Rib histomorphometry

García-Donas, Julieta G.; Bonicelli, Andrea; Scholl, Ashely Rose; Lill, Caroline; Paine, Robert R.; Kranioti, Elena F.

Published in:
Legal Medicine

DOI:
10.1016/j.legalmed.2020.101827

Publication date:
2021

Licence:
CC BY-NC-ND

Document Version
Peer reviewed version

Rib histomorphometry: A reliability and validation study with a critical review of histological techniques for forensic age estimation

Age estimation by bone histomorphometry is often used on fragmented human remains
Histomorphometry aging methods must be validated if they are to be used forensically
Methods reliability, accuracy and bias are tested on two Mediterranean samples
Age estimation errors might be related to intrinsic factors and methodological issues

Abstract

Fragmented human remains present a challenge for forensic experts as they attempt to identify individuals using standard forensic methods. Several histological age estimation techniques have been developed during the last fifty years to aid in this process. However, very few validation studies have been conducted in order to test their accuracy and bias, and thus, validation assessment is required as we employ them while testifying in court.

Histological variables are assessed from rib thin-sections from two Mediterranean samples; Cretan (N=41) and Cyprus (N=47). Intra and inter-observer errors are assessed through TEM analysis and Intra-class Correlation Coefficient by testing observers with different levels of experience as they collected data on osteon counts and area measurements. The relation between the variables and age is determined using correlation coefficients. Histomorphometric data are applied to four widely used age estimation formulas assessing the performance of the methods for the entire sample. Inaccuracy and bias are calculated with age estimations and known age tested for significance and proportional bias assessed.

Overall, histological parameters presented acceptable intra- and inter-observer errors. All variables exhibited statistically significant correlation with age ($P < 0.01$). For three of the techniques, data showed a systematic underestimation of age with an increase in inaccuracy in older individuals. One of the age estimation formulas produced overestimation of young individuals yet, it more accurately estimated the age of older individuals.

This validation study explores inter-population variation in bone remodeling dynamics and presents a critical evaluation on methodological issues that can affect the performance of existing histological techniques.

Keywords: age estimation, validation study, ribs, histomorphometry, Mediterranean population

1. Introduction

A forensic anthropologist is often tasked with the examination of decomposed human remains and the reconstruction of the biological profile for an unknown individual by using peer-reviewed methodologies [1]. Even with the use of DNA as a means for positively identifying the remains, osteological age estimation is still one of the first steps in the identification process. The choice of the method for doing so is crucial for ensuring an accurate result [2,3]. Occasionally, the fragmented nature of human remains makes histological methods one of the few tools available for estimating age-at-death (AAD) [4]. As with all forensic identification methods, validation studies for the application of the histomorphometric approach in courtroom are required to ensure accurate and reliable age estimation [5,6].

Bone histology applied in age estimation through the quantification of microscopic features is feasible due to cortical and trabecular remodeling occurring over time reflecting the chronological age of an individual [7–9]. Despite several microscopic methodological drawbacks such as specialized equipment and training [10], many researchers have been using histomorphometric assessment of bones and teeth for human identification purposes in both forensic anthropology and bio-archaeology research since 1965 [11–14]. Additionally, another concern specific to bone histological research is its destructive and invasive nature (i.e. the cutting and grinding of segments of bone). This issue was first addressed in the decision to use ribs in the histological evaluation of human remains [13]. The logic is that with twenty-four human ribs available for assessment the destructive analysis of a small fragment of a single rib is minimal in the overall need for gross anatomical assessment of the remains.

Numerous histological studies have examined the femur, tibia, and fibula, showing accuracy rates for age estimation from ± 5-10 years [7,15,16]. These methods rely on simple counts of specific micro-anatomical features, as for example intact secondary osteons. During the 1990's, ribs and clavicles were deemed suitable for histological assessment because they are not commonly used in standard osteological analysis and clinical data is available for developing comparative samples [13]. Quantitative methods included then cortical area measurements along with counting the number of intact and fragment secondary osteons. In this process, Osteon Population Density (OPD) was offered as a standard means for representing secondary osteon features employed in estimating AAD [13,17]. As aforementioned, in practical terms ribs are more likely to be recovered due to their number, and inter-costal element variability does not seem to dramatically affect the estimation of age [18]. Additionally, slight differences in the rib

sampling sites appear to produce low bias in age estimation compared to the bias emerging from the age estimation equations [19]. Lastly, there is nearly 60 years of analysis in both the medical research and the anthropological fields concerning rib microanatomy, and this body of knowledge makes rib thin sections an ideal skeletal element to work with [20–22].

As histomorphometry became more widely applied in the forensic assessment of human remains, researchers began to recognize that its application to samples not-closely related to the reference population might be a source of age estimation error. Such errors might result from genetic and environmental factors that differ between the target-sample (unknown individual) and the reference population [19,23–25]. Hence, quantitative bone histology studies have expanded the development of population specific equations in order to encompass the variation in bone remodeling dynamics between and within populations [26–28]. Moreover, inter-population variability issues for age estimation are not unusual even for macro-anatomical based methods, and they have been addressed numerous times by researchers [29,30].

Micro-anatomical studies of bone have demonstrated discrepancies between existing histological aging formulae depending on the method applied. Pavón et al. [31] tested two different techniques developed from Europeans and African-Americans on a Mayan population [13,32]. The main group and the control Mayan sample were tested against both methods reporting an overall higher accuracy for the mixed sample regression equation developed by Stout and Paine [13], than for the formulae generated by Cho et al. [32]. Thus, their results suggest that bone net formation rates can differ between the reference and the target samples. It is relevant to note, however, that Pfeiffer et al. [33] have described similar accuracy rates for population-specific and existing aging histological methods.

These discrepancies suggest that further research is needed. Hence, a twofold study design is presented here for this purpose. In the current study, the reliability of the histological methods has been evaluated through an assessment of histological parameters repeatability and reproducibility through intra- and inter-observer errors. A systematic analysis of the bias and inaccuracy of the existing histological formulae [13,21,32,34] is conducted on a Mediterranean sample to determine whether there is a need for Mediterranean population-specific standards for estimating age using ribs. This paper further investigates the nature of the errors produced by the validated methods discussing possible methodological issues, as well as the physiological and biological underlying factors behind cortical bone dynamics between reference and target samples.

2. Materials and Methods

A total of 88 individuals from Cretan and Greek-Cypriot origin were used in this study (**Table 1**) [19,35]. Selection of adult specimens from both collections was carried out based on specific known ages to cover a normal range of age for modern humans under-going autopsy procedures (19-100 years of age). The mid-shaft of the sixth rib of each specimen was collected, and when not available, the mid-shaft from another rib (ribs 4-9) was selected instead. This was done because previous studies reported no errors while collecting secondary osteon data from other ribs within the mentioned rib numbers [18].

The Cretan sample consists of 41 individuals collected from the Cretan Osteological Collection and from routine autopsies performed at the Forensic Medicine Unit (University of Crete). The Cretan osteological collection samples consist of 18 males and 16 females; demographic information was obtained from census records while sex was confirmed with the examination of pelvic morphology [36]. Cause of death was available for some of the individuals and cases with known or obvious pathologies and metabolic disturbances affecting bone remodeling rates were excluded. The Cretan autopsy sample consisted of seven individuals, five males and two females, with an age range of 20-69 years. For the autopsy samples, informed consent was acquired from next of kin in all cases.

The Greek-Cypriot Collection comprised 47 individuals (17 males and 30 females) whose remains were housed in the Limassol Municipal Ossuary inside St. Nicholas Cemetery in Limassol (Cyprus). These samples have known age and sex data which was gathered from cemetery records [37]. No clinical data specific to health status prior to death were available for this collection, but gross examination of all individuals was carried out in order to exclude those specimens with obvious pathological conditions.

Table 1. Sample demographics for the entire sample.

	N	Age Range	Mean Age	SD
Males	40	20-89	60.10	16.53
Females	48	19-100	60.52	19.11
Total	88	19-100	60.33	17.89

Thin-sections were prepared following published histological preparation procedures [10,38]. A Leica DM 750P research microscope fitted with a Leica MC 170 HD camera and the Leica Application Suite V4 software were used for the data acquisition and analysis. Between 30 and 80 single high-resolution microphotographs from each cross section were taken under 4x and 10x magnification and stitched together to obtain a complete cross-section montage (**Figure 1a**). Osteon counting data collection was performed while using the standard research light microscope. Both 10x and 20x objectives were used during the reading of thin sections to provide an accurate count of intact and fragmentary secondary osteons [4]. Single microphotographs were additionally used to keep a permanent record of the structures counted. Measurements were taken through the single microphotographs using the area and shape descriptors functions in ImageJ 1.48 software platform. All parameters were collected according to the original aging techniques, following the instructions and descriptions provided by the authors [13,21,32,34] (**Table 2**). Figure 1b illustrates examples of intact secondary osteons (red), fragmentary secondary osteons (blue) and osteon area and circularity measurements (dashed red) in a 20x image observed under semi-polarized light.

Table 2. Parameters under consideration and data collection methods.

Variable	Abbreviation	Calculation	Data acquisition
Total Area	Tt.Ar	*n/a*	Microphotographs - ImageJ software
Endosteal Area	Es.Ar	*n/a*	
Cortical Area	Ct.Ar	*Tt.Ar - Tr.Ar*	
Relative Cortical Area	Ct.Ar/Tt.Ar	*Ct.Ar./Tt.Ar.*	
*Intact Osteon Number**	N.On	*n/a*	Microscopy and microphotographs
*Fragmentary Osteon Number**	N.On.Fg	*n/a*	
Total Osteons	N.On.Tt	*N.On + N.On.Fg.*	
*Intact Osteon Density**	OPD(I)	*N.On / Ct. Ar.*	
*Fragmentary Osteon Density**	OPD(F)	*N.On.Fg / Ct.Ar*	
Total Visible Osteon Density	OPD	*N.On + N.On.Fg / Ct.Ar*	
Osteon Area*	On.Ar	*On.Ar*	Microphotographs - ImageJ software
Osteon Circularity	On.Cr	*(4π (area/perimeter^2))*	

*Description of the parameter or number of structures counted might differ slightly depending on the author

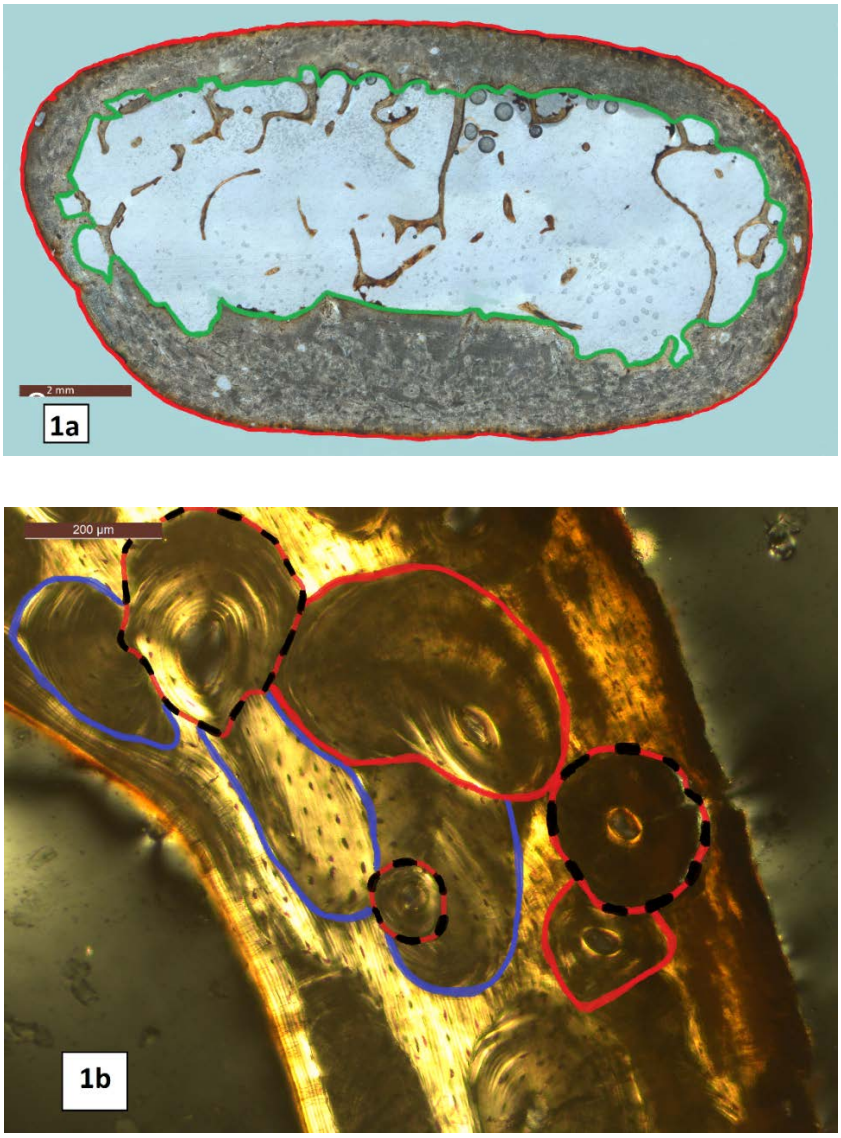

Figure 1. Example of histomorphometric parameters collected for the present study, 45 year old individual. ***1a***: stitched image indicating total area (red) and endosteal area (green). ***1b***: 20x semi-polarized light; intact secondary osteons (red), fragmentary secondary osteons (blue), osteon area and circularity measurements (dashed red).

A test for intra-observer error was performed on 22 random slides scored twice by the same observer who has advanced training and experience in histomorphometry for all variables on an interval of three months. Observer errors between the first author and a highly experienced histologist were already reported in a previous publication [19]. In this study, inter-observer error was assessed by the inclusion of two trained forensic anthropologists with different levels of histological experience. The thin-sections were assessed for osteon frequency

parameters by an observer with advanced histological training and experience; and a second observer, well-trained in bone histological methods but less experienced, assessed histomorphometric parameters. This approach was conducted following previous research reporting discrepancies for osteon counting due to problematic feature definitions and identification [39]. On the other hand, histomorphometric related parameters have proven less subjective as noted by former studies [40–42], suggesting that a less experienced observer might be capable of recording accurate measurements. Moreover, both raw and composite parameters (osteon counts and OPD) were subjected to this level of assessment to better understand the implications of adding specific variables when designing histological aging equations. Two statistical intra- and inter-rater error analyses were combined as recommended elsewhere [43,44]. Technical error of measurement (TEM), relative TEM (rTEM) and the coefficient of reliability (R) were applied in order to assess the reliability of the measurements [44]. Additionally, Intra-class Correlation Coefficient (ICC) was performed to assess the extent of consistency and agreement within and between researchers [45], as conducted in recent histomorphometric studies [22,42]. A two-way mixed effects model based on absolute agreement was selected with both ICC values and 95% confidence intervals over 0.80 and 0.90 being considered good and excellent agreement, respectively [46].

Descriptive statistics for the collected histological parameters were computed to explore the behavior of the variables and to assess normality through Shapiro-Wilk test. All parameters were tested for their correlation with age, sex and population sub-samples calculating either Pearson's two-tailed or Spearman's rho correlation coefficients. Additionally, R^2 was computed to examine the strength of the relationship between the histological parameters and age. The histological parameters obtained from the sample under study were inserted into the four existing published methods [13,21,32,34] (**Table 3**), and the age estimates and known age were tested for correlation. From the study conducted by Cho et al. [32], the three aging formulae provided by the authors were used in order to test if ethnicity/geographical point of origin had an impact on the accuracy of age estimation.

Table 3. Existing age prediction formulae and standards applied on the Mediterranean samples.

Author/year	Methodology	Formulae and Variables used	SEE
Stout and Paine (1992)	6th rib, middle third	Ln_Age= 2.343 + 0.050877(OPD)	3.9*
Stout et al. (1994)	4th rib, sternal sampling area	Age = 18.389 - 0.731(OPD) + 0.110 (OPD)²	10.43
Cho et al. (2002)	6th rib, middle third	<u>European-American</u> Age=38.029+1.603(OPD)-882.210(On.Ar x 1)-51.228(Ct.Ar/Tt. Ar)+57.441(Ct.Ar/Tt. Ar x 1)	12.22
		<u>Unknown Ethnicity</u> Age =29.524 + 1.560 (OPD) + 4.786 (Ct.Ar/Tt.Ar) - 592.899(On.Ar)	
		<u>African-American</u> Age=38.029+1.603(OPD)-51.228(Ct.Ar/Tt. Ar)	
Goliath et al. (2016)	Standard rib, middle third	Age= -472.331 + 591.369 (On.Cr)	6.06

*Absolute difference from known age to estimated age; SEE= Standard Error of the Estimate
OPD=Osteon population density, On.Ar= osteon area, Ct.Ar=cortical area, Tt.Ar= total area, On.Cr=osteon circularity, Ct.Ar/Tt. Ar =Relative Cortical Area

The four methods were assessed for accuracy and bias calculated according to Lovejoy and colleagues on the entire sample, the entire sample divided by age cohorts, and on sub-sample sets (both sex and population sub-groups, separately) [47]. To test whether the age estimates were significantly different from the known ages, Wilcoxon signed-rank test was performed on the entire sample and sub-sample sets [48]. Moreover, the relationship between the age estimates and known age (true value) was represented graphically by Bland and Altman (B&A) plots [49,50] following Crowder and Pfeiffer [51]. The mean difference is used to calculate the limits of agreement as two standard deviations plus or minus the mean difference. The line of equality – in which all values should fall if the age estimates were to be completely accurate – is placed in the plot to evaluate the direction of bias. The best line of fit was also represented to examine the average over- or under-estimation of the aging methods tested. Data analysis was carried out using SPSS 22.

3. Results

3.1 Intra and Inter-observer error

Observer errors are presented in **Table 4**. All secondary osteon frequencies and histomorphometric parameters show intra-observer agreement with rTEM within the limit of acceptance. Each variable reports R values over 0.97 indicating that 3% of the variance can be attributed to measurement error; the only parameter under the 5% accepted threshold [44] is On.Cr, with an R of 0.90.

Regarding inter-observer error, *rTEM* and associated R values indicate that overall high reliability is achieved for most of the parameters except for those parameters including only secondary fragmentary osteons. Both N.On.Fg and OPD(F) results suggest that more than 15% of the variance can be related to measurement error. Inter-observer error for most of the histomorphometric parameters falls within the 5% agreement threshold apart from On.Cr reporting only substantial agreement as indicated by the low *R* value (0.65).

For most of the parameters, Intra-class correlation coefficients (ICC) demonstrates overall excellent intra-observer agreement with ICC values over 0.92; 95% confidence intervals are all above 0.96, apart from On.Cr reporting a lower bound value of 0.84 indicating good agreement between first and second observation [46]. ICC inter-observer rates show excellent agreement with ICC values for intact and total osteon frequencies and rib area parameters ranging from a minimum of 0.92 to a maximum of 0.99. The 95% confidence intervals fall over the lower bound of 0.90, except for both N.On.Fg and OPD(F) that presented 95% confidence intervals with lower limits of 0.72 and 0.64, respectively. This indicates a 95% chance that the true value might fall within the fair agreement boundary. Osteon area and perimeter demonstrate excellent agreement except for On.Cr reporting an ICC value of 0.79 with 95% confidence interval ranging from 0.52 to 0.91.

Table 4. TEM results for intra- and inter-observer error for the histological parameters collected.

	Intra-observer			Inter-Observer		
	TEM	**Relative TEM**	**R**	**TEM**	**Relative TEM**	**R**
N.On	1.61	1.12	0.99	5.52	3.77	0.99
N.On.Fg	2.09	2.17	0.98	14.35	16.01	0.87
N.On.Tt	2.57	1.06	0.99	16.11	6.83	0.96
OPD(I)	0.130	1.49	0.99	0.376	4.24	0.97
OPD(F)	0.136	2.18	0.99	0.949	16.40	0.85
OPD	0.205	1.36	0.99	1.01	6.90	0.94
Ct.Ar	0.102	0.619	0.99	0.279	1.69	0.99
Ct.Ar/Tt.Ar	0.002	0.566	0.99	0.005	1.40	0.99
Tt.Ar	0.179	0.355	0.99	0.255	0.507	0.99
Es.Ar	0.156	0.462	0.99	0.225	0.667	0.99
On.Ar	0.001	2.80	0.99	0.001	3.02	0.99
On.Cr	0.009	0.966	0.90	0.013	1.50	0.65

3.2 Histological methods and age: total, sex and population sub-samples

Table 5 presents descriptive statistics for age and for the histological variables that are included in the four existing formulae (**Table 3**). These data are presented for the entire sample and the sample divided by sex and by population sample sub-groups. The correlation between age and the histological variables for the entire sample is statistically significant (P-values < 0.01) as indicated by Pearson's (r) and Spearman's coefficients (rho), with OPD and On.Cr parameters showing a positive correlation with age (r and *rho* are 0.71 and 0.67, respectively). The remaining parameters demonstrate a negative relation with age with correlation coefficients ranging from 0.55 to 0.64. R^2 values are calculated for each variable in relation to age for the entire sample, and for the sample divided into sex and populations. As seen in **Table 5**, overall, OPD and osteon circularity report the strongest correlation.

The entire sample is divided by sex as well as by sample population to investigate the relationship between the histological parameters for males and females, as well as for the Cretan and Greek-Cypriot sub-samples. Descriptive statistics and correlation coefficients as well as R^2 values are presented in **Table 5**. The histomorphometric parameters show the same positive or negative statistically significant relationship with age as seen for the entire sample, with r and rho coefficients ranging from 0.42 to 0.79 for sex and from 0.43 to 0.73 for population sub-

samples. The only parameter that does not show a statistically significant relationship is Ct.Ar/Tt.Ar for males (r=-0.27). As for R^2 values regarding the sub-samples, the strongest relation is reported for Ct.Ar in females, and for OPD and On.Ar for Greek-Cypriots.

Table 5. Descriptive statistics, normality values, correlation coefficients and R^2 for the entire sample, the sample divided by sex sub-samples (males and females) and by population sub-sample (Cretans and Greek-Cypriots).

		Min	Max	Mean	SE	SD	r /*rho*	R^2
TOTAL SAMPLE (88)	Known Age	19	100	60.33	1.91	17.89	N/A	
	OPD	4.49	25.62	15.44	0.46	4.35	0.71**	0.49
	Ct.Ar	6.38	44.77	19.21	0.85	7.98	*-0.58***	0.31
	Ct.Ar/Tt.Ar	0.091	0.596	0.322	0.01	0.12	*-0.55***	0.25
	On.Ar	0.015	0.052	0.032	0.001	0.01	-0.64**	0.40
	On.Cr	0.858	0.945	0.91	0.001	0.02	*0.67***	0.45
MALES (40) FEMALES (48)	Known Age	20	89	60.1	2.61	16.53	N/A	
		19	100	60.52	2.76	19.11		
	OPD	7.65	24.93	15.48	0.65	4.08	*0.79***	0.51
		4.49	25.62	15.42	0.66	4.61	0.69**	0.48
	Ct.Ar	9.26	44.77	21.16	1.37	8.64	*-0.42***	0.16
		6.38	39.95	17.59	1.02	7.07	*-0.73***	0.54
	Ct.Ar/Tt.Ar	0.091	0.572	0.304	0.02	0.12	-0.27	0.07
		0.152	0.596	0.336	0.02	0.13	*-0.70***	0.43
	On.Ar	0.016	0.052	0.032	0.001	0.01	-0.54**	0.29
		0.015	0.051	0.031	0.001	0.01	-0.69**	0.49
	On.Cr	0.858	0.945	0.913	0.003	0.02	0.67**	0.45
		0.859	0.942	0.908	0.002	0.02	0.68**	0.46
CRETAN (41) GREEK-CYPRIOT (47)	Known Age	19	98	57.49	3.31	21.17	N/A	
		20	100	62.81	2.07	14.2		
	OPD	4.49	24.93	14.5	0.7	4.5	0.68**	0.46
		8.03	25.62	16.26	0.6	4.09	0.73**	0.54
	Ct.Ar	8.17	42.11	17.74	1.19	7.63	*-0.68***	0.39
		6.38	44.77	20.49	1.19	8.13	*-0.58***	0.35
	Ct.Ar/Tt.Ar	0.091	0.596	0.316	0.02	0.13	-0.43**	0.18
		0.156	0.556	0.326	0.02	0.11	*-0.72***	0.42
	On.Ar	0.016	0.05	0.03	0.001	0.01	*-0.66***	0.41
		0.015	0.052	0.033	0.001	0.01	-0.73**	0.54
	On.Cr	0.858	0.942	0.905	0.003	0.02	0.68**	0.46
		0.859	0.945	0.913	0.002	0.02	0.64**	0.41

Grey rows indicate Male (sex sample) and Crete (population sample); ** Correlation significant at 0.01; italics indicates Spearman´s Rank (rho)

Existing histological age estimation methods on the Mediterranean sample

The histomorphometric parameters are inserted into the respective formulae and the age estimates are calculated. The entire dataset is used to explore the overall accuracy rates of the four methods and to examine the correlation and strength between age estimates and known age (**Table 6**).

Table 6. Age estimates (minimum, maximum, mean, SE and SD) produced by the four formulae on the entire sample.

N=88		Min	Max	Mean	SE	SD	r	R^2
Known Age		19	100	60.33	1.91	17.89	N/A	
Stout and Paine (1992)		13.03	33.68	22.47	0.46	4.3	0.69**	0.48
Stout et al. (1994)		17.32	71.87	35.39	1.27	11.95	0.67**	0.45
Cho et al. (2002)	European-American	7.6	66.2	36.63	1.43	13.37	0.75**	0.56
	Unknown Ethnicity	11.51	61.01	36.23	1.15	10.81	0.76**	0.57
	African-American	15.29	54.62	37.92	0.97	9.08	0.70**	0.49
Goliath et al. (2016)		35.05	86.42	65.94	1.36	12.76	0.67**	0.45

Min=minimum, Max=maximum, SE=standard error, SD=standard deviation. ** significant at 0.01

To further explore the histological parameters and their relation to age, the entire sample is divided by age cohorts taking into account the number of individuals for each age group, and is truncated at 60 years of age to consider OPD asymptote occurring at this age [17,52]. As seen in **Table 7**, there is a general underestimation trend for all methods except for Goliath et al. [34]. When the sample is divided into four age categories, the formulae developed by Stout and Paine [13], Stout et al. [21] and Cho et al. [32] show a noticeable increase in inaccuracy for individuals over the age of 40. The highest values of inaccuracy are seen for the individuals over the age of 80. The Stout and Paine [13] age estimation equation produces the highest error values. The opposite trend is observed in the application of Goliath et al. [34] formula for which the highest inaccuracy and bias values are observed in the youngest group (20-39 years old), while the lowest are evident in the 60-79 years age cohort. The only underestimation observed by using this method corresponds to the over 80 years of age group. Overall, there is an increase in inaccuracy and bias values when the sample is divided into under and over 60 years old for which the values almost double between the young and the old age cohorts. This pattern is observed in the

application of Stout and Paine [13], Stout et al. [21] and all three formulae developed by Cho et al. [32], suggesting a trend of increasing inaccuracy and bias in older age categories. Goliath et al. [34] follows the opposite pattern with decreasing inaccuracy and bias values with increasing age. For the entire sample, Stout and Paine [13] formula produce values that surpassed Stout et al. [21] and Cho et al. [32] methods in more than 10 years of difference being the least accurate method. From the three formulae applied from Cho et al. [32], the African-American ancestry one shows the lowest inaccuracy and bias values although a similar performance is observed by the three age predicting equations. Overall, Stout et al. [21] and Cho et al. [32] perform similarly. Comparing all four techniques, Goliath et al. [34] is the method that most accurately predicted age in the sample under study.

Table 7. Inaccuracy and bias values for the Mediterranean sample by age cohorts, </> 60 years and the entire sample.

						Cho et al. (2002)							
		Stout and Paine (1992)		**Stout et al. (1994)**		European formula		Unknown formula		African-American formula		**Goliath et al. (2016)**	
N	**Age range**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**
11	*19-39*	10.38	-10.38	6.19	-3.73	10.70	-5.90	8.72	-3.94	10.34	-0.36	22.52	22.52
28	*40-59*	31.22	-31.22	22.72	-22.72	22.87	-22.87	21.68	-21.68	18.72	-18.72	11.37	9.14
36	*60-79*	42.73	-42.73	27.17	-27.17	24.41	-24.41	25.85	-25.85	24.86	-24.86	6.87	4.05
13	*80*	61.94	-61.94	41.48	-41.48	38.58	-38.58	41.49	-41.49	42.24	-42.24	12.62	-12.01
39	*under 60*	25.34	-25.34	18.06	-17.36	19.44	-23.23	18.02	-16.68	16.35	-13.54	14.52	12.91
49	*over 60*	47.83	-47.83	30.97	-30.97	28.17	-28.17	30.00	-30.00	29.47	-29.47	8.40	-0.21
88	*All*	37.86	-37.86	25.24	-24.94	24.30	-23.70	24.69	-24.10	23.66	-22.41	11.11	5.61

When the sample is divided into subgroups (sex and population sub-samples), estimated age and known age correlations are all statistically significant at the P-value < 0.01 level with the highest coefficients (r >72) being provided by the European-African and unknown ethnicity formulae [32].

Further investigation of the age estimates produced by each method for males and females divided into age cohorts, under and over 60 years old, and the entire sex sub-samples demonstrate a similar inaccuracy and bias pattern as the one observed for the entire sample (see supplementary material). An increasing inaccuracy and bias values towards increasing age for all methods except for Goliath et al. [32] is reported. Stout and Paine formula [13] produces the highest inaccuracy and bias values and Stout et al. [21] and Cho et al. [32] perform similarly for the two sex sub-samples. When divided into age cohorts, females over 80 years of age produce higher inaccuracy and bias values than their male counterparts for all methods. This trend is also observed for Goliath et al. [34] method.

The Cretan and Greek-Cypriot samples produced similar values for the four age categories and the four aging methods, with slightly higher inaccuracy and bias values being reported for the Greek-Cypriot sample (see supplementary material). It must be noted that although sample sizes are uneven, similar results as the ones obtained for the entire data set and separated sex samples are produced. Again, Stout and Paine [13] method shows the highest bias and inaccuracy values. When all individuals included in each sample are compared, Greek-Cypriots show slightly higher inaccuracy and bias values than the Cretans. Among the three formulae of Cho et al. [32], the African-American formula perform slightly better than the others, although the difference is very small.

Further steps on the validation study are carried out on the entire data set based on the similar inaccuracy and bias trends reported for entire sample and for the sex and population sub-datasets. Wilcoxon signed-rank test is performed to verify whether the values for known age differed significantly from the estimated age for the entire sample and the sample divided into age cohorts (**Table 8**). High inaccuracy and bias values are observed for Stout and Paine [13] which underestimated age in all age cohorts with P-values less than the 0.05 threshold. Only individuals under 40 years of age showed non-statistically significant differences between estimated age and known age for Stout et al. [21] and for all the three formulae from Cho et al. [32], although this

outcome might be a result of the small sample size for this specific age cohort. When the sample is divided into 20 year age cohorts, Goliath et al. [34] method produces estimated ages that are statistically significantly different from zero for all the sub-groups. However, this method produces estimated ages not statistically significantly different from known age for over 60 years of age individuals, suggesting that it can accurately estimate age in old specimens.

Table 8. Wilcoxon paired test between known age and estimated age.

						Cho et al. (2002)							
		Stout and Paine (1992)		**Stout et al. (1994)**		European formula		Unknown formula		African-American formula		**Goliath et al. (2016)**	
N	**Age range**	**Z**	**p-value**	**Z**	**p-value**	**Z**	**p-value**	**Z**	**p-value**	**Z**	**p-value**	**Z**	**p-value**
11	*20-39*	-2.93	0.003	-1.33	**0.18**	-1.51	**0.13**	-1.24	**0.21**	-0.09	**0.93**	-2.93	< 0.001
28	*40-59*	-2.93	<0.001	-2.93	< 0.001	-2.93	< 0.001	-2.93	< 0.001	-2.93	< 0.001	-3.69	< 0.001
36	*60-79*	-5.23	<0.001	-5.23	< 0.001	-5.23	< 0.001	-5.23	< 0.001	-5.23	< 0.001	-5.23	< 0.001
13	*80*	-3.18	<0.001	-3.18	< 0.001	-3.18	< 0.001	-3.18	< 0.001	-3.18	< 0.001	-3.18	< 0.001
39	*under 60*	-5.44	<0.001	-5.2	< 0.001	-5.17	< 0.001	-5.17	< 0.001	-4.7	< 0.001	-4.77	< 0.001
49	*over 60*	-6.09	<0.001	-6.09	< 0.001	-6.09	< 0.001	-6.09	< 0.001	-6.09	< 0.001	-0.41	**0.68**
88	*ALL*	-8.14	<0.001	-8.07	< 0.001	-8.06	< 0.001	-8.07	< 0.001	-7.92	< 0.001	-3.88	< 0.001

In the final stage, B&A analysis is used to assess the agreement interval produced by the estimated ages; only those age categories that present non-statistically significant differences between estimated and known age are tested (**Table 8**). For Stout et al. [21] and Cho et al. [32] methods, plots are not presented due to the low number of cases (n = 11). A one sample T-test confirmed that the difference between estimated and know age is non-statistically significant (P-value > 0.05). The upper and lower limits of agreement for the formula developed by Stout et al. [21] range between 10.94 and -18.38 with a bias of -3.72. Cho et al. [32] European formula presents agreement levels of 17.7 and -27.7 with a bias of -5.9. Cho et al. [32] unknown formula limits of agreement range from 15.3 and -23.3 with a bias of -3.9; and the African-American formula presents 23.61 and -24.3 upper and lower limits of agreement with a bias of -0.36. All the cases fall within the limits of agreement calculated for each formula. Among these methods, Stout et al. [21] produced the narrowest limits of agreement.

Due to a larger sample size (n = 49), the over 60 years old age cohort was tested with B&A plot to graphically examine the agreement between estimated and known age for Goliath et al. [34] (**Table 8**). The difference between scores is not statistically significant (P-value > 0.05) as

indicated by sample t-test. **Figure 2** shows the upper and lower limits of agreement corresponding to 21.52 and -21.93 (solid lines). Only two cases fall outside the limits of agreement (4% of the total sample) with a bias of -0.21 (dotted line). The best-fit line (coloured line) indicates a general underestimation of individuals older than 73 years of age.

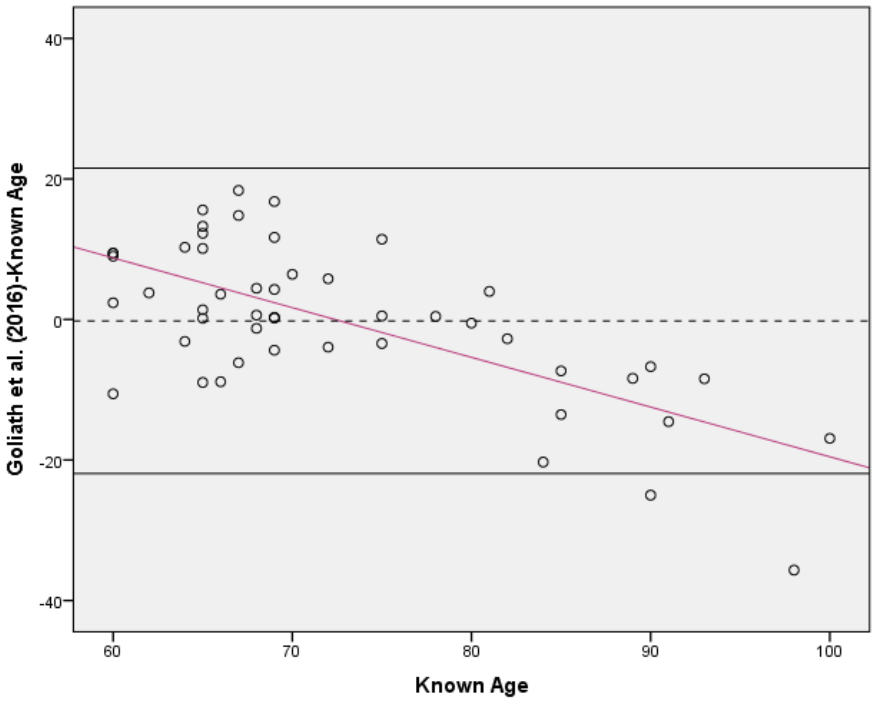

Figure 2. Difference in Goliath et al. (2016) estimated ages and known age against known age for the over 60-year-old sub-sample set.

4. Discussion

Age estimation is one of the first steps in the identification process. Aging methods applied by forensic anthropologists as they testify in court must be validated by the scientific community [2,53]. The methods commonly used to gather anthropological information from skeletal remains are scrutinized to evaluate measurement error, and to identify –when possible- the degree of error related to practitioners experience and/or to technical methodological aspects [54]. This is now accomplished by producing studies that show the rate of repeatability and validity of such research [19,22].

Repeatability and reproducibility are tested in this study through TEM analysis and Intra-class correlation coefficient [43]. Regarding secondary osteon frequency parameters, the results indicate acceptable levels of intra-observer error which is in agreement with other studies [34,55]. rTEM and R as well as ICC values for inter-observer errors ranged from fair to excellent agreement depending on the selected parameter. Osteon remodeling as expressed by OPD constitutes a critical variable in histological research [26,28]. One of its constituent parameters, N.On.Fg and consequently OPD(F) presented the higher inter-observer error in accordance with previous studies [39,56]. As originally expressed by Stout and Paine [13], the definition and the nature of the parameter states that at least 10% of the Haversian canal perimeter must exhibit evidence of resorption, implying the estimation of the missing percentage and possibly misleading intact and fragmentary osteons differentiation. Additionally, individuals from the old cohort present a less structured and more crowded rib cortex with microstructural markers such as fragmentary osteons and interstitial bone hampering the overall histological assessment. However, neither N.On or OPD(I) showed poor agreement between observers, and moreover, OPD showed overall acceptable levels of inter-observer errors as seen by TEM and ICC results. This indicates that the discrepancies regarding OPD(F) are compensated by the inclusion of OPD(I) as seen elsewhere [56,57].

Intra and inter-observer error values for the remaining histomorphometric parameters achieved precision and met overall repeatability standards, supporting that metrics on rib cortical bone might be easily performed [30,56,57]. The only exception is On.Cr, which demonstrated high within subjects' variance for both for intra- and inter-observer errors, as indicated by R, TEM and ICC values. This issue has also been noted by Lagacé et al. [22]. The assessment of On.Cr was performed through microphotographs using ImageJ software which calculates automatically the circularity index including area and perimeter values [60]. Other authors tested different On.Cr

acquisition protocols and highlighted some difficulties for data collection of this specific parameter [61]. A possible drawback in On.Cr data collection noted in this study entails the increasing overlapping of osteons as age increase, and thus, the difficulty of finding osteons with completely intact cement lines on the selected rib area [34]. This is an issue that has also been reported by other authors [62].

Histomorphometric methods are considered experience-based techniques [63]. Combining the information previously published by our group [19] and the observed errors presented in the current study, the same parameters demonstrated inconsistencies for inter-observer, regardless the levels of experience, which has been pointed out by other researchers [39]. This might suggest that factors other than practitioner expertise may play an important role in the method's reliability. Special attention must be placed on the definition of microstructures which may lead to subjective evaluation. Accounting for possible technical issues, the histological assessment was carried out through microscopy and microphotographs as suggested in other studies [18,64]. It is unlikely that the counting of features as fragmentary secondary osteons would be drastically altered due to the data acquisition protocol.

Inherent methodological limitations might have counted as sources of error [39]. As these methods are applied to the assessment of human remains associated with legal cases, the lack of agreement for some of the parameters between readers might indicate the need of further training in bone histology as well as clarification on parameters' description. Laboratory training with an experienced histologist is thus highly recommended.

The focus of future bone histological methods should contribute to standardization, improvement of data collection and reliability as most of the published studies either did not report observer error or only assessed intra-observer error [21,28,32,65,66]. This trend may decrease scientific reliability and limit the use of microscopic methods for expert witness testimony in court room, even if the method accurately estimates the age of unknown individuals [67].

The wide applicability of bone histological analysis for aging [57,65] allows practitioners to adapt histological methods to the case at hands to fit the purpose of the investigation. However, there are aspects such as inter-population variability in remodelling dynamics that may decrease the accuracy of a method when applied to a target population not closely related to the reference population [52]. Some researchers have developed population specific standards for estimating histological AAD to overcome this problem [26,31,32]. In this paper, a validation study to determine the performance and accuracy of four existing histological aging methods developed

from European and African Americans individuals [13,21,32,34] was conducted on two Mediterranean samples. The correlations between the variables and age as well as sex and population sub-samples were investigated, with bias and inaccuracy of the existing methods reported. Special attention must be placed to the main factors affecting the methods´ performance such as differences in demographic characteristics and biological affinity between reference and target samples, as well as inherent limitations related to histological age estimation techniques (i.e. OPD asymptote [17]) and methodological issues related to the statistical approach.

All the histomorphometric variables included in our study are statistically significant correlated to age and they behave as expected for the entire sample. The same can be said when the entire sample is divided by sex and by Cretans and Greek-Cypriots sub-samples. Inaccuracy and bias are commonly used for the evaluation of methods reliability [47]. The outcome from this research shows a general pattern of underestimation for each of the methods explored, except for Goliath et al.'s formula [34].

Regarding the entire Mediterranean sample, the results indicated that Stout and Paine method [13] produced inaccurate age estimates and this finding is in accordance with other studies [51,68–70]. Stout et al. [21] and Cho et al. [32] formulas perform similarly with a slightly lower accuracy and bias values seen for Cho et al.´s African-American equation. Goliath et al. [34] method provides the best performance overall, although it shows a tendency of overestimation of young individuals and underestimation of the oldest age cohort as seen somewhere else [22]. The same tendency is observed for the sub-samples by sex and by Cretans or Cypriots groups, with each age prediction equation displaying discrepancies for specific age categories. Overall, Stout and Paine [13], Stout et al. [21] and Cho et al. [32] methods show a gradual increase of inaccuracy and bias values with increasing individuals´ chronological age. When the sample and sub-samples are truncated at 60 years of age, a substantial decrease in accuracy and bias is seen for the formulas including OPD as a predictor. This observation has been reported by other studies [51] with the implication of OPD rib asymptote occurring at this age and provoking a dissociation between chronological and estimated age [8]. Furthermore, the under 60-years-old age cohort in the Mediterranean sample presented a closer mean age to the mean age of the reference samples explaining to some extent its higher accuracy. In reverse, Goliath et al. [34] formula only includes On.Cr as a single predictor. It was previously suggested that osteon circularity may be a mechanism for providing support to the loading forces and for preventing micro-damage, and thus, it is

positively correlated to age [34,71]. Hence, the better performance of this method is presumed since it is not affected by osteon densities plateau. The low number of young individuals included in Goliath et al. [34] original reference sample (7% of the total) could explain the higher inaccuracy and bias of this formula for the younger Mediterranean cohort.

The general inaccuracy and bias values produced by this validation study do not show major differences when sub-groups are compared (by sex and by Cretans and Greek-Cypriots sub-samples, separately). The fact that inherent sexual and genetic differences could not been discern within the sexes and within population specific sub-samples due to sample size and uneven number of individuals for each sub-group must be taken into consideration.

Male and female rib histomorphometry produced very similar results when the two sex groups were compared, although females exhibited slightly higher inaccuracy and bias values than males when the samples are divided into age categories. The differences between sexes are especially noticeable for the over 80 years old age group. For the three methods including OPD as predictor [13,17,33], this might suggest more variability in remodelling rates within the female group, along with a higher mean age for females (Females = 92 years old, Male = 83 years old). Moreover, the increment in bone remodelling rates related to post-menopause is not steady throughout time and it is also triggered by other diseases that are potentially appearing with advanced age [72]. A similar trend has been shown by other studies for the female group in their samples [56]. Regarding On.Ar, weight, strains and loading complexity might entail factors affecting osteon size [73–75], as well as genetics and physiological mechanisms playing an important role [76]. Additionally, postmenopausal osteoporosis disorder commonly observed in females over 60 years might have impacted osteonal structures morphology and the appearance of the rib cortex [26,77]. The higher variability observed in Goliath et al.'s formula [34] for the oldest Mediterranean female group might be due to intrinsic factors or to age related changes, and further examination of a larger sample size for this particular age cohort is required for confirmation.

Cretans and Cypriots show similar accuracy and bias values as for the entire sample and the sex sub-samples with slight differences observed by age cohorts (note the under-representation of Cypriot individuals younger than 40 years of age produce an unbalanced age distribution). However, the Cretan sample does not perform better. Hence, it seems that there might be a secondary reason, such as the overall skewed age distribution of the Mediterranean samples, possibly affecting the general performance of the methods. Based on the low accuracy of three of the four formulae applied [13,21,32], inter-population differences in bone remodeling between

the European and African-American reference samples and the Mediterranean samples might be considered as their biological affinity is not ensured [78–80]. This matter has been previously reported for other histological studies testing other populations [26]. Bone mass and bone structure have been shown to vary between populations [81–83], with discrepancies between known and estimated age being reported [26,31]. This could be a true statement for formulas accounting for remodeling rates variables such as OPD. It appears, however, that an alternative parameter like On.Cr might possibly minimize the inter-population differences. Recent research found inaccurate age estimates when methods based on degenerative articular changes were applied on the Cretan Collection, suggesting that age indicators may not be strongly correlated to age as they are in the reference sample [29].

Age and sex distribution differences between reference and target samples may have an impact on the methods reliability [23,84]. Stout and Paine [13], Stout et al.[21] and Cho et al. [32] samples differ from a minimum of 10 to a maximum of 30 years from the Mediterranean sample mean age (60 years), with Stout and Paine [13] accounting for the highest mean age difference between reference and target samples. According to B&A results, Stout et al. [21] procedure is recommended for younger adult individuals as assessed by the limits of agreement and bias. All formulae generated by Cho et al. [32] produced similar accuracy and prediction power; however, the ethnicity unknown equation performs the best as noticed by other studies [33]. Goliath et al. [34] sample age distribution matches the mean age of the sample under study supporting its overall good performance. This method estimated age within the limits of agreement for 96% of individuals older than 60 years of age. Regarding sex distribution, Stout and Paine [13] is skewed towards males while Goliath et al .[34] accounts for a more even sex distribution which might also influence our validation study results.

Methodological issues need to be considered when revising the performance of the applied histological methods. One of the underlying reasons for the poor performance for Stout et al. [21] could be explained by sampling area and/or rib number. A previously published sampling error pilot study [19] suggested that other inherent factor as inter-population variability rather than rib topographical location could be the major causes of the reported errors. Regarding rib number, ribs from the Mediterranean samples consist mostly of 6^{th} ribs and it could be thought that certain bias is introduced by the fact that Stout et al. [21] used the 4^{th} costal element. As noted above, standard ribs do not seem to introduce major bias on the methods reliability [18]. Indeed, recent histological aging studies do not specify or report rib number in their standards [33,34].

Differences between histological variables values among the samples are examined for their

potential impact on the methods age estimates. The Mediterranean mean OPD value is of the same general magnitude as those reported by Stout and Paine [13] and Stout et al. [21]. However, the Mediterranean sample presents a higher mean age. A larger difference is noted for mean OPD, On.Ar and Ct.Ar/Tt.Ar values between European and African-American ethnicity groups and the study sample [32]. Moreover, a larger variation was accounted for among Cretans and African-Americans. On.Cr mean value for Goliath et al. [34] reference sample is comparable to that obtained by the study sample supporting a better method's performance (0.905 versus 0.910). Excluding possible inter-observer scoring differences among authors, variability in bone remodeling rates between populations might be the origin of the discrepancies as suggested elsewhere [85].

The last argument concerns the statistical approach undertaken by the four aging prediction methods tested in this research. Stout and Paine [13] method applies logarithmic transformed data and its systematic underestimation of ribs from other communities has been reported by several studies [56,69,86]. Transformation from logarithmic data into arithmetic units might induce an underestimation bias due to the geometric mean used for making the age predictions [87]. Some authors have demonstrated that least square linear regression produces systematic under and over-estimation issues. The use of classical calibration (dependent variable being the age indicator) has been presented as an alternative [88], although confidence intervals and source of errors would be more difficult to be established [89]. The application of Bayes´ theorem should be explored as a possible alternative to avoid systemic under and over-estimation [90].

Conclusions

The reliability and accuracy of four existing age histological methods on two contemporary Mediterranean Cretan and Greek-Cypriot samples are tested. Based on our results, the level of error introduced by observers is dependent of the nature of the histomorphometric variables examined. A similar error pattern is reported for the four aging methods tested producing a systemic under or over estimation of individuals. This inaccuracy could be explained due to inter-population variation in bone remodelling rates, although other methodological issues such as sample demographic characteristics or statistical approach must be taken into consideration as well. This study supports the fact that none of these methods are deemed reliable for broad forensic application when the case under investigation is not-closely related to the reference sample from which the prediction equation is derived from. However, some of the methods might

be applied to those cases in which the individual can be placed in young or old age cohorts, preferably when combined with other available age estimation methods. As a result of our findings, we encourage further histological research for population-specific samples, so additional references samples can be made available for use in international forensic investigation.

Ethics statement: Ethic approval for this research was given by the Ethics committee at the University Hospital of Heraklion Crete and the Ethics committee at the University of Edinburgh.

Acknowledgements: The authors would like to acknowledge the Orthodox Church of Limasol for offering permission for the study of the Greek-Cyprus material. Special thanks to Dr. Despoina Nathena and Mr. Stratos Kougios for collecting the samples and providing the demographic information or the Cretan autopsy material. We also would like to thank the reviewers of our manuscript for their thoughtful comments which guided us as we improved this paper.

References

[1] S. Byers, Introduction to Forensic Anthropology: A textbook, 4th edn., Routledge, New York, 2011.

[2] C. Merritt, Testing the accuracy of adult skeletal age estimation methods: original methods versus revised and newer methods, Vis-à-Vis Explor. Anthropol. 12 (2013) 102–119. https://doi.org/http://vav.library.utoronto.ca/index.php/vav/article/view/19146.

[3] N. Márquez-Grant, An overview of age estimation in forensic anthropology: perspectives and practical considerations, Ann. Hum. Biol. 42 (2015) 308–322. https://doi.org/10.3109/03014460.2015.1048288.

[4] C.M. Crowder, J. Heinrich, S. Stout, Rib Histomorphometry for Adult Age Estimation, in: L. Bell (Ed.), Forensic Microsc. Skelet. Tissues. Methods Protoc., Humana Press, London, 2012: pp. 109–128.

[5] S.D. Stout, M.A. Porro, B. Perotti, Brief Communication: A test and correction of the clavicle method for histological age determination of skeletal remains, Am. J. Phys. Anthropol. 100 (1996) 139–142.

[6] C.M. Crowder, Estimation of age at death using cortical bone histomorphometry, New York, NY, 2013.

[7] E.R. Kerley, The microscopic determination of age in human bone, Am. J. Phys. Anthropol. 23 (1965) 149–164.

[8] M. Streeter, Histological Age-at-death estimation, in: C.M. Crowder, S.D. Stout (Eds.), B. Histol. An Anthropol. Perspect., CRC Press, U.S.A., 2012: pp. 135–152.

[9] H. Mcgivern, C. Greenwood, N. Márquez-grant, E.F. Kranioti, B. Xhemali, P. Zioupos, Age-Related Trends in the Trabecular Micro-Architecture of the Medial Clavicle : Is It of Use in Forensic Science ?, Front. Bioeng. Biotechnol. 7 (2020) 1–8. https://doi.org/10.3389/fbioe.2019.00467.

[10] R.R. Paine, How to equip a basic histological lab for the anthropological assessment of human bone and teeth, J. Anthropol. Sci. 85 (2007) 213–219.

[11] H.I. Hollund, N. Arts, M.M.E. Jans, H. Kars, Are teeth better? histological characterization of diagenesis in archaeological bone-tooth pairs and a discussion of the consequences for archaeometric sample selection and analyses, Int. J. Osteoarchaeol. 25 (2015) 901–911. https://doi.org/10.1002/oa.2376.

[12] R. Pitfield, C. Deter, P. Mahoney, Bone histomorphometric measures of physical activity in children from Medieval England, Am. J. Phys. Anthropol. (2019) 1–17.

[13] S.D. Stout, R.R. Paine, Brief communication: Histological age estimation using rib and clavicle, Am. J. Phys. Anthropol. 87 (1992) 111–115.

[14] R.S. Weinstein, D.J. Simmons, C.O. Lovejoy, Ancient bone disease in a Peruvian mummy revealed by quantitative skeletal histomorphometry, Am. J. Phys. Anthropol. 54 (1981) 321–326.

[15] I.J. Singh, D.L. Gunberg, Estimation of age at death in human males from quantitative histology of bone fragments, Am. J. Phys. Anthropol. 33 (1970) 373–381. https://doi.org/10.1002/ajpa.1330330311.

[16] D.D. Thompson, C.A. Galvin, Estimation of Age At Death By Tibial Osteon Remodeling in an Autopsy Series, Forensic Sci. Int. 22 (1983) 203–211. https://doi.org/10.1016/0379-0738(83)90015-4.

[17] S.D. Stout, R.R. Paine, Brief communication: Bone Remodelling rates: A test of an algorith for estimating missing osteons, Am. J. Phys. Anthropol. (1994) 123–129.

[18] C.M. Crowder, L. Rosella, Assessment of intra- and intercostal variation in rib histomorphometry: Its impact on evidentiary examination, J. Forensic Sci. 52 (2007) 271–

276. https://doi.org/10.1111/j.1556-4029.2007.00388.x.
[19] J.G. García-Donas, J. Dyke, R.R. Paine, D. Nathena, E. Kranioti, Accuracy and sampling error of two age estimation techniques using rib histomorphometry on a modern sample, J. Forensic Leg. Med. 38 (2016) 28–35. https://doi.org/10.1016/j.jflm.2015.11.012.
[20] E.D. Sedlin, H.M. Frost, A.R. Villanueva, Age changes in resorption in human rib cortex, J. Gerontol. 18 (1963) 345–349. https://doi.org/10.1093/geronj/18.4.345.
[21] S.D. Stout, W.H. Dietze, M.Y. Iscan, S.R. Loth, Estimation of age at death using cortical histomorphometry of the sternal end of the fourth rib, J. Forensic Sci. 39 (1994) 778–784.
[22] F. Lagacé, E. Verna, P. Adalian, E. Baccino, L. Martrille, Testing the accuracy of a new histomorphometric method for age-at-death estimation, Forensic Sci. Int. 296 (2019) 48–52. https://doi.org/10.1016/j.forsciint.2019.01.020.
[23] M. Bouvier, D.H. Ubelaker, A comparison of two methods for the microscopic determination of age at death., Am. J. Phys. Anthropol. 46 (1977) 391–4. https://doi.org/10.1002/ajpa.1330460303.
[24] H. Cho, S.D. Stout, T.A. Bishop, Cortical bone remodeling rates in a sample of African American and European American descent groups from the American midwest: Comparisons of age and sex in ribs, Am. J. Phys. Anthropol. 130 (2006) 214–226. https://doi.org/10.1002/ajpa.20312.
[25] R.R. Paine, B.P. Brenton, The paleopathology of pellagra: investigating the impact of prehistoric and historical dietary transitions to maize, J. Anthropol. Sci. 84 (2006) 125–135.
[26] Y.S. Kim, D.I. Kim, D.K. Park, J.H. Lee, N.E. Chung, W.T. Lee, S.H. Han, Assessment of histomorphological features of the sternal end of the fourth rib for age estimation in Koreans, J. Forensic Sci. 52 (2007) 1237–1242. https://doi.org/10.1111/j.1556-4029.2007.00566.x.
[27] P. Beauchesne, S. Saunders, A test of the revised Frost's "rapid manual method" for the preparation of bone thin sections, Int. J. Osteoarchaeol. 16 (2006) 82–87. https://doi.org/10.1002/oa.804.
[28] Y. Watanabe, M. Konishi, M. Shimada, H. Ohara, S. Iwamoto, Estimation of age from the femur of Japanese cadavers, Forensic Sci. Int. 98 (1998) 55–65. https://doi.org/10.1016/S0379-0738(98)00136-4.
[29] E. Michopoulou, P. Negre, E. Nikita, E.F. Kranioti, The auricular surface as age indicator in a modern Greek sample: A test of two qualitative methods, Forensic Sci. Int. 280 (2017) 246.e1-246.e7. https://doi.org/10.1016/j.forsciint.2017.08.004.
[30] S. Hisham, N. Abdullah, M.H. Mohamad Noor, D. Franklin, Quantification of Pubic Symphysis Metamorphosis Based on the Analysis of Clinical MDCT Scans in a Contemporary Malaysian Population, J. Forensic Sci. 64 (2019) 1803–1811. https://doi.org/10.1111/1556-4029.14125.
[31] M.V. Pavón, A. Cucina, V. Tiesler, New formulas to estimate age at death in Maya populations using histomorphological changes in the fourth human rib, J. Forensic Sci. 55 (2010) 473–477. https://doi.org/10.1111/j.1556-4029.2009.01265.x.
[32] H. Cho, S.D. Stout, R.W. Madsen, M. a Streeter, Population-specific histological age-estimating method: a model for known African-American and European-American skeletal remains, J. Forensic Sci. 47 (2002) 12–18.
[33] S. Pfeiffer, J. Heinrich, A. Beresheim, M. Alblas, Cortical bone histomorphology of known-age skeletons from the Kirsten collection, Stellenbosch university, South Africa, Am. J. Phys. Anthropol. 147 (2016) 137–147. https://doi.org/10.1002/ajpa.22951.
[34] J. Goliath, M. Stewart, S. Stout, Variation in osteon histomorphometrics and their impact on age-at-death estimation in older individuals, Forensic Sci. Int. 262 (2016) 282.e1-

282.e6. https://doi.org/10.1016/j.forsciint.2016.02.053.
[35] J.G. Garcia-Donas, Age estimation on two Mediterranean samples using rib histomorphometry, University of Edinburgh, 2018.
[36] E.F. Kranioti, M.Y. Işcan, M. Michalodimitrakis, Craniometric analysis of the modern Cretan population, Forensic Sci. Int. 180 (2008) 1–5. https://doi.org/10.1016/j.forsciint.2008.06.018.
[37] 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, Forensic Sci. Int. 271 (2017). https://doi.org/10.1016/j.forsciint.2016.11.018.
[38] J.G. García-Donas, A. Dalton, I. Chaplin, E.F. Kranioti, A revised method for the preparation of dry bone samples used in histological examination: Five simple steps, HOMO- J. Comp. Hum. Biol. 68 (2017). https://doi.org/10.1016/j.jchb.2017.07.001.
[39] N. Lynnerup, J.L. Thomsen, B. Frohlich, Intra- and inter-observer variation in histological criteria used in age at death determination based on femoral cortical bone, Forensic Sci. Int. 91 (1998) 219–230. https://doi.org/10.1016/S0379-0738(97)00197-7.
[40] S. Pfeiffer, C.M. Crowder, L. Harrington, M. Brown, Secondary osteon and Haversian canal dimensions as behavioural indicators, Am. J. Phys. Anthr. 131 (2006) 460–468. https://doi.org/10.1002/ajpa.
[41] M.C. Stewart, L.E. McCormick, J.R. Goliath, P.W. Sciulli, S.D. Stout, A Comparison of Histomorphometric Data Collection Methods, J. Forensic Sci. 58 (2013) 109–113. https://doi.org/10.1111/j.1556-4029.2012.02195.x.
[42] M. Cummaudo, C. Raffone, A. Cappella, N. Márquez-Grant, C. Cattaneo, Histomorphometric analysis of the variability of the human skeleton: Forensic implications, Leg. Med. 45 (2020) 101711. https://doi.org/10.1016/j.legalmed.2020.101711.
[43] M.A. Bujang, N. Baharum, A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review, Arch. Orofac. Sci. 12 (2017) 1–11.
[44] S.J. Ulijaszek, D.A. Kerr, Anthropometric measurement error and the assessment of nutritional status, Br. J. Nutr. 44 (1999) 165–177.
[45] M.J. Haber, H.X. Barnhart, Coefficients of agreement for fixed observers, Stat. Methods Med. Res. 15 (2006) 255–271. https://doi.org/10.1191/0962280206sm441oa.
[46] G. Perinetti, StaTips Part IV: Selection, interpretation and reporting of the intraclass correlation coefficient, South Eur. J. Orthod. Dentofac. Res. 5 (2018) 3–5. https://doi.org/10.5937/sejodr5-17434.
[47] C. Lovejoy, S. Meindl, R. Mensforth, T. Barton, Multifactoral Determination of Skeletal Age at Death: A Method and Blind Test of its Accuracy, Am. J. Phys. Anthropol. 68 (1985) 1–14. https://doi.org/10.1002/ajpa.1330680102.
[48] J.H. McDonald, Handbook of biological statistics, 3rd edn., Sparky House Publishing, Maryland, Baltimore, 2014.
[49] J. Martin Bland, D. Altman, Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement, Lancet. 327 (1986) 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8.
[50] J.M. Bland, D.G. Altman, Comparing methods of measurement: why plotting difference against standard method is misleading, Lancet. 346 (1995) 1085–1087. https://doi.org/10.1016/S0140-6736(95)91748-9.
[51] C.M. Crowder, S. Pfeiffer, The application of cortical bone histomorphometry to estimate age at death, in: K. Latham, M. Finnegan (Eds.), Age Estim. Hum. Skelet., Charles C

Thomas, 2010: pp. 193–215.
[52] A.G. Robling, S.D. Stout, Histomorphometry of human cortical bone: applications to age estimation, in: M. and S.S. Katzenberg (Ed.), Biol. Anthropol. Hum. Skelet., 2nd edn., Wiley-Liss, New Yersey, 2008: pp. 149–182.
[53] A.M. Christensen, C.M. Crowder, Evidentiary standards for forensic anthropology, J. Forensic Sci. 54 (2009) 1211–1216. https://doi.org/10.1111/j.1556-4029.2009.01176.x.
[54] T. Krishan, T. Kanchan, Measurement Error in Anthropometric Studies and its Significance in Forensic Casework, Ann. Med. Heal. Sci. Res. 1 (2016) 62–63. https://doi.org/10.4103/2141-9248.180277.
[55] N. Keough, E.N. L'Abbé, M. Steyn, The evaluation of age-related histomorphometric variables in a cadaver sample of lower socioeconomic status: implications for estimating age at death, Forensic Sci. Int. 191 (2009) 12–15. https://doi.org/10.1016/j.forsciint.2009.07.012.
[56] C.M. Crowder, Evaluating the use of quantitative bone histology to estimate adult age at death, University of Toronto, 2005.
[57] C. Cannet, J.P. Baraybar, M. Kolopp, P. Meyer, B. Ludes, Histomorphometric estimation of age in paraffin-embedded ribs: A feasibility study, Int. J. Legal Med. 125 (2011) 493–502. https://doi.org/10.1007/s00414-010-0444-6.
[58] M. Cummaudo, A. Cappella, M. Biraghi, C. Raffone, N. Màrquez-Grant, C. Cattaneo, Histomorphological analysis of the variability of the human skeleton: forensic implications, Int. J. Legal Med. 132 (2018). https://doi.org/10.1007/s00414-018-1781-0.
[59] M.C. Stewart, L.E. McCormick, J.R. Goliath, P.W. Sciulli, S.D. Stout, A {Comparison} of {Histomorphometric} {Data} {Collection} {Methods}*: {COMPARISON} {OF} {HISTOMORPHOMETRIC} {METHODS}, J Forensic Sci. 58 (2013) 109–113. https://doi.org/10.1111/j.1556-4029.2012.02195.x.
[60] W. Rasband, https://imagej.nih.gov/ij/plugins/circularity.html, (2000).
[61] C.S. Mears, S.M. Litton, C.M. Phippen, T.D. Langston, K.E. Keenan, J.G. Skedros, Improving accuracy , precision , and efficiency in analysis of osteon cross-sectional shape', Am. J. Phys. Anthropol. 154 (2014) 223.
[62] J. Goliath, Variation in osteon circularity and its impact on estimating age at death, The Ohio State University, 2010.
[63] H.M. Garvin, N. V. Passalacqua, Current practices by forensic anthropologists in adult skeletal age estimation, J. Forensic Sci. 57 (2012) 427–433. https://doi.org/10.1111/j.1556-4029.2011.01979.x.
[64] K. Absolonova, P. Veleminsky, M. Dobisikova, M. Beran, J. Zocova, Histological Estimation of Age at Death from the Compact Bone of Burned and Unburned Human Ribs, J. Forensic Sci. 58 (2013) 135–145. https://doi.org/10.1111/j.1556-4029.2012.02303.x.
[65] S.H. Han, S.H. Kim, Y.W. Ahn, G.Y. Huh, D.S. Kwak, D.K. Park, U.Y. Lee, Y.S. Kim, Microscopic age estimation from the anterior cortex of the femur in Korean adults, J. Forensic Sci. 54 (2009) 519–522. https://doi.org/10.1111/j.1556-4029.2009.01003.x.
[66] U.Y. Lee, G.U. Jung, S.G. Choi, Y.S. Kim, Anthropological age estimation with bone histomorphometry from the human clavicle, Anthropologist. 17 (2014) 929–936.
[67] A.M. Christensen, C.M. Crowder, S.D. Ousley, M.M. Houck, Error and its meaning in forensic science, J. Forensic Sci. 59 (2014) 123–126. https://doi.org/10.1111/1556-4029.12275.
[68] J.C. Dudar, S. Pfeiffer, S.R. Saunders, Evaluation of morphological and histological adult skeletal age-at-death estimation techniques using ribs., J. Forensic Sci. 38 (1993) 677–85.
[69] D.G. Pratte, S. Pfeiffer, Histological age estimation of a cadaveral sample of diverse origins, J. Can. Soc. Forensic Sci. 32 (1999) 155–167.

https://doi.org/10.1080/00085030.1999.10757496.
[70] R.R. Paine, B.P. Brenton, Dietary health does affect histological age assessment: An evaluation of the Stout and Paine (1992) age estimation equation using secondary osteons from the rib, J. Forensic Sci. 51 (2006) 489–492. https://doi.org/10.1111/j.1556-4029.2006.00118.x.
[71] M.B. Schaffler, K. Choi, C. Milgrom, Aging and matrix microdamage accumulation in human compact bone, Bone. 17 (1995) 521–525. https://doi.org/10.1016/8756-3282(95)00370-3.
[72] R. Recker, J. Lappe, K.M. Davies, R. Heaney, Bone Remodeling Increases Substantially in the Years After Menopause and Remains Increased in Older Osteoporosis Patients, J. Bone Miner. Res. 19 (2004) 1628–1633. https://doi.org/10.1359/JBMR.040710.
[73] J.D. Currey, Some effects of ageing in human haversian systems, J. Anat. 98 (1964) 69–75.
[74] R.F.M. Van Oers, R. Ruimerman, B. van Rietbergen, P.A.J. Hilbers, R. Huiskes, Relating osteon diameter to strain, Bone. 43 (2008) 476–482. https://doi.org/10.1016/j.bone.2008.05.015.
[75] H.M. Britz, C.D.L. Thomas, J.G. Clement, D.M.L. Cooper, The relation of femoral osteon geometry to age, sex, height and weight, Bone. 45 (2009) 77–83. https://doi.org/10.1016/j.bone.2009.03.654.
[76] V.M. Dominguez, A.M. Agnew, Examination of factors potentially influencing osteon size in the human rib', Anat. Rec. 299 (2016) 313–324. https://doi.org/10.1002/ar.23305.
[77] H.M. Frost, Bone's mechanostat: A 2003 update, Anat. Rec. Part Ahe Anat. Rec. Part A. 275 (2003) 1081–1101. https://doi.org/10.1002/ar.a.10119.
[78] J.R. Hughey, P. Paschou, P. Drineas, D. Mastropaolo, D.M. Lotakis, P.A. Navas, M. Michalodimitrakis, J.A. Stamatoyannopoulos, G. Stamatoyannopoulos, A European population in minoan Bronze age crete, Arch. Hell. Med. 30 (2013) 456–466. https://doi.org/10.1038/ncomms2871.
[79] G. Hellenthal, G.B.J. Busby, G. Band, J.F. Wilson, C. Capelli, D. Falush, S. Myers, A genetic atlas of human admixture history, Science (80-.). 343 (2014) 747–751. https://doi.org/10.1126/science.1243518.
[80] A. Heraclides, E. Bashiardes, E. Fernández-Domínguez, S. Bertoncini, M. Chimonas, V. Christofi, J. King, B. Budowle, P. Manoli, M.A. Cariolou, Y-chromosomal analysis of Greek Cypriots reveals a primarily common pre-Ottoman paternal ancestry with Turkish Cypriots, PLoS One. 12 (2017) 1–26. https://doi.org/10.1371/journal.pone.0179474.
[81] R. Weinstein, N. Bell, Diminished rates of bone formation in normal black adults, N. Engl. J. Med. 319 (1988) 1698–1701.
[82] C.M. Schnitzler, Bone quality: A determinant for certain risk factors for bone fragility, Calcif. Tissue Int. 53 (1993). https://doi.org/10.1007/BF01673398.
[83] M.F. Ericksen, Histologic estimation of age at death using the anterior cortex of the femur, Am. J. Phys. Anthropol. 84 (1991) 171–179.
[84] L.C. Aiello, T. Molleson, Are microscopic ageing techniques more accurate than macroscopic ageing techniques?, J. Archaeol. Sci. 20 (1993) 689–704. https://doi.org/10.1006/jasc.1993.1043.
[85] D. Botha, N. Lynnerup, M. Steyn, Inter-population variation of histomorphometric variables used in the estimation of age-at-death, Int. J. Legal Med. (2019) 709–719. https://doi.org/10.1007/s00414-019-02048-7.
[86] J.C. Dudar, S. Pfeiffer, S.R. Saunders, Evaluation of morphological and histological adult skeletal age-at-death estimation techniques using ribs, J. Forensic Sci. 38 (1993) 677–85.
[87] R.J. Smith, Logarithmic transformation bias in allometry, Am. J. Phys. Anthropol. 90 (1993) 215–228. https://doi.org/10.1002/ajpa.1330900208.

[88] N. Rogers, S. Stout, Selecting a calibration method to reduce bias in histological aging methods, Am. J. Phys. Anthropol. 105(s26) (1998) 191.
[89] R.G. Aykroyd, D. Lucy, Technical note: regression analysis in adult age estimation, Am. J. Phys. Anthropol. 104 (1997) 259–265.
[90] R. Samworth, R. Gowland, Estimation of adult skeletal age-at-death: Statistical assumptions and applications, Int. J. Osteoarchaeol. 17 (2007) 174–188. https://doi.org/10.1002/oa.867.

Supplementary material

Inaccuracy and bias values for the Mediterranean sample divided by sexes: age cohorts, </> 60 years and the entire sample (male and female, respectively).

							Cho et al. (2002)							
			Stout and Paine (1992)		**Stout et al. (1994)**		European Formula		Unknown Formula		African-American Formula		**Goliath et al. (2016)**	
	N	**Age range**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**
MALES (n=40)	5	20-39	10.39	-10.39	6.97	-4.43	13.17	-6.86	10.57	-4.46	12.59	2.14	22.24	22.24
	12	40-59	32.43	-32.43	24.79	-24.79	23.40	-23.40	22.48	-22.48	17.13	-17.13	12.92	9.98
	17	60-79	42.62	-42.62	27.39	-27.39	25.42	-25.42	26.55	-26.55	24.93	-24.93	7.46	6.58
	6	80	56.61	-56.61	35.08	-35.08	34.99	-34.99	37.24	-37.24	38.53	-38.53	8.24	-6.91
	17	<60	25.95	-25.95	19.55	-18.80	20.39	-18.54	18.97	-17.18	15.80	-11.47	15.66	13.59
	23	>60	46.27	-46.27	29.40	-29.40	27.92	-27.92	29.34	-29.34	28.48	-28.48	7.66	3.06
	40	ALL	37.63	-37.63	25.21	-24.89	24.72	-23.93	24.93	-24.17	23.09	-21.25	11.06	7.54
FEMALES (n=48)	6	19-39	10.37	-10.37	5.54	-3.14	8.65	-5.11	7.17	-3.51	8.47	-2.45	22.76	22.76
	16	40-59	30.31	-30.31	21.16	-21.16	22.47	-22.47	21.08	-21.08	19.90	-19.90	10.21	8.50
	19	60-79	42.84	-42.84	26.98	-26.98	23.51	-23.51	25.23	-25.23	24.80	-24.80	6.35	1.79
	7	80	66.51	-66.51	46.96	-46.96	41.65	-41.65	45.13	-45.13	45.41	-45.41	16.38	-16.38
	22	<60	24.87	-24.87	16.90	-16.25	18.70	-17.74	17.28	-16.29	16.78	-15.14	13.63	12.39
	26	>60	49.21	-49.21	32.36	-32.36	28.40	-28.40	30.59	-30.59	30.35	-30.35	9.05	-3.10
	48	ALL	38.05	-38.05	25.27	-24.97	23.95	-23.51	24.49	-24.04	24.13	-23.38	11.15	4.00

Inaccuracy and bias values for the Mediterranean sample divided by population sample: age cohorts, </> 60 years and the entire sample (Cretans and Cypriots, respectively).

							Cho et al. (2002)							
			Stout and Paine (1992)		Stout et al. (1994)		European Formula		Unknown Formula		African-American Formula		Goliath et al. (2016)	
	N	Age range	Inaccuracy	Bias	Inaccuracy	Bias	Inaccuracy	Bias	Inaccuracy	Bias	Inaccuracy	Bias	Inaccuracy	Bias
CRETE (n=41)	10	19-39	11.19	-11.19	6.51	-4.39	11.75	-6.52	9.29	-4.63	10.71	-1.06	21.51	21.51
	11	40-59	30.94	-30.94	22.08	-22.08	19.65	-19.65	19.41	-19.41	15.93	-15.93	10.65	8.78
	12	60-79	43.42	-43.42	29.14	-29.14	21.07	-21.07	23.96	-23.96	24.48	-24.48	4.77	2.36
	8	80	62.44	-62.44	45.45	-45.45	40.83	-40.83	43.32	-43.32	44.03	-44.03	12.64	-11.65
	21	under 60	21.54	-21.54	14.67	-13.66	15.89	-13.40	14.59	-12.38	13.45	-8.85	15.82	14.84
	20	over 60	51.02	-51.02	35.66	-35.66	28.97	-28.97	31.70	-31.70	32.30	-32.30	7.92	-3.24
	41	ALL	35.92	-35.92	24.91	-24.39	22.27	-21.00	22.94	-21.80	22.64	-20.29	11.97	6.02
CYPRUS (n=47)	1	20-39	2.26	-2.26	2.94	2.94	0.25	0.25	2.94	2.94	6.64	6.64	32.67	32.67
	17	40-59	31.40	-31.40	23.13	-23.13	24.95	-24.95	23.14	-23.14	20.52	-20.52	11.84	9.36
	24	60-79	42.39	-42.39	26.19	-26.19	26.09	-26.09	26.80	-26.80	25.05	-25.05	7.92	4.90
	5	80	61.16	-61.16	35.12	-35.12	34.97	-34.97	38.56	-38.56	39.37	-39.37	12.59	-12.59
	18	under 60	29.78	-29.78	22.01	-21.68	23.58	-23.55	22.02	-21.69	19.75	-19.01	12.99	10.66
	29	over 60	45.63	-45.63	27.73	-27.73	27.62	-27.62	28.83	-28.83	27.52	-27.52	8.72	1.88
	47	ALL	39.56	-39.56	25.54	-25.41	26.07	-26.06	26.22	-26.10	24.54	-24.26	10.36	5.25