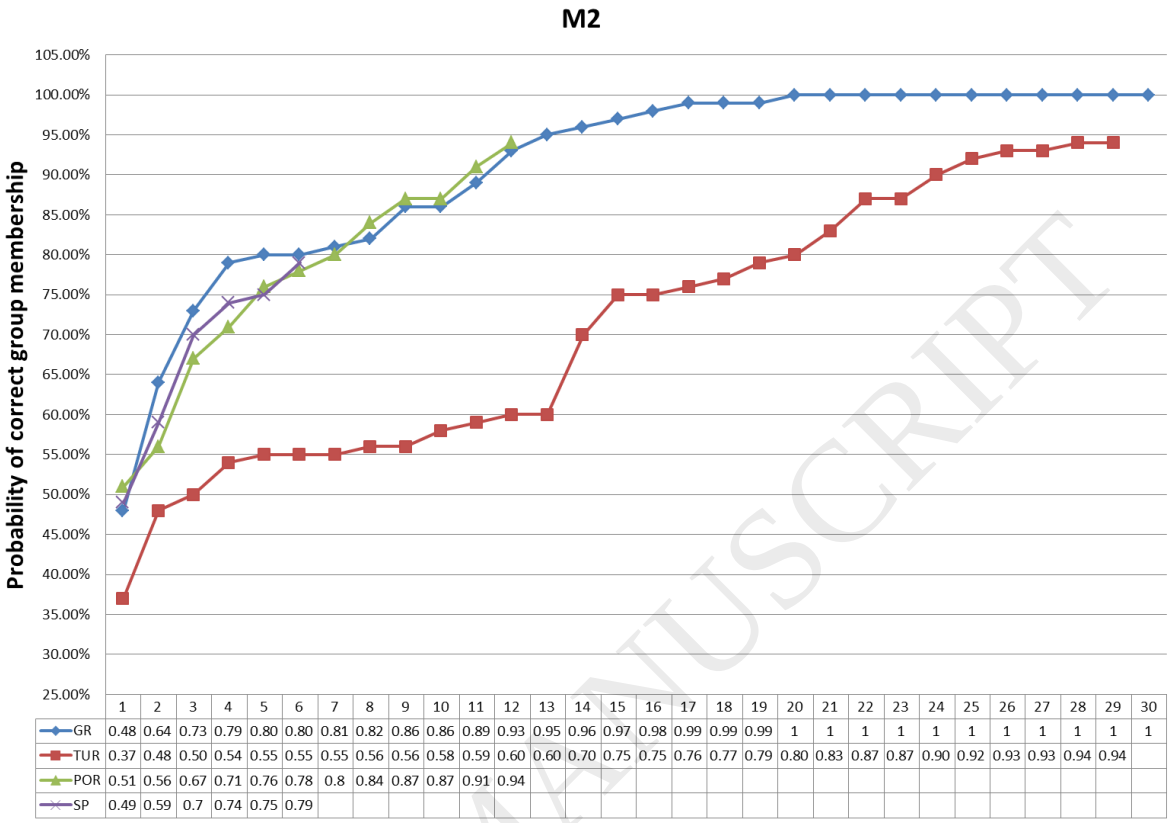
1. Probabilities of correct classification for each ethnic group on the validation sample for M1.
2. Probabilities of correct classification for each ethnic group on the validation sample for M2.
3. Unknown human remains recovered on the North-East part of the island of Crete. Photo Credit: Antonis Papadomanolakis, Head of the Division of Forensic Pathology in Crete, Ministry of Justice, Transparency and Human Rights.

Fig 1`



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
GR	0.44	0.46	0.49	0.50	0.51	0.54	0.55	0.56	0.57	0.58	0.60	0.61	0.62	0.63	0.66	0.67	0.70	0.70	0.74	0.76	0.78	0.79	0.86	0.86
TUR	0.36	0.39	0.41	0.46	0.46	0.48	0.51	0.51	0.51	0.52	0.55	0.56	0.58	0.58	0.59	0.60	0.65	0.66	0.67	0.68	0.68	0.70	0.71	0.79
POR	0.27	0.34	0.41	0.43	0.53	0.58	0.58	0.59	0.63	0.64														
IT	0.30	0.32	0.36	0.36	0.39	0.39	0.40	0.45	0.47	0.47	0.59	0.63												
SP	0.33	0.49																						

Fig 2



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GR	0.48	0.64	0.73	0.79	0.80	0.80	0.81	0.82	0.86	0.86	0.89	0.93	0.95	0.96	0.97	0.98	0.99	0.99	0.99	1	1	1	1	1	1	1	1	1	1	1
TUR	0.37	0.48	0.50	0.54	0.55	0.55	0.55	0.56	0.56	0.58	0.59	0.60	0.60	0.70	0.75	0.75	0.76	0.77	0.79	0.80	0.83	0.87	0.87	0.90	0.92	0.93	0.93	0.94	0.94	
POR	0.51	0.56	0.67	0.71	0.76	0.78	0.8	0.84	0.87	0.87	0.91	0.94																		
SP	0.49	0.59	0.7	0.74	0.75	0.79																								

Fig 3

Tables

Table 1. Demographic information of the sample

	Population	Cretans	Cypriots	Italians	Spanish	Portuguese	Turkish
Male	**N**	86	70	81	47	62	124
	Mean Age	69.2	74.4	70	70	70.6	60
	SD	14.3	11.5	NA	15.1	16.9	14.3
Female	**N**	72	62	109	58	59	79
	Mean Age	71.8	73.6	75	74.6	76.8	60.2
	SD	17.7	20	NA	10	14.6	14.5
Total	**N**	158	114	190	105	121	203
	Mean Age	70.4	73.9	71.4	63.8	73.8	60
	SD	15.9	16.9	17.2	26.7	16	13.2
	Age Range	19-94	20-100	18-104	30-97	25-98	18-92
	Side	Left	Left	*Mean	Left	Left	Left
	Date of Death	1968-1998	1976-2003	1970-1990	1975-1985	1995-2008	Examinations in 2014-2016

Table 2. Frequencies of sex and population groups for the reference and validation sample

	Reference Sample			**Validation Sample**		
Population	M	F	T	M	F	T
Greeks	136	114	250	20	20	40
Italians	71	95	166	10	14	24
Spanish	40	48	88	7	10	17
Portuguese	50	50	100	12	9	21
Turkish	104	62	166	20	17	37
Total	**401**	**369**	**770**	**69**	**70**	**139**

Table 3. Kruskal-Wallis test for independent samples between the groups for males and females.

Test Statistics[a,b]

Sex		ML	UB	LB	NFmaxL	NFminL	NFCircL	MinCirc
Male	Chi-Square	16.664	46.424	276.235	73.032	65.092	58.261	7.389
	Asymp. Sig.	.005	.000	.000	.000	.000	.000	.025
	Monte Carlo Sig.	.006[c]	.000[c]	.000[c]	.000[c]	.000[c]	.000[c]	.020[c]
Female	Chi-Square	30.997	21.716	242.222	40.369	54.064	34.544	9.684
	Asymp. Sig.	.000	.001	.000	.000	.000	.000	.008
	Monte Carlo Sig.	.000[c]	.001[c]	.000[c]	.000[c]	.000[c]	.000[c]	.007[c]

[a]Kruskal-Wallis Test, [b]Grouping Variable: origin, [c]Based on 10000 sampled tables with starting seed 743671174.

Table 4. Multinomial logistic regression equations for Model 1 and Model 2

		Model 1			Model 2		
		B	**P-value**	**β1**	**B**	**P-value**	**β1**
TUR	**ML**	-0.044	**0.001**	0.957	-0.076	**0.001**	0.927
	UB	-0.332	**0.001**	0.717	-0.409	**0.001**	0.665
	LB	1.238	**0.001**	3.449	1.489	**0.001**	4.434
	Ntap				0.009	0.939	1.009
	NFtrsv				0.205	0.210	1.228
	NFCirc				-0.183	0.372	0.833
	Sex (=1)	-2.064	**0.001**	0.127	-2.534	**0.001**	0.079
	Intercept	-17.545	**0.001**		-15.712	**0.003**	
POR	**ML**	-0.01	**0.001**	0.99	-0.021	**0.048**	0.979
	UB	-0.402	**0.001**	0.669	-0.553	**0.001**	0.575
	LB	0.826	**0.001**	2.284	1.001	**0.001**	2.722
	Ntap				-0.251	**0.005**	0.778
	NFtrsv				-0.183	0.280	0.833
	NFCirc				1.154	**0.001**	3.169
	Sex (=1)	-0.769	**0.001**	0.463	-0.170	0.763	0.844
	Intercept	-5.393	0.072		1.412	0.746	
SP	**ML**	-0.068	0.139	0.934	-0.094	**0.001**	0.911
	UB	-0.444	**0.001**	0.642	-0.385	**0.002**	0.681
	LB	1.416	**0.001**	4.119	1.709	**0.001**	5.525
	Ntap				0.019	0.885	1.019

	NFtrsv				-0.288	0.174	0.75
	NFCirc				-0.249	0.378	0.78
	Sex (=1)	-1.922	**0.002**	0.146	-1.581	0.076	0.206
	Intercept	-10.634	**0.007**		-6.284	0.261	
IT	**ML**	-0.024	**0.001**	0.976			
	UB	0.054	**0.001**	1.055			
	LB	0.212	**0.001**	1.236			
	Sex (=1)	-1.095	**0.001**	0.334			
	Intercept	-4.548	**0.028**				

Table 5. Classification accuracy for M1 and M2 for both reference and validation sample

Observed	Model 1													
	Predicted for Reference Sample							Predicted for Validation Sample						
	GR	**TUR**	**POR**	**SP**	**IT**	**Total**	**%**	**GR**	**TUR**	**POR**	**SP**	**IT**	**Total**	**%**
GR	191	5	4	0	37	237	**80.6**	24	1	2	0	7	34	**70.6**
TUR	1	129	19	7	10	166	**77.7**	1	24	7	2	3	37	**64.9**
POR	18	34	29	3	16	100	**29.0**	1	4	12	2	2	21	**57.1**
SP	0	48	7	22	2	79	**27.8**	0	9	2	2	0	13	**15.4**
IT	82	12	10	4	55	163	**33.7**	8	1	4	0	10	23	**43.5**
Total	292	228	69	36	120	745	**57.2**	34	38	25	6	15	128	**56.3**
	Model 2													
	Predicted for Reference Sample							**Predicted for Validation Sample**						
	GR	**TUR**	**POR**	**SP**	**IT**	**Total**	**%**	**GR**	**TUR**	**POR**	**SP**	**IT**	**Total**	**%**
GR	225	6	5	1	-	237	**94.9**	30	2	1	0	-	33	**90.9**
TUR	9	131	8	17	-	165	**79.4**	2	29	3	3	-	37	**78.4**
POR	15	10	70	5	-	100	**70.0**	3	4	21	2	-	30	**57.1**
SP	0	32	9	38	-	79	**48.1**	0	3	4	12	-	19	**41.7**
Total	249	179	92	61	-	581	**79.9**	35	38	29	17	-	119	**73.8**