Journal of Archaeological Science 171 (2024) 106080

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: www.elsevier.com/locate/jas

'JSDNE': A novel R package for estimating age quantitatively with the auricular surface by Dirichlet normal energy

Jisun Jang [a,*], Enrico Mariconti [b], Rebecca Watts [a]

[a] Institute of Archaeology, University College London, 31-34 Gordon Sq, London, WC1H 0PY, United Kingdom
[b] Department of Security and Crime Science, University College London, 35 Tavistock Sq, London, WC1H 9EZ, United Kingdom

ARTICLE INFO

Keywords:
Age estimation method
JSDNE package
R package
Curvature analysis
Auricular surface
Dirichlet normal energy
Quantitative method
Computational age estimation

ABSTRACT

Age estimation plays a significant role in forensic anthropology and bioarchaeology. However, widely-used traditional methods involving macroscopic observation suffer from subjectivity and statistical bias. The present research aims to minimize both issues by applying computational and mathematical approaches. A laser scanner was used to reconstruct 890 auricular surfaces of adult individuals from three known-age European skeletal collections. Dirichlet Normal Energy (DNE) was applied to assess the curvature of the auricular surface and its relationship with known age-at-death. Ten variables had high correlations, including total DNE per Total polygon faces, Mean value of DNE on apex, proportion of polygon faces with DNE of less than 0.0001 and proportion of polygon faces with DNE of over 0.6. The variables were used to develop age prediction models which are freely available in a novel R package, JSDNE. The package predicts age mathematically, objectively, and user-independently. It includes three functions: principal component quadratic discriminant analysis (PCQDA), principal component regression analysis (PCR), and principal component logistic regression analysis (PCLR), which produce age estimates with 91%, 76%, and 92.9% levels of accuracy, respectively. JSDNE package (https://cran.r-project.org/package=JSDNE) can be downloaded automatically using install.packages("JSDNE"). The detailed code and the raw data of this study are openly available at https://github.com/jisunjang19/cran-JSDNE, doi: 10.5281/zenodo.12708779 or 'JSDNE' package.

1. Introduction

Age estimation, together with sex and stature estimation, is crucial in constructing the biological profile of human remains. In bioarchaeology, the biological profile is the foundation for reconstructing past demographic information, leading to the proper understanding of human life history. Forensic anthropology uses the biological profile to identify unknown individuals, helping to investigate cases of missing persons, mass disasters, genocide, and violent crime (Márquez-Grant, 2015; Moraitis et al., 2014; Rissech et al., 2012). With increased globalisation, natural disasters, and international terrorism, attention has been paid to developing more accurate and objective age estimation methods (Zhang et al., 2022).

Age estimation methods have been developed based on observing bone changes that occur throughout life (Byers, 2017). These bone changes can be divided into developmental and degenerative changes (Moraitis et al., 2014); developmental changes are based on the maturation rate of bones, such as epiphyseal ossification (Schmeling et al., 2004), and tooth eruption (Kumar and Sridhar, 1990). These are used to estimate subadult age. As for adults who have completed growth, age is estimated by examining the degree of degenerative change (Schwartz, 1995), including the pubic symphysis (Brooks and Suchey, 1990), the sternal ends of the ribs (Işcan et al., 1989), and the auricular surface (Lovejoy et al., 1985). Among these, the auricular surface is considered one of the most significant indicators as it is dense, highly durable and often well-preserved against decomposition (Rivera-Sandoval et al., 2018). Moreover, it is argued that degenerative changes on the auricular surface continue after 60 years of age, making it possible to estimate older adults' age more accurately than other age indicators (Rissech et al., 2012). Even though age estimation methods have been developed and revised over several decades, there is still no way to precisely interpret the speed and degree of change for different individuals. This is because bone changes are affected by intrinsic and extrinsic factors, such as genes, diet, socioeconomics, and activity throughout the life, with rates of change varying between individuals. Therefore, estimating adult age can be complicated and inaccurate (Miranker, 2016). Age

* Corresponding author.
E-mail addresses: jisun.jang.19@ucl.ac.uk (J. Jang), e.mariconti@ucl.ac.uk (E. Mariconti), r.watts@ucl.ac.uk (R. Watts).

https://doi.org/10.1016/j.jas.2024.106080
Received 21 February 2024; Received in revised form 5 September 2024; Accepted 6 September 2024
Available online 10 September 2024

estimation methods traditionally suffer from two common issues: subjectivity and statistical bias.

Most age estimation methods are based on visual assessment, which increases subjectivity. When assessing degenerative changes observers rely on their own expertise and experience to interpret the morphological features. Thus, the interpretation can differ depending on the observer, and the estimated result also varies. Even if the same observer assesses the same bone on different occasions, the results can be different. These are called interobserver and intraobserver errors, respectively (Garvin and Passalqua, 2012; Moraitis et al., 2014). Statistical bias occurs due to differences in the age structure of the reference and target samples. Fundamentally, the estimated results for a target sample can be biased by the spread of ages of individuals that made up the reference sample. This is known as age structure mimicry (Mensforth, 1990:90). Regional and diachronic variations can also influence the degree of degeneration displayed by different morphological features, biasing results towards the reference sample and causing inaccurate and unreliable results. Overall, these issues impact the accuracy, reliability, reproducibility and repeatability of results. To address these limitations, more objective and less observer-dependent age estimation methods have been sought (Botha et al., 2019; Curate et al., 2022).

Recently, virtual anthropology has gained attention due to its potential to improve and revise traditional approaches through the application of computational imaging technology, such as computed tomography and laser scanning (Warrier et al., 2023). This allows a mathematical approach to extract subtle surface changes and more complex visual information not limited by human vision (Zhang et al., 2022; Warrier et al., 2023). It is also more user-independent, so the extracted results are objective, reproducible, and repeatable (Zhang et al., 2022). Currently, the most widely used mathematical approach in age estimation is curvature analysis of the joint surface. Curvature analysis is used to evaluate the degree to which surfaces are undulated or flat, and can therefore be used to analyse surface texture irregularities (Villa et al., 2015). To date, most research has focused on the pubic symphysis (Biwasaka et al., 2013; Stoyanova et al., 2015; Stoyanova et al., 2017; Slice and Algee-Hewitt, 2015), with only three papers exploring curvature analysis on the auricular surface (Jang et al., 2024; Štepanovský et al., 2023; Villa et al., 2015). Villa et al. (2015) used Gaussian curvature analysis to examine the correlations between five curvature variables and age: mean, highest, lowest, convex, and flat curvatures. The variables had moderate correlations with age, though it was noted that both young and older adults displayed high curvature values; young individuals because of transverse organisation, and older individuals due to increased macroporosity. Middle-aged individuals, who generally showed no transverse organisation or macroporosity (reflecting a flatter auricular surface), had lower curvature values. However, due to the sensitivity limitations of the equipment, detailed surface features such as surface texture and microporosity could not be analysed (Villa et al., 2015). While Villa et al. (2015) were able to show a correlation between curvature values and adult age, Štepanovský et al. (2023) developed this method further to produce age estimation software, CoxAGE3D, which uses an automatic data mining model. CoxAGE3D is fully automated, with age estimations produced after the user uploads a 3D laser scan of an auricular surface to the program. Calculations are based on multi-population auricular surfaces from five known-age Asian and European collections, reducing statistical bias levels. While this method appears to be able to identify a number of age-related coefficients, accuracy rates were not provided and it is yet to be subjected to independent validation.

Jang et al. (2024) examined age-related change to the auricular surface using a different type of curvature analysis, Dirichlet Normal Energy (DNE). This requires a surface to be converted into polygon faces before DNE is used to measure how much the surface bends, which it does by quantifying the deviation of surface energy from a stable state, which is 0 or flat. Each polygon face (ρ) has two principal curvatures: the largest curvature (ν) and the smallest curvature (υ), and the points at which these meet are called vertices (Fig. 1). Each vertex has a normal (red arrow), which is the average value of the adjacent polygonal normal vectors (blue arrows). A polygon is formed by translating the normal vectors to a common vertex. This polygon has edge vectors (n_ν and n_υ), which are derivatives of ν and υ (Bunn et al., 2011; Pampush et al., 2016; Winchester et al., 2014). The Dirichlet Energy Density (DED) is obtained from the sum of square roots of edge vectors. Then, DNE is calculated as the total sum of the values multiplied by DNE and area (Pampush et al., 2016:398):

$$\mathrm{DED}(\rho) = (\boldsymbol{n}_\nu)^2 + (\boldsymbol{n}_\upsilon)^2$$

$$\mathrm{DNE} = \sum \mathrm{DED}(\rho) \text{ x } \mathrm{area}(\rho)$$

This makes it possible to identify subtle changes in surface curvature more sensitively than other methods of curvature analysis (Bunn et al., 2011; Godfrey et al., 2012). As Jang et al. (2024) tested this method on archaeological skeletons with no biological information, age was estimated using the Lovejoy et al. (1985) and Buckberry and Chamberlain (2002) methods. It was found that four DNE variables correlated with the estimated age phases, those relating to surface undulation, apical activity, macroporosity, and the total DNE per total number of polygon faces. In terms of total DNE per total number of polygon faces, younger adults had higher values than middle-aged adults due to the presence of transverse organisation (surface undulation), which occurred more frequently in the former. Younger adults had the highest DNE values for surface undulation, with these values falling with increasing age. Middle-aged adults had lower total DNE values because, as initially noted by Villa et al. (2015), their auricular surfaces were flatter due to a lack of transverse organisation or macroporosity. Older adults had the highest total DNE values and the highest DNE values for macroporosity and apical activity, which both showed a positive correlation with age phase (Jang et al., 2024). Jang et al. (2024) tested the levels of intra-observer error associated with this method and found very high levels of agreement. However, there was still a degree of user dependency as the user must manually select a region of interest (ROI) before DNE values can be calculated (Jang et al., 2024). Moreover, as this study used archaeological skeletons, the age of the individuals was not documented, potentially obscuring relationships between DNE values and age at death.

To address these limitations, the current paper builds on the work of Jang et al. (2024) by reducing the need to select a ROI manually, and examining the relationship between DNE values and age-at-death in known-age individuals. Age-related variables will be identified via three statistical analysis methods: principal component discriminant analysis (PCQDA), principal component regression analysis (PCR), and principal component logistic regression analysis (PCLR). The results have been used to develop an R package, JSDNE, which can be used to estimate age-at-death in unknown individuals.

2. Materials

Data for the current study was collected from three European anatomical skeletal collections; the Luís Lopes Skeletal Collection, the 21st Century Identified Skeletal Collection, and the CAL Milano Cemetery Skeletal Collection. The Luís Lopes Skeletal Collection is curated by the National Museum of Natural History and Science (MUHNAC), Lisbon, Portugal. It contains 1692 skeletons in total, with biographical information available for 699 individuals, including age-at-death, place of birth and residence, and cause of death. The remains were collected from a modern cemetery in Lisbon between 19th and 20th centuries (Cardoso, 2006). The 21st Century Identified Skeletal Collection is curated by the University of Coimbra, Portugal. This contained 159 adult individuals, ranging in age from 29 to 99 years at death. Most individuals were Portuguese, dying between 1995 and 2008. Biographical information, including sex and age-at-death, is available for

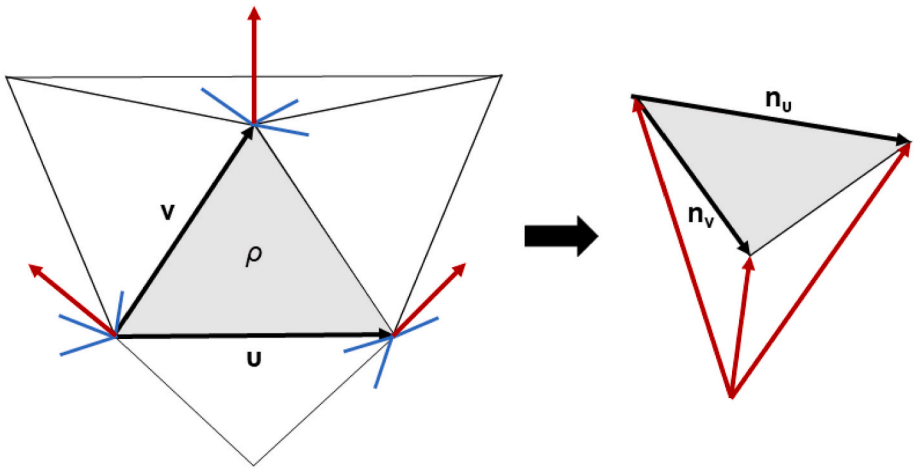

Fig. 1. Dirichlet normal energy (Jang et al., 2024).

the entire collection, with documentation relating to the cause of death available for 12 individuals (Ferreira et al., 2014). Lastly, the CAL Milano Cemetery Skeletal Collection is curated by the Laboratorio di Antropologia e Odontologia Forense (LABANOF) in the Department of Biomedical Sciences for Health of the University of Milan (Italy) and part of the Collezione Antropologica LABANOF (CAL), Italy. It contains 2127 skeletons from Cimitero Maggiore, Cimitero di Lambrate, and Cimitero di Baggio. Biographical information, such as sex, age-at-death, date of birth and death, cause of death, and pathology (if it is specified), were all documented. The individuals died between 1910 and 2001, but most of them died after 1980 (Cattaneo et al., 2018). The current research consists of 890 (402 females, 488 males) adult individuals with a known age-at-death from these three collections (Table 1 and Fig. 2). All 890 individuals had fully preserved auricular surfaces that displayed no taphonomic damage or pathological conditions.

3. Methods

3.1. Preparation and importing of reconstructed models

Auricular surfaces were scanned using an EXScan SP laser scanner (resolution: 1.3 megapixels). Each surface was scanned from a minimum of five different angles and the resultant images were combined using the software EXScanS_v3.1.0.1 to produce a single 3D model which was saved as a.ply file. Each.ply file was then edited manually in Meshmixer to produce two ROIs; the auricular surface and the apical area. The auricular surface was isolated using the outer margin as a guide (following the criterion of Jang et al., 2024). The apical area was delineated by measuring polygons within a 1 cm distance from both the left and right sides of the terminated point of the arcuate line. The 3D models were imported into R studio with the vcgPlyRead function from Rvcg package (Schlager, 2017).

Table 1
The number of auricular surfaces for each collection.

Collection	Female auricular surface	Male auricular surface	Total
21st Century Identified Skeletal Collection	66	83	149
CAL Milano Cemetery Skeletal Collection	134	173	307
Luis Lopes Skeletal Collection	202	232	434
Total	402	488	890

3.2. Analysis

Principal component analysis (PCA), quadratic discriminant analysis (PCDQA), and regression analysis (PCLR) were used to identify DNE variables that were highly correlated with known age-at-death. These correlations were then used to develop a model that estimates an unknown auricular surface's age-at-death. In order to make this model available to other researchers, an R package, JSDNE, was created, which requires R version 4.3.2 (R Core Team, 2024). JSDNE can be installed automatically by operating the install function, install.packages ("JSDNE"). Once the package is installed, it can be imported by the library function, library(JSDNE). Four packages are automatically imported to the user's library: dplyr (Wickham et al., 2019), MASS (Venables and Ripley, 2002), molaR (Pampush et al., 2016), and nnet (Venables and Ripley, 2002). After importing the.ply files with the vcgPlyRead function, the auricular surface.ply file should be replaced with x and the apical area with y. If oppositely applied, incorrect results will be calculated.

The JSDNE package consists of three functions to estimate age (Table 2). The PCQDA function can assign an individual to one of four broad age ranges (Table 4). The PCR function provides an output of estimated age in years and the standard error. The PCLR function also assigns an individual to a broad age range, but with a more limited range of under 67 years or over 63 years.

4. Results

4.1. Age-related variables

Initial analyses identified ten DNE variables that were highly correlated with age-at-death in the known-age sample (Table 3 and Fig. 3). Nine of these variables (total DNE per total number of polygon faces, median value of DNE on whole surface, interquartile range (IQR) on whole surface, mean value of DNE on convex and concave surfaces, proportion of polygon faces with DNE of over 0.6, and mean, median and IQR of DNE on apex) had high positive correlations, while the proportion of polygon faces with DNE of less than 0.0001, had a high negative correlation.

4.2. Identifying the suitable variables and the cluster

The best combination of variables was chosen on the basis of three criteria: (1) high total accuracy, (2) high accuracy in younger auricular surfaces and older auricular surfaces, and (3) low bias. For PCR analysis, an additional criterion was used: low standard error. For the PCQDA analysis, four variables provided the best combination: proportion of

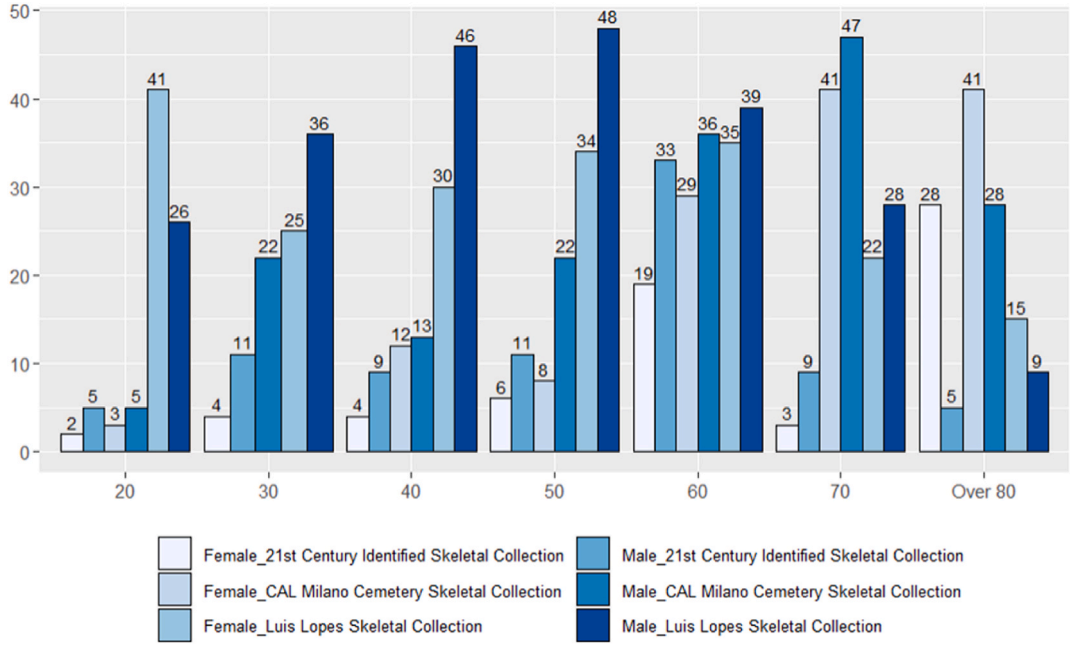

Fig. 2. Distribution of the number of auricular surfaces according to decade of age-at-death and sex.

Table 2
Function in the JSDNE package.

Function	Description
PCQDA(x,y)	Estimates the age phase through the PCQDA analysis x: whole auricular surface, y: apex Output: estimated age phase and age range
PCR(x,y)	Estimates the age through the PCR analysis x: whole auricular surface, y: apex Output: estimated age and standard error (8.8 yrs)
PCLR(x,y)	Estimates the age phase through the PCLR analysis x: whole auricular surface, y: apex Output: estimated age phase and age range

Table 3
Pearson's correlation for ten DNE variables and known age-at-death. *** means *p*-value is under 0.001.

Variables	Pearson's correlation r-value	Interpretation
Mean value of DNE on apex	0.710***	High positive correlation
Median value of DNE on apex	0.772***	High positive correlation
IQR of DNE on apex	0.727***	High positive correlation
Total DNE per total polygon faces	0.753***	High positive correlation
Median value of DNE on whole surface	0.763***	High positive correlation
IQR of DNE on whole surface	0.789***	High positive correlation
Mean value of DNE on convex surface	0.726***	High positive correlation
Mean value of DNE on concave surface	0.732***	High positive correlation
Proportion of polygon faces with DNE of less than 0.0001	−0.726***	High negative correlation
Proportion of polygon faces with DNE of over 0.6	0.751***	High positive correlation

Table 4
Age range, mean, median values of each age phase of PCQDA.

Age phase	Min	Max	Mean	Median
1	20	44	25.75	25.5
2	31	74	54.09	55.0
3	63	93	78.57	78.0
4	76	96	86.56	85.0

polygon faces with DNE of less than 0.0001, proportion of polygon faces with DNE of over 0.6, mean value of DNE on apex, and total DNE/total polygon faces. These variables grouped estimated ages into four age phases (Table 4 and Fig. 4). The age intervals of each phase are 24 years (phase 1), 43 years (phase 2) and 30 years (phase 3), with an open-ended phase 4. For PCR analysis, five variables provided the best combination: total DNE per total number of polygon faces, mean value of DNE on apex, IQR value of DNE on apex, mean value of DNE on convex surface, proportion of polygon faces with DNE of less than 0.0001. As this is a regression model no age phase was produced. Instead, the output provides an estimated age in years with a standard error of ±8.8 yrs. Finally, the PCLR analysis was developed using five variables: total DNE per total number of polygon faces, median value of DNE on whole surface, IQR of DNE on whole surface, mean value of DNE on convex surface, and mean value of DNE on apex. Estimated age was divided into two phases to determine if the auricular surface is under 67 years or over 63 years (Table 5 and Fig. 5). The age phases for each of the three types of analysis were designated in order of the lowest minimum value of the age range. The minimum values of age phases 1,2,3 and 4 in PCQDA were 20,31,63 and 76, respectively, and minimum values of age phases 1 and 2 in PCLR were 20 and 63, respectively. When converting the variables to principal components, the first and second principal components contained 98%, 96%, and 96% of the age-related variation in the PCQDA, PCLR, and PCR methods, respectively. Thus, the first and second principal components were used as variables. Then, the statistical models were developed using cross-validation. The training and test sets contained 80% and 20% of the data set randomly.

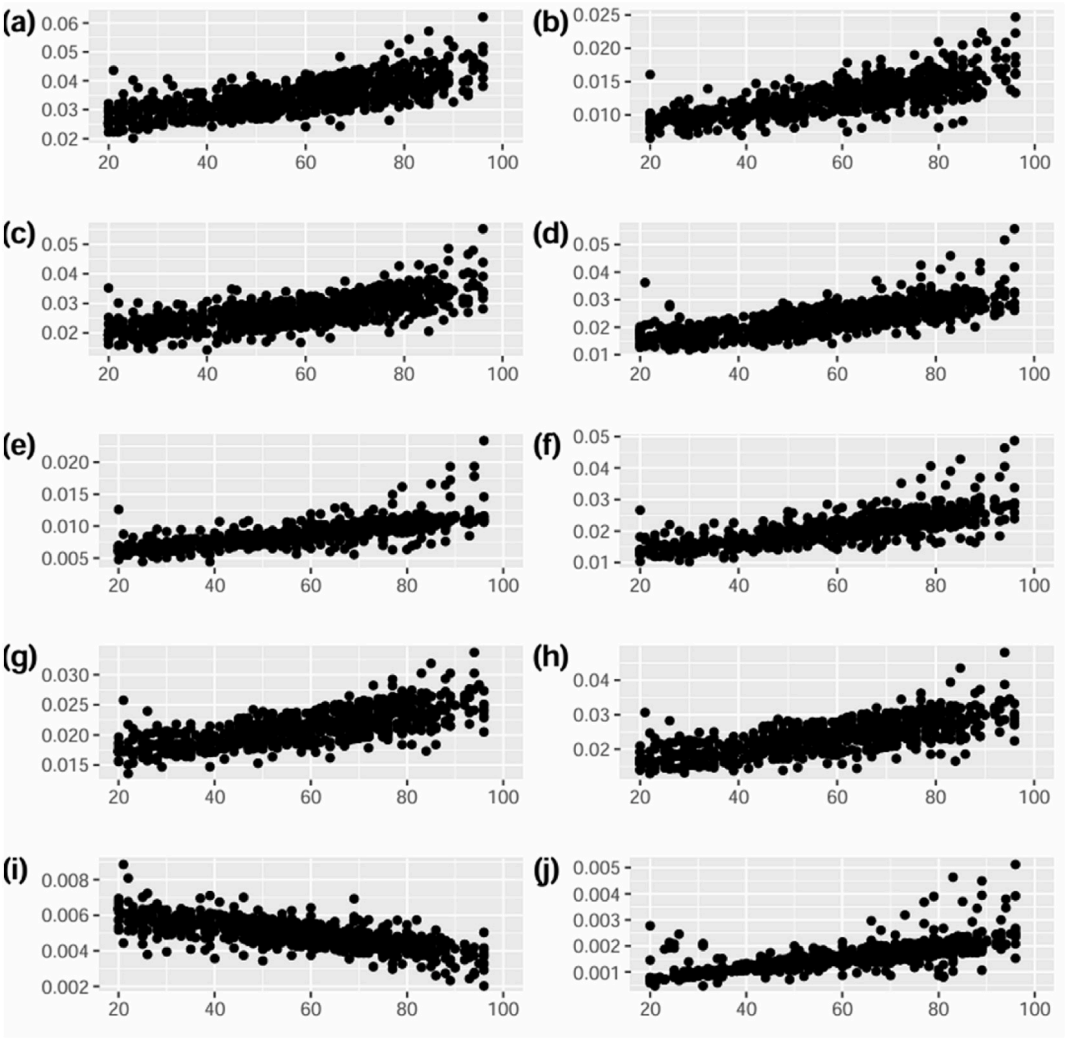

Fig. 3. Correlation graphs of each 10 DNE variables with age: (a) Mean value of DNE on apex; (b) Median value of DNE on apex; (c) IQR of DNE on apex; (d) Total DNE per Total polygon faces; (e) Median value of DNE on whole surface; (f) IQR of DNE on whole surface; (g) Mean value of DNE on convex surface; (h) Mean value of DNE on concave surface; (i) Proportion of polygon faces with DNE of less than 0.0001; (j) Proportion of polygon faces with DNE of over 0.6.

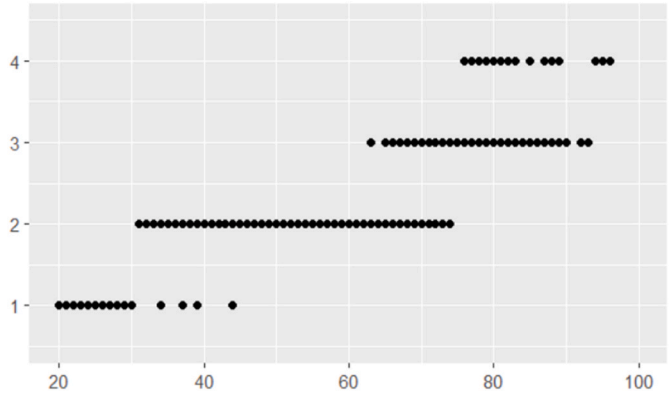

Fig. 4. Distribution graph of each age phase of PCQDA.

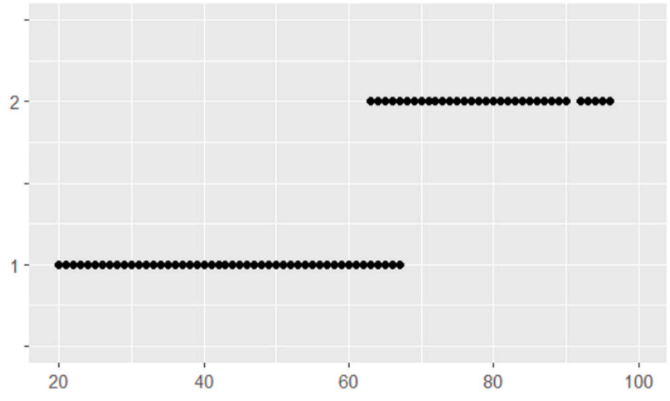

Fig. 5. Distribution graph of each age phase of PCLR.

Table 5
Age range, mean, median values of each age phase of PCLR.

Age phase	Min	Max	Mean	Median
1	20	67	44.29	46
2	63	96	75.68	75

4.3. Accuracy, bias, and the average absolute deviation of the JSDNE age estimation package

The training set, test set, and combined set showed similar results. Accordingly, the results were described based on the combined set. Accuracy was defined as JSDNE age estimates that fell within the age

phase that contained the known age-at-death. Bias was calculated as the mean value of the subtracted values between the mean age of the estimated age phase and the known age at death. The average absolute deviation (AAD) was the mean value of the absolute values subtracted between the mean age of the estimated age phase and the known age at death:

$$\text{Bias} = \sum(\text{mean age of the estimated age phase} - \text{known age at death})/N$$

$$\text{AAD} = \sum|\text{mean age of the estimated age phase} - \text{known age at death}|/N$$

4.3.1. PCQDA

The total accuracy and the average absolute deviation (AAD) are 91% and 9.58 years, respectively (Table 6). This demonstrates that the PCQDA function estimates age with a high level of accuracy and precision. Phase 2, which estimated individuals to be over 31 and under 74 years of age, had the lowest accuracy (88.8%) and the highest AAD (11.52 years). The highest accuracy and the lowest AAD were shown in phase 1 (20–44 years of age), with 97.8% and 5.53 years, respectively. While the cluster analysis identified this four-phase grouping system as the best fit for the prediction model, levels of accuracy were also examined using 10-year age intervals (Table 7), in order to compare the relative applicability of each of the functions across age cohorts. This showed that there was a tendency to underestimate the age of older individuals (those in their 60s and older) and overestimate the age of younger and middle aged individuals (in their 50s and younger). Individuals in their 20s were estimated with the lowest level of accuracy (65.9%). The highest AAD (15.98 years) was for individuals in their 30s. The highest accuracy levels were seen for individuals in their 40s (98.2%).

4.3.2. PCR

Accuracy was defined as JSDNE age estimates that fell within the standard error (8.8yrs) of the known age-at-death. The total accuracy and ADD are 76% and 6.46 years, respectively (Table 8). When examined in 10-year age intervals, the accuracy levels range from poor to moderate, but this is due to the high precision of the age estimate in the majority of intervals. Individuals in their 20s and over 80 were predicted with the lowest levels of accuracy (52.4% and 56.3%), and highest levels of AAD (9.37 years and 9.40 years). The highest accuracy (92.2%) was found for individuals in their 50s. As with the PCQDA function, there was a tendency for younger and middle aged adults (those in their 40s and younger) to be overestimated, while the age of individuals over 50 was underestimated.

4.3.3. PCLR

Accuracy was defined as JSDNE age estimates that fell within the age phase that contained the known age-at-death. PCLR produced a high level of accuracy overall (92.9%), and high precision (9.93 years) (Table 9). Phase 1 has higher accuracy (95.2%) than phase 2 (90.2%). However, phase 2 has higher precision than phase 1, indicated by the lower AAD value (11.48 years vs. 8.09 years). When classifying individuals with 10-year age intervals (Table 10), the highest accuracy was shown for individuals in their 30s (100%), and the lowest accuracy for individuals in their 60s (80.6%). In terms of precision, individuals in 20s had the lowest precision (ADD: 20.45 years), and those in their 40s had the highest precision (ADD: 2.68 years). Yet again, individuals under 40 years of age tended to be overestimated, while individuals over 40 were underestimated.

5. Discussion

Initial analyses involving the DNE values and known age-at-death revealed ten DNE variables that were highly correlated with age (Table 3 and Fig. 3). The majority of these variables had positive correlations, demonstrating that the auricular surface becomes more curved with increasing adult age. The only variable to show a negative correlation was the proportion of polygon faces with DNE of less than 0.0001, which demonstrates that flatter areas of the auricular surface (the number of polygons that are mathematically flat as a product of n_v and n_u), become reduced with increasing age as the surface becomes more irregular. These findings are in agreement with traditional methods of age estimation, which have long been aware that areas of macroporosity and osteophytic growth become more visible on the auricular surfaces of older individuals (Buckberry and Chamberlain, 2002; Lovejoy et al., 1985). Previous research with curvature analysis has suggested that young adults have high curvature values (Jang et al., 2024; Villa et al., 2015), a finding which is not echoed here. It is well known that the auricular surfaces of younger adults display distinctive areas of transverse organization over a large proportion of the surface (Buckberry and Chamberlain, 2002; Lovejoy et al., 1985) and it was previously thought that this feature contributed to the high levels of curvature reported by Jang et al. (2024) and Villa et al. (2015). However, while transverse organization is visible on the reconstructed models used in the current study, it appears to have had no measurable effect on curvature. It is possible that results for younger adults in Jang et al. (2024) and Villa et al. (2015) are highly affected by the presence of subchondral defects (which produce very high DNE values (>0.6)). While subchondral defects are present in all age groups, they have a limited on impact on the variance and median DNE values when the sample is large. However, when the sample is small, the inclusion of several individuals with subchondral defects can increase the variance dramatically and raise median DNE values for the entire group, leading to improper generalisation due to the small sample size. These results demonstrate that curvature variables relating to a number of the morphological, degenerative features used in traditional age estimation methods can be identified through DNE analysis. However, instead of relying on the experience of the observer, DNE can quantify mathematical aspects of surface geometry that are correlated with age, reducing the subjectivity and user-dependence.

The JSDNE R package uses DNE variables that are correlated with age-at-death to estimate age for unknown auricular surfaces. Three functions are available; PCQDA, PCR, and PCLR. The total accuracy of the PCQDA function produces a high level of accuracy (91%), likely a result of wide age ranges of each age phase (Millán et al., 2013). The PCLR function also produces the highest level of accuracy (92.9%), but is used to determine whether an individual is under 67 or over 63 years of age. The PCR function produces a narrower age range (SE: 8.8yrs) with a reduced level of accuracy (76%). These results are similar to the

Table 6
PCQDA results of each age phase. Acc. = Accuracy, AAD. = Average absolute deviation (Year), Accuracy is in percent (%) and Bias, AAD are in years.

	Test set			Train set			Combined set		
	ACC.	Bias	AAD.	ACC.	Bias	AAD.	ACC.	Bias	AAD.
Total	89.2	0.52	10.16	90.5	−0.48	9.58	91.0	−0.27	9.58
Phase 1	95.5	−6.80	7.64	98.5	−2.12	5.19	97.8	−2.96	5.53
Phase 2	88.6	0.99	12.24	87.7	−0.37	11.48	88.8	−0.06	11.52
Phase 3	87.5	2.84	7.50	94.2	−0.24	6.72	93.4	0.26	6.75
Phase 4	100.0	−5.77	5.77	94.1	0.80	5.85	95.0	−0.14	5.84

Table 7
PCQDA results within 10-year intervals. Acc. = Accuracy, AAD. = Average absolute deviation (Year), Accuracy is in percent (%) and Bias, AAD are in years.

	Test set			Train set			Combined set		
	ACC.	Bias	AAD.	ACC.	Bias	AAD.	ACC.	Bias	AAD.
Total	89.2	0.52	10.16	90.5	−0.48	9.58	91.0	−0.27	9.58
20s	53.3	13.04	14.81	65.7	11.17	11.72	65.9	10.82	11.60
30s	96.2	6.13	14.66	94.4	12.22	16.10	94.9	11.47	15.98
40s	96.3	6.58	9.14	98.9	6.73	8.80	98.2	6.69	8.88
50s	89.5	1.35	4.51	98.2	0.45	2.83	97.7	0.40	2.95
60s	93.0	−4.86	11.61	98.6	−6.54	10.80	97.9	−6.42	10.97
70s	90.9	−2.35	8.47	81.2	−6.91	11.01	85.3	−5.58	10.07
Over80	87.0	−7.63	7.63	87.4	−8.01	8.14	87.3	−7.94	8.05

Table 8
PCR results (SE: 8.8 yrs) were obtained within 10-year intervals. Acc. = Accuracy, AAD. = Average absolute deviation (Year), Accuracy is in percent (%) and Bias, AAD are in years.

	Test set			Train set			Combined set		
	ACC.	Bias	AAD.	ACC.	Bias	AAD.	ACC.	Bias	AAD.
Total	79.8	−0.07	5.94	74.5	0.00	6.61	76.0	−0.02	6.46
20s	57.1	6.57	7.96	51.5	9.25	9.80	52.4	8.75	9.37
30s	88.2	3.97	5.07	67.9	5.35	6.60	73.5	5.09	6.37
40s	84.8	2.53	4.75	79.0	1.12	5.49	81.6	1.56	5.33
50s	85.2	−0.97	5.38	94.1	0.05	4.43	92.2	−0.16	4.61
60s	87.8	−0.38	4.86	86.7	−2.26	5.21	88.5	−1.89	5.13
70s	79.4	−4.78	6.13	75.0	−3.08	6.71	76.0	−3.47	6.62
Over80	59.1	−2.39	9.55	53.8	−4.45	9.43	56.3	−4.02	9.40

Table 9
PCLR results of each age phase. Acc. = Accuracy, AAD. = Average absolute deviation (Year), Accuracy is in percent (%) and Bias, AAD are in years.

	Training set			Test set			Combined set		
	ACC.	Bias	AAD.	ACC.	Bias	AAD.	ACC.	Bias	AAD.
Total	92.7	−0.17	10.21	92.8	0.27	9.88	92.9	0.18	9.93
Phase 1	97.2	−1.80	11.25	94.7	−0.47	11.57	95.2	−0.75	11.48
Phase 2	87.1	1.87	8.92	90.7	1.14	7.90	90.2	1.27	8.09

Table 10
PCLR results within 10-year intervals. Acc. = Accuracy, AAD. = Average absolute deviation (Year), Accuracy is in percent (%) and Bias, AAD are in years.

	Test set			Train set			Combined set		
	ACC.	Bias	AAD.	ACC.	Bias	AAD.	ACC.	Bias	AAD.
Total	92.7	−0.17	10.21	92.8	0.27	9.88	92.9	0.18	9.93
20s	93.3	23.45	23.45	98.5	19.77	19.77	97.6	20.45	20.45
30s	100.0	10.46	10.46	100.0	9.61	9.61	100.0	9.77	9.77
40s	100.0	−1.33	2.49	98.9	−1.02	2.74	99.1	−1.09	2.68
50s	96.9	−8.23	9.77	88.7	−6.67	11.13	91.5	−7.30	10.75
60s	79.6	1.15	13.06	81.0	−0.12	13.27	80.6	0.37	13.17
70s	91.3	−2.18	4.89	92.9	−0.88	4.48	92.7	−1.08	4.54
Over80	100.0	−10.03	10.03	99.0	−10.60	10.60	99.2	−10.47	10.47

accuracy rates gained by the majority of researchers who carry out validation studies of the Buckberry and Chamberlain (94%–86%) and Lovejoy (63%–27%) methods of traditional analysis (Hughes et al., 2023; Millán et al., 2013; Rissech et al., 2012). However, while the traditional methods rely on subjectivity and the researchers own experience in assigning an individual to an age phase (Štepanovský et al., 2023), the JSDNE results are completely objective and have excellent levels of repeatability. When choosing which function to use, the nature of the research must be considered. PCQDA and PCLR may be more suitable for estimating age in archaeological individuals as these functions can identify broad age ranges that correspond to various stages of the lifecourse with moderate to high levels of accuracy. For forensic cases, which require a more precise estimate of age-at-death, the PCR function may be more appropriate. It is also possible to use all three functions in conjunction with one another.

Although able to provide an objective and user-independent estimate of age-at-death, these results still demonstrate a degree of bias, as all of the functions tended to overestimate the age of younger individuals and underestimate the age of older individuals. This is a common tendency that has been observed in other studies (Falys et al., 2006; Moraitis et al., 2014; Navitainuck et al., 2022; Rissech et al., 2012; Rivera-Sandoval et al., 2018; Štepanovský et al., 2023), and relates to the age structure of skeletal samples used to develop the methods. Generally, the number of younger and older adults is fewer than those who are middle-aged, meaning that a wider range of variation is captured for the middle-aged auricular surfaces and the resulting age estimates are also biased towards this age group. There is currently no way to alleviate this problem with the known-age skeletal collections that are available worldwide, although it may be possible to reduce its effect through tightly controlled statistical analyses. Methods such as profile likelihood

analysis could be a good alternative to reduce the underestimation and overestimation issue with predicted ages, albeit with broader prediction intervals, as this does not require age to be regressed onto curvature variables. However, there is currently no R package for profile likelihood, so this method of analysis would have to be performed in another statistical programme.

6. Conclusion

The present research successfully identified ten geometric variables of the auricular surface that were associated with known age-at-death and used these to create an objective, mathematical age estimation method that is available as a novel R package, JSDNE (https://cran.r-project.org/package=JSDNE). The detailed code and the raw data of this study are openly available at https://github.com/jisunjang19/cran-JSDNE, https://doi.org/10.5281/zenodo.12708779, or 'JSDNE' package. This estimates age using three functions which offers high levels of accuracy (76–92.9%). This method does not require a great degree of experience to gather and interpret the results, only familiarity with R studio and access to a laser scanner. Moreover, the results can be applied to research in bioarchaeology and forensic anthropology.

CRediT authorship contribution statement

Jisun Jang: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Conceptualization. **Enrico Mariconti:** Writing – review & editing, Methodology. **Rebecca Watts:** Writing – review & editing, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr. Maria Susana De Jesus Garcia with the Luís Lopes Skeletal Collection, curated by the National Museum of Natural History and Science (MUHNAC), in Lisbon, Portugal, Dr. Maria Teresa Ferreira with the 21st Century Identified Skeletal Collection, curated at the University of Coimbra, in Portugal, and Dr. Mirko Mattia and Dr. Cristina Cattaneo with the CAL Milano Cemetery Skeletal Collection, curated at the University of Milan, in Italy, for allowing to access to their valuable collections.

Reproducibility Report

The Associate Editor for Reproducibility downloaded all materials and could reproduce the results presented by the authors.

References

Biwasaka, H., Sato, K., Aoki, Y., Kato, H., Maeno, Y., Tanijiri, T., Fujita, S., Dewa, K., 2013. Three dimensional surface analyses of pubic symphyseal faces of contemporary Japanese reconstructed with 3D digitized scanner. Leg. Med. 15/5, 264–268. https://doi.org/10.1016/j.legalmed.2013.02.003.

Botha, D., Lynnerup, N., Steyn, M., 2019. Age estimation using bone mineral density in South Africans. Forensic Sci. Int. https://doi.org/10.1016/j.forsciint.2019.02.020.

Brooks, S., Suchey, J.M., 1990. Skeletal age determination based on the os pubis: a comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. Hum. Evol. (5/3), 227–238.

Buckberry, J.L., Chamberlain, A.T., 2002. Age estimation from the auricular surface of the ilium: a revised method. Am. J. Phys. Anthropol. 119/3, 231–239. https://doi.org/10.1002/ajpa.10130.

Bunn, J.M., Boyer, D.M., Lipman, Y., StClair, E.M., Jernvall, J., Daubechies, I., 2011. Comparing Dirichlet normal surface energy of tooth crowns, a new technique of molar shape quantification for dietary inference, with previous methods in isolation and in combination. Am. J. Phys. Anthropol. 145/2, 247–261. https://doi.org/10.1002/ajpa.21489.

Byers, S.N., 2017. Introduction to Forensic Anthropology, fifth ed. Routledge, Oxon.

Cardoso, F.V., 2006. Brief communication: the collection of identified human skeletons housed at the bocage Museum (National Museum of natural history), Lisbon, Portugal. Am. J. Phys. Anthropol. 129 https://doi.org/10.1002/ajpa.20228, 173-173.

Cattaneo, C., Mazzarelli, D., Cappella, A., Castoldi, E., Mattia, M., Poppa, P., De Angelis, D., Vitello, A., Biehler-Gomez, L., 2018. A modern documented Italian identified skeletal of 2127 skeletons: the CAL Milano Cemetery Skeletal Collection. Forensic Sci. Int. 287, 219.e1–219.e5. https://doi.org/10.1016/j.forsciint.2018.03.041.

Curate, F., Navega, D., Cunha, E., Coelho, J.O., 2022. Dxage 2.0 - adult age at death estimation using bone loss in the proximal femur and the second metacarpal. Int. J. Leg. Med. 136 (5), 1483–1494. https://doi.org/10.1007/s00414-022-02840-y.

Falys, C., Schutkowski, H., Weston, D.A., 2006. Auricular surface aging: worse than expected? A test of the revised method on a documented historic skeletal assemblage. Am. J. Phys. Anthropol. 130 (4), 508–513. https://doi.org/10.1002/ajpa.20382.

Ferreira, M.T., Vincente, R., Navega, D., Gonçalves, D., Curate, F., Cunha, E., 2014. A new forensic collection housed at the University of Coimbra, Portugal: the 21st century identified skeletal collection. Forensic Sci. Int. 245, 202.e1–202.e5. https://doi.org/10.1016/j.forsciint.2014.09.021.

Garvin, H.M., Passalacqua, N.V., 2012. Current practices by forensic anthropologists in adult skeletal age estimation. J. Forensic Sci. 57/2, 427–433. https://doi.org/10.1111/j.1556-4029.2011.01979.x.

Godfrey, L.R., Winchester, J.M., King, S.J., Boyer, D.M., Jernvall, J., 2012. Dental topography indicates ecological contraction of lemur communities. Am. J. Phys. Anthropol. (148/2), 215–227. https://doi.org/10.1002/ajpa.21615.

Hughes, C., Yim, A.-D., Juarez, C., 2023. Considerations for age estimation accuracy: method-derived outcomes and practitioner interpretations. Journal of Forensic Science 1–10. https://doi.org/10.1111/1556-4029.15505, 00.

Işcan, M.Y., Loth, S.R., 1989. Osteological manifestations of age in the adult. In: Iscan, M. Y., Kennedy, L.A.R. (Eds.), Skeletal Biology of Past Peoples: Research Methods. Wiley-Liss, New York, pp. 189–224.

Jang, J., Mariconti, E., Watts, R., 2024. Technical note: analysis of the auricular surface for age estimation using dirichlet normal energy. Forensic Imaging 36. https://doi.org/10.1016/j.fri.2024.200579.

Kumar, C.L., Sridhar, M.S., 1990. Estimation of the age of an individual based on times of eruption of permanent teeth. Forensic Sci. Int. 48/1, 1–7. https://doi.org/10.1016/0379-0738(90)90266-2.

Lovejoy, C.O., Meindl, R.S., Pryzbeck, T.R., Mensforth, R.P., 1985. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am. J. Phys. Anthropol. 68 (1), 15–28. https://doi.org/10.1002/ajpa.1330680103.

Márquez-Grant, N., 2015. An overview of age estimation in forensic anthropology: perspectives and practical considerations. Ann. Hum. Biol. 42 (4), 308–322. https://doi.org/10.3109/03014460.2015.1048288.

Mensforth, 1990. Paleodemography of the carlston annis (Bt-5) late archaic skeletal population. Am. J. Phys. Anthropol. (82/1), 81–99. https://doi.org/10.1002/ajpa.1330820110.

Millán, M.S., Rissech, C., Turbón, D., 2013. A test of Suchey–Brooks (pubic symphysis) and Buckberry–Chamberlain (auricular surface) methods on an identified Spanish sample: paleodemographic implications. J. Archaeol. Sci. 40 (4), 1743–1751. https://doi.org/10.1016/j.jas.2012.11.021.

Miranker, M., 2016. A comparison of different age estimation methods of the adult pelvis. J. Forensic Sci. 61/5, 1173–1179, 101111/1556-402913130.

Moraitis, K., Zorba, E., Eliopoulos, C., Fox, S.C., 2014. A test of the revised auricular surface aging method on a modern European population. J. Forensic Sci. 59/1, 188–194. https://doi.org/10.1111/1556-4029.12303.

Navitainuck, D.U., Vach, W., Pichler, S.L., Alt, K.W., 2022. Age-at-death estimation in archaeological samples: differences in population means resulting from different aging methods can be predicted from the mean ages of method-specific reference samples. Int. J. Osteoarchaeol. 32 (6), 1226–1237. https://doi.org/10.1002/oa.3157.

Pampush, J.D., Winchester, J.M., Morse, P.E., Vining, A.Q., Boyer, D.M., Kay, R.F., 2016. Introducing molaR: a new R package for quantitative topographic analysis of teeth (and other topographic surfaces). J. Mamm. Evol. (23/4), 397–412. https://doi.org/10.1007/s10914-016-9326-0.

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. http://www.R-project.org/.

Rissech, C., Wilson, J., Winburn, A.P., Turbón, D., Steadman, D., 2012. A comparison of three established age estimation methods on an adult Spanish sample. Int. J. Leg. Med. 126 (1), 145–155. https://doi.org/10.1007/s00414-011-0586-1.

Rivera-Sandoval, J., Monsalve, T., Cattaneo, C., 2018. A test of four innominate bone age assessment methods in a modern skeletal collection from Medellin, Colombia. Forensic Sci. Int. 282, 232.e1–232.e8.

Schlager, S., 2017. Morpho and Rvcg – shape analysis in R. In: Zheng, G., Li, S., Szekely, G. (Eds.), Statistical Shape and Deformation Analysis. Academic Press, Massachusetts, pp. 217–256.

Schmeling, A., Schulz, R., Reisinger, W., Mühler, M., Wernecke, K.-D., Geserick, G., 2004. Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. Int. J. Leg. Med. 118 (1), 5–8. https://doi.org/10.1007/s00414-003-0404-5.

Schwartz, J.H., 1995. Skeletal Keys an Introduction to Human Skeletal Morphology, Development, and Analysis. Oxford University Press, New York.

Slice, D.E., Algee-Hewitt, B.F.B., 2015. Modeling bone surface morphology: a fully quantitative method for age-at-death estimation using the pubic symphysis. J. Forensic Sci. 60/4, 835–843. https://doi.org/10.1111/1556-4029.12778.

Štepanovský, M., Buk, Z., Kotěrová, A.P., Brůžek, J., Bejdová, Š., Techataweewan, N., Velemínská, J., 2023. Automated age-at-death estimation from 3D surface scans of the facies auricularis of the pelvic bone. Forensic Sci. Int. 349 https://doi.org/10.1016/j.forsciint.2023.111765.

Stoyanova, D., Algee-Hewitt, B.F.B., Slice, D.E., 2015. An enhanced computational method for age-at-death estimation based on the pubic symphysis using 3D laser scans and thin plate splines. Am. J. Phys. Anthropol. 158 (3), 431–440. https://doi.org/10.1002/ajpa.22797.

Stoyanova, D.K., Algee-Hewitt, B.F.B., Kim, J., Slice, D.E., 2017. A computational framework for age-at-death estimation from the skeleton: surface and outline analysis of 3D laser scans of the adult pubic symphysis. J. Forensic Sci. 62/6, 1434–1444. https://doi.org/10.1111/1556-4029.13439.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, fourth ed. Springer, New York. Available at: https://www.stats.ox.ac.uk/pub/MASS4/.

Villa, C., Buckberry, J., Cattaneo, C., Frohlich, B., Lynnerup, N., 2015. Quantitative analysis of the morphological changes of the pubic symphyseal face and the auricular surface and implications for age at death estimation. J. Forensic Sci. 60 (3), 556–565. https://doi.org/10.1111/1556-4029.12689.

Warrier, V., Shedge, R., Garg, P.K., Dixit, S.G., Krishan, K., Kanchan, T., 2023. Age estimation from iliac auricular surface using Bayesian inference and principal component analysis: a CT-based study in an Indian population. Forensic Sci. Med. Pathol. https://doi.org/10.1007/s12024-023-00637-y.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4/43, 1686. https://doi.org/10.21105/joss.01686.

Winchester, J.M., Boyer, D.M., St Clair, E.M., Gosselin-Ildari, A.D., Cooke, S.B., Ledogar, J.A., 2014. Dental topography of platyrrhines and prosimians: convergence and contrasts. Am. J. Phys. Anthropol. 153 (1), 29–44. https://doi.org/10.1002/ajpa.2239.

Zhang, Y., Wang, Z., Liao, Y., Li, T., Xu, X., Wu, W., Zhou, J., Huang, W., Luo, S., Chen, F., 2022. A machine-learning approach using pubic CT based on radiomics to estimate adult ages. Eur. J. Radiol. https://doi.org/10.1016/j.ejrad.2022.110516.