Charles University

Faculty of Science

Study programme:
Anthropology and Human Genetics

Mgr. Barbora Suchá

SEX CLASSIFICATION USING 3D SKULL MODELLING IN MULTI-POPULATION SAMPLE

Klasifikace pohlaví na základě 3D modelování lebky v multipopulačním souboru

Doctoral thesis

Supervisor:
prof. RNDr. Jana Velemínská, Ph.D.

Prague, 2024

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne

……………………

Mgr. Barbora Suchá

I want to dedicate this work to my beloved friend Martin. I know you would be proud.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all those who have supported and guided me throughout the course of this research project.

First and foremost, I am profoundly thankful to my supervisor, prof. RNDr. Jana Velemínská, Ph.D. Her positive mindset, great mentoring, and empathetic approach have been invaluable. Her constant encouragement and insightful feedback have guided me through this journey, making the process both rewarding and enriching.

I am also extremely grateful to my consultant, prof. RNDr. Jaroslav Brůžek, CSc., for his experienced advice and original research idea. His extensive knowledge in the field has been crucial to the development and success of this study. His guidance has helped shape the direction of this research and has greatly enhanced its quality.

A special thanks to Mgr. Ján Dupej, Ph.D., our programmer, whose visionary approach and exceptional programming skills made this project possible. His unique combination of technical expertise and understanding of anthropology has been indispensable. He has always been helpful and willing to go the extra mile to ensure the success of this project.

I would also like to extend my heartfelt thanks to my dearest friend Alex, who studied with me from our first day at the department to the very last. My university experience and life would not have been the same without her by my side.

My deepest gratitude goes to my amazing family and friends, with special thanks to my loving husband Vítek. He has been a steadfast support, handling my frustrations with patience, offering his IT expertise, and continually encouraging me.

Lastly, this research was supported by funding from GAUK No. 1590218.

This work would not have been possible without the contributions of each and every one of you. Thank you.

ABSTRAKT

Identifikace kosterních pozůstatků představuje jeden z nejdůležitějších aspektů biologické antropologie. Přesná identifikace jedince je nezbytná pro forenzní antropologii a vysoká úspěšnost a spolehlivost metod jsou předpoklady pro právní využití výsledků. Tato disertační práce představuje inovativní metodologii pro odhad pohlaví pomocí exokraniálního povrchu lebek, která byla důkladně testována a ověřena na různých populacích. Studie navrhla metodu, která dosáhla vysoké úspěšnosti a spolehlivosti, srovnatelné s již zavedenými technikami využívajícími jak exokraniální, tak endokraniální povrchy. Výsledky potvrdily, že nová metoda dosáhla 91% úspěšnosti ve francouzském vzorku populace, což demonstruje její účinnost. Další testování na větším a rozmanitějším souboru dat odhalilo vysokou spolehlivost metody, zejména u evropských populací. Konkrétně metoda dosáhla 96 % úspěšnosti v českém vzorku a 92 % ve slovenském vzorku, což podporuje její robustnost a použitelnost v těchto kontextech. Výzkum rovněž zkoumal změny v pohlavním dimorfismu kraniofaciální morfologie související s věkem, přičemž zjistil významné snížení dimorfních oblastí s rostoucím věkem. Podíl významně odlišných oblastí v kraniální formě se snížil o 12 %, z 94,8 % u mladších jedinců na 82,6 % u starších jedinců. Tyto nálezy zdůrazňují dynamickou povahu kraniofaciální morfologie během života. Nicméně použitelnost metody napříč různými populacemi vykázala nekonzistentní výsledky. Zatímco byla vysoce spolehlivá pro nevzdálené evropské populace, jako jsou česká, slovenská a francouzská skupina, prokázala nižší míru úspěšnosti klasifikace u egyptské (82 %) a dánské (80 %) populace. To naznačuje potřebu dalšího zdokonalení, aby se zvýšila spolehlivost metody u různorodějších a geograficky vzdálenějších populací.

Tato disertační práce předkládá podrobné výsledky a představuje potenciál nové metody odhadu pohlaví podle lebky, zároveň však poukazuje na oblasti, které je třeba zlepšit.

Klíčová slova: forenzní antropologie, virtuální antropologie, pohlavní dimorfismus, odhad pohlaví, biologický profil, geometrická morfometrie, 3D zobrazovací metody

ABSTRACT

The identification of skeletal remains represents one of the most critical aspects of biological anthropology. Accurate individual identification is indispensable for forensic anthropology, and high method success and reliability are prerequisites for legal utilisation of results. This doctoral thesis presents an innovative methodology for sex estimation using the exocranial surface of skulls, rigorously tested and validated across diverse populations. The study aimed to develop a method that would achieve high accuracy and reliability, comparable to established techniques using both exocranial and endocranial surfaces. The results confirmed that the new method achieved an impressive 91% accuracy in the French population sample, demonstrating its effectiveness. Further testing on a larger and more diverse dataset revealed the method's high reliability and accuracy, particularly in European populations. Specifically, the method achieved 96% accuracy in the Czech sample and 92% in the Slovak sample, supporting its robustness and applicability in these contexts. The research also explored age-related changes in craniofacial sexual dimorphism, finding a significant reduction in dimorphic areas with increasing age. The proportion of significantly different areas in cranial form decreased by 12%, from 94.8% in younger individuals to 82.6% in older individuals. These findings highlight the dynamic nature of craniofacial morphology over the lifespan. However, the method's applicability across multiple populations showed mixed results. While it was highly reliable for closely related European populations, such as the Czech, Slovak, and French groups, it demonstrated lower accuracy rates for the Egyptian (82%) and Danish (80%) populations.

In conclusion, this doctoral thesis presents detailed results and underscores the potential of the novel sex estimation method, while also highlighting areas for improvement.

Key words: forensic anthropology, virtual anthropology, sexual dimorphism, sex estimation, biological profile, geometric morphometry, 3D imaging

TABLE OF CONTENT

1. INTRODUCTION

Osteological analysis plays a significant part in both forensic and bioarchaeological contexts, typically involving the estimation of sex, age, stature, and ancestry. Various methods, including morphoscopic and morphometric approaches, have been developed for each aspect of the identification process. However, both visual and metric assessments are frequently influenced by the expertise of the examiner and the level of sexual dimorphism present in the population being studied (Buikstra & Ubelaker, 1994; Curate, 2022). Virtual methodologies represent a significant advancement in the field by offering standardized, objective, and efficient tools for analysing skeletal remains. These technologies enhance the accuracy and reproducibility of sex, age, stature, and ancestry estimations (Omari, Hunt, Coumbaros, & Chapman, 2021; Simmons-Ehrhardt, Ehrhardt, & Monson, 2019; Wilkinson, Liu, Shrimpton, & Greenway, 2024).

Accurate sex estimation from skeletal remains is fundamental aspect of the biological profile in forensic investigations and identification of unknown human remains. Sexual dimorphism, the physical differences between males and females, plays a key role in sex estimation. It relies on the understanding that there are significant size and shape distinctions in the skeletal structure of males and females, both within populations and across different populations. The most pronounced sex differences in the skeleton are observed in the pelvis (Braun, Schwendener, Kanz, Lösch, & Milella, 2024; Brůžek & Murail, 2006; Phenice, 1969) and the skull (Walrath et al. 2004; Garvin, Sholts, and Mosca 2014; Del Bove et al. 2023)

In recent decades the utilisation of modern imaging technologies led to development of new approaches and methodologies within forensic anthropology (Braun et al., 2024; Cao et al., 2021; Mamabolo, Alblas, & Brits, 2020; Rowbotham & Blau, 2020; Thali et al., 2005). Geometric morphometric methods use three-dimensional coordinates of specific cranial landmarks to analyse shape variation. Statistical techniques such as principal component analysis (PCA) and discriminant function analysis can be applied to these landmark coordinates to identify patterns of sexual dimorphism in cranial shape. Machine learning algorithms, such as support vector machines (SVM) or artificial neural networks, can be trained on three-dimensional cranial data to develop predictive models for sex estimation. These algorithms learn patterns and relationships from a large dataset of known crania of known sex and can then be used to predict the sex of unknown individuals based on their cranial features (Abdel Fatah, Shirley, Jantz, & Mahfouz, 2014; Boucherie, Chapman, García-Martínez, Polet, & Vercauteren, 2022; Čechová et al., 2019; D. Franklin, Freedman, Milne, & Oxnard, 2006; Garvin & Ruff, 2012; Kranioti, İşcan, & Michalodimitrakis, 2008; Meinerová et al., 2023).

Building on years of personal research, this thesis is organized around four primary objectives that collectively advance our understanding of sex assessment methodologies from skull. The primary (1) objective of this doctoral thesis was to introduce, validate, and enhance an innovative sex classification system based on the analysis of exocranial form and shape. This system leverages advanced geometric morphometric techniques to achieve high accuracy and reliability in sex estimation (2). For better understanding of method's applicability across different subpopulations, other analyses were conducted. This additional research included a comprehensive analysis of changes in sexual dimorphism with age (3), contributing to the understanding of craniofacial morphological transformations. Testing and validation of the method are crucially dependent on a large and diverse dataset of individuals to ensure its robustness. To achieve this, the study included individuals from a wide geographic range within Europe, encompassing central, eastern, northern, and Mediterranean regions, as well as from North Africa (4). These objectives are closely interpreted in section 4. Hypotheses.

2. ESTABLISHING BIOLOGICAL PROFILE

The biological profile is a fundamental analysis conducted during the examination of skeletal remains. The profile encompasses the estimation of sex, skeletal age, ancestry, and stature – commonly referred to as the "big four." These parameters are essential for individual identification within forensic anthropology and for accurate data processing in bioarchaeology (Spradley, Jantz, Robinson, & Peccerelli, 2008). The specific sequence of these parameters is not firmly established. In the construction of a biological profile within forensic anthropology, significant emphasis is placed on the individual's ancestry (referred to as the "American school"). Conversely, European biological and forensic anthropology prioritize sex estimation as the initial step (Guyomarc'h & Brůžek, 2011). To ensure inclusion in a forensic anthropologist's expert witness testimony, methodologies employed for deducing the biological profile must align with contemporary legal proceedings (Lesciotto, 2023). Modern genetic techniques are, in many instances, used to test the accuracy of traditional physical methods. The alignment between physical and molecular analyses in sex estimation were confirmed by several studies (e.g. Pilli et al. 2023; R. Thomas, Parks, and Richard 2017).

When skeletal remains are uncovered, the initial inquiry typically raised by authorities is whether the bones belong to a human or an animal. Identifying human bones relies heavily on the expertise of the forensic osteologist, as well as the specific elements of the skeleton that are recovered. The subsequent inquiry is to ascertain whether the remains are of forensic or archaeological significance (Scheuer, 2002). A different approach is used when identifying adults compared to subadults.

The process of identifying juvenile remains involves a multifaceted approach that integrates various scientific methods and forensic techniques. It typically includes anthropological examination, where forensic anthropologists analyse the size, shape, and development of bones to estimate age, sex, ancestry, and stature. This may involve comparing the remains to reference databases and population-specific standards (Ritz & Schütz, 1993; Scheuer & Black, 2004; Wilson, MacLeod, & Humphrey, 2008; Wood & Cunningham, 2000). Dental analysis is also crucial, as dental features such as tooth eruption patterns, dental development, and dental anomalies are used to assess age at death (Karadayı, Afşin, Ozaslan, & Karadayı, 2014; Liversidge, 2015; Moorrees, 1965). Additionally, radiographic imaging techniques, including X-rays or CT scans, may be used to visualise skeletal development and identify fractures or pathologies (Azarfar, Ko, Adams, & Babyn, 2024; Brough, Rutty, Black, & Morgan, 2012; Stock, Stull, Garvin, & Klales, 2016). DNA analysis is conducted when possible. The DNA samples are extracted from the remains and compared to reference samples or databases to confirm identity or familial relationships (Padmanabhan & Sapna, 2023; Tierney & Bird, 2015).

Overall, the identification of subadult remains necessitates a specialized approach distinct from that used for adults. The following chapters will focus exclusively on the identification of adult remains, as this thesis pertains to individuals aged 18 years and older.

2.1 Estimation of age at death

The development of an individual is strongly influenced by both genetic predispositions and environmental factors. When estimating age, it is important to remember that for unidentified individuals, we can only determine what is known as skeletal or biological age. In contrast, the term used for individuals of known age is chronological age. Excessive or inadequate stress on the musculoskeletal system, or conversely, slowed growth, can lead to significant discrepancies between chronological age and skeletal age (Garvin et al., 2012). An extensive amount of age assessment methods was published, predominantly founded on either microscopic examination of dental and bone tissues or macroscopic observation of degenerative changes impacting specific skeletal regions (Franklin 2010; Algee-Hewitt and Kim 2021).

On the skeleton, physiological developmental features diminish with advancing age, and subsequently degenerative features, which begin to prevail upon reaching adulthood as a reflection of aging processes. Therefore, age estimation is approached differently in juveniles and adults. Age estimation in non-adult individuals is most commonly based on dental mineralization (Demirjian, Goldstein, & Tanner, 1973; Moorrees, 1965; Willems et al., 2001) and the degree of ossification of individual bones (Scheuer & Black, 2004).

Estimation of age in adults relies on the assessment of degenerative changes, which become more pronounced with increasing chronological age. Methods are also influenced by the skeletal elements accessible for analysis; certain bones are more resistant to damage from taphonomic processes. Hence the substantial number of methods published using various skeletal indicators, for example:

- Morphology of pubic symphyses (Klepinger et al. 1992; Buckberry and Chamberlain 2002; Brooks and Suchey 1990; Kotěrová et al. 2022; Schanandore, Wolden, and Smart 2022)
- Morphology of sternal rib ends (Christensen, 2023; E. Nikita, 2012; Paulson, 2023; Yoder, Ubelaker, & Powell, 2001)
- Trabecular bone loss in long bones (Walker and Lovejoy 1985)
- Cranial suture obliteration (Figure 1) (Meindl & Lovejoy, 1985; Ruengdit, Case, & Mahakkanukrauh, 2020)

- Dental wear (Demirjian et al., 1973; Esan, Yengopal, & Schepartz, 2017; Vystrčilová & Novotný, 2000)
- Auricular surface of the ilium (Buckberry & Chamberlain, 2002; Falys, Schutkowski, & Weston, 2006; Mulhern & Jones, 2005; Ost, 2022)

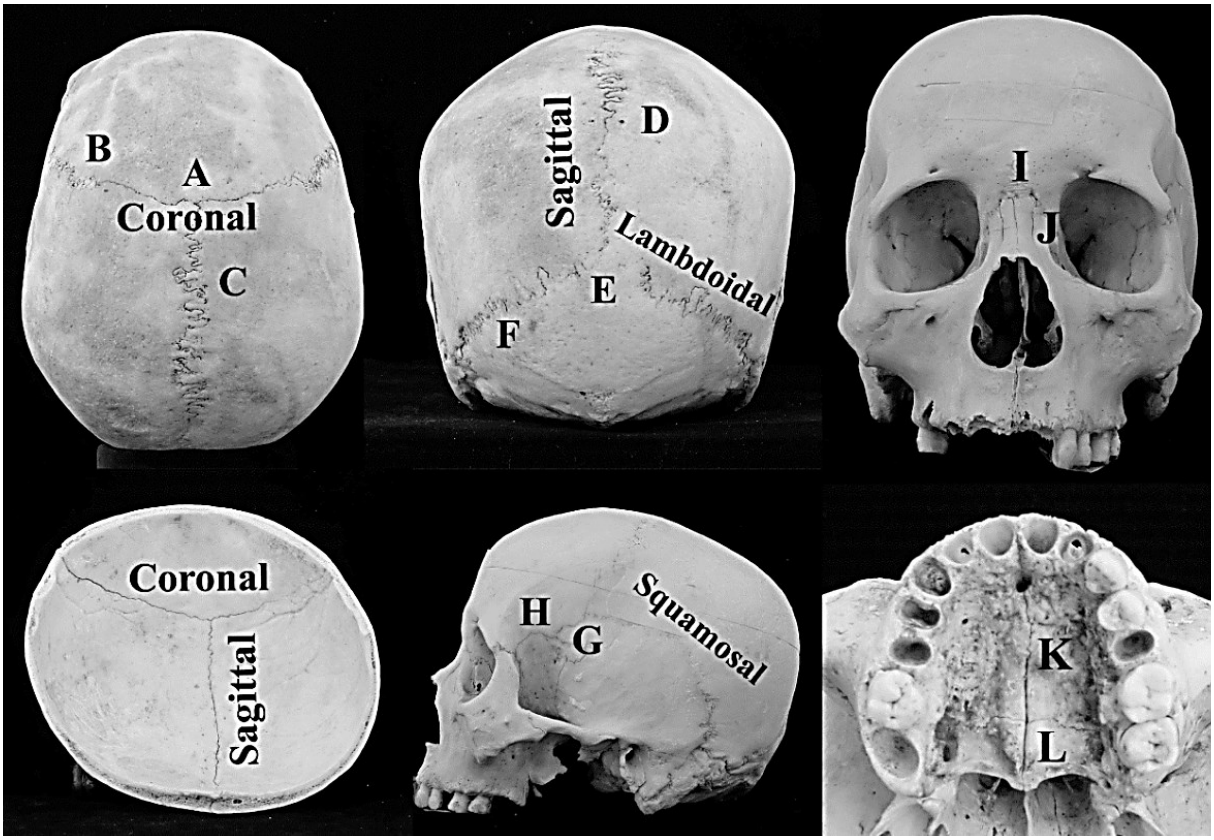


Figure 1. Examples of cranial sutures explored in key publications (Ruengdit et al.,2020).

Age can also be determined using non-morphological methods, such as metric analysis of bone dimensions (Merritt, 2015), histological examination of bone tissue structure (Kerley, 1965; Lee, Jung, Choi, & Kim, 2014), biochemical analysis of amino acid racemization (Ohtani, 1995; Ritz & Schütz, 1993), or molecular techniques assessing telomere shortening (Meissner & Ritz-Timme, 2010; Tsuji, Ishiko, Takasaki, & Ikeda, 2002). The accuracy of age estimation tends to decline as age increases. For optimal classification, segmenting into three age intervals (≤ 29, 30–69, and ≥ 70 years) proves most reliable. Moreover, individuals aged 80 years and older exhibit enhanced classification reliability when stratified into < 80 years and ≥ 80 years intervals (Castillo, Galtés, Crespo, & Jordana, 2021; Kotěrová et al., 2022).

Estimation of age in juveniles is relatively accurate and reliable, but as growth and development reach completion, accuracy and reliability diminish rapidly (Cunha et al., 2009; Iscan & Steyn, 2013).

Biological sex also plays significant role as male and female growth patterns diverge due to hormonal influences. The timing of epiphyseal fusion (the process where the ends of long bones fuse with the bone shaft) can differ between males and females. Typically, females reach skeletal maturity earlier than males (Humphrey, 1998; Scheuer & Black, 2004). As individuals age, they experience degenerative changes in their skeletons, such as osteoarthritis or changes in vertebral morphology. These changes can be sex-specific, as men and women might experience different rates or patterns of degeneration due to differences in physical activity, hormonal changes, and other factors. The impact of senescence on sex estimation is more detailed in Chapter 2.4.1.

2.2 Estimation of stature

An important parameter for biological identification is the estimation of stature. This can subsequently aid in estimating the individual's health status, fitness, and body dimensions. Stature is characterized as a relative ratio of genetic factors and environmental influences (Lettre, 2009). Stature estimation can be approached through two main methods: the mathematical method and the anatomical method. The mathematical method relies on regression formulas derived from correlations of typically long bones, which are strongly associated with total stature. The anatomical method consolidates the lengths and heights of all skeletal elements contributing to total height, and incorporates mass of soft tissue (Raxter, Auerbach, & Ruff, 2006). The anatomical approach can be employed when a whole skeleton is accessible for forensic examination. For stature estimation of incomplete skeletons, the mathematical techniques like regression analysis and the multiplication factor method are used. The forensic relevance of these mathematical methods stems from the strong linear correlation observed between an individual's stature and the length of the body part or bone (Krishan, Kanchan, & Sharma, 2012; Verma, Krishan, Rani, Kumar, & Sharma, 2020).

The longest bones of the upper and lower limbs exhibit the highest correlation with stature (Trotter & Gleser, 1952). According to a study by Mall et al., the long bones of the hand reliably provide stature estimation in almost 95% of cases, with the humeral head individually accounting for 90.41% and the length of the radius for 89.13% (Mall et al., 2001). Raxter et al. (2006) revised method proposed by Fully et al. (1956) by incorporating new soft-tissue correction factors. The original method contains measurement of basion-bregma height of the cranium; maximum height of the corpus of the C2–L5 vertebra measured separately; anterior height of the first sacral segment; physiological length of the femur; maximum length of the tibia without the spine and including the malleolus; articulated height of the talus and calcaneus (Figure 2). Stature estimates are accurate within 4.5 cm in 95% of individuals in our sample, without any directional bias. According to Raxter et al. (2006) sex and ancestry do not influence stature prediction (Raxter et al., 2006). Conversely, Zeman and Benus (2020) recommend use of least squares regression equations specific to that population, if the population origin of the individual in question is known (or can be reliably assumed). However, when the skeleton is incomplete and the population origin is unknown (or cannot be reliably estimated), only methods based on direct proportionality are viable. Nonetheless, estimates derived from direct proportionality are not precise and should be approached with caution (Zeman & Beňuš, 2020).

Studies have shown positive correlations between stature and footprints (Babladi, Pejavar, & Reddy, 2014), foot length (Pandey, Roshan, Kharate, & Sonawane, 2014), or measurements in the craniofacial region (Krishan, 2008).

More advanced 3D methods can also be utilised for stature estimation - estimation from postmortem CT pelvic scans significantly correlates with stature (Torimitsu, Nishida, Takano, & Koizumi, 2014). Femoral measurements obtained from post-mortem computed tomography were used to formulate novel equations for estimating adult stature in Danish forensic cases (Zhang et al. 2020)

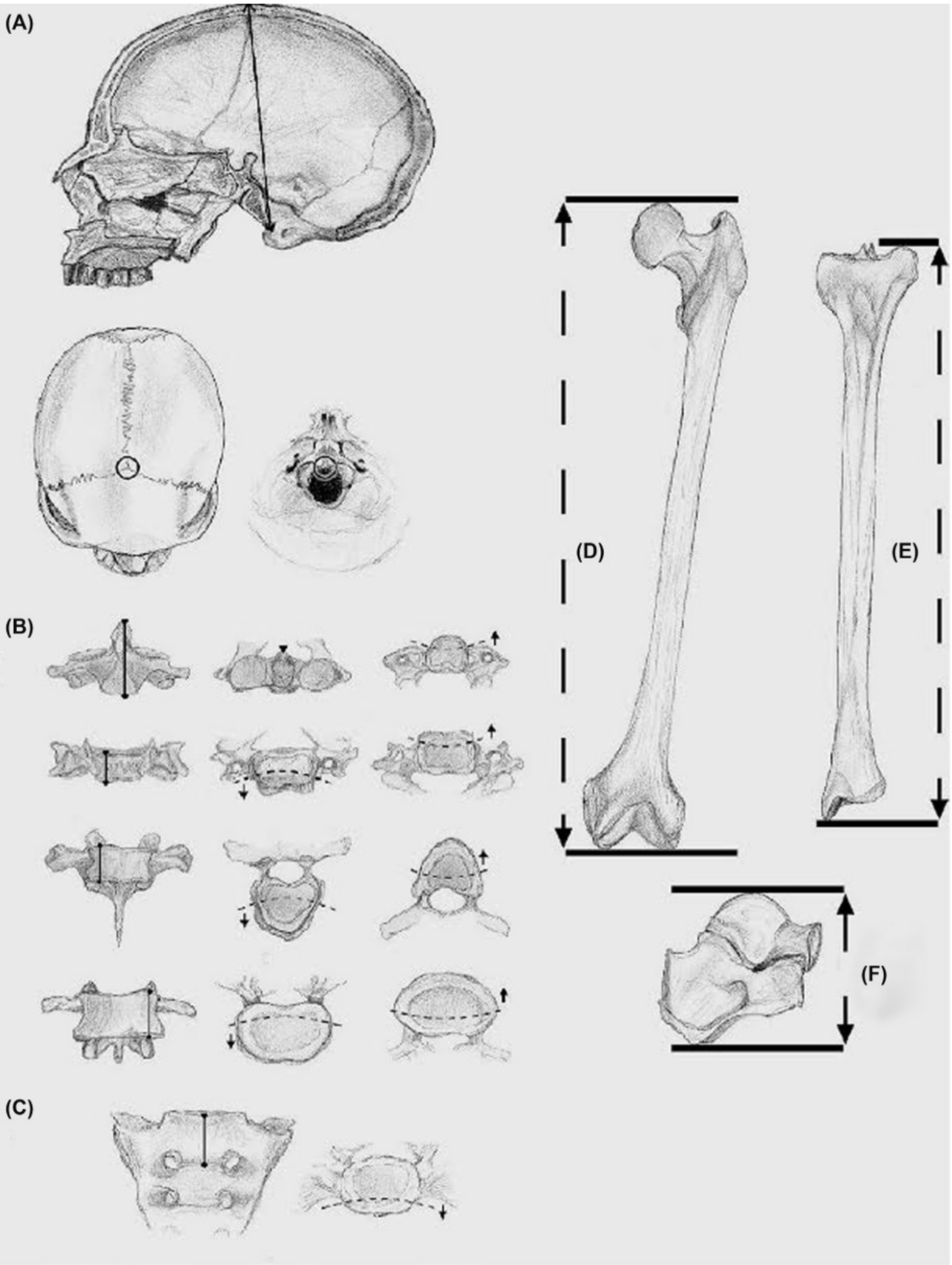


Figure 2. Illustration of measurements. A) The basion–bregma height of the cranium. (B) Heights of vertebral bodies: C2 (top) measured from the odontoid process to the inferior anterior rim; C3–L5 measured from the maximum anterior to the pedicles. (C) The height of the first sacral segment. (D) The physiological length of the femur. (E) The length of the tibia, from the lateral condyle to the medial malleolus. (F) The height of the articulated talocalcaneal joint. (Raxter et al., 2006).

Also determining someone's BMI (Body Mass Index) category, whether they are underweight, normal, overweight, or obese, could aid in identification. An individual BMI is recognised to impact the estimation of age-at-death from skeletal remains (Jeanson et al., 2017). Body mass predictions generally exhibit lower accuracy compared to stature estimations, even when employing population-specific equations (Sládek, Macháček, Makajevová, Přichystalová, & Hora, 2018).

Also, the estimation of stature is sex specific and the use of sex-specific equations is recommended whenever feasible (Sládek et al., 2018). This is because males and females generally have different average heights, body proportions or bone robusticity (Ruff 2000). Hence, different standards should be used. The equations are derived from statistical analyses of skeletal measurements from large samples of known-sex individuals. Separate formulas for males and females increase the accuracy of stature estimations (Uzun, Yegİnoglu, Ertemoğlu Öksüz, Kalkışım, & Zihni, 2019).

For archaeological and paleontological purposes, the estimation of body size and stature is an important component as it can serve as an indicator of health conditions, sexual dimorphism, environmental background, or social status (Ruff et al. 2012). Various formulas have been applied to estimate stature from long bone measurements in European archaeological findings (Giannecchini & Moggi-Cecchi, 2008; Sjøvold, 1990; Vercellotti, 2012; Vercellotti et al., 2014). Using ancient DNA and computing polygenic scores is another approach used to estimate height. It can explain portion of the variability in height among ancient individuals, but the score itself is insufficient for making reliable predictions of individual phenotypes (Cox et al., 2022).

2.3 Estimation of ancestry

The population specificity or origin of an individual can be considered a somewhat controversial area of research. However, for forensic anthropologists, it is one of the most important indicators and a crucial point of biological identification (DiGangi & Hefner, 2013). It refers to information about an individual's origin and their classification into a population group characterized by geography and culture (Konigsberg, Algee-Hewitt, & Steadman, 2009).

Populations manifest various degrees of sexual dimorphism , mainly due to different socio-economic and environmental influences, as well as unique genetic background, leading to variability in human development (Ubelaker & DeGaglia, 2017).

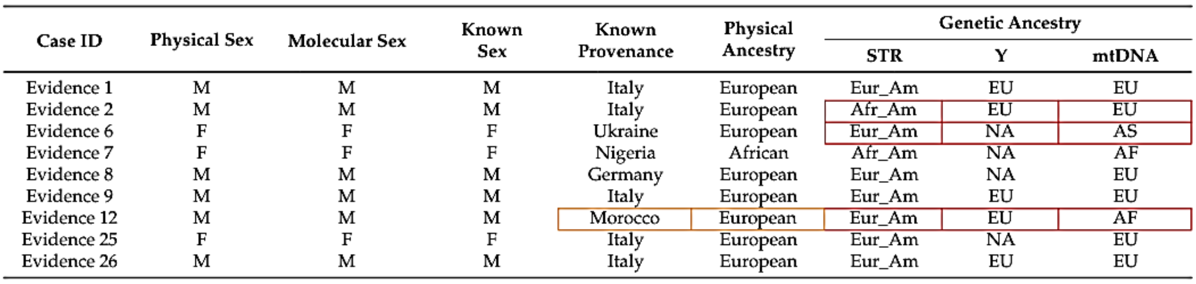
Knowing the origin of remains significantly narrows down the search in the missing persons database and expedites the successful identification of the individual (Ross, Slice, & Williams, 2010). Therefore, estimating the origin becomes a necessary but challenging task, especially given the current rate of globalization (Brůžek & Murail, 2006). In forensic practice, population affinity is often estimated as the first component of identification, because, for example, sexual dimorphism is highly population-specific (Milella, Franklin, Belcastro, & Cardini, 2021). In practice, this would mean that due to interpopulation differences in size and shape, a skull belonging, for example, to a Hispanic male might be incorrectly identified as a female skull of Caucasian origin (Spradley et al., 2008). The study by Dudzik and Jantz (2016) found significant misclassification issues when using Fordisc 3.1, a computer program used for estimating ancestry. The software often misclassified Hispanic individuals as Asian, especially Japanese, with accuracy rates for Hispanics being notably low. The study found that over 50% of Hispanic samples were misclassified. This misclassification is due to similarities in craniometric dimensions between these groups (Dudzik & Jantz, 2016). The effectiveness of methods initially designed for one population may vary greatly when applied to another population, leading to unpredictable variations (for more details see Chapter 2.4.2) (Franklin and Flavel 2019).

The study by Elliot and Collard (2009) assessed the effectiveness of Fordisc using 200 specimens with known ancestry. The analysis was performed under multiple scenarios: both including and excluding the test specimen's source population in the program's reference sample, as well as with and without specifying the sex of the test specimen. The program can correctly attribute ancestry only when the unidentified specimen belongs to one of the populations represented in the program's reference samples and the remains are relatively complete. Even under these optimal conditions, Fordisc can be expected to classify no more than 1 percent of specimens with high confidence (Elliott & Collard, 2009).

Artificial intelligence (AI) has increasingly been employed to address questions related to ancestry classification, leveraging its ability to analyse complex patterns and large datasets with high precision. Hefner et al. (2014) explored the use of AI for classifying ancestry based on cranial measurements, demonstrating the potential of machine learning in improving accuracy over traditional methods (Hefner, Spradley, & Anderson, 2014). In study by Navega et al. (2015) AI was used to analyse skeletal remains, showing it can effectively distinguish between different ancestral groups by identifying subtle morphological differences (Navega, Coelho, & Vicente, 2015).

The study by Pilli et al. (2023) aims to compare the effectiveness of genetic and physical anthropological methods in establishing the biological profile of unidentified and identified bodies in Milan. Genetic analysis offered more precise ancestry estimation, while physical anthropological methods relied heavily on morphological traits, which can be ambiguous (Table 1). The study concludes that combining genetic and physical anthropological analyses provides a more comprehensive and accurate biological profile for both unidentified and identified bodies (Pilli et al., 2023).

Table 1 Comparison of physical and genetic analysis outcomes in developing the biological profiles for the nine identified cases. Known sex and origin represent the antemortem data following confirmed identification. Discrepancies between physical and molecular analyses, as well as among genetic markers, are highlighted in red and orange. NA stands for not available. Eur_Am denotes European American, EU represents European, Afr_Am stands for African American, AS is Asian, and AF signifies African (Pilli et al., 2023).



Case ID	Physical Sex	Molecular Sex	Known Sex	Known Provenance	Physical Ancestry	Genetic Ancestry		
						STR	**Y**	**mtDNA**
Evidence 1	M	M	M	Italy	European	Eur_Am	EU	EU
Evidence 2	M	M	M	Italy	European	Afr_Am	EU	EU
Evidence 6	F	F	F	Ukraine	European	Eur_Am	NA	AS
Evidence 7	F	F	F	Nigeria	African	Afr_Am	NA	AF
Evidence 8	M	M	M	Germany	European	Eur_Am	NA	EU
Evidence 9	M	M	M	Italy	European	Eur_Am	EU	EU
Evidence 12	M	M	M	Morocco	European	Eur_Am	EU	AF
Evidence 25	F	F	F	Italy	European	Eur_Am	NA	EU
Evidence 26	M	M	M	Italy	European	Eur_Am	EU	EU

2.4 Estimation of biological sex

Estimating sex is essential in archaeological, anthropological, and forensic research. There are three possible approaches for this determination: osteology, genomics, and proteomics. The first two are traditional methods, while proteomics is a relatively new method that has been gaining increasing attention from scientists (Mikšík, Morvan, & Brůžek, 2023). Three approaches will be described in the following subchapters, with the primary focus on osteology.

Biological sex of individuals is inferred by forensic anthropologists from their skeletal biology, influenced by genetic sex chromosomes and bodily hormones, though not solely determined by them (Cabo, Brewster, & Luengo Azpiazu, 2012). It is important to differentiate between terms sex and gender as they are two distinct concepts that are often conflated. While they are frequently interchanged in conversation, documentation, and scientific literature, they represent different aspects of human identity and biology. Sexual dimorphism should be distinguished from gender, which encompasses the social and cultural roles and behaviours attributed to individuals. Gender is a socially constructed phenomenon influenced by elements like cultural norms, historical background, ethnicity, religious beliefs, and sexual orientation (Frayer & Wolpoff, 1985). In this thesis, the term "sex" will be used exclusively.

Also, there is a need for clarification regarding the nuances of sex estimation terminology. The term sex assessment describes historical approaches, often invalid and unreliable and also lacking the scientific rigor. Spradley and Jantz (2011) defined sex assessment as the utilisation of "morphological traits with no estimable error rates, classification rates, or any associated statistics" (Spradley & Jantz, 2011). Hence, it's advisable to move away from both the practice of assessment and the utilisation of the term itself. In contrast, sex determination implies a level of certainty approaching 100% accuracy, suggesting that sex can be regarded as a confirmed criterion. This applies for DNA analysis, but it is not advised to use in osteological studies "until accuracy rates consistently reach 100% (which will likely never happen due to human variation)" (Moore et al., 2013). As advised by many studies, the term sex estimation should be used in any process of inferring sex from skeletal parameters, along with associated metrics such as classification accuracy and error rates (Garvin & Klales, 2020; Moore et al., 2013; Spradley & Jantz, 2011; P. L. Walker, 2008).

Various parameters can be examined in order to estimate sex in biological anthropology (*Figure 3.*) as will be revealed in following chapters.

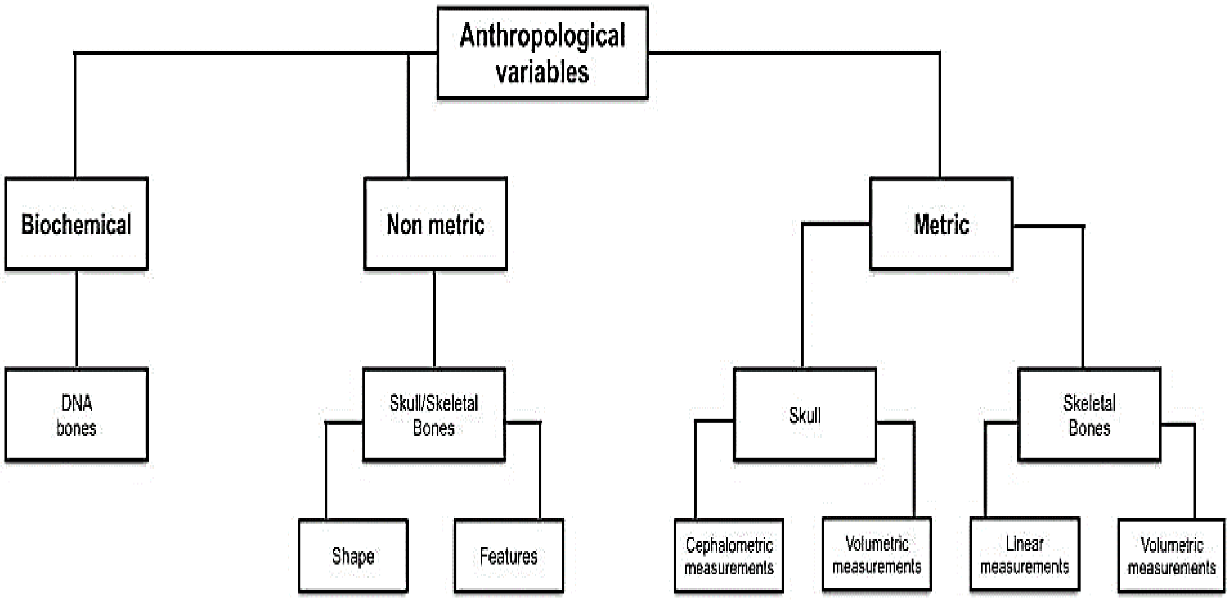


Figure 3. Flowchart of potential parameters for anthropological estimation of sex (Capitaneanu, Willems, & Thevissen, 2017).

2.4.1 Genomics and proteomics

Sex estimation of skeletal remains have evolved significantly with the advent of DNA isolation and amplification from bone (Hagelberg & Clegg, 1991; Hagelberg, Sykes, & Hedges, 1989). The introduction of DNA sequencing, known as Sanger sequencing or chain-termination method, was developed by Frederick Sanger and his colleagues in 1975, revolutionised the ability to determine the sex (Sanger & Coulson, 1975). It also expanded the scope of sex estimations to include fragmentary, pathological, and degraded skeletal materials, even estimating sex of very young individuals. Due to taphonomic degradation or poor preservation, DNA extraction is not always possible. In these instances, traditional anthropological methods remain the sole approach for determining the individual's sex (Thomas 2020). Genetics can employ ancient DNA to determine the biological sex of a skeleton at the individual level, infer phenotypic traits from an individual's genotype, and detect specific pathogens in an infected individual. However, aDNA research cannot be conducted in standard molecular biology laboratories, as it requires specific methodology (Raff, 2019).

The advancement in molecular methods is attributed to the identification of sequence variations in the amelogenin gene, which is responsible for producing one of the organic components of tooth enamel on the X and Y chromosomes. Stone et al. (1996) developed a method employing molecular genetic techniques, successfully identifying the sex of 19 out of 20 skeletons with known sex (Stone, Milner, Paäbo, & Stoneking, 1996). Forensic sex determination using PCR analysis utilises molecular markers such as amelogenin, SRY, DXYS156, and TSPY (Butler & Li, 2014). Generally, PCR-based methods, which focused on sex-specific molecular markers such as the amelogenin gene family, often encountered interference from contemporary contamination. That can be avoided by using of shotgun DNA sequencing, which is capable to identify chemical alterations distinctive of ancient DNA and discern exogenous DNA contamination. Furthermore, the recent development of massively parallel DNA sequencing has significantly enhanced genome coverage, further advancing sex estimation capabilities (Buonasera et al., 2020). For reliable sex estimations, Skoglund et al. (2012) proposed that a minimum of 100,000 total chromosome reads, mapped to the human reference genome (or 3,000 reads mapped to sex chromosomes), were required (Skoglund et al., 2012). Proposing an alternative method for sex estimation, Mittnik and colleagues (2016) employed high throughput shotgun sequencing of DNA. This method assesses sex by comparing the proportion of DNA reads mapped to the human X chromosome with the proportion of reads mapped to each of the autosomal chromosomes. With as few as 1,000 human genome reads, the method yielded confident results (Mittnik, Wang, Svoboda, & Krause, 2016). Compared to non-metric and metric measurements, DNA remains the most reliable and accurate method for sex determination (Capitaneanu et al., 2017).

Teeth, being the hardest tissue in the body, frequently stay intact after death, making them valuable for forensic identification (Silva, Sales-Peres, Oliveira, Oliveira, & Sales-Peres, 2007). In recent years, proteomics has emerged as a compelling method for investigating human, animal, and biological profiles, and origins. It serves as an alternative to DNA analysis, which faces constraints such as contamination risks, limited preservation of nuclear DNA or high costs (Gasparini et al., 2022; Parker et al., 2019). Innovative approach by Brůžek et al (2024) using proteomic sex analysis is based on the identification of two sex-dependent forms of the amelogenin protein in tooth enamel. The proteomic method for sex estimation demonstrated 100% accuracy in a sample of 60 teeth, comprising 32 males and 28 females. It proved to be a minimally invasive and dependable approach applicable to both contemporary and historical populations (Brůžek et al., 2024).

In comparative study by Buonasera et al. (2020), three independent methods of sex estimation were tested against each other (proteomic analysis of amelogenin peptides, shotgun-sequenced DNA, and standard osteological methods). In the archaeological sample, biological sex estimation was achievable for 100% of individuals using proteomics, for 91% using genomics, and for 51% using osteology. For optimal coverage and confidence they recommend a comprehensive approach combining proteomic, osteological, and genomic methods (Buonasera et al., 2020).

2.4.2 Osteology

Sexual dimorphism in human skeletal remains is evident in both size and shape, reflecting a mix of genetic and environmental influences typically manifesting around puberty (Moore et al., 2013). The success of methods for sex estimation is linked to the extent and manifestation of sexual differences in the examined skeletal structure. Sexual dimorphism is particularly evident in the postcranial skeleton (Papaloucas, Fiska, & Demetriou, 2008; Peckmann, Orr, Meek, & Manolis, 2015), most prominently in the pelvis (Decker, Davy-jow, & Hilbelink, 2011), where physiological demands for efficient bipedal walking (Lovejoy, 1988) collide with the requirements of giving birth to a relatively large-headed infant (Rosenberg & Trevathan, 2002; Rosenberg & Trevathan, 1996; Tague & Lovejoy, 1998; Trevathan & Rosenberg, 2000). As a result, the pelvis exhibits relatively high reliability for sex estimation (Brůžek, 2002). In terms of morphology, the most pronounced sex differences are observed in pelvic shape, primarily influenced by the functional demands and adaptations related to childbirth in females. Females typically exhibit a transversely wider pelvis, characterized by broader sciatic notches, relatively longer pubic lengths, greater subpubic angles, and sharper ischiopubic rami (Arsuaga, Lorenzo, & Carretero, 1995).

Phenice's method (1969) includes the ventral arc, subpubic concavity, and the medial aspect of the ischiopubic ramus, along with the greater sciatic notch and preauricular sulcus, for estimating sex from the pelvis, as endorsed by Buikstra and Ubelaker in their 1994 Standards for Data Collection from Human Remains (Buikstra & Ubelaker, 1994; Phenice, 1969). The presence of these traits is typically found in females and generally absent in males. The method was independently tested and showed an accuracy of 94.5%, thus it was deemed reliable and suitable for sex estimation (Klales, Ousley, & Vollner, 2012). However, another study that verified the reliability of the Phenice method on a sample with known sex achieved a classification accuracy of only 59% (MacLaughlin & Bruce, 1990). Also, validation study by Lovell (1989) reported an accuracy rate of 83 %. However, Lovell notes that age could be a factor, given that the individuals in this study were older than those in the original study (Lovell, 1989).

Currently, Brůžek's method (2002) is the only widely utilised technique that combines both morphological traits with a quasi-metric approach for estimating sex from the pelvis. The success rate of method, when the entire pelvis was used, was approximately 98%. Five regions were selected to reflect functional morphology: preauricular area, greater sciatic notch, composite arch, inferior pelvis, and ischiopubic proportion. Each trait is defined by one to three conditions and is evaluated as either female, male, or indeterminate. Based on the prevalence of female/male traits, the observer decides the sex. If the number of female and male traits is equal, the sex is evaluated as indeterminate (Brůžek, 2002). Validation study by Listi and Bassett (2006) achieved accurate classifications in 90 and 92 percent of cases (Listi & Elizabeth Bassett, 2006).

In study by Cao et al. (2021) they combined deep learning algorithm with three-dimensional surface scanning technology to create a method for estimating pelvic sex. Convolutional neural networks models achieved high performance using CT-based images of ventral pubis, dorsal pubis, and greater sciatic notch (Figure 4) with accuracies of 98.0%, 98.5%, and 94.0%. The feasibility was evaluated by applying well-trained convolutional neural networks models, originally developed with CT data, to predict images from 3D surface scans, resulting in satisfactory accuracies of 96.2% for ventral and dorsal pubis images. When these models were used for greater sciatic notch images, the accuracy dropped to 73.3%. This integration highlights the potential of combining deep learning algorithms with advanced imaging techniques for reliable sex estimation in forensic analysis (Cao et al., 2021).

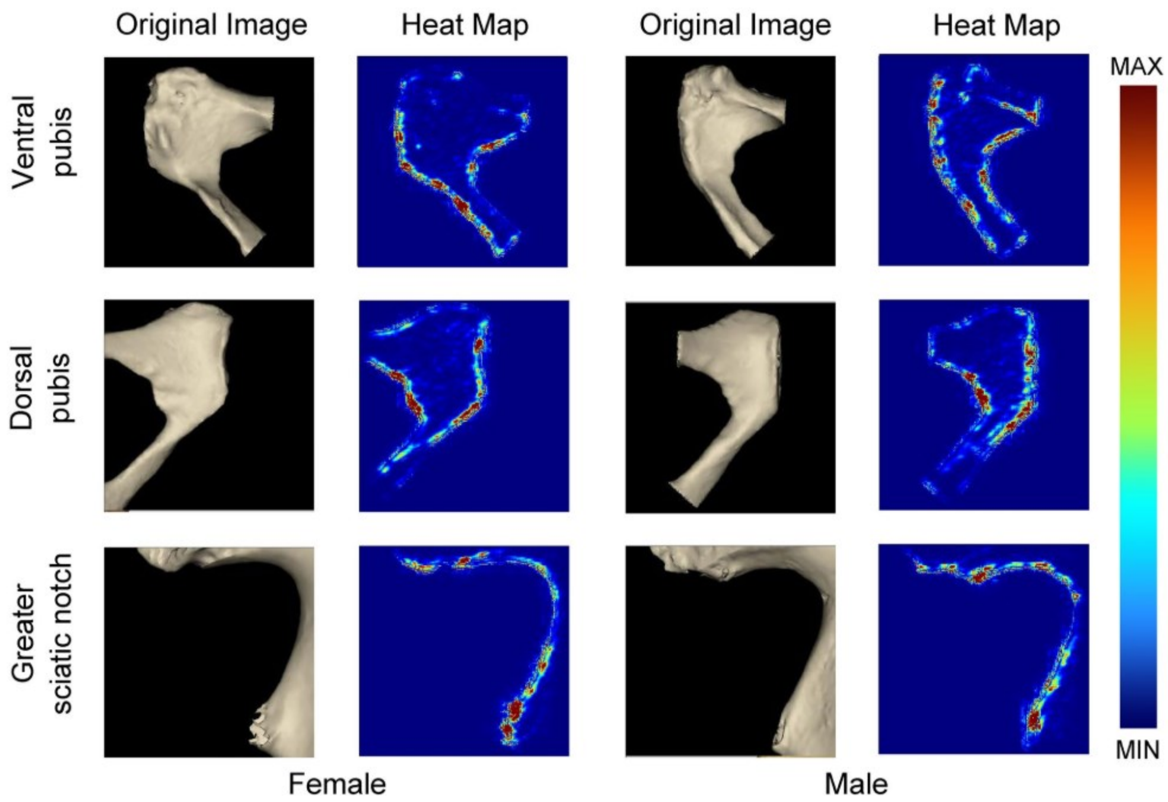


Figure 4. Colour-coded maps of the convolutional neural network models applied to CT scans display the ventral pubis, dorsal pubis, and greater sciatic notch. Red areas indicate significant contributions to sex estimation, while blue areas indicate minimal contributions (Cao et al., 2021).

However, pelvis is often missing, or found incomplete in a significant number of cases and identification is then subsequently based on the skull (Abdel Fatah et al., 2014; Buikstra & Ubelaker, 1994; Ferembach, Schwidetzky, & Stloukal, 1979; Garvin & Ruff, 2012). The human skull reflects the most significant evolutionary changes, such as exclusive bipedalism, encephalisation, specific neural trends, speech capability, and the ability to recognise individuals based on facial features (Lieberman, McBratney, & Krovitz, 2002; Martínez-Abadías, Paschetta, de Azevedo, Esparza, & González-José, 2009). These developmental trends have led to a series of morphological changes in the face, as well as the preference for partners based on facial or skeletal characteristics (Little, Jones, Penton-Voak, Burt, & Perrett, 2002; Lurdes et al., 2016). Consequently, these structures become primary indicators for sex estimation (Cosmides & Tooby, 1994; Thornhill & Gangestad, 1999).

The skull is a complex system of bones, which is frequently employed to estimate the biological profiles of individuals when the entire body is unavailable or poorly preserved due to postmortem changes or environmental factors (Abdel Fatah et al., 2014; Buikstra & Ubelaker, 1994; Garvin & Ruff, 2012). This thesis is dedicated to sex estimation using skull and the details will be revealed in following chapters. Many other studies focus on sexual dimorphism in the postcranial skeleton. For example, sexual differences have been observed in patella (Bidmos, Steinberg, & Kuykendall, 2005; Introna, Di Vella, & Campobasso, 1998; Mahfouz et al., 2007), clavicle (Alcina, Rissech, Clavero, & Turbón, 2015; Mohanraj, 2020; Panuganti, Reddy, Poojari, Shrish, & Kumar, 2022), humerus (Berner, Sládek, Holt, Niskanen, & Ruff, 2018; Işcan, Loth, King, Shihai, & Yoshino, 1998; Tallman & Blanton, 2020), femur (Cavaignac et al., 2016; Mahfouz, Merkl, Abdel Fatah, Booth Jr, & Argenson, 2007; Pandya et al., 2011), talus (Dagar, Sharma, & Khanna, 2019; Sorrentino et al., 2020), and tibia (Brzobohatá, Krajíček, Horák, & Velemínská, 2016; Kotěrová et al., 2017; Kranioti & Apostol, 2015).

Forensic anthropologists employ various methods to estimate sex accurately. The estimation of sex from skeletal remains is often challenging due to the frequent damage or incompleteness of the bones. The most significant sexual dimorphism occurs in the pelvic and cranial regions, making them the most suitable for developing a method for sex estimation (Scheuer, 2002). These methodologies may include morphological analysis, metric measurements, and statistical models based on population-specific data. Combining multiple approaches and considering regional variations have potential to enhance the reliability of sex estimation. Knowing individuals' ancestry significantly increases accuracy of sex estimation and using software that includes both parameters proved to be beneficial in process of identifying unknown human remains (Ousley & Jantz, 2013). Despite the distinct goals of forensic anthropology and bioarchaeology, both disciplines necessitate a high reliability of methods employed in constructing the biological profile. However, these methods encounter biological and methodological limitations, such as the degree of variation in sexual dimorphism across different populations (Kotěrová et al., 2017). For methodology review see Chapter 2.3.

2.4.3 Sexual dimorphism on skull

Sexual dimorphism refers to all differences between males and females and results from a combination of genetic factors and environmental influences. Sexual differences are quite common, across the animal kingdom and there are several reasons for this phenomenon. One of the most prominent causes is sexual selection, and ecological factors also play a role (Emlen & Oring, 1977; Slatkin, 1984). A component of sexual selection is female choice, which has shaped certain phenotypic traits in males, especially in facial features (Rhodes, Hickford, & Jeffery, 2000) and voice (O'Connor et al., 2014). A series of events in human evolution, such as the transition to obligate bipedalism, encephalisation, specific dietary changes, the development of speech, and the ability to recognise faces, likely caused a set of specific morphological changes in the skull. The morphological variations in the human skull are thus probably adaptations to all these evolutionary events (Lieberman et al., 2002). The mere existence of sexual dimorphism in the skull is a prerequisite for designing any methods for sex estimation.

It is essential to recognise the complexity of the human skull. The skull consists of 22 bones, not counting the three pairs of ear ossicles and the hyoid bone. Most of these bones develop through intramembranous ossification, while the base of the skull forms via endochondral ossification. The bones interlock with adjacent cranial bones through sawtooth or zipper-like articulations known as sutures. Sutures, made of fibrous tissue, connect the cranial bones and are key sites for bone growth along the bone margins during craniofacial development, particularly during rapid growth phases (Ruengdit et al., 2020). Anatomically, the skull can be divided into the splanchnocranium and neurocranium. Neurocranium is composed of eight bones: the frontal, parietal, temporal, occipital, sphenoid, and ethmoid bones. It serves to multiple vital functions, including protecting the brain and sensory organs, providing structural support for the face, and housing the initial segments of the digestive and respiratory systems (Morriss-Kay & Wilkie, 2005; Tiede et al., 1993). The morphology of neurocranium is influenced by various factors such as age, ancestry, diet, epigenetic traits, pathology, and sex (Axelsson, Kjær, Heiberg, Bjørnland, & Storhaug, 2005; Gonzales, Bernal, & Perez, 2009; Lieberman, Pearson, & Mowbray, 2000; Lillie, Urban, Lynch, Weaver, & Stitzel, 2016; Patel, Fernandez-Miranda, Wang, & Wang, 2016). Viscerocranium, also known as splanchnocranium, comprises of 14 bones that forms the face. This part includes the nasal bones, maxillae, zygomatic bones, mandible, palatine bones, vomer, inferior nasal conchae, and lacrimal bones. Face and facial expression carries a significant importance in partner selection so the splanchnocranium area is, unsurprisingly, heavily influenced by sexual selection. Facial attractiveness may thus be an indicator of health and fertility (Barber, 1995; Thornhill & Gangestad, 1999).

The size of neurocranium corresponds with the size of brain. Even in brain size, intersexual differences can be observed, with the male brain being on average 8-10% larger than the female brain (Filipek, Richelme, Kennedy, & Caviness Jr., 1994). The development of the brain and the shaping of the skull are two closely related processes involving a series of morphogenetic events. Therefore, the face, the cranial base, and the neurocranium do not behave as entirely independent functional units during ontogenetic development. To the contrary, their development is demonstrably interconnected (Martínez-Abadías et al., 2009). This principle aligns with the functional matrix theory – skull growth is a passive process dependent on the growth of the brain, meninges, cerebrospinal fluid quantity, etc. (Moss, 1968). In the context of sexual dimorphism, there is no need to anatomically divide the skull; instead, it is advisable to consider it as a whole.

There are some general anatomical differences between male and female skulls, although it's important to note that individual variation exists, and not all characteristics apply to every person. Some of the key differences are included in Figure 5 and Table 2.

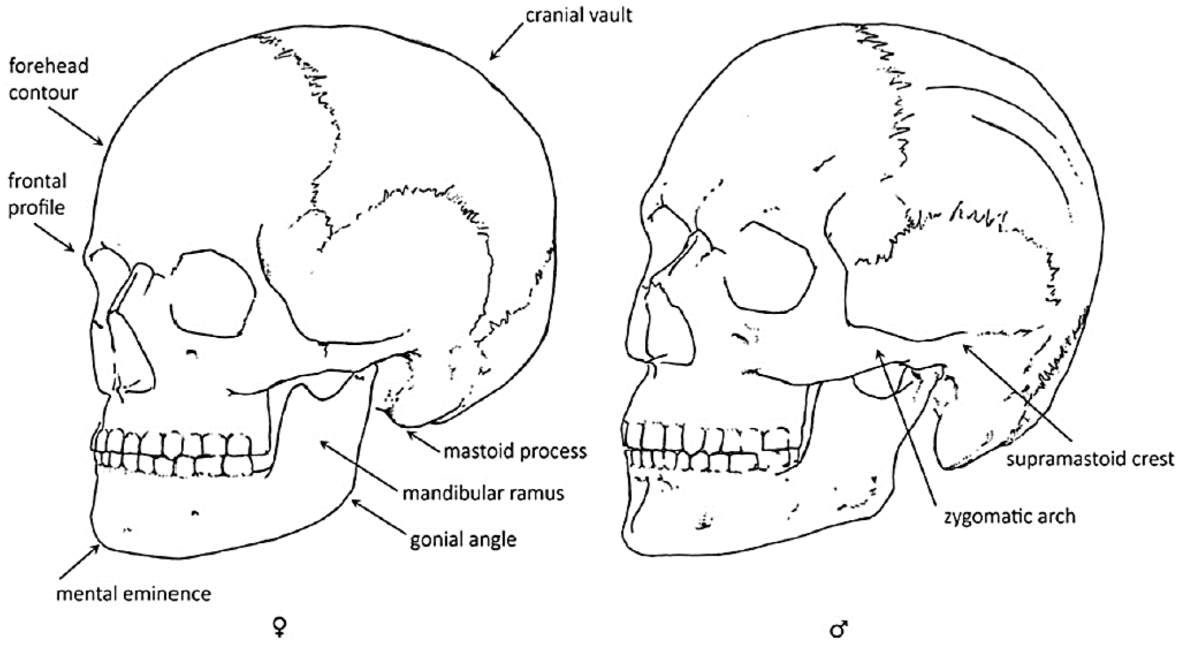


Figure 5. Average morphological sex differences between female skull (left) and male skull (right). Illustration drawn by A. DesMarais (Franklin, 2023).

Sexual dimorphism is evident in skull size, with males typically exhibiting larger and more robust crania (Best, Garvin, & Cabo, 2018; Kimmerle, Ross, & Slice, 2008). Cranial morphology differences between females and males are not solely dependent on size. Allometric traits are often associated with shape characteristics that are closely linked to size (Profico et al., 2017). This is particularly manifested in the development of the brow ridge, mandible, and mastoid processes. Females tend to present with a more rounded and smoother frontal bone, while males often display a more prominent supraorbital region. Males typically featuring a more robust and square-shaped mandibular morphology, whereas females generally exhibit a smaller and more rounded mandible. The external occipital protuberance may be more prominent in males. Notably, the glabellar and supraorbital areas emerged as two of the most sexually dimorphic regions in the study. These areas are considered crucial traits in the midfacial region, highly influenced by facial expressions (Velemínská et al. 2012; Hennessy et al. 2005). Males exhibited wider and more robust zygomatic arches, a trait often considered reliable for sex determination (Franklin et al., 2006). The mastoids and the nuchal region, including the nuchal crest, were found to be more prominent and robust in males due to muscle attachment (Figure 6) (Boucherie et al., 2022).

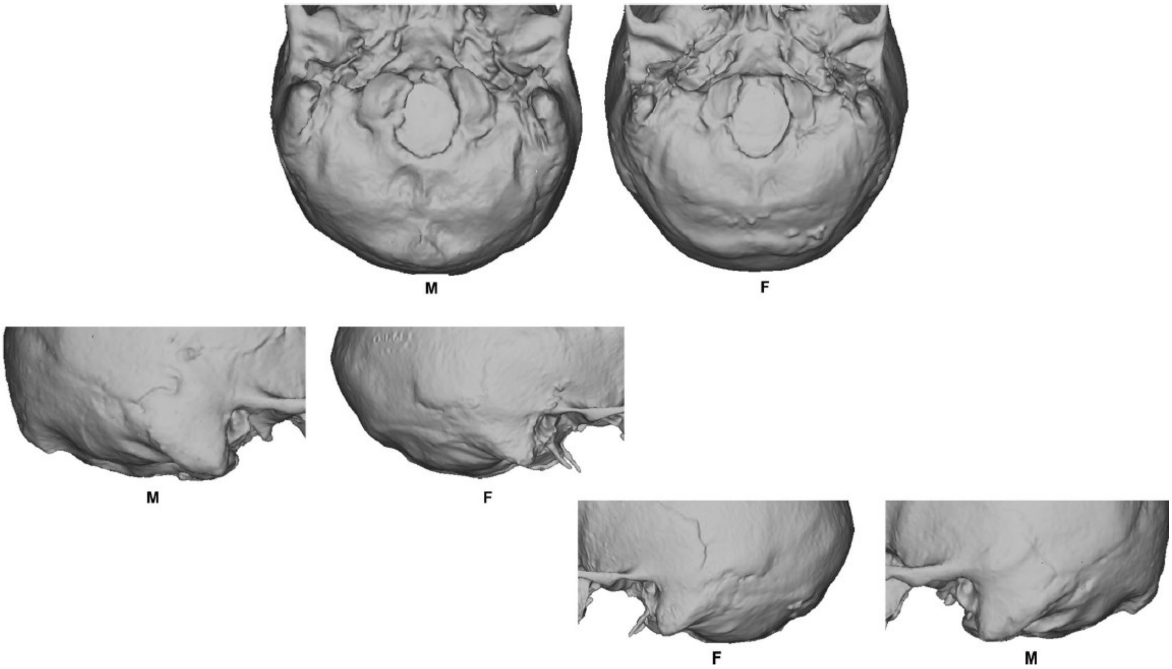
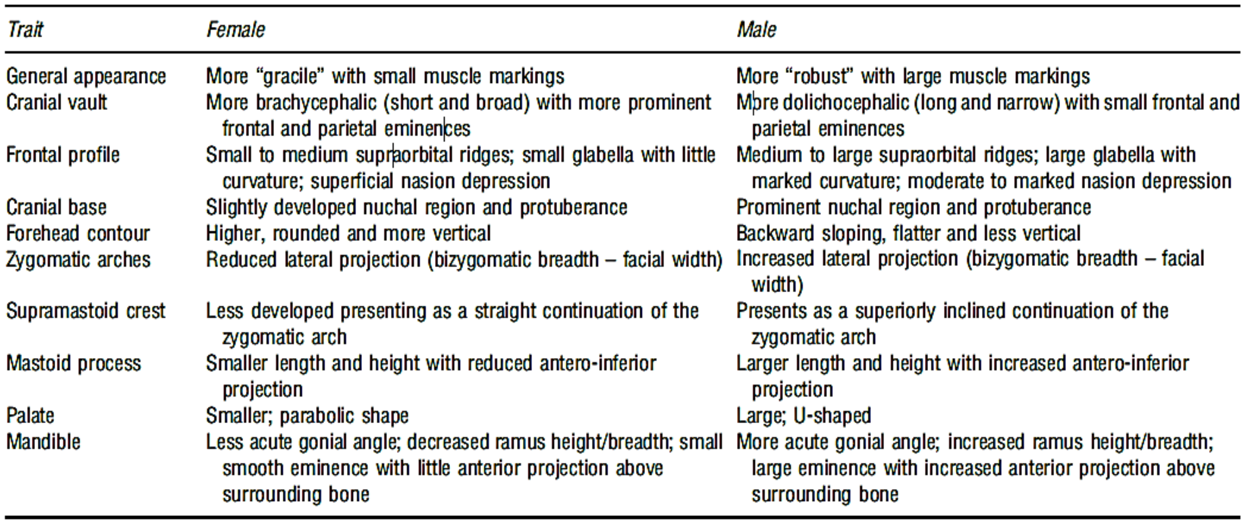


Figure 6. The primary sexual differences in occipital and mastoidal area. Surface models of two "extreme" individuals, one male (M, 91 years old) and one female (F, exact age unknown) (Boucherie et al., 2022)

Other distinct characteristics in males included a bevelled forehead, more prominent nasal bones, and a nasal spine with a wider nasal aperture. These features correlate with the longer growth period for the nasal parts in males compared to females (Bigoni, Velemínská, & Brůžek, 2010). It was observed that male cranial vaults were considerably thicker than those of females in all regions, except at the metopion and vertex points, where females exhibited a thicker frontal bone area than males (Abdel Fatah et al., 2014). Overall, males were observed to have larger, wider, and more robust skulls compared to females, attributed to differences in the masticatory system (Franklin et al. 2006).

Table 2. Cranial sex differences (Franklin 2023).



Trait	*Female*	*Male*
General appearance	More "gracile" with small muscle markings	More "robust" with large muscle markings
Cranial vault	More brachycephalic (short and broad) with more prominent frontal and parietal eminences	More dolichocephalic (long and narrow) with small frontal and parietal eminences
Frontal profile	Small to medium supraorbital ridges; small glabella with little curvature; superficial nasion depression	Medium to large supraorbital ridges; large glabella with marked curvature; moderate to marked nasion depression
Cranial base	Slightly developed nuchal region and protuberance	Prominent nuchal region and protuberance
Forehead contour	Higher, rounded and more vertical	Backward sloping, flatter and less vertical
Zygomatic arches	Reduced lateral projection (bizygomatic breadth – facial width)	Increased lateral projection (bizygomatic breadth – facial width)
Supramastoid crest	Less developed presenting as a straight continuation of the zygomatic arch	Presents as a superiorly inclined continuation of the zygomatic arch
Mastoid process	Smaller length and height with reduced antero-inferior projection	Larger length and height with increased antero-inferior projection
Palate	Smaller; parabolic shape	Large; U-shaped
Mandible	Less acute gonial angle; decreased ramus height/breadth; small smooth eminence with little anterior projection above surrounding bone	More acute gonial angle; increased ramus height/breadth; large eminence with increased anterior projection above surrounding bone

These morphological differences are subject to variation and should be interpreted with an understanding of the inherent variability within each sex, as well as the influence of factors such as age, ancestry, and individual genetic diversity. Sexual dimorphism arises from distinct patterns of growth and development, which are often influenced by geographic variability in genetic and environmental factors (Franklin & Blau, 2020; Ubelaker & DeGaglia, 2017). This has prompted numerous studies to examine the extent and manifestation of sexual dimorphism across various human populations - examples include research on North Americans (Kimmerle et al., 2008), on South Africans (Pretorius, Steyn, & Scholtz, 2006), and on various European populations (Gonzales et al., 2009; Rosas & Bastir, 2002).

2.3 Sex estimation on skull and methodological background

Many of the cranial traits that exhibit sexual dimorphism are not strictly discrete or categorical; rather, they exist along a spectrum of expression across male and female individuals (although practitioners may categorise them discretely in some methods). With advancing technology, more quantitative methods are emerging to analyse these traditionally "nonmetric" traits (Garvin & Klales, 2020). In the following chapters, the key methods will be presented and divided into three subgroups based on their approach: morphoscopic, morphometric, and virtual.

2.3.1 Morphoscopic methods

Morphoscopic methods for sex assessment involve the visual inspection and analysis of morphological features of skeletal remains to determine the biological sex of an individual. Anthropologists visually inspect the skeletal remains for the presence of sexually dimorphic features. This involves a detailed examination of the skull when features are compared to established standards and reference collections that include known males and females. These reference collections help in making informed comparisons.

Acsadi and Nemeskeri (1970) introduced a method for sex identification through the visual examination of skeletal elements. The Acsadi-Nemeskeri method employs a scoring system that assesses the degree of expression or the presence/absence of specific cranial features to determine an individual's sex (Figure 7). Morphological traits must be evaluated based on their degree of development, which varies between sexes. The method also includes detailed guidelines for positioning the skull prior to assessment (Acsádi & Nemeskéri, 1970).Two decades following the introduction of the Acsadi & Nemeskeri method, another landmark publication in physical anthropology emerged: "The Standards for Data Collection from Human Skeletal Remains," edited by Jane E. Buikstra and Douglas H. Ubelaker in 1994. They selected five morphological areas exhibiting high sexual dimorphism in shape, size, or robustness for cranial sex estimation (Figure 8). These include the robustness of the nuchal crest, size of the mastoid process, sharpness of the supraorbital ridge, prominence of the glabella, and appearance of the mental eminence. Each trait is evaluated on a scale from 1 (most feminine/gracile) to 5 (most masculine/robust) (Buikstra & Ubelaker, 1994).

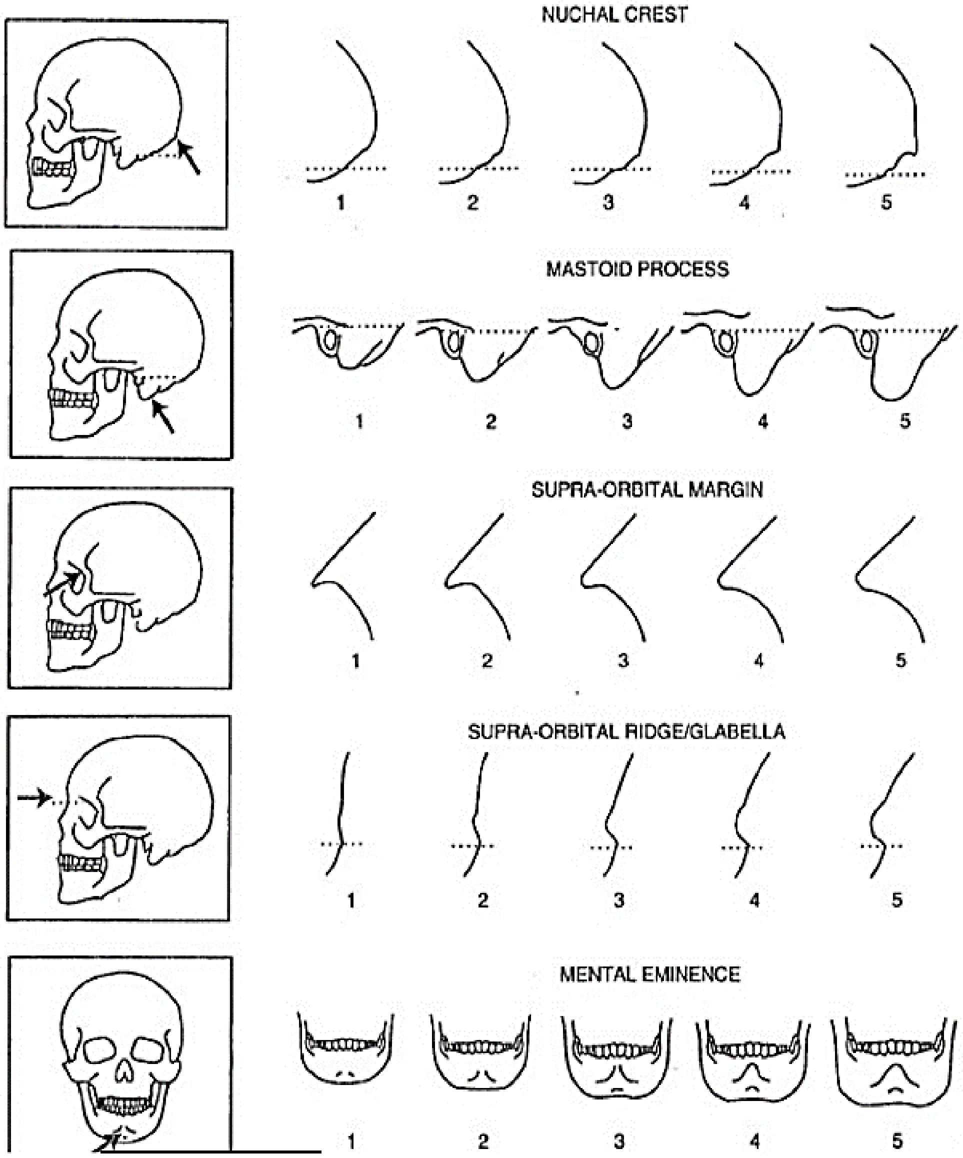
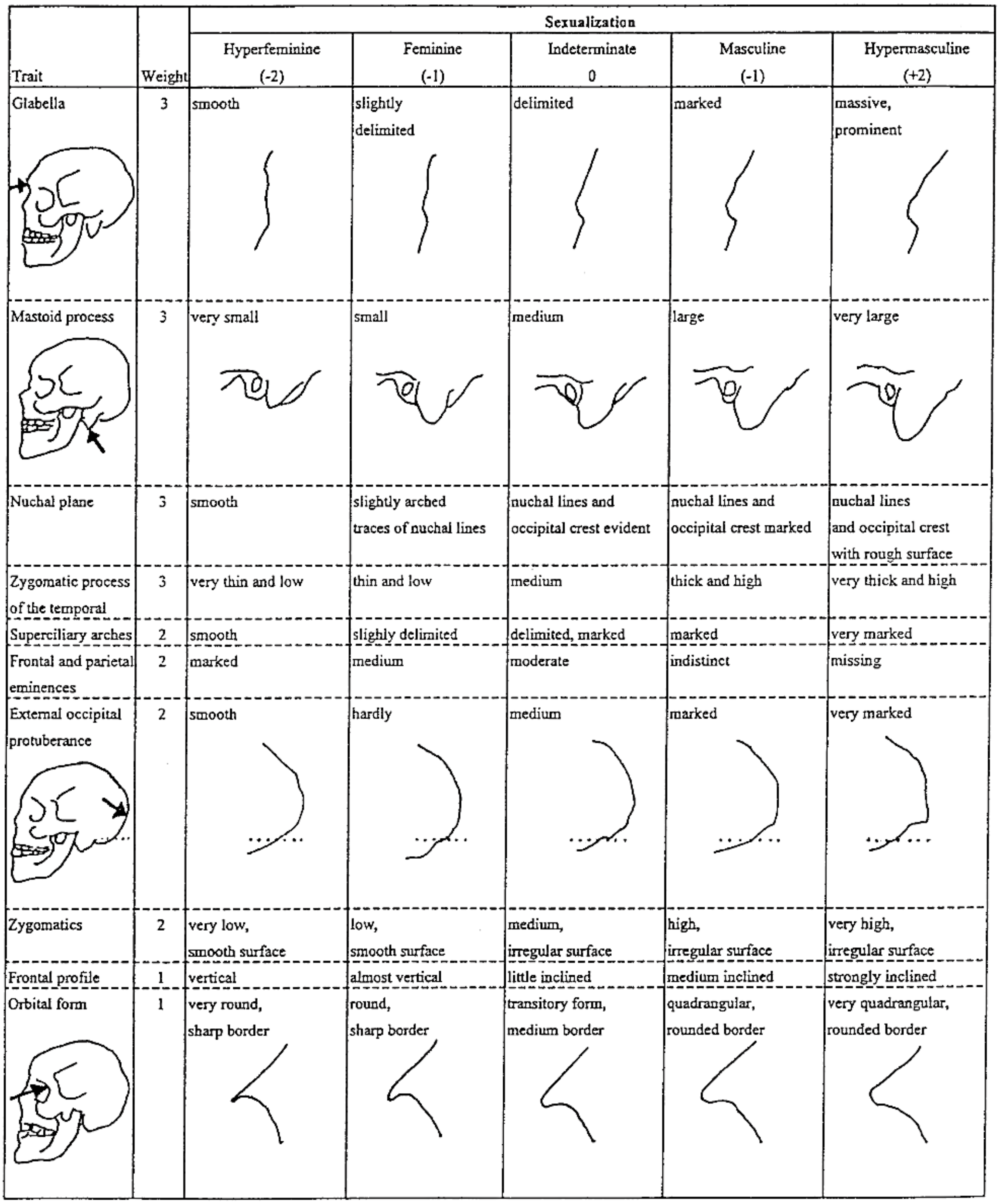


Figure 7. Scoring system for sex estimation using morphoscopic traits. Traits with the lowest expression are considered feminine traits, while the highest number indicates the greatest expression of the trait – a masculine trait (Acsádi and Nemeskéri 1970).

Walker verified this in his study by applying method to different populations. Utilising all traits, he achieved an accuracy rate of 88% in correctly identifying skulls. Considering age and population affinity did not significantly increase the accuracy of the method (Walker, 2008). Modified method was presented by Walrath (2004), where 10 cranial traits were visually evaluated by two observers. The study illustrates that not all osteological fragments are equally useful for sex estimation, regardless of their preservation state. Furthermore, it is crucial to thoroughly test the reliability of sexing methodologies, as well as using appropriate cross-population comparisons (Walrath et al., 2004).



		Sexualization				
Trait	Weight	Hyperfeminine (-2)	Feminine (-1)	Indeterminate 0	Masculine (-1)	Hypermasculine (+2)
Glabella	3	smooth	slightly delimited	delimited	marked	massive, prominent
Mastoid process	3	very small	small	medium	large	very large
Nuchal plane	3	smooth	slightly arched traces of nuchal lines	nuchal lines and occipital crest evident	nuchal lines and occipital crest marked	nuchal lines and occipital crest with rough surface
Zygomatic process of the temporal	3	very thin and low	thin and low	medium	thick and high	very thick and high
Superciliary arches	2	smooth	slighly delimited	delimited, marked	marked	very marked
Frontal and parietal eminences	2	marked	medium	moderate	indistinct	missing
External occipital protuberance	2	smooth	hardly	medium	marked	very marked
Zygomatics	2	very low, smooth surface	low, smooth surface	medium, irregular surface	high, irregular surface	very high, irregular surface
Frontal profile	1	vertical	almost vertical	little inclined	medium inclined	strongly inclined
Orbital form	1	very round, sharp border	round, sharp border	transitory form, medium border	quadrangular, rounded border	very quadrangular, rounded border

Figure 8. Characteristics for visually assessing sex determination of the cranium. Adapted from Ferembach et al. (1980) and Buikstra et al. (1994) (Walrath, 2004).

Williams and Rogers defined 21 cranial traits in their study. These include features such as the size of the mastoid processes, the size of the orbits, overall cranial size and structure, size and shape of the nasal aperture, and the gonial angle – these traits are assessed dichotomously, as either present/absent or pronounced/unpronounced, etc. Using 20 of these traits, they achieved an accuracy rate of 96%. Using only the six top-performing indicators (size and architecture, supraorbital ridge, nasal aperture, zygomatic extension, mastoid size, and gonial angle) resulted in a 94 percent accuracy rate for sex estimation (Williams & Rogers, 2006).

Morphoscopic methods are prone to significant observer error, which has driven the legal requirement for statistical accuracy and led to the adoption of metric methods. However, traditional approaches have identified the most sexually dimorphic areas, and these remain integral to modern methodologies.

2.3.2 Morphometric methods

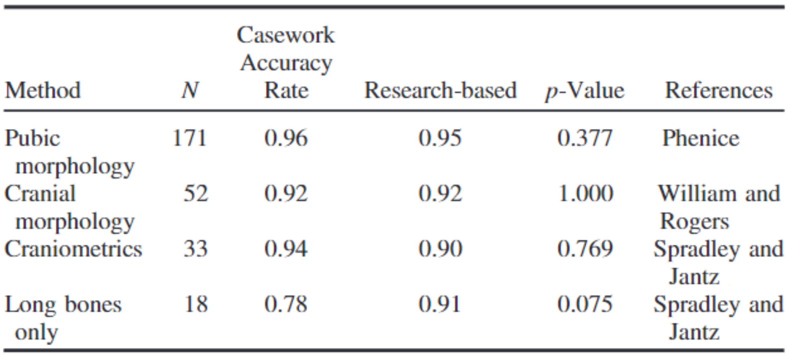
Morphometric methods are based on the quantitative measurement and analysis of anatomical features. Anthropometric measurements involve taking precise measurements of specific landmarks on the skull. Some methods use statistical models to assign probabilities of correct sex estimation. The measurement process is relatively economical, requiring only the use of a calliper or digital landmarks for acquisition. However, it can be challenging trying to convert the multidimensional concept of three-dimensional form into a quantification limited to two dimensions, specifically linear measurements (Paulinus et al., 2019).

The initial effort to employ morphometrics for sex identification from a skeletal collection of known sex was undertaken by Gilles & Elliot (1963). To classify sex based on linear cranial measurements, they introduced a mathematical model, specifically a multivariate linear discriminant analysis. Various measurements, such as cranial capacity, bizygomatic diameter, and basion-bregma height, were examined in order to develop a predictive model. Relying on objective measurements instead of subjective assessments of anatomical features, marked a significant improvement over previously developed qualitative methods. Despite its progress, the method faced limitations, including the requirement for a substantial sample size of skulls with known sex and the possibility of population-specific variations affecting prediction accuracy (Giles & Elliot, 1963). Gapert and Last (2009) selected the occipital region because it contains several prominent anatomical structures, including the foramen magnum and occipital condyles, which also exhibit sexual and population differences. Based on 12 measurements (e.g., width and length of the foramen magnum, width, height, and length of the occipital condyle, bicondylar width), they concluded that the occipital region is one of the suitable areas for further investigation of sexual dimorphism (Gapert, Black, & Last, 2009). Ali and Nakib conducted measurements on X-ray images using classical landmarks, further defined linear measurements and angle measurements. Significant differences between male and female skulls were found in all selected measurements. Using 11 measurements in a discriminant function analysis, they achieved an estimation accuracy of 86.7% (Ali & Al-Nakib, 2012). Ogawa and colleagues, for instance, worked on measuring the entire skull. On 113 modern Japanese skulls, they measured 10 traits (skull length, cranial base length, skull width, forehead width, etc.). They used these measurements in a discriminant function analysis, which achieved an accuracy rate ranging from 79% to 89.9% (Ogawa, Imaizumi, Miyasaka, & Yoshino, 2013). The mastoid triangle, delineated by the points porion, mastoidale, and asterion, has also been employed for sex estimation. De Paiva and Segre (2003) found that the mastoid triangles in males typically measure greater than or equal to 1447.40 mm^2, whereas those in females tend to be less than or equal to 1260.36 mm^2 (de Paiva & Segre, 2003).

Widely used software Fordisc is a tool used in forensic anthropology to help identify skeletal remains by estimating sex, ancestry, and stature from skeletal measurements (S. Ousley & Jantz, 2013). The user measures various dimensions of skeletal elements (such as the skull, long bones, etc.) and enters these measurements into Fordisc where statistical algorithms, primarily discriminant function analysis, are used to compare the entered measurements against reference databases (e.g., Forensic Data Bank, Howells Craniometric Data Set) of known individuals. Based on the comparisons, it classifies the unknown individual into categories for sex, ancestry, and sometimes stature and provides a classification probability (Dudzik & Jantz, 2016; Elliott & Collard, 2009; Flaherty, Byrnes, & Maddalena, 2023; Guyomarc'h & Brůžek, 2011; L'Abbé, Van Rooyen, Nawrocki, & Becker, 2011; Manthey & Jantz, 2020).

SexEst is another software tool used in forensic anthropology and bioarchaeology for estimating the sex of skeletal remains. Developed by the Cyprus Institute, it utilizes a combination of traditional morphological methods and modern statistical techniques to improve the accuracy of sex estimation. The software supports the analysis of various skeletal elements, allowing researchers to input measurements and receive sex estimation results based on validated algorithms. In study by Nikita et al. (2024) the highest classification accuracy, reaching 89.67%, was achieved by using a combination of postcranial variables with a posterior probability threshold set at 0.65 (Nikita et al., 2024). Another study reports classification accuracies for postcranial data ranging from 80.8% to 89.5%, and for cranial data ranging from 81.2% to 87.7% (Constantinou & Nikita, 2022).

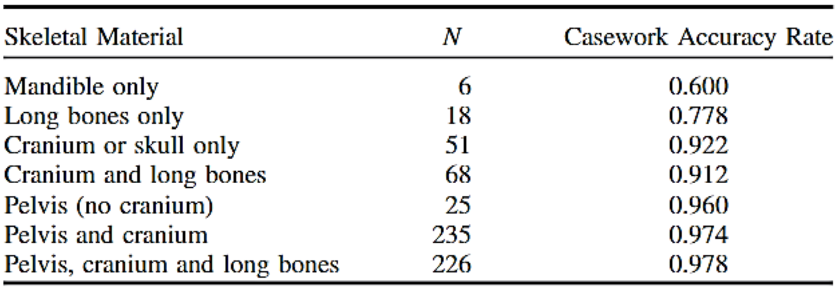
Table 3. Accuracy rates of osteological methods in comparison with DNA sex estimation (Thomas, Parks, and Richard 2016).



Method	*N*	Casework Accuracy Rate	Research-based	*p*-Value	References
Pubic morphology	171	0.96	0.95	0.377	Phenice
Cranial morphology	52	0.92	0.92	1.000	William and Rogers
Craniometrics	33	0.94	0.90	0.769	Spradley and Jantz
Long bones only	18	0.78	0.91	0.075	Spradley and Jantz

In study conducted by Thomas et al. (2016) the accuracy rates of four methods for sex estimation were compared with DNA analysis (Table 3). The case files at the FBI Laboratory, which contain both DNA-confirmed sex and anthropologically estimated sex, provide an opportunity for a comprehensive, longitudinal analysis of the accuracy rates of sex estimation in forensic anthropological casework. As anticipated, the sex estimation accuracy rate increased as more skeletal material became available for analysis. The highest rate (97.8%) was observed when a nearly complete skeleton was available, while the lowest rate (60.0%) was found when only a mandible was examined (Table 4) (Thomas, Parks, & Richard, 2016).

Table 4. Accuracy rates in forensic casework based on available skeletal material (Thomas, Parks, and Richard 2016).



Skeletal Material	*N*	Casework Accuracy Rate
Mandible only	6	0.600
Long bones only	18	0.778
Cranium or skull only	51	0.922
Cranium and long bones	68	0.912
Pelvis (no cranium)	25	0.960
Pelvis and cranium	235	0.974
Pelvis, cranium and long bones	226	0.978

2.3.3 Geometric morphometry, innovative 3D approaches and machine learning

Biological anthropology has undergone significant transformation with the advent of new investigative methods. Since the inception of virtual anthropology, there has been a surge in innovative strategies for sex identification, including geometric morphometrics methods, the application of machine and deep learning, 3D modelling techniques, and linear measurements (Čechová et al., 2019; Gao, Geng, & Yang, 2018; Santos, Guyomarc'h, & Brůžek, 2014; Toneva, Nikolova, Tasheva-Terzieva, Zlatareva, & Lazarov, 2022).

The research evolved by the advancements in modern imaging and analytical techniques, including computerized tomography, magnetic resonance imaging, and geometric morphometrics. These sophisticated methods have significantly contributed to the enhancement of research capabilities, enabling more precise and comprehensive analyses in various scientific domains (Zhang, 2024). Partnerships with hospitals and medical examiner's offices have facilitated the creation of extensive computed tomography databases, driving the advancement of sophisticated techniques (Ramsthaler, Kettner, Gehl, & Verhoff, 2010; Verhoff et al., 2008). Incorporating various data modes offers a more comprehensive representation of the diverse population, consequently influencing the reference collections used by forensic anthropologists (Spradley & Stull, 2018).

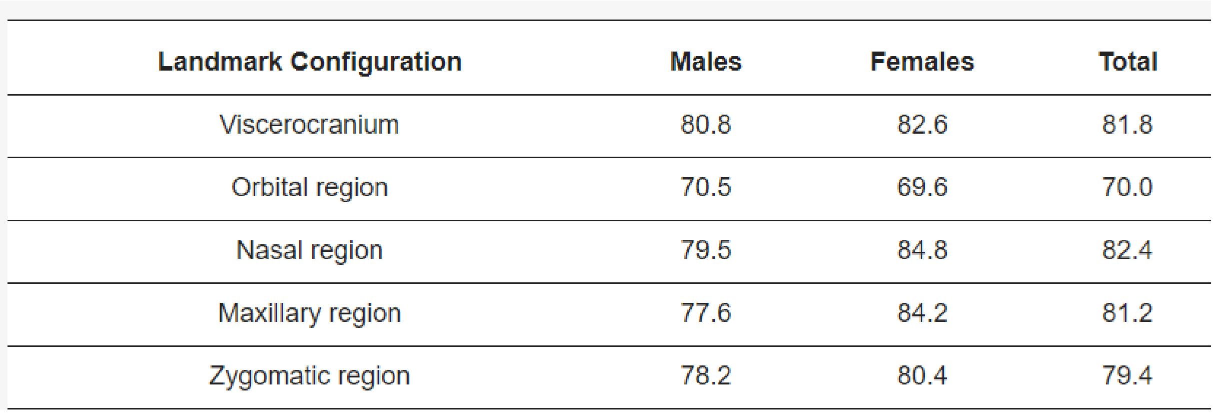
Geometric Morphometrics (GM) arose from the fusion of morphological shape description and statistical analysis (Adams, Rohlf, & Slice, 2004; Mitteroecker & Gunz, 2009; Slice, 2007; Zelditch, Swiderski, & Sheets, 2012). This integration allows for the visualisation of shape variation patterns among various shapes or groups of shapes. In mathematical terms, shape encompasses all geometric features of an object excluding its size, position, and orientation (Dryden & Mardia, 1998; Kendall, 1977). Various supervised machine learning classifiers are available, including k-nearest neighbours, Bayesian methods (such as Naïve Bayes and Gaussian process), Trees (including Decision Trees and ensemble methods like AdaBoost and RandomForest), and functions like Quadratic Discriminant Analysis and support vector machine (SVM). These approaches offer flexibility and robustness in handling complex classification tasks by adapting to diverse datasets and circumventing potential limitations of traditional methods (Mennatt, Attia, Tarek Farghaly, Ahmed El-Sayed Abulnoor, & Curate, 2022).

Traditional classification methods may encounter challenges in maintaining classification accuracies when faced with group differences, potentially due to bending or breaking theoretical or mathematical assumptions. In contrast, machine learning simplifies pattern recognition in previously unseen data and enables to maintain high classification accuracy. The evolution of research methodologies is evident - transitioning from craniometric examinations of complete crania (e.g. Franklin et al. 2006; Kranioti, İşcan, and Michalodimitrakis 2008; Saini et al. 2011) to more targeted investigations of particular traits. Recent studies have focused on various cranial regions: occipital and temporal bones (Boucherie et al., 2022), frontal bone (Čechová et al., 2019; Del Bove, Profico, Riga, Bucchi, & Lorenzo, 2020), brow ridge and chin (Garvin & Ruff, 2012), cranial base (Lestrel, Cesar, Takahashi, & Kanazawa, 2005), foramen magnum (Kanchan, Gupta, & Krishan, 2013), palate and cranial base (Chovalopoulou, Valakos, & Manolis, 2013), craniofacial morphology (Bejdová, Krajíček, Velemínská, Horák, & Velemínský, 2013; Bigoni et al., 2010), cortical thickness of cranial bones (Lillie et al., 2016) and neurocranium (Chovalopoulou, Valakos, & Manolis, 2016). Other researchers have applied 3D modelling to a wide range of cranial features, enhancing the precision and accuracy of cranial studies (Abdel Fatah et al., 2014; Dereli, 2018; Green & Curnoe, 2009; Imaizumi et al., 2020; Kelley & Tallman, 2022)

Bertsatos (2020) proposed an automated approach for estimating sex from cranial traits such as glabella, supraorbital ridge, occipital protuberance, and mastoid process. Classification performance exceeded 91% for the entire sample when combining all sex diagnostic traits in multi-feature sex estimation. Additionally, approximately three-fourths of the sample could be estimated with 100% accuracy based on posterior probability (Bertsatos, Chovalopoulou, Brůžek, & Bejdová, 2020).

The study by Del Bove and Veneziano (2022) introduces a new protocol for sex estimation based on cranial metric traits using a neural network. This method focuses on a small set of ten cranial measurements and aims to enhance the accuracy and robustness of sex estimation by preventing overfitting and ensuring broad population inclusiveness. The neural network model achieved a cross-validation accuracy of 86.7% (Del Bove & Veneziano, 2022). The differences in the size and shape of the viscerocranium between males and females using geometric morphometric techniques were investigated in study by Toneva et al. (2022). They studied images of 156 males and 184 females by using of 31 cranial landmarks and the highest accuracy for sex classification was obtained when considering both size and shape of the entire viscerocranium (Table 5). The study concludes that size is a more reliable indicator of sex than shape in the viscerocranium and the combination of size and shape yields the best results in sex estimation (Toneva et al., 2022).

Table 5. Classification accuracy (in %) based on the centroid size (Toneva, 2022).



Landmark Configuration	Males	Females	Total
Viscerocranium	80.8	82.6	81.8
Orbital region	70.5	69.6	70.0
Nasal region	79.5	84.8	82.4
Maxillary region	77.6	84.2	81.2
Zygomatic region	78.2	80.4	79.4

This research was originally inspired by study Abdel Fatah et al. (2014) based on the measurement of exocranial and endocranial dimensions from 222 skulls of Caucasian Americans. They used landmarks such as orbitale and porion to derive the Frankfurt horizontal plane and basion, glabella, inion, mastoidale, metopion, nasion, opisthocranion, opisthion, sella, supraglabella, vertex, and zygion were marked for registration. After setting the landmarks, a PCA was conducted. Shape analysis of the skull revealed the glabellar region and mastoid processes as the most significantly sexually dimorphic areas. The form analysis, which also considered size, identified the glabellar and occipital regions, mastoid processes, and zygomatic bones as the most significant areas. Based on these dimorphic regions a second analysis was conducted, which utilised measurements of distances, angles, and cranial vault thickness. It was found that male cranial vaults were significantly thicker than those of females in all regions, except at the metopion and vertex points, where females had a thicker frontal bone area than males. Subsequently, a linear discriminant analysis (LDA with leave-one-out cross-validation was performed. The highest accuracy in sex classification was achieved using 11 variables, with a success rate of 97.3% (Abdel Fatah et al., 2014).

Virtual methodologies can integrate data from various sources, including different populations and demographic groups, to create more comprehensive and robust models (Weber & Bookstein, 2011). This integration improves the generalisability of the methods across diverse populations, addressing the issue of varying sexual dimorphism (Franklin, Swift, & Flavel, 2016). Digital models and datasets can be easily stored, shared, and accessed, facilitating collaborative research and long-term preservation of data (Ramsthaler et al., 2010). This accessibility enhances the ability to validate and refine methodologies over time, contributing to continuous improvements in the field (Verhoff et al., 2008).

The integration of virtual methodologies in osteological analysis is demonstrated in this doctoral research. A novel virtual method for sex estimation was developed based on the exocranial surface of skulls and this study aims to assess its accuracy and prove its reliability.

2.4 Limitations and challenges

2.4.1 Senescence of human skull

Craniofacial development undergoes changes throughout adulthood (Mendelson, Hartley, Scott, McNab, & Granzow, 2007). These changes are attributed to a gradual and continuous process involving appositional growth, resorption, and remodelling. Bone remodelling is regulated by the endocrine system, with hormones such as parathyroid and thyroid hormones playing crucial roles. As individuals age, alterations in the endocrine system can impact bone structure and remodelling dynamics (Oddie et al. 1966).

Bone density and mass are critical aspects of skeletal health that differ significantly between males and females. Generally, males have greater bone density and mass compared to females. As men age, they experience a gradual decline in bone density. This reduction can lead to conditions such as osteoporosis, although typically at a slower rate compared to women. The slow progression in bone density loss means that men often maintain a higher bone mass for a longer period. In contrast, women experience a more rapid decline in bone density, particularly after menopause. The decrease in estrogen levels during this period accelerates bone loss, making women more susceptible to osteoporosis and related structural changes, leading to an increased risk of fractures and other bone-related issues (Kanis & Kanis, 1994; Kanis, Melton III, Christiansen, Johnston, & Khaltaev, 1994).

Craniofacial features undergo noticeable changes with age, influenced by both bone resorption and remodelling (Hohlweg-Majert, Schmelzeisen, Pfeiffer, & Schneider, 2006). In males, prominent craniofacial features such as brow ridges and larger jaws are characteristic. However, despite these changes, the craniofacial features in males often remain more pronounced compared to females. This persistence in prominent features can be attributed to the slower rate of bone density loss in males. For females, craniofacial features may undergo more significant changes with age. The decline in bone density, coupled with changes in soft tissue, can lead to alterations in facial structure. These changes can diminish the differences in facial features between older males and post-menopausal females. The loss of bone mass in areas such as the jaw and cheekbones (Habets, Bras, & Borgmeyer-Hoelen, 1988) can result in a less defined facial structure in elderly women, contributing to a convergence in the craniofacial appearance of the sexes as they age (Coleman & Grover, 2006; B. Mendelson & Wong, 2012).

The shape and size of the skull are significantly sexually dimorphic (Nikita, 2012). In study Ross et al. (1998) on changes in skull bone thickness, was concluded that thickness increases with advancing age. However, this manifests differently in men and women. In women, a slow increase in bone tissue can be observed in the frontal region around the age of 65 (hyperostosis frontalis interna), whereas in men, a slow decrease in bone tissue was observed in the frontal region (Ross et al., 2010). Recent research utilising 3D imaging methods confirmed the theory that the cranial vault is highly dimorphic – on average, women have a significantly thicker frontal region, while men tend to have a thicker occipital region (Abdel Fatah et al., 2014). Additionally, sexual dimorphism may be evident in the length of the middle part of the cranial base or the curvature of the supramastoid crest (Franklin et al., 2006). Internal structures, such as the frontal sinus, continues to enlarge throughout adulthood, particularly in females, which is likely linked to bone resorption (Greening et al., 2024).

Guglielmi et al. conducted an analysis, utilising quantitative computed tomography, on various aspects including the density of spinal trabecular bones within an Italian population. Their findings revealed a bone loss rate of 1.7% per year for females and 1.5% for males, covering an age range from 30 to 85 years. Notably, there was a more pronounced loss of trabecular bone material in both sexes after the age of 50, with females experiencing a more significant decline than males. This heightened bone loss is attributed to estrogen-related changes in the female body, coinciding with menopause around the age of 50 (Guglielmi et al., 1995).

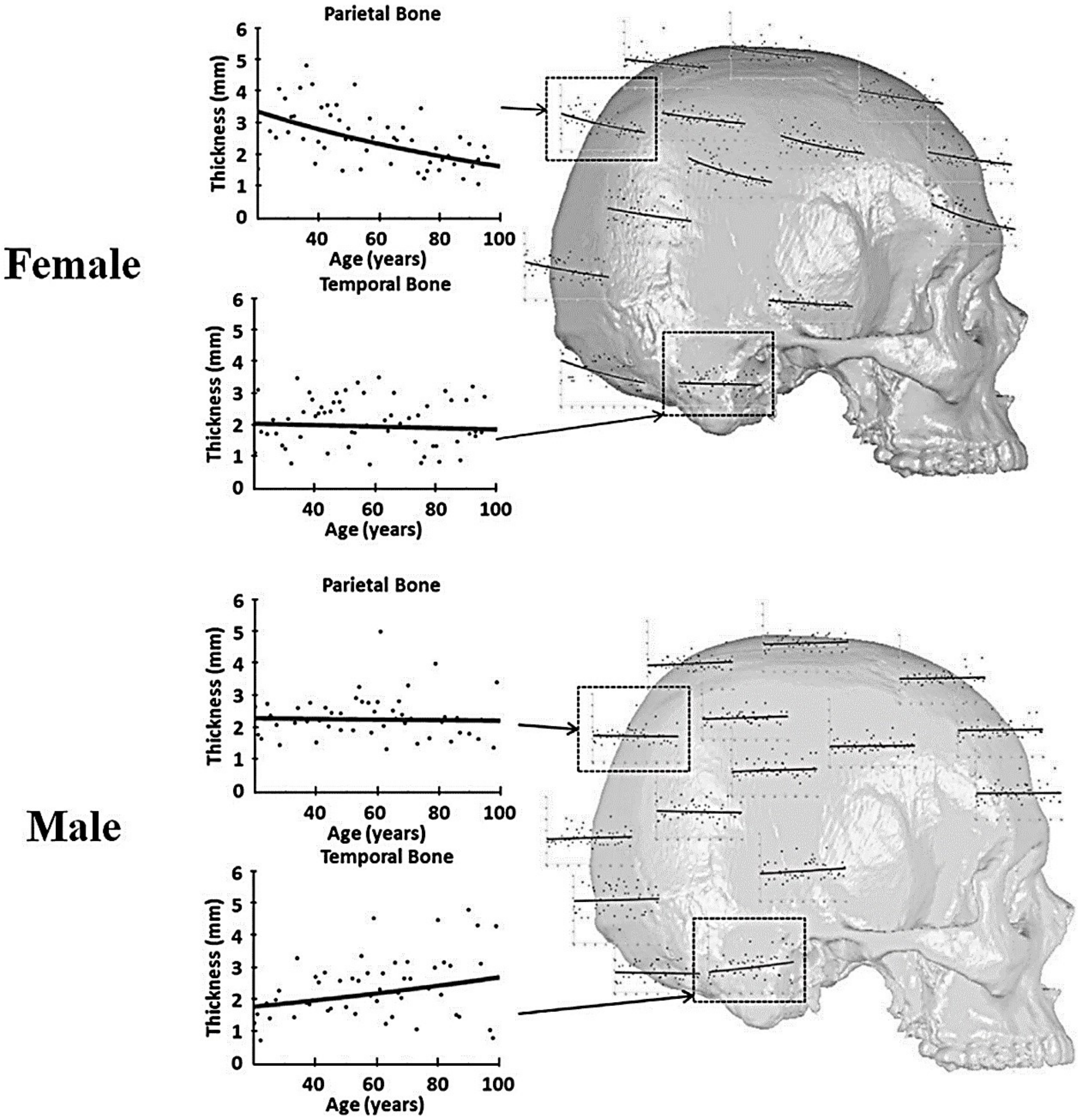


Figure 9. Regression trends in bone thickness of female (top) and male skull (bottom) (Lillie et al. 2016).

The study by Lillie et al. (2016) identified a notable correlation between age and cortical thinning in specific regions of the skull, including the frontal, occipital, and parietal bones. In females, there was a significant relationship, leading to a substantial decrease ranging from 36% to 60% in cortical thickness between ages 20 and 100 years. Intriguingly, males did not exhibit significant changes in cortical thickness across the same age range. The female parietal bone exhibited a notable reduction in cortical thickness with age, contrasting with the relatively stable thickness observed in males (Figure 9). Conversely, in the mastoid of the temporal bone, females showed minimal changes, while males displayed a slight increase in cortical thickness over time (Lillie et al., 2016).

Age-related changes in bone density and craniofacial features illustrate the complex interplay of genetic, hormonal, and environmental factors influencing skeletal health. Understanding these changes is important for interpretation of our results.

2.4.2 Population variation

Knowing the origin of remains significantly narrows down the search in the missing persons database and expedites the successful identification of the individual (Ross et al., 2010). Therefore, estimating the origin becomes a necessary but challenging task, especially given the current rate of globalization. In forensic practice, population affinity is often estimated as the first component of identification, because sexual dimorphism is highly population-specific (Brůžek & Murail, 2006). In practice, disregarding interpopulation differences might lead to incorrect sex estimation as discussed in chapter 2.3. Estimation of ancestry (Spradley et al., 2008). Kotěrová and colleagues (2017) empirically assessed the extent of error when population-specific discriminant functions, based on tibia measurements, were utilised in a Czech sample. They applied published classification functions designed for populations from Portugal, Spain, Italy, Greece, and North America to a set of 56 virtual models representing the modern Czech population. The findings indicate a clear failure in sex estimation when discriminant functions developed from geographically and temporally diverse populations were applied to the modern Czech population (Kotěrová et al., 2017).

Various studies in the literature highlight the population-specific nature of sexual dimorphism. Some articles demonstrate population-specific methods (Celbis & Agritmis, 2006; Colman et al., 2018; Spradley, 2021; Spradley, Anderson, & Tise, 2015), while others illustrate the diminished performance of methods derived from one population when applied to another (Bidmos & Dayal, 2004; Guyomarc'h & Brůžek, 2011; Kotěrová et al., 2017).

Urbanová et al. conducted a study to test the reliability of various software tools, including FORDISC 2.0, FORDISC 3.1.293, COLIPR 1.5.2, and 3D-ID 1.0, for sex and ancestry estimation using a sample of 174 human crania from a multi-ancestral Brazilian population.

The study aimed to assess the accuracy of these tools in classifying individuals from diverse ancestral backgrounds such as European Brazilians, Afro-Brazilians, Japanese Brazilians, and individuals of admixed ancestry. The software tools were able to correctly classify approximately 50% of the Brazilian specimens into their appropriate major reference groups. The study highlighted that while these software tools can be useful, their accuracy is significantly affected by the diversity and representativeness of the reference populations. The reliability of sex assessments varied, with an overall accuracy ranging from 60% to 71% across the software tools. COLIPR software showed the highest and most balanced accuracy rate for sex estimation, correctly classifying 78% of the crania. The results emphasise the importance of complementing software-based estimates with expert anthropological judgment and other lines of evidence (Urbanová, Ross, Jurda, & Nogueira, 2014).

When the population affinity is established before estimating age, sex, and stature, it's advisable to utilise population-specific standards if they are accessible for the particular population. The utilisation of suitable standards, leading to precise biological profile estimation, is accepted, and endorsed (Franklin, Cardini, Flavel, & Kuliukas, 2013; Ubelaker & DeGaglia, 2017). Unfortunately, it is not feasible to have a reference dataset for every population. Therefore, it is essential to develop a sufficiently robust and sophisticated method that can operate effectively across different populations. According to Sierp and Henneberg (2016), sex estimation methods would benefit from focusing on traits that directly reflect the influence of gonadal steroid hormones on skeletal morphology, regardless of the individual's size. However, size remains a crucial aspect of sexual dimorphism and cannot be entirely disregarded. Geometric morphometrics offers the separate examination of sex differences in both size and shape and provides a nuanced understanding of sexual dimorphism by analysing the complex interplay between these factors (Sierp & Henneberg, 2016).

2.4.3 Variations in sexual development and misgendering

The traditional binary classification of biological sex into male and female categories fails to encompass the complexity and variability inherent in human sexual development. While societal and medical definitions often rely on clear distinctions, biological sex is a spectrum influenced by various genetic, hormonal, and environmental factors. The variations in sexual development highlight the diversity and fluidity in the conventional male/female dichotomy (Blackless et al., 2000). Biological sex is determined by a combination of chromosomal configurations, hormonal profiles, and physical characteristics, all of which contribute to an individual’s sex phenotype. Typically, males are characterized by the presence of XY chromosomes, while females have XX chromosomes, but there are variations important to mention, such as Klinefelter syndrome (XXY), Turner syndrome (XO) or androgen insensitivity syndrome (Cools & Köhler, 2019).

The knowledge about sex estimation in intersex or transgender individuals is also limited, and there is a lack of reference data to address this issue (Buchanan, 2014; Tallman, Kincer, & Plemons, 2021). The complexity of gender identities has resulted in the use of various terms within academic discourse to describe this community. These terms include transgender and gender non-conforming, transgender and gender diverse, transgender and gender expansive, and transgender and gender variant (Isa, Michael, Blatt, & Flaherty, 2022).

The case study by Flaherty et al., (2023) explores the challenges faced by forensic anthropologists in accurately determining the sex of transgender individuals using traditional forensic tools like FORDISC 3.1. The study sample was transgender woman (male at birth) who had undergone extensive gender-affirming medical care, including facial feminisation surgeries. Testosterone-based hormonal therapy started at age 11 and from that underwent series of facial feminisation surgeries that change the skull morphology (e.g. rhinoplasty, gonial angle reduction, cheekbone reduction, orbital rim shaving, genioplasty) as well as hormonal treatment. Based on her 3D printed skull (Figure 10), the estimation of sex was conducted by standard craniometric assessment in FORDISC 3.1 and the results indicated that the skull belonged to a "White Male." This classification was inaccurate, as subject was a transgender woman. This demonstrates significant limitations in accurately identifying the sex of individuals who have undergone gender-affirming surgeries (Flaherty et al., 2023).

The misclassification highlights the challenges and potential inaccuracies of automatized methods. An expert would note the evidence of surgical interventions and gender-affirming procedures on subject's skull. Documenting and analysing this evidence would enhance the anthropologist's understanding of the individual both medically and socially.

Figure 10. Tested skull of transgender woman after facial feminisation surgeries. A) Anterior view; B) Inferior view; C) Lateral right view; D) Lateral left view (Flaherty et al.,2023).

Forensic anthropologists should intensify their efforts to determine if individuals have undergone gender-affirming medical procedures, especially given the current increase in the number of individuals undergoing gender transition (Weiss, 2024). Although this complex problem has not been widely explored in forensic anthropology literature, initiatives like the Trans Doe Task Force (http://transdoetaskforce.org/) are beginning to raise awareness about the challenges of identifying transgender decedents.

2.4.4 Observer error

The traditional osteological methods rely heavily on the skill and experience of the examiner (Franklin, Cardini, Flavel, Kuliukas, et al., 2013). Visual assessments can be subjective, leading to variability in results between different practitioners. Metric methods, while more objective, still require accurate landmark identification and consistent measurement techniques, which can vary based on the examiner's proficiency (Buikstra & Ubelaker, 1994; Pilmann Kotěrová et al., 2024).

Virtual methods reduce subjectivity by using consistent, automated processes for data collection and analysis. This standardisation minimises human error and variability, leading to more reliable and reproducible results (Mitteroecker & Gunz, 2009). Techniques such as 3D imaging, CT scans, and MRI can capture detailed and precise representations of skeletal features (Alves Proença, Slavicek, Cunha, & Sato, 2014; Lentini, Kasahara, Arver, & Savic, 2013; Verhoff et al., 2008; Villa, Lynnerup, & Buckberry, 2019). These digital models allow for comprehensive analysis of bone structures without the need for physical handling, which can sometimes lead to damage or misinterpretation (Slice, 2007). By applying geometric morphometric techniques, virtual methodologies can analyse shape and form with high precision (Adams et al., 2004). These techniques use mathematical models to quantify the geometry of skeletal structures, providing detailed insights into morphological differences that might be subtle or difficult to measure with traditional methods (Franklin et al., 2016). Automated processes and machine learning algorithms can handle large datasets efficiently, identifying patterns and making classifications with minimal human intervention (Martínez-Abadías et al., 2009). This efficiency is particularly valuable in forensic contexts, where timely and accurate identification is crucial (Avent, Hughes, & Garvin, 2022).

This research introduces an automated approach to sex estimation from cranial measurements, effectively reducing observer error. By leveraging advanced machine learning techniques and three-dimensional imaging, the method ensures high accuracy and reliability, thereby minimising the subjective biases typically associated with manual assessment.

3. THE AIMS OF THE RESEARCH

This thesis encompasses years of research regarding skull morphology, sexual dimorphism, and sex estimation. Based on the previous chapters and the scope of research, four main objectives were established.

1) Introducing new method for sex estimation using exocranial surface

The first objective was to assess accuracy of virtual methodology created in the 3D Imaging Methods Laboratory at the Faculty of Science, Charles University, for the author's diploma thesis. Developed methodology was inspired by study of Abdel Fatah et al. (2014), which using both exocranial and endocranial surface networks of skulls, achieved a high success rate in sex classification (97%).

Presented method modifies the approach to only exocranial surface and simplifies the accessibility of method by more automated process. The classifier relied on geometric morphometric analyses (CPD-DCA, PCA, SVM), achieving a sex estimation accuracy surpassing 90% for recent French population sample.

2) Analysing the reliability of method by broadening the dataset

The second objective was to assess reliability of previously developed method. The accuracy refers to the percentage of individuals correctly classified from the given sample. However, for sex estimation in bioarchaeology or forensic investigations, high method reliability is crucial. The reliability of the method is established through further testing on independent populations (Brůžek & Murail, 2006).

Method was successfully validated using a cohort of 208 individuals from two contemporary European populations. Incorporating data from the Czech population further enhanced the method's reliability, resulting in the highest accuracy of 96.2%. With the combined dataset, the method demonstrated reliability of 91.8%.

3) Observe and describe changes in sexual dimorphism with increasing age

The third objective was to visualise senescence and estimate changes in sexually dimorphic areas using colour-coded maps. Throughout the lifespan, individuals of both sexes typically undergo a transformation in craniofacial morphology, tending to adopt a more masculine character. This observation, as noted by Meindl et al. in 1985, suggests that certain morphological features associated with masculinity become more prominent as individuals age (Meindl et al. 1985).

By applying the presented method reduction in sexual dimorphism between the younger and the older subgroup was observed. Specifically, the proportion of significantly distinct areas decreased in cranial form by 12% (from 94.8% to 82.6%).

4) Verify the method applicability in multi-population sample

The fourth objective was to apply our method to large multi-population sample in order to examine reliability of sex estimation. Sex estimation standards typically vary by population, but it is argued that machine learning techniques could enhance the accuracy of biological sex estimation across different populations (Mennatt et al., 2022).

The final testing phase was enhanced by incorporating data from Slovak, Danish, and Egyptian populations into the original dataset, which initially included individuals from France and the Czech Republic. Further on, the unpublished findings from support vector machine and posterior probability analyses are presented.

4. HYPOTHESES

Hypothesis 1: The novel methodology developed for sex estimation using the exocranial surface of skulls will achieve an accuracy rate comparable to or higher than the established methods, specifically surpassing 90% accuracy in a sample population.

Hypothesis 2: The sex estimation method will demonstrate high reliability and accuracy when tested on a larger and more diverse dataset, maintaining an accuracy rate above 90% across different contemporary European populations.

Hypothesis 3: Significant changes in sexual dimorphism of craniofacial morphology will be observed with increasing age, with a notable reduction in the proportion of sexually dimorphic areas in older individuals compared to younger ones.

Hypothesis 4: The developed sex estimation method will be applicable and reliable across multiple populations, demonstrating high accuracy rates and robustness when applied to diverse samples, including individuals from Slovak, Danish, Egyptian, French, and Czech populations.

5. MATERIALS AND METHODS

5.1 Materials

5.1.1 Ethical statement

All experiments were conducted in accordance with relevant guidelines and regulations. Relevant Ethics Committee approved the study. Patient consent was not required by the ethics committee as all data was pre-existing and de-identified.

5.1.2 Study population

Material includes five datasets (Table 6) that contain individuals from different regions of Europe, namely central Europe (Czech population, Slovak population), northern Europe (Denmark population), south Europe (French population), and North Africa (Egypt population). CT scans of skulls were collected from a total of 619 individuals. The inclusion of individuals from diverse regions allows for the investigation of physical variations among these populations.

Table 6. Populations used in the study and demographic details

Population (country of dataset collection)	Number of individuals	Number of women	Number of men
Czech	143	59	84
Slovak	92	55	37
Denmark	180	89	91
French	103	51	52
Egypt	101	50	51
Total	619	304	315

The first investigated dataset consists of 103 CT images of a Mediterranean southern European population from France. There are 51 women in the group in the age range of 28-90 years, when the average age is 58 years; and 52 men in the age range of 18-92 years, when the average age is 52 years. These data were collected and anonymised with the approval of the Ethics

Committee of the University of Aix-Marseille: Faculty of Medicine in Marseille, France. Scans were acquired using a Siemens Sensation 64 scanner (Erlangen, Germany) at the Department of Radiology at the North Hospital in Marseille, France.

The second examined dataset consists of 143 CT images of the Central European population from the Czech Republic. There are 59 women in the group aged 21-77, with an average age of 51; and 84 men in the age range of 25-84 years, when the average age is 48 years. These data were collected and anonymised with the approval of the ethics committee at the Radiology Department at Na Homolce Hospital in Prague. Scans were acquired using a Siemens Somatom Sensation 16 scanner (Erlangen, Germany).

The third examined dataset consists of 92 CT images of the Central European population from Slovakia. There are 55 women in the group in the age range of 27-88 years, when the average age is 65 years; and 37 men in the age range of 19-97 years, when the average age is 65 years. This data was collected and anonymised with the approval of the Charles University ethics committee.

The fourth investigated dataset consists of 101 CT images of a North Africa population from Egypt. There are 50 women and 51 men in the group, aged 20–86 years. This data was collected and anonymised with the approval of the ethics committee at the Alexandria Faculty of Medicine (review report number 0304454/11/2019). Scans were acquired using an Aquilion 64, Toshiba Medical Systems (Nasu, Japan).

The fifth investigated dataset consists of 180 CT images of the Northern European population from Denmark. There are 89 women and 91 men in this dataset. This data was anonymised and divided into three age subgroups as follows: Group 1 (ages 20-39), Group 2 (ages 40-59), and Group 3 (ages 60 and above). The distribution of individuals in each group was as follows: 59 individuals in Group 1, 60 individuals in Group 2, and 60 individuals in Group 3.

5.2 Methods

5.2.1 Image processing

CT scans were manually converted into surfaces using Amira and/or Avizo (Thermo Fisher Sciences Inc., Waltham, Massachusetts). Triangle meshes were trimmed in Rapidform XOS (INUS Technology, Inc., Seoul, South Korea) and/or MeshLab (ISTI - CNR, Pisa, Italy). In this step, the process involved manually eliminating the spine and unrelated objects from the dataset. Additionally, the jaw and teeth were excluded due to their inherent variability, and any shadow artifacts arising from dental fillings were also removed. Despite these adjustments, the surface model retained the endocranium and other internal structures within the skull.

To optimise the computational efficiency and reduce complexity, the surfaces were then simplified to approximately 50,000 triangles. This simplification aimed to strike a balance between preserving essential details and ensuring manageable computational loads for subsequent analyses or applications.

Exocranial surfaces were acquired in Morphome3cs (Charles University, Prague, Czech Republic). An automated procedure was implemented to minimise user interaction. This involved constructing an axis-aligned cube around the skull, with a side length set to 100 times the distance between the two most distant points of the skull. The cube's centre and the vertex centroid of the skull were aligned. The set Q consisted of 26 points, including the 8 cube vertices, 12 edge mid-points, and six face centres. The exterior surface of the skull was then generated by removing triangles not visible from at least two points of Q. A point, such as point a, was considered visible from point b if no triangle intersected the line segment ab. This automated processing, conducted in Morphome3cs, took only seconds per skull, significantly reducing the manual workload while efficiently producing a simplified exterior surface of the skull.

Eight landmarks (Table 7) were manually placed on cranial surface in Morphome3cs by an expert anthropologist (Figure 11).

Table 7. List of used landmarks with Martin's handbook definitions (Martin & Saller, 1957).

Landmark	Definition
Glabella	Most anteriorly projecting point in the midsagittal plane at the lower margin of the frontal bone, which lies above the nasal root and between the superciliary arches
Inion	Most prominent point of the external occipital protuberance
Opisthion	The point at which the midsagittal plane intersects the posterior margin of the *foramen magnum*
Nasospinale	The point at which the midsagittal plane intersects lowest point of lower margin of *apertura piriformis*
Mastoidale dx, sin	Lowest point of the mastoid process (*processus mastoideus*)
Zygomaticofrontale dx, sin	The point at which the inner orbital margin intersects the *sutura zygomaticofrontalis*

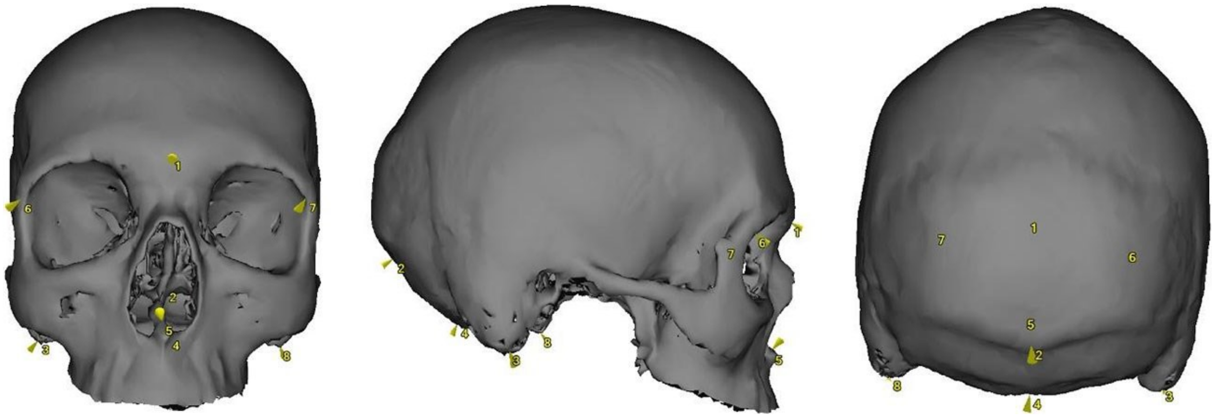


Figure 11. Cranial landmarks. Eight landmarks specifically selected on exocranial surface in anterior, lateral, and posterior views: glabella (1), opisthion (2), inion (4), mastoidale right (3) and left (8), nasospinale (5), zygomaticofrontale right (6) and left (Velemínská et al., 2021).

Before engaging in statistical processing, it was imperative to ensure vertex homology among the surfaces. Among the various applicable procedures, the chosen method was coherent point drift-dense correspondence analysis (CPD-DCA) (Dupej, Krajíček, & Pelikán, 2015). This approach involves an initial rigid alignment of the meshes using generalised Procrustes analysis (GPA) on landmarks. Subsequently, an automatic nonrigid registration algorithm, coherent point drift (CPD), is applied to compensate for deformations between the studied meshes and a template, referred to as the base mesh. Following registration, a closest-point search is employed to identify corresponding vertices, termed quasilandmarks. Any vertices that cannot be matched, for any reason, are excluded from further processing to prevent unwanted variability from contaminating the results. To reduce sensitivity to landmark placement errors, the surfaces undergo another round of rigid alignment through GPA, this time involving all quasilandmarks. This multi-step process ensures that the surfaces are aligned and comparable for subsequent statistical analyses.

Given the substantial dimensionality of the quasilandmark coordinate matrix, approximately 100k dimensions, a high-dimensional principal component analysis (Bishop, 2006) was conducted on these coordinates. The aim of this analysis was to reduce the dimensionality while retaining most of the form and shape variability present in the dataset. PCA is a statistical technique commonly used for dimensionality reduction, allowing for a more manageable representation of the data while preserving its essential variation.

5.2.2 Classification using support vector machine learning and cross-validation

In the final step of the analysis, support vector machines (SVM), a robust supervised machine learning approach, were employed. The goal was to train a classifier using the principal component scores and the known sex of the skulls. The SVM was trained with a radial kernel, and the training process was repeated for both form and shape scores, considering different numbers of principal components (ranging from 1 to 30).

To evaluate the performance of the classifier and identify potential overfitting, leave-one-out cross-validation was employed. This process assessed how well the classifier generalised the information by leaving out one observation at a time for testing while training on the remaining data. For both form and shape, the analysis involved selecting the lowest number of principal components that produced the highest cross-validation success rate.

All these statistical and machine learning procedures were carried out using Morphome3cs. Additionally, the R programming language, specifically the package e1071 (Dimitriadou et al., 2010), was utilised for statistical computing throughout the analysis.

5.2.3 Performance of classification models - confusion matrices

Confusion matrices serve as a valuable tool in machine learning and statistical analysis for assessing the performance of classification models. They are especially beneficial in binary classification scenarios, where the outcome can be categorised into two distinct classes, such as true or false, positive, or negative, and so on. By comparing the predicted classifications with the actual outcomes, confusion matrices provide a comprehensive overview of the model's performance, including measures like accuracy or precision (Susmaga, 2004). This detailed evaluation aids in understanding the strengths and weaknesses of the classification model and guides potential improvements.

TP (True Positive): correctly classified males

FN (False Negative): males incorrectly classified as females

FP (False Positive): females incorrectly classified as males

TN (True Negative): correctly classified females

To calculate the bias, we first need to understand what it represents in the context of a confusion matrix. In the context of classification, bias refers to the tendency of a model to systematically predict one class over the other.

Bias can be calculated using the formula:

$$\text{Bias} = \frac{FP}{FP+TN}$$

The ROC (Receiver Operating Characteristic) area, also known as the Area Under the ROC Curve (AUC), is a measure used to evaluate the performance of a binary classification model. The ROC curve is a graphical plot that illustrates the ability of classifier and consists of two axes – true positive rate and false positive rate. Both axes are calculated using results from confusion matrices. AUC is a value that summarizes the performance of a classifier, the higher the AUC the better the model is at distinguishing between the positive and negative classes (Bradley, 1997).

5.2.4 Posterior probability

The posterior probability is then calculated by applying Bayesian statistics, which combine the prior probability (prevalence of males and females in the population) and the likelihood of the observed measurements given each sex. It provides an estimate of the probability that the skull belongs to a male or female based on the available evidence.

Based on the findings of Avent et al. (2021), the recommendation is made that posterior probabilities (PPs) below 0.75 should not be used to assign a sex estimation in forensic casework (although this can differ depending on which part of the skeleton is being analysed). Instead, these cases should be labelled as "indeterminate." This suggestion stems from the observation that accuracy rates within the 0.50–0.64 and 0.65–0.74 PP intervals did not significantly differ from random chance. Furthermore, the study indicates that cases with PPs equal to or greater than 0.85 resulted in significantly higher accuracy rates compared to the 0.75–0.84 PP intervals (Avent et al., 2022). Therefore, practitioners are advised to place more confidence in sex estimates when the posterior probabilities are 0.85 or greater. This recommendation aims to enhance the reliability of sex estimations in forensic casework by establishing a threshold for acceptable posterior probabilities (Pilmann Kotěrová et al., 2024).

6. RESULTS

6.1 Introducing new method for sex estimation using exocranial surface

The pivotal study (Musilová, Dupej, Velemínská, Chaumoitre, & Brůžek, 2016) aimed to develop a robust classifier for sex determination using virtual anthropology. The entire cranial surface of 103 individuals from a recent southern French population was analysed using coherent point drift-dense correspondence analysis. Classification was performed using a support vector machine with a radial kernel to minimise subjective error. Leave-one-out cross-validation was applied for validation. Detailed methodology is described in chapter 5.2 Methods.

The process of sex classification by form was performed using representations of specimens ranging from 1 to 20 first principal components. We determined the optimal representation based on the success rate of cross-validation. The resulting accuracies, including both posterior and cross-validation rates, are depicted in Figure 12. According to our criteria, the most effective classification by form was achieved using 14 first principal components, resulting in an accuracy of 90.3%. The corresponding area under the ROC curve was 0.971.

The optimal shape classifier was chosen and assessed in a similar manner. Accuracy based on the count of principal components is illustrated in Figure 12. The most effective classification following cross-validation was attained with 16 first principal components, resulting in an accuracy of 87.4%. As anticipated, this accuracy is slightly lower than that achieved with form analysis, indicating that size might account for some of the sexual dimorphism observed in the skull. The corresponding area under the ROC curve was 0.962.

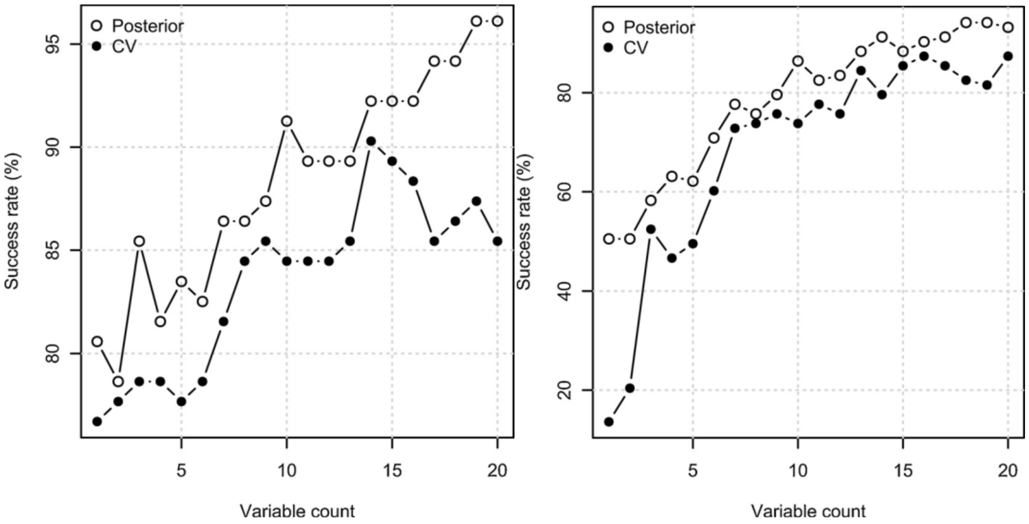


Figure 12. The graph illustrates posterior and cross-validation accuracies plotted against the number of principal components utilised for both form (left) and shape (right). The optimal representation, determined by cross-validation accuracy, yielded an accuracy of 92.2% (90.3% with cross-validation) for form and 90.3% (87.4% with cross-validation) for shape (Musilová et al., 2016).

Confusion matrices was used to provide insights into the performance of the SVM classifier in accurately predicting the sex of individuals based on their cranial form and shape. The classification based on form showed slightly higher success rates for males compared to females (Table 8), indicating a minor sex bias of -3.73%. Conversely, the classification using shape variables yielded more accurate results for females, resulting in a sex bias of 7.95%.

Table 8. Confusion matrices for the optimal Support Vector Machine (SVM) classifier based on form and shape. M – male, F – female (Musilová et al., 2016).

		Predicted by form		Predicted by shape	
		M	F	M	F
Actual	M	47	5	49	3
	F	3	48	7	44

Using colour-coded maps (Figure 13), the analysis reveals that certain features, such as supraorbital ridges, zygomatic arches, and nasal aperture, are more prominent in males, while females lack prominent traits in terms of size. Additionally, the entire skull surface shows significant differences between sexes. Shape analysis, after size normalisation, highlights positive projections in males for certain features like supraorbital ridges, while females exhibit more convex foreheads and less prominent nasal regions. Overall, females tend to have more globular skulls, while males show more elongated skulls with prominent features. The study emphasises that size normalisation impacts sexual dimorphism, with fewer significant differences observed in cranial shape compared to cranial form.

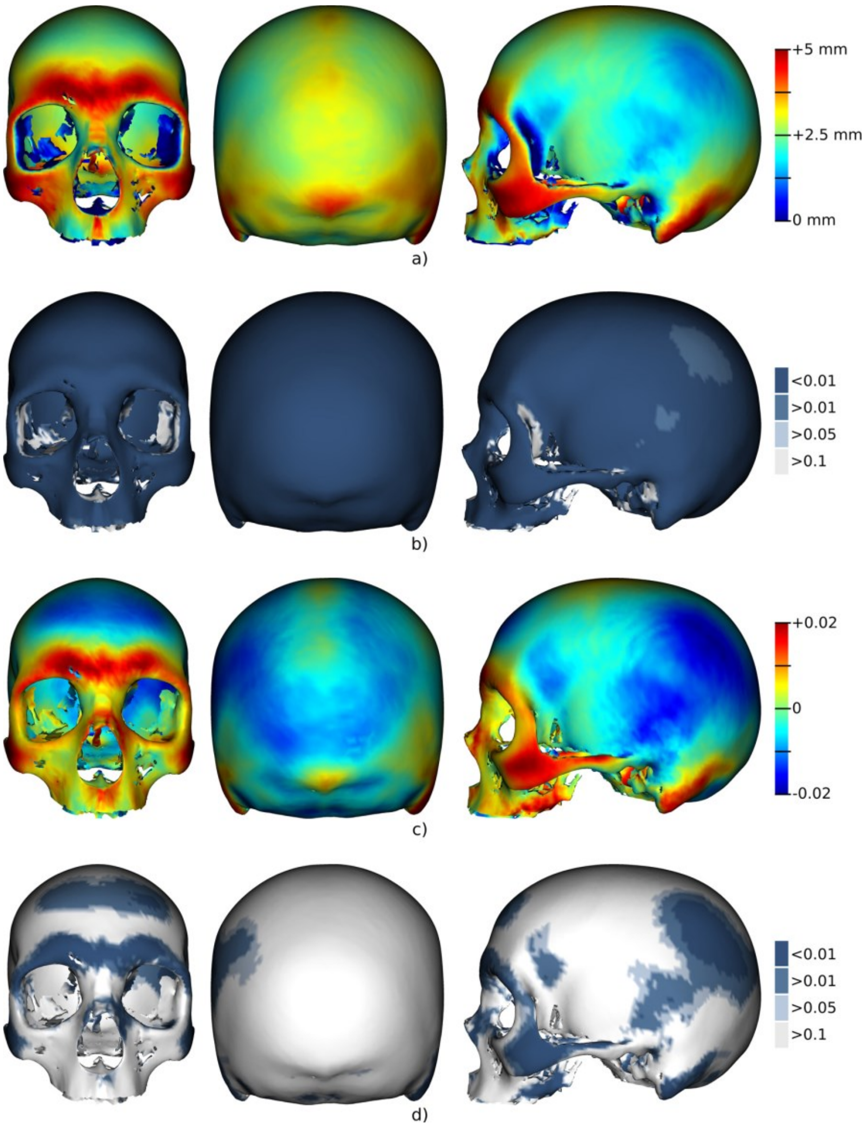


Figure 13. The visual representation of sexual differences between males and females in form and shape. Form: (a) Size differences are depicted, with red indicating areas larger in males compared to females, measured in millimetres. (b) The scale evaluates p-values, with darker colours indicating lower p-values and higher significance. Shape: (c) Relative differences are illustrated, with red representing areas more prominent in males compared to females. (d) Similar to the form visualisation, the scale evaluates p-values, with darker colours indicating lower p-values and higher significance (Musilová et al., 2016).

6.2 Analysing the reliability of method by broadening the dataset.

An extended analysis was done to examine the reliability of method presented in previous chapter. The approach was tested on a sample of 208 individuals from two recent European populations (Musilová, Dupej, Brůžek, Bejdová, & Velemínská, 2019). To enhance the method's reliability, the dataset was expanded to include a Czech population sample, resulting in the highest accuracy of 96.2% in form (Table 9). The achieved accuracy was 6% higher than the accuracy previously obtained for French skulls (90.3 %). The most successful classification by shape for the Czech dataset achieved only 78.9 % (8% lower than the accuracy obtained for French skulls). When utilising the combined (pooled) dataset, the overall reliability of the method reached 91.8% in form and 83.2% in shape.

Using the same classifier, we extended its application to test the estimation of population affinity in the pooled sample, employing support vector machine analysis with cross-validation. In the analysis of form, 90.9% of individuals were accurately classified to their respective population groups. Moreover, an even higher accuracy was achieved after eliminating size from consideration. In the shape analysis, 92.8% of individuals from both the Czech and French populations were correctly classified (Table 9).

Table 9. The performance of a classifier based on form and shape for sex and population estimation. The accuracies are presented per class, and for classifications on the pooled sample (Sex pooled and Population), additional rows provide subsample accuracies. FR – French sample, CZ – Czech sample, F – females, M – males (Musilová et al., 2019).

	Sex						Population	
	FR		CZ		Pooled			
FORM		90.29 %		96.15 %		91.83 %		90.87 %
	F	86.27 %	F	95.83 %	F	89.90 %	FR	88.46 %
	M	94.23 %	M	96.43 %	M	93.58 %	CZ	93.27 %
					FR	92.31 %	F	88.89 %
					CZ	91.35 %	M	92.66 %
SHAPE		87.38 %		78.85 %		83.17 %		92.79 %
	F	86.27 %	F	77.08 %	F	83.84 %	FR	89.42 %
	M	88.46 %	M	80.36 %	M	82.57 %	CZ	96.15 %
					FR	83.65 %	F	91.92 %
					CZ	82.69 %	M	93.58 %

In both populations, the expression of sexual dimorphism appeared strikingly similar; however, notable distinctions in cranial shape were observed, highlighting the high population specificity of cranial morphology (Figure 14). The exocranial form (Fig. 14a) and shape (Fig. 14b) were studied using principal component analysis. The variation in sexual dimorphism was primarily captured along the first principal component axis (PC1), which predominantly represented size combined with typical dimorphic traits. The second principal component axis (PC2) represented shape distinctions between skulls from the two populations.

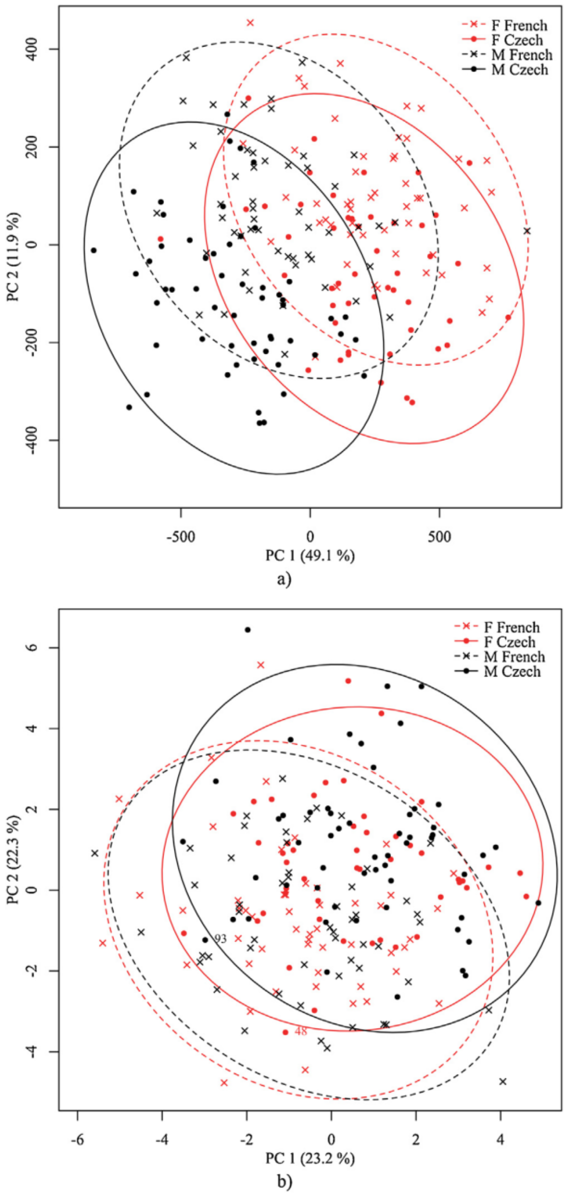


Figure 14. Relation between exocranial form (a) and shape (b), sex and population affinity. Visualisation of the first two principal components. F — females, M — males (Musilová et al, 2019).

We also focused on differences in exocranial shape between the Czech and French populations, individually examining females and males (Figure 15). Our shape analysis, which considered the p-values from per-vertex t-tests conducted separately for each subgroup, revealed that nearly identical differences were deemed significant for both sexes. Only in the frontal view did we observe slightly less pronounced discrepancies in females, particularly in the forehead and nasal areas, compared to males.

Czech males exhibited notable distinctions from French males (Fig. 15a), showcasing more rounded skulls with widened dimensions and prominent parietal and temporal bones, including the mastoids on both sides. Their facial structure also appeared broader, attributed to the pronounced lateral protrusion of the zygomatic bones and wider posterior segments of the alveolar process. Conversely, French males tended to possess elongated and narrower skulls characterized by a more prominent forehead and occipital bone, along with the entire facial region, including the orbits and nasal aperture. Significant differences between the male groups were observed in areas such as the supraorbital ridges, forehead, posterior portions of the orbit, nasal root, subnasal region, and the posterior segments of the cranium, including the parietal and temporal bones with mastoids (Fig. 15b).

Similar patterns in skull shape were noted among females. Czech women displayed distinct cranial roundness, with their parietal and temporal bones positioned more laterally, along with increased facial width marked by prominent zygomatic arches and widened posterior segments of the alveolar process. Additionally, Czech females exhibited thicker infraorbital margins. Conversely, French women tended to have longer crania and a more protruding facial structure (Fig. 15c). Significant disparities between Czech and French females were evident in the forehead area, nasal aperture, and the lateral and posterior regions of the neurocranium (Fig. 15d).

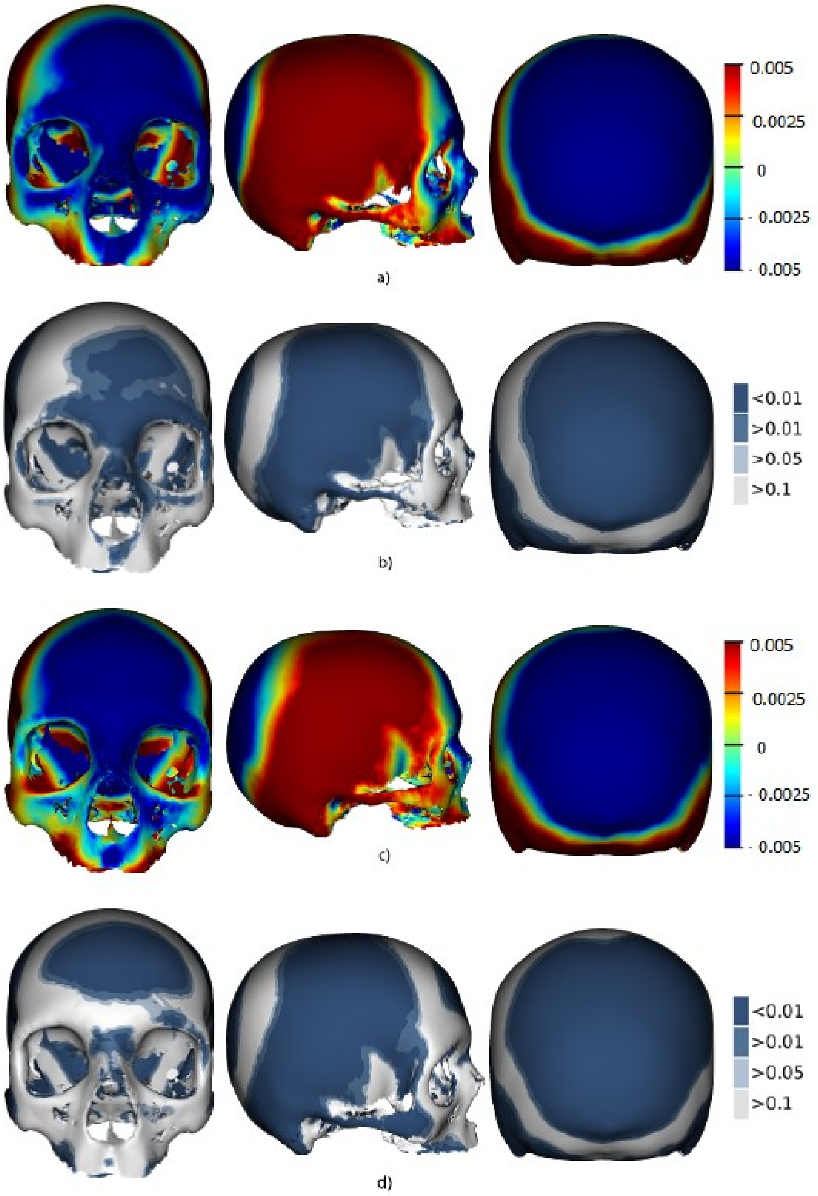


Figure 15. Visualised differences in shape between Czech and French males and females. Males: (a) The scale depicts relative differences, with red indicating the most prominent areas in Czech males compared to French males. Note that this scale does not represent actual dimensions. (b) The scale represents p-values, where darker colours indicate lower p-values and greater significance, highlighting the areas of significant difference between Czech and French males. Females: (c) Similar to males, the scale illustrates relative differences, with red indicating the most prominent areas in Czech females compared to French females. Again, this scale does not represent actual dimensions. (d) The scale evaluates p-values, with darker colours indicating lower p-values and greater significance, illustrating the areas of significant difference between Czech and French females (Musilová et al., 2019).

6.3 Observe and describe changes in sexual dimorphism with increasing age.

Senescence affects sexual dimorphism through changes in bone density, muscle mass, fat distribution, craniofacial structure, hormonal levels, and overall body composition. While some aspects of sexual dimorphism may persist into older age, the differences between males and females can become less pronounced due to the various physiological changes that occur with aging. The aging of the skull can significantly influence the morphological traits used in sex estimation, potentially reducing the sexual dimorphism that these methods depend on. Therefore, understanding these changes is vital for interpreting and applying sex estimation methods accurately, especially in older populations.

To assess the extent of sexual dimorphism concerning age, we created colour scale maps (Figure 16) using p-values from per-vertex t-tests conducted independently on two subgroups: the adult subgroup (MF0, age < 60 years) and the senile subgroup (MF1, age ≥ 60 years). The quantification of sexual dimorphism was achieved by determining the percentage of the cranial surface area that displayed statistically significant differences, with a significance level set at $\alpha = 0.05$. This approach allowed us to analyse and compare the distribution of sexual dimorphism across different age categories.

In cranial form, there was a 12% decrease in the portion of significantly different areas dropping from 94.8% to 82.6%. The most notable reduction occurred in the frontal and occipital regions. Changes in shape were particularly evident in the supraorbital, frontal, and nuchal regions, where sexual dimorphism diminished with age. Conversely, differences in zygomatic arches became significant in the senile group. Overall, in the senile group, the significantly different area of cranial shape was lower by 4.1%, decreasing from 11.6% to 7.1%.

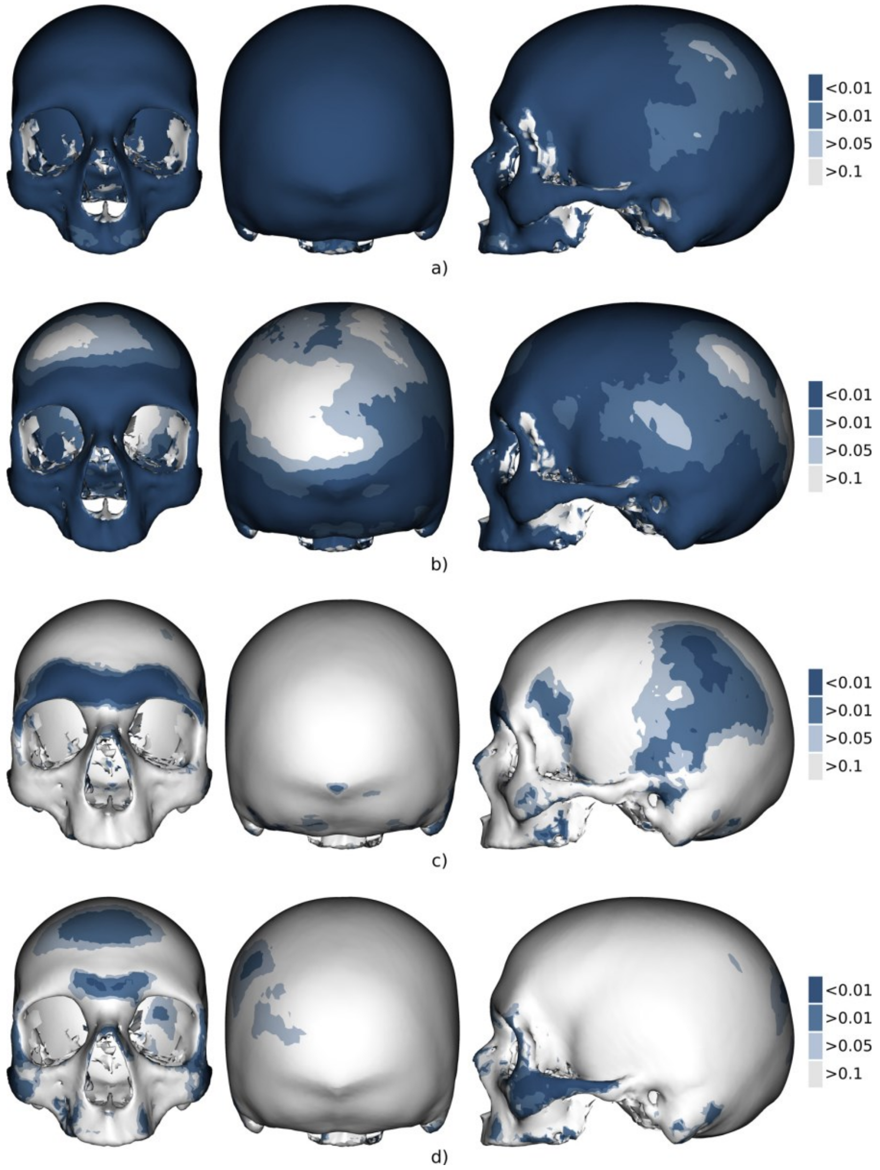


Figure 16. Visualised sexual differences between adults and elders in form and shape. The scale maps p-values to colours; darker shades of blue represent statistically significant local sexual dimorphism at a lower level. (a) Form in adults, (b) form in senile group, (c) shape in adults, (d) shape in senile group (Musilová et al., 2016).

In the co-author study by Velemínská et al (2021) the age-related differences in sexual dimorphism were studied on European crania. The research dataset included 245 CT scans of individuals from contemporary Czech (83 males and 59 females) and French (52 males and 51 females) populations. Utilising geometric morphometrics, virtual scans covering ages from 18 to 92 years were analysed. In terms of cranial form, males consistently exhibited significantly greater measurements across all age categories. However, when it comes to cranial shape (Figure 17), sexual dimorphism in regions such as the frontal, occipital, and zygomatic areas tended to decrease in the elderly. The most prominent changes associated with aging included the widening of the neurocranium and the retrusion of the face, particularly noticeable after the age of 60 in both males and females. Notably, cranial senescence contributed to a slight decrease in the accuracy of sex classification, ranging from 2% to 3%.

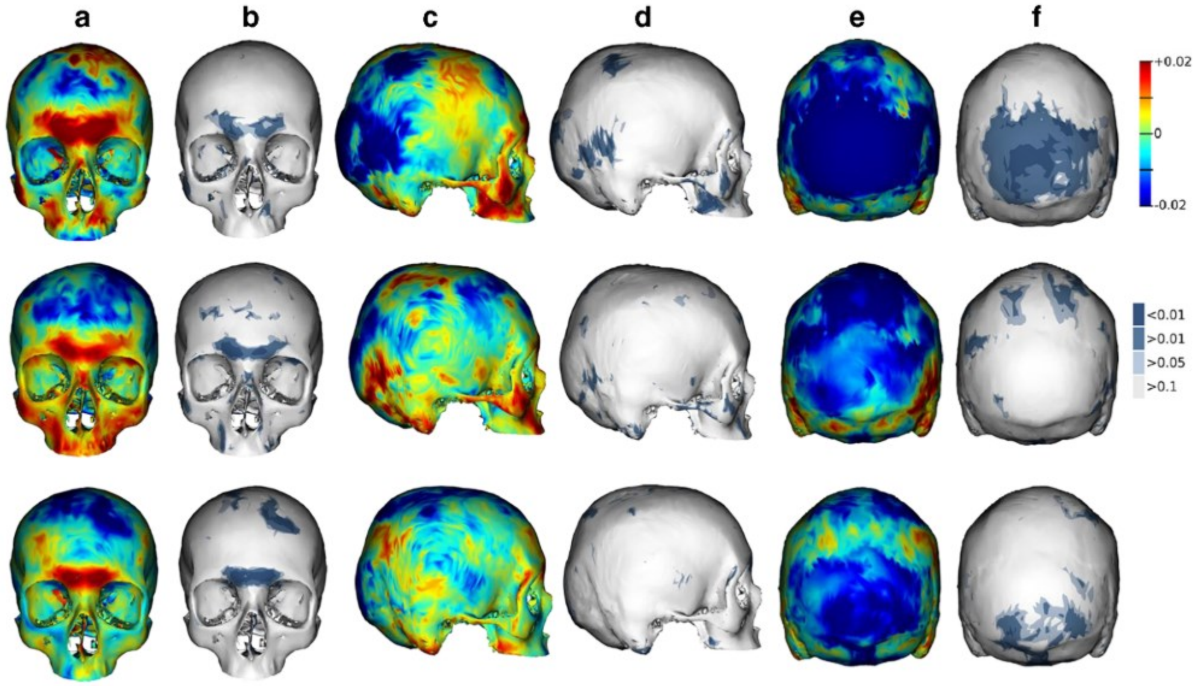


Figure 17. Cranial shape. Sexual dimorphism patterns in the Czech sample across different age groups: young adults (upper row), middle-aged adults (middle row), and elderly adults (lower row) in various skull positions. Columns A, C, and E show coloured maps depicting the superposition between average male and female shapes of the skull. The scale represents relative differences, with red indicating the most prominent areas in males compared to females. Columns B, D, and F show maps of significance, where the scale evaluates p-values, with darker colours indicating lower p-values and higher significance (Velemínská et al., 2022).

The relationship between skull aging and facial aging is an important area of study in anthropology and forensic science. Skull aging and facial aging are closely connected, as changes in the underlying bone structure of the skull directly influence the appearance of the face. The aging of the skull directly impacts facial aging through changes in bone structure, thickness, craniofacial proportions, and dental health. These skeletal changes provide the foundation for the visible signs of aging in the face, influencing how individuals appear as they grow older.

In following co-author study, the facial aging and sexual dimorphism was analysed using form and shape analysis of three-dimensional facial surface models of 456 individuals aged 14–83 years (Figure 18). As individuals age, both sexes exhibit common traits such as increased facial roundness, reduced facial convexity, sagging soft tissue, smaller visible areas of the eyes, larger noses, and thinner lips. However, male faces tend to increase in size until around 30 years of age. After the age of 70, while female facial size remains stagnant, male facial size may decrease slightly. Additionally, sexual dimorphic traits tend to diminish in the frontal and orbitonasal areas but increase in the gonial area with age.

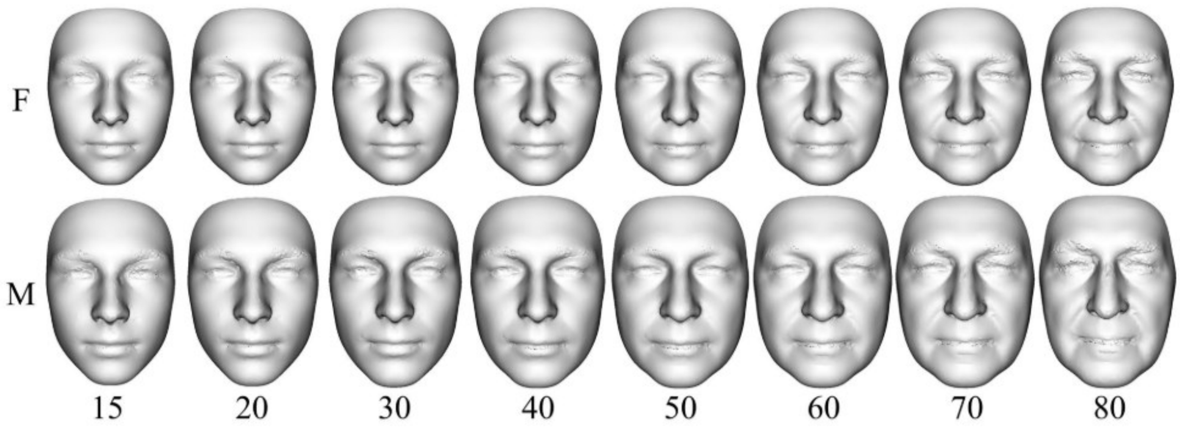


Figure 18. Predicted synthesized facial models for ages 15, 20, 30, 40, 50, 60, 70, and 80 were generated using 200 principal component contributions to the mean form for both females (F) and males (M) (Velemínská et al., 2022)

6.4 Verifying the method applicability in multi-population sample

The final enhancement over a previously developed virtual method was to broaden the dataset. As previously stated in Material, section five subdatasets were studied in final analysis. Each dataset represents a specific population from a particular geographic region (CZE - Czech population, SVK - Slovak population, DEN - Denmark population, FRA - French population, EGY - Egypt population).

Disclosure: It is important to note that the dataset of the Danish population was provided solely for testing purposes and must not be used as input data for the classifier's training mechanism.

The classifier performed very well on Czech dataset, with an accuracy of 96.15%, the highest among the individual populations. The accuracy for Slovaks is also high at 92.25%. The classifier accuracy for the French population is 90.29%, indicating good performance. When trained on CZE+SVK+FRA dataset, the classifier achieved an accuracy of 82.18% on Egyptian dataset. And on Danish dataset the classifier achieved an accuracy of 80.45%, which is the lowest among the individual populations (Table 10).

The prediction using PCA followed by SVM has a cross-validation accuracy of 91.74% on the combined data from Czech, Slovak, and French populations. If we use this model to classify Egyptian skulls, the accuracy is 82.18% and to classify skulls from Danish dataset, the accuracy is 80.45%. The model based on data from Czech, Slovak, French, and Egyptian populations, which also uses information about ancestry during prediction, is not better than one that does not directly know about ancestry. The best-performing model achieved a cross-validation accuracy of 89.8% with 32 variables (4 of which are ancestry, 28 PCA components). Without ancestry, we achieved a maximum of 89.5% (28 variables). Other parameters of this model are the same, so it is the same model with or without known ancestry. This is surprising, but it indicates that SVM is aware of ancestry even when it is not explicitly provided, as 28 variables provide enough information.

*Table 10. Sex classification accuracy. * Classifier trained on CZE+SVK+FRA dataset. ** Model considers the origin of the dataset. *** Model does not consider the origin of the dataset.*

Sample	SVM + CV
EGY	82,18 % *
CZE	96,15 %
SVK	92,25 %
FRA	90,29 %
DEN	80,45 % *
CZE+SVK+FRA	93,74 %
CZE+SVK+EGY	89,80 % **
CZE+SVK+EGY	89,50 % ***

According to posterior probability (Table 11) the classifier is more successful in correctly classifying Egyptian samples (56.4%) compared to Danish samples (43.6%). The misclassification rates are relatively low for both populations, with Egyptians at 3.0% and Danes at 2.8%, indicating that when the model does make an error, it is not dramatically high. The classifier shows moderate performance with a higher correct classification rate for Egyptians than Danes. However, a substantial number of samples from both populations could not be classified, indicating potential limitations in the model's ability to handle these datasets. The low misclassification rates are promising, suggesting that the model is relatively reliable when it does make a classification decision.

Table 11. Posterior probability results with high threshold (pp 0.90) (Avent et al.,2021).

** Classifier trained on CZE+SVK+FRA dataset.*

	EGY	DEN
Correctly classified	56,4 %	43,6 %
Not classified	40,6 %	53,6 %
Incorrectly classified	3,0 %	2,8 %

Density plots show the probability distributions of a variable for two populations: Egyptians (EGY) and Danes (DEN). The Figure 19 depicts the estimated probability that an individual is male. X-Axis represents the probability of being classified as male. Values range from 0 to 1, where 0 indicates a very low probability of being classified as male and 1 indicates a very high probability. Y-Axis represents the density of the data points at different probability values. Higher peaks indicate more frequent occurrences of certain probabilities within the sample. Red

Density Curve shows the distribution of probabilities for females; Blue Density Curve shows the distribution of probabilities for males. The small ticks at the bottom (rug plots) indicate individual data points. Each tick represents a specimen's probability of being classified as male.

In Egyptian dataset the density of the red curve is highest at the lower end of the probability scale (around 0.0), indicating that most females have a low probability of being classified as male. The density of the blue curve is higher across a range of probabilities but peaks around 0.1 and 0.5, indicating some males are being classified with moderate probability, but some have higher probabilities. There is significant overlap between the red and blue curves, suggesting some ambiguity in classification. The model has difficulty distinguishing between males and females in the Egyptian population.

In Danish dataset the red curve peaks at the lower end (around 0.1), showing that most females are classified with a low probability of being male, which is similar to the Egyptian plot. The blue curve peaks around 0.0 to 0.1, with a small secondary peak at higher probabilities, indicating a slightly more distinct separation than in the Egyptian population. There is also overlap here, but less pronounced than in the Egyptian plot. This suggests the classifier performs slightly better for the Danish population but still has some misclassifications. These density plots suggest that while the classifier can generally distinguish between sexes, there are significant instances of overlap where the classifier is less certain, especially in the Egyptian population.

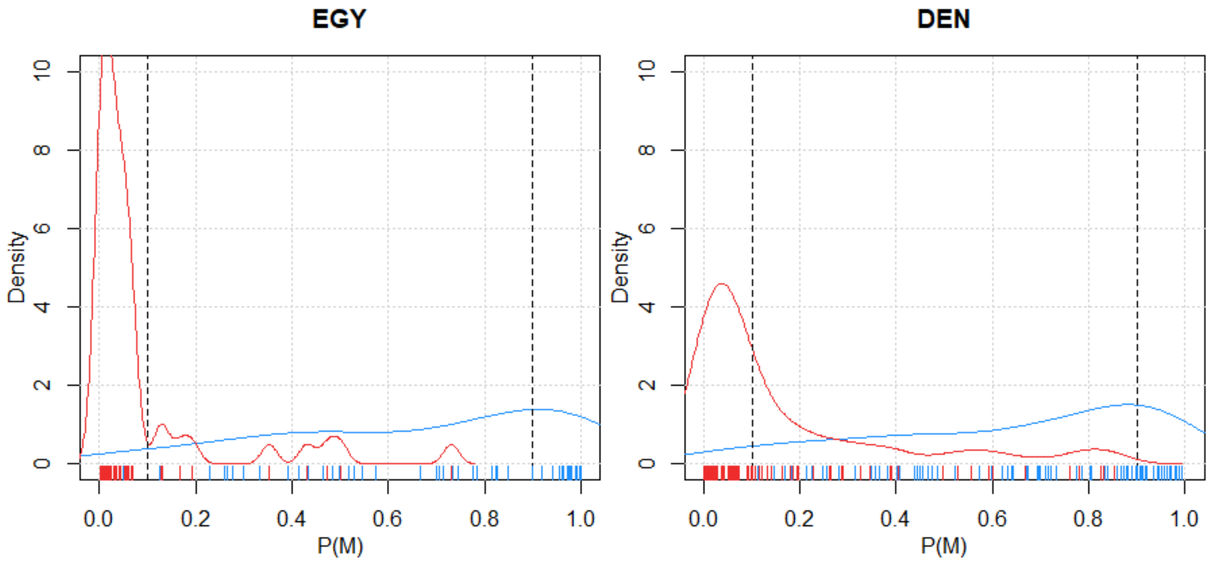


*Figure 19. Density based on the estimated probability that an individual is male in Egypt and Denmark sample. * Classifier trained on CZE+SVK+FRA.*

Centroid size provides a way to standardise shape data by removing the effect of size, allowing us to focus on shape when comparing differences between populations. Centroid size is calculated as the square root of the sum of squared distances between each landmark point and the centroid of the configuration. The centroid is essentially the geometric center of the shape. Therefore, centroid size is a measure of the extent of the shape from its centroid. In Figure 20 the median centroid sizes for Czech, Denmark, France, and Slovakia are quite similar, mostly between 83 and 85, with Egypt being slightly lower around 82. The variability within each group is also quite similar, indicating a consistent spread of centroid sizes within these populations. Outliers are present in the Czech and Slovak populations, indicating there are a few specimens with significantly smaller centroid sizes in these groups. Egypt shows a slightly lower median centroid size compared to the other groups, which might suggest a smaller overall size for specimens from this population.

This plot shows that while there are some differences in centroid sizes among the groups, particularly with Egypt having a slightly smaller median size, the overall size distributions are fairly consistent. The presence of outliers in some groups suggests individual variability that may warrant further investigation.

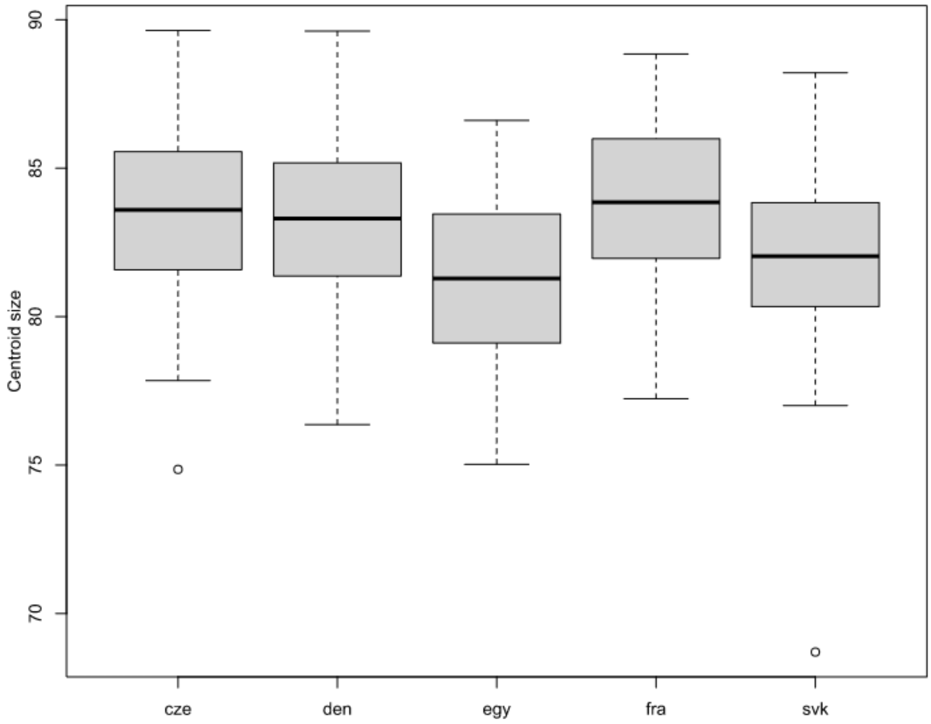


Figure 20. Centroid size representing the five datasets. The y-axis represents the centroid size, which is a measure of overall size for the specimens in each group. The x-axis lists different groups or populations labeled as "cze" (Czech), "den" (Denmark), "egy" (Egypt), "fra" (France), and "svk" (Slovakia).

The second box plot (Figure 21) compares the centroid sizes of male and female skulls across five different regions (CZE, DEN, EGY, FRA, SVK). Across all regions, males generally have larger centroid sizes than females, indicated by higher median values. Both males and females show variability within regions, but the range tends to be wider for males. The plot indicates that, on average, male skulls tend to have larger centroid sizes compared to female skulls across all five regions. Most regions have median centroid sizes for both males and females over 80, except for females in Egypt, who have a median around 78. Which indicates that skulls from Egyptian sample are generally smaller than rest of the dataset, also sexual dimorphism is less pronounced in Danish dataset especially in size. There is consistent variability within each group and some outliers, suggesting that while most of the data is clustered, there are some extreme values.

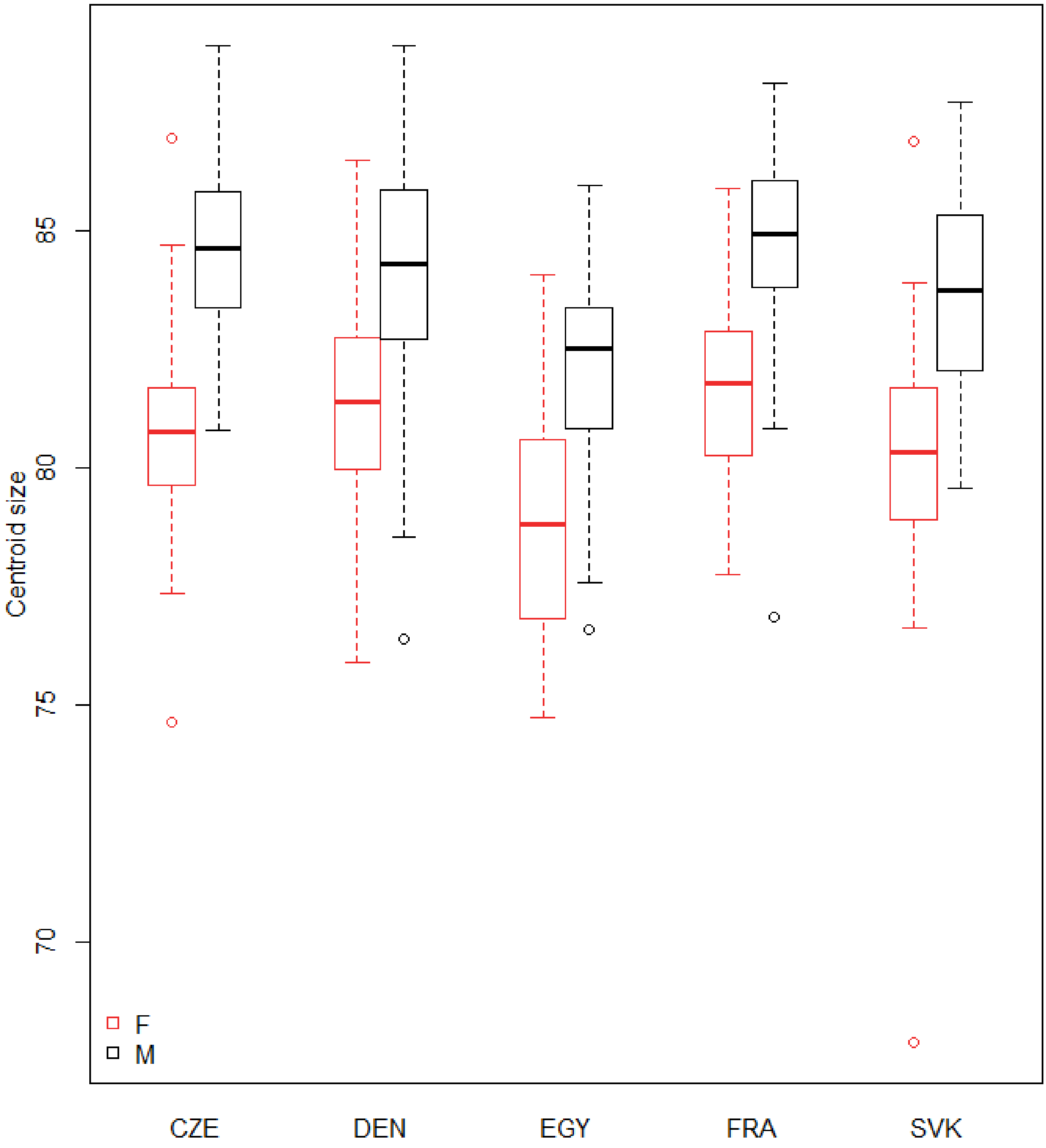


Figure 21. Box plot comparing the centroid sizes of male and female skulls across five different regions (CZE, DEN, EGY, FRA, SVK).

Results were quantified by ANOVA (Analysis of Variance) test, which is used to determine if there are significant differences between groups (Table 12). The p-value for sex and sample are extremely low, indicating that the effect of sex and population on the dependent variable are highly significant. There are statistically significant differences in the dependent variable between males and females as well as across different populations. Surprisingly, the interaction of sex and population is not statistically significant. In other words, the combined effect of sex and population on the dependent variable does not differ significantly from what would be expected if their effects were independent of each other.

Table 12. The p-values for the main effects of sex, sample, and their interaction from the ANOVA test. ('' The notation "< 2e-16" means that the p-value is less than 2 times 10 to the power of -16, or 0.0000000000000002. This is an extremely small number, indicating a highly significant result.*

Source	p-value
Sex	< 2e-16*
Sample	< 2e-16*
Sex:sample	0.1576

The Tukey HSD (Honestly Significant Difference) test results for male cranial samples from different populations are interpreted in Table 13. Significant differences in male cranial dimensions are observed mainly between Egyptian and other populations, indicating distinct cranial features.

Table 13. Table shows the p-values for each pairwise comparison of male cranial samples from different populations. Level of statistical significance: $p < 0.0010$ *'***';* $p < 0.01$ *'**';* $p < 0.05$ *'*'*

	CZE	**DEN**	**EGY**	**FRA**	**SVK**
CZE	-	0.3564188	0.0000000***	0.9998990	0.1059087
DEN	0.3564188	-	0.0000003***	0.4028712	0.8327400
EGY	0.0000000***	0.0000003***	-	0.0000000***	0.0025796**
FRA	0.9998990	0.4028712	0.0000000***	-	0.1256752
SVK	0.1059087	0.8327400	0.0025796**	0.1256752	-

Using Tukey HSD test, the significant differences in female cranial dimensions are interpreted in Table 14. Highly significant differences were mainly observed between Egyptian and other populations, indicating distinct cranial features. The other pairwise comparisons mostly show no significant differences, except for the Slovakian and Danish, Slovakian and Egyptian, and Slovakian and French samples, which display notable differences.

*Table 14. Table shows the p-values for each pairwise comparison of female cranial samples from different populations. Level of statistical significance: $p < 0.0010$ ‘***’; $p < 0.01$ ‘**’; $p < 0.05$ ‘*’*

	CZE	**DEN**	**EGY**	**FRA**	**SVK**
CZE	-	0.4377284	0.0000933***	0.2189353	0.6380567
DEN	0.4377284	-	0.0000000***	0.9618181	0.0142667*
EGY	0.0000933***	0.0000000***	-	0.0000000***	0.0178075*
FRA	0.2189353	0.9618181	0.0000000***	-	0.0066034**
SVK	0.6380567	0.0142667*	0.0178075*	0.0066034**	-

Principal component analysis (PCA) biplot (Figure 22) was used to visualise the variation in a multivariate dataset. PC1 explains 25.9% of the variance in the data and PC2 explains 8.1% of the variance in the data, while each point represents an individual specimen. There is a significant overlap between the groups, indicating that while there are some differences, the groups are not completely distinct in terms of the principal components analysed. Egypt (green) shows the most distinct clustering, particularly towards the positive end of PC1 and the positive end of PC2. Czech (blue) and Slovakia (black) have considerable overlap, suggesting these populations have similar characteristics in the dimensions captured by PC1 and PC2. Denmark (red) and France (yellow) are somewhat intermixed but show a broader spread along PC1 and PC2. Egypt (green) has a more spread-out ellipse along PC2, indicating greater variability in this principal component compared to other groups. Slovakia (black) and Czech (blue) have elongated ellipses, suggesting some variability along both PC1 and PC2 but not as spread out as Egypt. A few outliers can be noted, such as some Egyptian specimens (green) that are far from the main cluster, indicating unique characteristics in those samples.

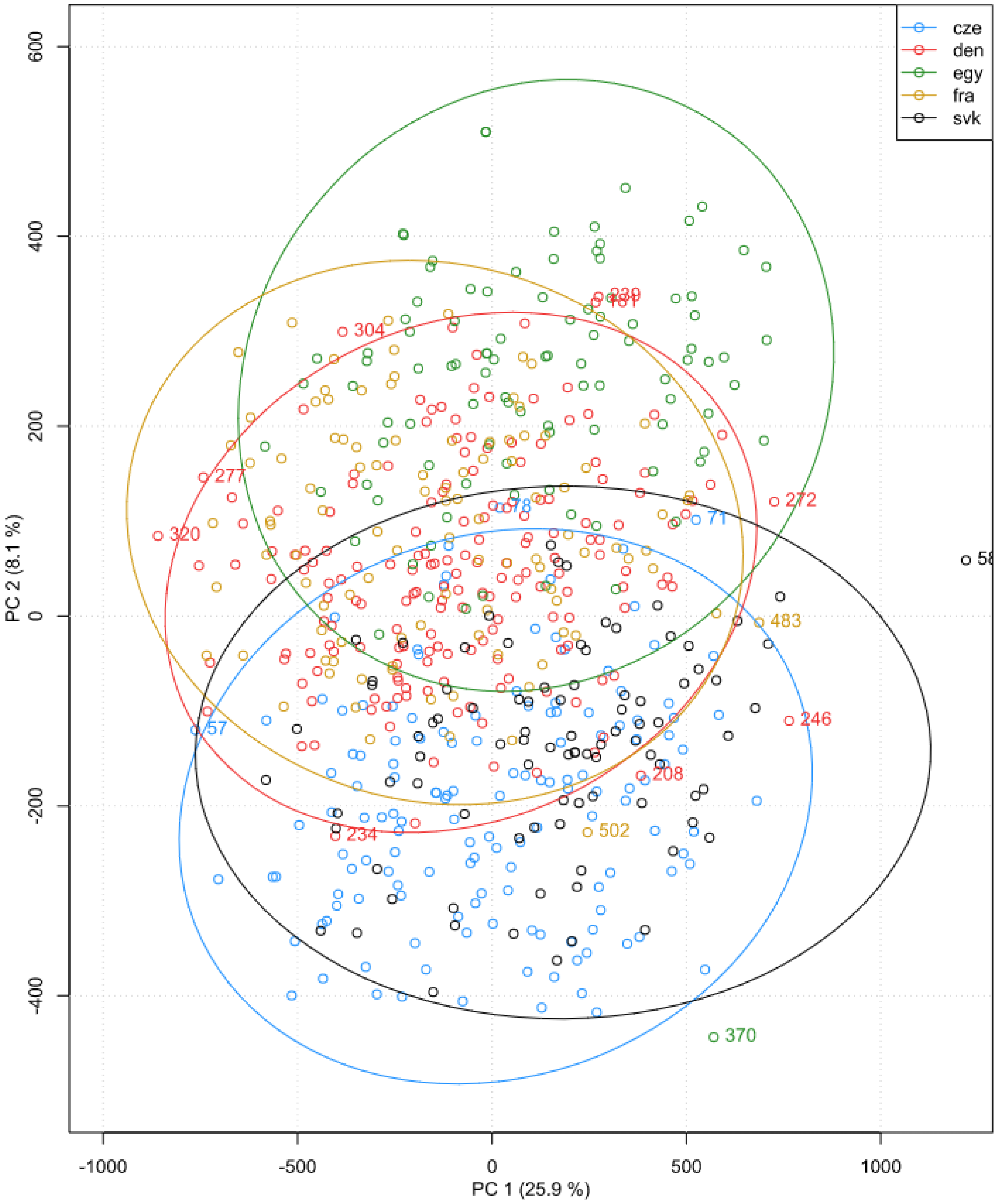


Figure 22. PCA scatter plot depicting all studied datasets. Distribution of the sample in the first two principal components in form. The colours and shapes of the points correspond to different groups (populations) as indicated by the legend: Blue circles (cze): Czech; Red circles (den): Denmark; Green circles (egy): Egypt; Yellow circles (fra): France; Black circles (svk): Slovakia.

Figure 23 consists of comparative visualisations of skulls from different regions, showcasing both form and shape differences. The regions represented are Czech Republic (CZE), Denmark (DEN), Egypt (EGY), France (FRA) and Slovakia (SVK). Each row of skulls represents a different aspect of comparison: Form (top row) and Shape (bottom row). CZE and EGY sample show more intense deviations in both form and shape compared to DEN, FRA and SVK samples. DEN, FRA and SVK samples have more moderate deviations, with a focus on specific regions of the skull. The frontal bone (especially glabellar region) consistently shows significant (dark red) areas across all regions in terms of form. Shape differences are more varied, with notable positive projections in the nasal and zygomatic bones, and negative projections in the frontal and temporal bones

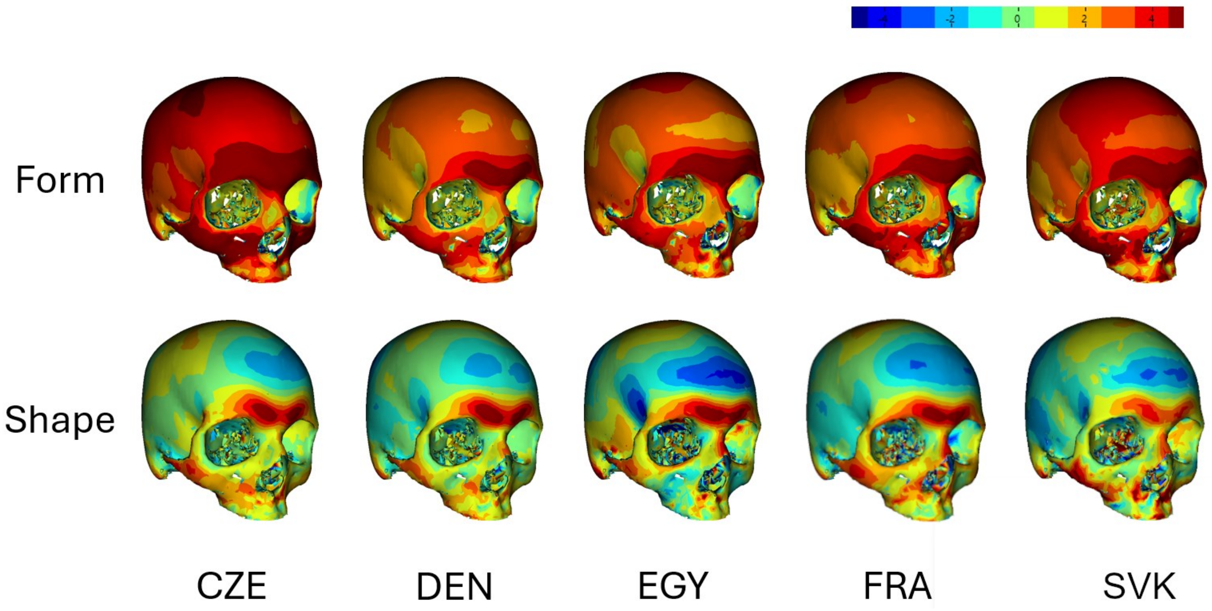


Figure 23. Comparative visualisations of skulls from different regions, showcasing both form and shape differences. Czech Republic (CZE), Denmark (DEN), Egypt (EGY), France (FRA) and Slovakia (SVK).

These results suggest that variations in classification accuracy are likely driven by differences in skull size rather than shape. Hence it might be beneficial to focus future research on shape analysis.

7. DISCUSSION

The primary goal of this study was to develop a swift and dependable method for estimating sex from unidentified crania. Efficiency was a key focus, achieved through the implementation of software tools that leverage advanced imaging technologies and geometric morphometric analyses. By utilising 3D imaging, the study aimed to create detailed digital models of cranial features, which were then analysed using automated processes and machine learning algorithms. These virtual methodologies offer several advantages over traditional approaches. First, they minimise subjectivity by employing standardised, automated processes for data collection and analysis, leading to more reliable and reproducible results (Mitteroecker & Gunz, 2009). Second, they enhance the accuracy of sex estimation by precisely quantifying cranial shape and form through mathematical models (Bertsatos et al., 2020; M. Zhang, 2024). Third, the use of large datasets from diverse populations allows the development of robust models that generalise well across different demographic groups (Garvin & Klales, 2020). In our case, reliability was pursued and largely attained through two main strategies. The incorporation of landmark optimisation algorithms minimises observer error, ensuring high reproducibility of extracted morphometric features and the subsequent sex estimation. Secondly, employing a comprehensive scheme of statistical analysis ensures that reported accuracy is not inflated by sample limitations. Additionally, this approach provides a posterior probability estimate, further enhancing the reliability of sex estimation for unidentified cases.

This goal is encompassed within the first two hypotheses. The first hypothesis proposed that the new methodology would achieve an accuracy rate comparable to or higher than established methods using both exocranial and endocranial surfaces, specifically surpassing 90% accuracy in a sample population. The second hypothesis posited that the sex estimation method would demonstrate high reliability and accuracy when tested on a larger and more diverse dataset, maintaining an accuracy rate above 90% across different contemporary European populations. The results confirmed both hypotheses, demonstrating the method's efficacy with a 91% accuracy rate in the French population and high reliability in the Czech (96%) and Slovak (92%) samples. These findings will be further discussed in first part of this chapter.

The third hypothesis and its results provided valuable insights into the variability within our dataset and opened new horizons for further research. By observing significant changes in craniofacial sexual dimorphism with increasing age, we identified a 12% reduction in dimorphic areas from younger to older individuals. This understanding of age-related morphological changes enhances our ability to interpret the data more accurately and will be encompassed in second part of Discussion.

The final phase of the study revealed important limitations in the method's applicability across multiple populations. While the sex estimation method was highly reliable for closely related European populations, it showed lower accuracy rates for the Egyptian (82%) and Danish (80%) populations. These mixed results suggest that the classifier's effectiveness diminishes when applied to more diverse and geographically distant populations.

a. The new method and how does it stand amongst others

Sex estimation from skeletal remains is a crucial component of establishing the biological profile within forensic contexts (Krishan et al., 2016). An inaccurate sex assessment can hinder the identification process, potentially preventing an individual from ever being identified (Spradley & Stull, 2018). This estimation relies on sexual dimorphism, which refers to the differences in size and shape between the skeletal elements of females and males. Sex estimation has been the subject of many studies, each presenting different approaches. These methods can be morphoscopic (Buikstra & Ubelaker, 1994; Ferembach et al., 1979; Pilmann Kotěrová et al., 2024; Walrath et al., 2004), morphometric (Constantinou & Nikita, 2022; Ogawa et al., 2013; Ramsthaler, Kreutz, & Verhoff, 2007; Walker, 2008), or 3D virtual methods (Abdel Fatah et al., 2014; Bertsatos et al., 2020; Del Bove & Veneziano, 2022; Franklin, Cardini, Flavel, Kuliukas, et al., 2013; Kimmerle et al., 2008). Regardless of the input information the method utilises, achieving maximum accuracy is always required.

Initial inspiration for this work was the outcome of the grant project by Jantz and colleagues (2013), subsequently published in the Journal of Forensic Sciences (Abdel Fatah et al., 2014), where they introduced a method achieving a success rate of 97.3%. This success was achieved through a 3D analysis of both endocranial and exocranial skull dimensions. After thorough review of the subject matter, we decided to work solely with the surface of the exocranium - the external structure of the skull. Such a method brings significant advantages and becomes usable even for scans obtained by surface scanners, not just for volumetric data from computer tomography.

In terms of accuracy the study demonstrated an accuracy of 90.3% in pivotal study population sample. This level of correct classification rate falls within the results interval of 86% to 92% reported by other authors (Ali & Al-Nakib, 2012; Ogawa et al., 2013; P. L. Walker, 2008; Williams & Rogers, 2006).

Reasons for potential reduction in success rate and suggestions for improvement include the advanced age of studied specimens, which led to a 12% decrease in sexual dimorphism in our dataset (more details in Chapter 6.3). Sexual dimorphic traits diminish with age due to hormonal changes causing bone tissue remodelling (Schulte-Geers et al., 2011; Wells, Treleaven, & Cole, 2007). Additionally, learning systems such as support vector machines (SVM) could be enhanced by integrating data from additional populations, which should reduce the population-specific bias of the classifier (Mennatt et al., 2022). Another suggestion is to utilise shape analysis instead of form analysis, as sexual dimorphism in skull size varies among populations. Shape analysis might be a more suitable approach because it eliminates the influence of size, which is more influenced by population-specific factors (Bigoni et al., 2010; Gapert et al., 2009; Kimmerle et al., 2008).

The classifier showed a sex bias of -3.73% in form. Conversely, shape variables yielded a more precise classification for females, leading to a sex bias of 7.95%. The absolute sex bias should not surpass 5% (Franklin et al. 2013), a criterion met by our form classifier. Presented method eliminates subjectivity, as the classification process is fully automatic. In our study, setting landmarks does not necessarily require precision, as these landmarks were utilised solely for rigid registration purposes. This reduces the potential for human error and ensures consistency in the classification process, enhancing the reliability and reproducibility of our results.

Achieving high accuracy is not sufficient for the general acceptance of a method. A method becomes reliable through further testing on independent populations. Therefore, when discussing forensic practice, the reliability of the method should exceed 95% (Brůžek & Murail, 2006). Hence, the reliability was tested on another dataset. For the Czech population sample, this study demonstrates a higher accuracy of 96.2%. Similarly, the combined sample of French and Czech individuals shows an accuracy of 91.8%, which remains notably high. Differences between populations were more pronounced after eliminating size, with a specificity rate of 92.8%. The shape differences between the French and Czech populations were much more marked compared to those between sexes. In other words, while the manifestation of sexual dimorphism was very similar between the two populations, the cranial shape showed high population specificity. Consequently, the approach based on pooled populations proves to be more robust than testing single populations separately. Our findings are consistent with the results of study by Milella et al. (2021). It explored cranial sexual dimorphism among modern humans, focusing on intra-sex variability. The study analysed a homogeneous sample of Italian adult crania from the 19th and 20th centuries using 3D landmarks. The findings revealed that males exhibited more variance in cranial size and shape, though this was statistically significant only for total cranial size. Shape differences were more prominent between populations than between sexes, highlighting the importance of population-specific models for sex estimation in forensic contexts (Milella et al., 2021).

The final results of the study present varying degrees of accuracy in sex estimation across different populations, demonstrating the effectiveness of the developed method in form. The Czech sample achieved the highest accuracy at 96.15%, suggesting that the method is particularly effective for this group. This high level of accuracy indicates that the cranial morphological features used for sex estimation are well-defined and distinct in the Czech population, making it easier to distinguish between sexes. The Slovak sample showed a high accuracy of 92.25%, indicating that the method is robust and reliable for this population as well. This result highlights the method's consistency and applicability across different Central European populations. The French sample had an accuracy of 90.29%, demonstrating the method's effectiveness in South European contexts. The sex estimation accuracy for the Egyptian sample was 82.18%. While this accuracy is lower compared to other populations in the study, it still indicates a reasonably reliable performance. The Danish sample exhibited the lowest accuracy at 80.45%. Similar to the Egyptian population, both samples were tested on CZE, SVK, FRA datasets and was not a part of classifier learning system. This exclusion from the training process may have contributed to the lower accuracy observed in these populations (Del Bove & Veneziano, 2022). In comparison to our findings, previous studies have demonstrated varying degrees of accuracy in sex estimation using different cranial parameters and methodologies. Kimmerle (2008) reported an accuracy rate of 87% - 90% after incorporating centroid size analysis (Kimmerle et al., 2008). Similarly, Green and Curnoe (2009) achieved an accuracy of 86.8% in sex estimation for Southeast Asian skulls when both shape and centroid size were included in the discriminant analysis (Green & Curnoe, 2009). Gillet (2020) demonstrated a significant improvement in the accuracy of identifying sexual dimorphism through advanced analysis methods. The study achieved a remarkable accuracy rate of 97.7% for the cranium and 84.2% for the mandible, surpassing the results obtained through traditional metric analysis (Gillet et al., 2020).

The posterior probability provides an estimate of the probability that the skull belongs to a male or female based on the available evidence (Johnson & Wichern, 2007). It's important to note that while discriminant function analysis and posterior probability calculations can be useful tools in determining sexual dimorphism in skulls, they are not infallible (Franklin, O'Higgins, Oxnard, & Dadour, 2008). There can be individual variation and overlap in measurements between males and females, which can affect the accuracy of the classification (Konigsberg et al., 2009). Additionally, these techniques rely on reference samples that may not represent the entire human population accurately. Therefore, caution and additional methods are often employed when analysing sexual dimorphism in the human skull (Steyn & İşcan, 2008). The rates for correctly classified samples (56.4% for Egyptians and 43.6% for Danes) are relatively low. The high non-classification rates (40.6% for Egyptians and 53.6% for Danes) are caused by our predetermined level of the low misclassification rates (3.0% and 2.8%) based on high posterior probability threshold (pp 0.9). The low misclassification rates suggest that the method is precise when it does make a classification, but the frequent inability to classify reduces overall effectiveness. It is important to note that these two datasets were tested against training dataset concluding from CZE, FRA and SVK samples, which may result in less precise outcomes. For comparison, study by Pilmann Kotěrová et al. (2024) utilised a higher posterior probability threshold (pp 0.95), which resulted in sex estimation for 54% of individuals for Observer A and 66% for Observer B (Pilmann Kotěrová et al., 2024).

b. Variance in sexual dimorphism and how population variability can affect classifier

We opted for colour coded maps for visualisation of sexual dimorphism between selected groups. First analysis revealed significant differences between the sexes in Czech sample across the entire surface of the skull in terms of form. Males exhibited larger, wider, and overall, more robust skulls compared to females, primarily due to differences in the masticatory system (Franklin et al., 2006). Specifically, certain regions such as the glabellar and supraorbital areas showed notable dimorphism, highlighting their importance in the midfacial region, which is heavily influenced by facial expressions (Hennessy et al., 2005; J. Velemínská et al., 2012). Males typically have wider and more robust zygomatic arches, making them one of the most reliable traits for sex determination. Additionally, regions such as the mastoids and the nuchal region with the nuchal crest are more prominent and robust in males due to muscle attachment (Bigoni et al., 2010; D. Franklin et al., 2006).

According to our additional results, the average male and female cranial models within the Czech and French groups show only slight variations. French males and females generally have skulls that are elongated from front to back, featuring prominent frontal and occipital bones. On the other hand, Czech males and females typically possess wider, more globular skulls, with notable prominence in the temporal and parietal bones. This finding is in concordance with Bergmann's thermoregulatory rule, when the shape and size of the modern human skull align well with climate models and reflect (Harvati & Weaver, 2006). According to this rule, populations adapted to warmer or hotter climates tend to have smaller, narrower cranial vaults, while those adapted to cooler climates have larger, broader skulls. This variation arises from the need for better heat retention in the head and brain (Beals et al., 1984; C. B. Ruff, 1994; C. Ruff, Trinkaus, & Holliday, 1997). The French samples were collected in Marseille, located in the south of France, the climate is Mediterranean, characterized by hot, dry summers and mild, wet winters. The average annual temperature is around 16°C[1]. In contrast, the other samples were collected in Prague, situated in the Czech Republic, experiences a temperate continental climate with cold winters and warm summers. The average annual temperature in Prague is approximately 9°C[2]. According to these findings, it is evident that exocranial shape differences are more pronounced between populations than between sexes. The average male and female cranial models within the Czech and French groups show only slight differences. This aligns with a previous 2D geometric morphometric study on cranial variation in the American population (Murphy & Garvin, 2018). Also, the colour-coded maps reveal that French males and females tend to have greater facial prognathism, characterized by more pronounced maxillary, nasal, and frontal bones. Conversely, Czech males and females exhibit a more orthognathic facial structure, with prominent zygomatic bones and a typically brachycephalic skull. These anatomical differences likely stem from distinct biomechanical stresses and subsistence strategies, resulting in more gracile and narrow facial dimensions closely linked to

---

[1] Climate & Weather Averages in Marseille, Provence-Alpes-Côte-d'Azur, France. (n.d.). Retrieved June 2, 2024, from Time and Date website: https://www.timeanddate.com/weather/france/marseille/climate

[2] Climate & Weather Averages in Prague, Czechia. (n.d.). Retrieved June 2, 2024, from Time and Date website: https://www.timeanddate.com/weather/czech-republic/prague/climate

the energy demands of masticatory muscles and structures involved in chewing (Bejdová et al., 2013; Ibrová et al., 2017; Lieberman et al., 2002).

The colour-coded map for all five datasets generally showed larger significance in form, indicating size differences in specific regions of the skull. Shape analysis showed a more differentiated range of significance, revealing the areas that are more about the geometric configuration rather than size. These results are consistent with other studies, that advocate for using shape instead of size (Adams et al., 2004; Del Bove et al., 2020). These findings were expanded of results from other analyses, such as PCA and centroid size plots resulting in better understanding of how classifier performs. The Egyptian skulls are generally smaller, and the Danish skulls appear to exhibit less pronounced sexual dimorphism in size. This variability in skull size and sexual dimorphism across different populations presents a challenge for the model in making accurate classifications, particularly with the Danish samples.

Furthermore, based on the results from classifier as well as from additional statistical analyses, it was inevitable to consider the interactions between sex and ancestry factors. While ancestry estimation plays a crucial role in the identification process, it is being criticized it as perpetuating the notion of biological races - a concept rooted in early biological anthropological research that inaccurately linked racial hierarchies to morphological differences and nonphysical traits (Quintyn, 2010). Presently, the majority of researchers, explicitly reject typological research and the idea of races, issuing statements to distance their field from such earlier work and so do we (Armelagos & Goodman, 1998; Stephen Ousley, Jantz, & Freid, 2009; Sauer, 1992; Smay & Armelagos, 2000). Despite the rejection of the concept of races, forensic anthropologists grapple with a paradox when estimating ancestry. This complex estimation becomes possible due to a non-zero correlation between social race, skeletal morphology, and geographic origin. The process of estimating ancestry is complex and more challenging than other components of the biological profile (Elliott & Collard, 2009).

In forensic anthropology, the accurate classification of skeletal remains often relies on population-specific data to account for variations in bone structure and size. When working with diverse populations, such as the Egyptian and Danish samples in this study, the model must contend with differing morphological characteristics. Egyptian skulls tend to be smaller on average, which can influence the model's ability to correctly identify sex based on size-related features. Similarly, Danish skulls show reduced sexual dimorphism in size, meaning the differences between male and female skulls are less pronounced. This reduced dimorphism makes it more challenging for the model to distinguish between sexes accurately. The Tukey HSD test results reveal that significant differences in cranial measurements are predominantly found between Egyptian samples and other populations, for both males and females. This indicates distinct cranial features in the Egyptian population compared to others. The findings suggest that while the sex estimation method is highly accurate within certain populations, there are notable variations when applied across different geographical groups. These results emphasise the importance of considering population-specific traits in forensic, and they support the ongoing development of more robust and adaptable methods for sex estimation (Garvin & Klales, 2020).

For instance, including data from both Egyptian and Danish populations during the training phase could improve the model's ability to generalise and maintain higher accuracy across diverse groups (Del Bove & Veneziano, 2022). However, this approach also necessitates careful balancing to avoid overfitting to specific population characteristics, which could reduce the model's overall robustness. Adding the Egyptian data introduces another layer of complexity. When Egyptian male skulls are included, which are only slightly larger than French female skulls, it further complicates the model's decision-making process. The overlap in size between Egyptian males and French females makes it more difficult for the model to differentiate between sexes based solely on skull size. The model appears to rely heavily on size-related features when classifying skulls, with smaller skulls being more likely to be classified as female and larger skulls as male. This scenario demonstrates the need for a multifaceted approach that integrates other morphological features or advanced statistical techniques to enhance classification accuracy (Adams et al., 2004).

In forensic anthropology, global equations are often inadequate for sex estimation aimed at individual identification. Although they can be improved by using large, diverse reference samples to ensure broad representation and reduce bias (Kenyhercz, Klales, Stull, McCormick, & Cole, 2017). While global equations may achieve acceptable accuracy rates, they do not accurately represent the sexual dimorphism of each population. Therefore, for precise individual identification, it is preferable to use population-specific equations when available. Although it is unrealistic to develop population-specific standards for every method and bone, research can prioritize the most reliable skeletal elements (Garvin & Klales, 2020). One way to address population variation in sexual dimorphism is to create global databases of trait scores and skeletal measurements, and to ensure these databases and the resulting equations are accessible to everyone (e.g. FORDISC 3.1, Osteoware, SexEST or MorphoPASSE).

Also, accurate sex estimation often hinges on the consideration of both size and shape characteristics of skeletal remains. While shape can provide valuable information, such as sexually dimorphic features in facial structure or cranial morphology, the current findings suggest that size may have a more substantial influence on classification outcomes in this particular model. To improve the accuracy and reliability of sex estimation models, it is essential to develop approaches that account for population-specific size variations while also considering shape-related features. This may involve refining existing models to incorporate additional parameters or implementing advanced statistical techniques to better capture the complexities of size distributions across diverse populations (Adams et al., 2004; Del Bove et al., 2020; Kelley & Tallman, 2022; Nikita et al., 2024).

8. PERSPECTIVES

The sex estimation method, originally developed for forensic purposes using contemporary specimens with known origins and sex, shows promise for bioarchaeological applications. This potential was explored by Meinerová et al. (2023), who tested the classifier on skulls from both contemporary Egyptians and those from the Egyptian Old Kingdom Period (2700–2180 BC). The study revealed high accuracy for the contemporary sample (89.59%) but showed lower accuracy (61.11%) for the ancient skulls (Meinerová et al., 2023). These findings suggest the classifier's robustness in contemporary contexts and highlight the need for further refinement to improve accuracy in bioarchaeological applications.

Given the lower accuracy rates observed in the Egyptian and Danish populations, future research should focus on refining the sex estimation method to improve its reliability across a broader range of populations. This could involve developing population-specific models or integrating additional morphological features that are more universally applicable.

Considering that the current findings suggest size may have a more substantial influence on classification outcomes, future research should emphasise shape analysis rather than form. By prioritising shape, the method may achieve more accurate and reliable classification outcomes. To further enhance accuracy, future studies could explore the inclusion of additional skeletal features beyond the exocranial surface. Incorporating other structures might provide complementary data that enhance the overall reliability of sex estimation.

Additionally, developing software based on our classifier could significantly enhance its accessibility and usability for practitioners in the field. Such software should be designed to facilitate easy input of data, provide clear instructions for users, and offer robust analytical capabilities, making the methodology widely available for both forensic and bioarchaeological applications.

By addressing these research directions, future studies can build on the findings of this thesis to develop even more accurate, reliable, and universally applicable methods for sex estimation from skeletal remains.

9. CONCLUSIONS

In this doctoral thesis, the new method was rigorously tested on a broad dataset that included individuals from central6, northern, and southern Europe, as well as North Africa. This extensive dataset was carefully curated to encompass a wide range of genetic backgrounds and environmental influences, providing a comprehensive evaluation of the method's applicability and reliability across diverse populations. By including samples from regions such as the Czech Republic, Slovakia, Denmark, France, and Egypt, the study aimed to assess the robustness of the sex estimation methodology in varying demographic and geographic contexts.

The results of this study confirm Hypothesis 1. The novel methodology for sex estimation, which focuses solely on the exocranial surface of skulls, achieved an accuracy rate of 91% in the French population sample. This accuracy rate not only meets but surpasses the 90% threshold set forth in the hypothesis. This demonstrates that the developed method is not only comparable to but also competitive with established methods.

The findings of this research support Hypothesis 2. The sex estimation method demonstrated high reliability and accuracy when applied to a larger and more diverse dataset. Specifically, the method achieved an accuracy rate of 96% in the Czech sample and 92% in the Slovak sample. These results confirm that the methodology maintains an accuracy rate well above the 90% threshold across different contemporary European populations.

The results of this study confirm Hypothesis 3. Significant changes in sexual dimorphism of craniofacial morphology were observed with increasing age. Specifically, in terms of cranial form, there was a 12% decrease in the proportion of significantly different areas, dropping from 94.8% in younger individuals to 82.6% in older individuals. Additionally, in the elderly group, the significantly different area of cranial shape was lower by 4.1%, decreasing from 11.6% to 7.1%. These findings demonstrate a clear reduction in sexual dimorphism with age, supporting the hypothesis and highlighting the dynamic nature of craniofacial morphology over the lifespan.

The results of this study partially support Hypothesis 4. The sex estimation method proved to be highly reliable and accurate for closely related European populations, achieving high accuracy rates for the Czech (96%), Slovak (92%), and French (91%) samples. However, the method did not perform as well for the Egyptian (82%) and Danish (80%) populations. These lower accuracy rates indicate that the classifier's reliability diminishes when applied to more diverse and geographically distant populations. Thus, while the method is robust and reliable within closely related European groups, it requires further refinement to achieve similar reliability across a broader range of populations.

This study set out to develop and validate a novel methodology for sex estimation using the exocranial surface of skulls, aiming to achieve high accuracy and reliability across diverse populations. Through rigorous testing, the method demonstrated exceptional performance in certain contexts but revealed limitations in others. Additionally, significant changes in craniofacial sexual dimorphism with age were observed, further enhancing our understanding of morphological variations. These findings underscore the potential and challenges of applying advanced forensic methodologies across varied demographic contexts, highlighting areas for future improvement and research.

SOUHRN (CZECH SUMMARY)

V této disertační práci byla testována nová metoda pro odhad pohlaví na souboru dat, který zahrnoval jedince ze střední, severní a jižní Evropy a ze severní Afriky. Tento soubor dat byl sestaven tak, aby poskytl komplexní hodnocení úspěšnosti a spolehlivosti metody napříč různými populacemi. Zařazením vzorků z regionů jako Česká republika, Slovensko, Dánsko, Francie a Egypt měla studie za cíl posoudit robustnost metodologie odhadu pohlaví v různých demografických a geografických kontextech.

Nová metodologie pro odhad pohlaví, která se zaměřuje výhradně na exokraniální povrch lebek, dosáhla míry přesnosti 91 % ve vzorku francouzské populace. Tato míra přesnosti nejen splňuje, ale i překračuje hranici 90 % stanovenou v hypotéze. To ukazuje, že vyvinutá metoda je nejen srovnatelná, ale také konkurenceschopná se zavedenými metodami. Vysoká míra dosažené úspěšnosti dokládá potenciál této nové metodologie pro odhad pohlaví. Metoda prokázala vysokou spolehlivost a úspěšnost při aplikaci na větší a rozmanitější soubor dat. Konkrétně metoda dosáhla úspěšnosti 96 % ve vzorku české populace a 92 % ve vzorku slovenské populace. Tyto výsledky potvrzují, že metodologie udržuje míru úspěšnosti nad hranicí 90 % napříč různými současnými evropskými populacemi, což odpovídá robustní metodě, potenciálně využitelné pro forenzní účely. Tyto výsledky jsou v pozitivním souladu s prvními dvěma hypotézami.

Současně byly pozorovány významné změny v pohlavním dimorfismu kraniofaciální morfologie s rostoucím věkem. Konkrétně v rámci kraniální formy došlo k 12% poklesu podílu významně odlišných oblastí, což představuje snížení z 94,8 % u mladších jedinců na 82,6 % u starších jedinců. Kromě toho ve skupině seniorů byl podíl významně odlišných oblastí kraniálního tvaru nižší o 4,1 %, klesl z 11,6 % na 7,1 %. Tyto výsledky ukazují na jasné snížení pohlavního dimorfismu s věkem a podporují třetí hypotézu.

Čtvrtá hypotéza byla podpořena pouze částečně. Metoda odhadu pohlaví se ukázala být spolehlivá a úspěšná pro blízce příbuzné evropské populace – populace z České republiky (96 %), Slovenska (92 %) a Francie (91 %). Nicméně metoda nefungovala tak dobře u egyptské (82 %) a dánské (80 %) populace. Tato nižší úspěšnost naznačuje, že spolehlivost klasifikátoru klesá při aplikaci na geograficky vzdálenější populace. Proto, i když je metoda robustní a spolehlivá v rámci blízce příbuzných evropských skupin, vyžaduje další zdokonalení, aby dosáhla podobné spolehlivosti napříč širším spektrem populací.

Tato disertační práce si kladla za cíl vyvinout a ověřit novou metodologii pro odhad pohlaví pomocí exokraniálního povrchu lebek, s cílem dosáhnout vysoké úspěšnosti a spolehlivosti napříč různými populacemi. Prostřednictvím důkladného testování metoda prokázala výjimečný výkon v určitých kontextech, ale odhalila také nedostatky v jiných. Kromě toho byly pozorovány významné změny v kraniofaciálním sexuálním dimorfismu s věkem, což dále prohlubuje naše chápání morfologických variací. Tyto výsledky podtrhují potenciál a výzvy při aplikaci pokročilých forenzních metodik napříč různými demografickými kontexty a zdůrazňují oblasti pro budoucí zlepšení a výzkum.

BIBLIOGRAPHY

Abdel Fatah, E. E., Shirley, N. R., Jantz, R. L., & Mahfouz, M. R. (2014). Improving sex estimation from crania using a novel three-dimensional quantitative method. *Journal of Forensic Sciences*, *59*(3), 590–600.

Acsádi, G. Y., & Nemeskéri, J. (1970). *History of human life span*. JSTOR.

Adams, D. C., Rohlf, F. J., & Slice, D. E. (2004). Geometric morphometrics: ten years of progress following the 'revolution.' *Italian Journal of Zoology*, *71*(1), 5–16.

Alcina, M., Rissech, C., Clavero, A., & Turbón, D. (2015). Sexual dimorphism of the clavicle in a modern Spanish sample. *European Journal of Anatomy*, *19*(1), 73–83.

Algee-Hewitt, B. F. B., & Kim, J. (2021). *Remodeling Forensic Skeletal Age: Modern Applications and New Research Directions*. Academic Press.

Ali, A. R., & Al-Nakib, L. H. (2012). The Value of lateral cephalometric image in sex identification. *Journal Bagh College Denstistry*, *25*(June), 54–58.

Alves Proença, H. H. F., Slavicek, R., Cunha, E., & Sato, S. (2014). A 3D computerized tomography study of changes in craniofacial morphology of Portuguese skulls from the eighteenth century to the present. *International Journal of Stomatology & Occlusion Medicine*, *7*(2), 33–45. https://doi.org/10.1007/s12548-014-0105-3

Armelagos, G. J., & Goodman, A. H. (1998). Race, racism, and anthropology. *Building a New Biocultural Synthesis: Political-Economic Perspectives on Human Biology*, 359–377.

Arsuaga, J. L., Lorenzo, C., & Carretero, J. M. (1995). Sexual dimorphism of the hip bone in the Coimbra population (Portugal). *Antropologia Portuguesa*, *13*, 171–191.

Avent, P. R., Hughes, C. E., & Garvin, H. M. (2022). Applying posterior probability informed thresholds to traditional cranial trait sex estimation methods. *Journal of Forensic Sciences*, *67*(2), 440–449.

Axelsson, S., Kjær, I., Heiberg, A., Bjørnland, T., & Storhaug, K. (2005). Neurocranial morphology and growth in Williams syndrome. *The European Journal of Orthodontics*, *27*(1), 32–47.

Azarfar, G., Ko, S.-B., Adams, S. J., & Babyn, P. S. (2024). Deep learning-based age estimation from chest CT scans. *International Journal of Computer Assisted Radiology and Surgery*, *19*(1), 119–127.

Babladi, P. I., Pejavar, S. R., & Reddy, M. (2014). A Study of Estimation of Stature from Footprint Length. *Indian Journal of Forensic Medicine and Toxicology*, *8*(1), 176–181.

Barber, N. (1995). The evolutionary psychology of physical attractiveness: Sexual selection and human morphology. *Ethology and Sociobiology*, *16*, 395–424. https://doi.org/10.1016/0162-3095(95)00068-2

Beals, K. L., Smith, C. L., Dodd, S. M., Angel, J. L., Armstrong, E., Blumenberg, B., … Trinkaus, E. (1984). Brain Size, Cranial Morphology, Climate, and Time Machines [and Comments and Reply]. *Current Anthropology*, *25*(3), 301–330. https://doi.org/10.1086/203138

Bejdová, Š., Krajíček, V., Velemínská, J., Horák, M., & Velemínský, P. (2013). Changes in the sexual dimorphism of the human mandible during the last 1200 years in Central Europe. *HOMO- Journal of Comparative Human Biology*, *64*(6), 437–453.

Berner, M., Sládek, V., Holt, B., Niskanen, M., & Ruff, C. B. (2018). Sexual dimorphism. *Skeletal Variation and Adaptation in Europeans: Upper Paleolithic to the Twentieth Century*, 133–161.

Bertsatos, A., Chovalopoulou, M. E., Brůžek, J., & Bejdová, Š. (2020). Advanced procedures for skull sex estimation using sexually dimorphic morphometric features. *International Journal of Legal Medicine*, *134*(5), 1927–1937. https://doi.org/10.1007/s00414-020-02334-9

Best, K. C., Garvin, H. M., & Cabo, L. L. (2018). An Investigation into the Relationship between Human Cranial and Pelvic Sexual Dimorphism. *Journal of Forensic Sciences*, *63*(4), 990–1000. https://doi.org/10.1111/1556-4029.13669

Bidmos, M. A., Steinberg, N., & Kuykendall, K. L. (2005). Patella measurements of South African whites as sex assessors. *HOMO- Journal of Comparative Human Biology*, *56*, 69–74. https://doi.org/10.1016/j.jchb.2004.10.002

Bidmos, Mubarak A, & Dayal, M. R. (2004). Further evidence to show population specificity of discriminant function equations for sex determination using the talus of South African blacks. *Journal of Forensic Sciences*, *49*(6), JFS2003431-6.

Bigoni, L., Velemínská, J., & Brůžek, J. (2010). Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a Central European sample of known sex. *HOMO - Journal of Comparative Human Biology*, *61*(1), 16–32.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Blackless, M., Charuvastra, A., Derryck, A., Fausto-Sterling, A., Lauzanne, K., & Lee, E. (2000). How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology: The Official Journal of the Human Biology Association*, *12*(2), 151–166.

Boucherie, A., Chapman, T., García-Martínez, D., Polet, C., & Vercauteren, M. (2022). Exploring sexual dimorphism of human occipital and temporal bones through geometric morphometrics in an identified Western-European sample. *American Journal of Biological Anthropology*, *178*(1), 54–68.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*(7), 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2

Braun, S., Schwendener, N., Kanz, F., Lösch, S., & Milella, M. (2024). What we see is what we touch? Sex estimation on the skull in virtual anthropology. *International Journal of*

Legal Medicine, 1–13.

Brooks, S., & Suchey, J. M. (1990). Skeletal age determination based on the os pubis: a comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, *5*, 227–238.

Brough, A. L., Rutty, G. N., Black, S., & Morgan, B. (2012). Post-mortem computed tomography and 3D imaging: anthropological applications for juvenile remains. *Forensic Science, Medicine, and Pathology*, *8*, 270–279.

Brůžek, J. (2002). A method for visual determination of sex, using the human hip bone. *American Journal of Physical Anthropology*, *117*(September 1997), 157–168.

Brůžek, J., Mikšík, I., Pilmann Kotěrová, A., Morvan, M., Drtikolová Kaupová, S., Santos, F., … Velemínský, P. (2024). Undertaking the biological sex assessment of human remains: The applicability of minimally-invasive methods for proteomic sex estimation from enamel peptides. *Journal of Cultural Heritage*, *66*, 204–214. https://doi.org/https://doi.org/10.1016/j.culher.2023.11.021

Brůžek, J., & Murail, P. (2006). Methodology and reliability of sex determination from the skeleton. In A. Schmitt, E. Cunha, & J. Pinheiro (Eds.), *Forensic Anthropology and Medicine: Complementary Sciences From Recovery to Cause of Death*.

Brzobohatá, H., Krajíèek, V., Horák, Z., & Velemínská, J. (2016). Sexual dimorphism of the human tibia through time: Insights into shape variation using a surface-based approach. *PLoS ONE*, *11*(11), 1–22. https://doi.org/10.1371/journal.pone.0166461

Buchanan, S. (2014). *Bone modification in male to female transgender surgeries: considerations for the Forensic Anthropologist*. Louisiana State University and Agricultural & Mechanical College.

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

Buikstra, J., & Ubelaker, D. (1994). Standards for data collection from human skeletal remains. In *Fayetteville Arkansas Archaeological Survey*.

Buonasera, T., Eerkens, J., de Flamingh, A., Engbring, L., Yip, J., Li, H., … Parker, G. (2020). A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation. *Scientific Reports*, *10*(1), 1–15. https://doi.org/10.1038/s41598-020-68550-w

Butler, E., & Li, R. (2014). *Genetic markers for sex identification in forensic DNA analysis*.

Cabo, L. L., Brewster, C. P., & Luengo Azpiazu, J. (2012). Sexual dimorphism: interpreting sex markers. *A Companion to Forensic Anthropology*, 248–286.

Cao, Y., Ma, Y., Vieira, D. N., Guo, Y., Wang, Y., Deng, K., … Chen, F. (2021). A potential method for sex estimation of human skeletons using deep learning and three-dimensional surface scanning. *International Journal of Legal Medicine*, *135*(6), 2409–2421.

Capitaneanu, C., Willems, G., & Thevissen, P. (2017). A systematic review of odontological sex estimation methods. *The Journal of Forensic Odonto-Stomatology*, *35*(2), 1.

Castillo, A., Galtés, I., Crespo, S., & Jordana, X. (2021). preliminary insight into a new method for age-at-death estimation from the pubic symphysis. *International Journal of Legal Medicine*, *135*(3), 929–937.

Cavaignac, E., Savall, F., Faruch, M., Reina, N., Chiron, P., & Telmon, N. (2016). Geometric morphometric analysis reveals sexual dimorphism in the distal femur. *Forensic Science International*, *259*, 246-e1.

Čechová, M., Dupej, J., Brůžek, J., Bejdová, Š., Horák, M., & Velemínská, J. (2019). Sex estimation using external morphology of the frontal bone and frontal sinuses in a contemporary Czech population. *International Journal of Legal Medicine*, *133*, 1285–1294.

Celbis, O., & Agritmis, H. (2006). Estimation of stature and determination of sex from radial and ulnar bone lengths in a Turkish corpse sample. *Forensic Science International*, *158*(2–3), 135–139.

Chovalopoulou, M.-E., Valakos, E. D., & Manolis, S. K. (2013). Sex determination by three-dimensional geometric morphometrics of the palate and cranial base. *Anthropologischer Anzeiger*, Vol. 70, pp. 407–425. https://doi.org/10.1127/0003-5548/2013/0363

Chovalopoulou, M.-E., Valakos, E. D., & Manolis, S. K. (2016). Sex determination by three-dimensional geometric morphometrics of craniofacial form. *Anthropologischer Anzeiger*, *73*(3), 195–206. https://doi.org/10.1127/anthranz/2016/0470

Christensen, A. F. (2023). Sternal Rib Ends as a Method of Age Estimation at the CIL: A Brief Note. *Forensic Sciences*, *3*(4), 576–581.

Coleman, S. R., & Grover, R. (2006). The anatomy of the aging face: Volume loss and changes in 3-dimensional topography. *Aesthetic Surgery Journal*, *26*(1 SUPPL.). https://doi.org/10.1016/j.asj.2005.09.012

Colman, K. L., Janssen, M. C. L., Stull, K. E., Van Rijn, R. R., Oostra, R. J., De Boer, H. H., & Van Der Merwe, A. E. (2018). Dutch population specific sex estimation formulae using the proximal femur. *Forensic Science International*, *286*, 268-e1.

Constantinou, C., & Nikita, E. (2022). SexEst: An open access web application for metric skeletal sex estimation. *International Journal of Osteoarchaeology*, *32*(4), 832–844.

Cools, M., & Köhler, B. (2019). Disorders of sex development. *Brook's Clinical Pediatric Endocrinology*, 105–131.

Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: toward an evolutionarily rigorous cognitive science. *Cognition*, *50*(1–3), 41–77.

Cox, S. L., Moots, H. M., Stock, J. T., Shbat, A., Bitarello, B. D., Nicklisch, N., … Mathieson, I. (2022). Predicting skeletal stature using ancient DNA. *American Journal of Biological Anthropology*, *177*(1), 162–174. https://doi.org/10.1002/ajpa.24426

Cunha, E., Baccino, E., Martrille, L., Ramsthaler, F., Prieto, J., Schuliar, Y., … Cattaneo, C. (2009). The problem of aging human remains and living individuals: a review. *Forensic Science International*, *193*(1–3), 1–13.

Curate, F. (2022). The estimation of sex of human skeletal remains in the Portuguese identified collections: history and prospects. *Forensic Sciences*, *2*(1), 272–286.

Dagar, T., Sharma, L., & Khanna, K. (2019). Sexual dimorphism: metric measurements based study in human talus bone. *Int J Res Med Sci*, *7*(8), 3070.

de Paiva, L. A. S., & Segre, M. (2003). Sexing the human skull through the mastoid process. *Revista Do Hospital Das Clinicas*, *58*(1), 15–20.

Decker, S. J., Davy-jow, S., & Hilbelink, D. R. (2011). Virtual determination of sex: metric and nonmetric traits of the adult pelvis from 3D computed tomography models. *Journal of Forensic Sciences*.

Del Bove, A., Menéndez, L., Manzi, G., Moggi-Cecchi, J., Lorenzo, C., & Profico, A. (2023). Mapping sexual dimorphism signal in the human cranium. *Scientific Reports*, *13*(1), 16847.

Del Bove, A., Profico, A., Riga, A., Bucchi, A., & Lorenzo, C. (2020). A geometric morphometric approach to the study of sexual dimorphism in the modern human frontal bone. *American Journal of Physical Anthropology*, *173*(4), 643–654.

Del Bove, A., & Veneziano, A. (2022). A generalised neural network model to estimate sex from cranial metric traits: A robust training and testing approach. *Applied Sciences*, *12*(18), 9285.

Demirjian, A., Goldstein, H., & Tanner, J. M. (1973). A new system of dental age assessment. *Human Biology*, 211–227.

Dereli, A. K. (2018). Sex determination with morphological characteristics of the skull by using 3D modeling techniques in computerized tomography. *Forensic Science, Medicine and Pathology*, 16–20.

DiGangi, E. A., & Hefner, J. T. (2013). Ancestry Estimation. In *Research Methods in Human Skeletal Biology*. https://doi.org/10.1016/B978-0-12-385189-5.00005-4

Dimitriadou, A. E., Hornik, K., Leisch, F., Meyer, D., Weingessel, A., & Friedrichleischr-projectorg, M. F. L. (2010). *Misc functions of the Department of Statistics (e1071), TU Wien*.

Dryden, I. L., & Mardia, K. V. (1998). Statistical shape analysis. *Statistics in Medicine'(John Wiley & Sons, Ltd., 2000)*, *19*(19), 2716–2717.

Dudzik, B., & Jantz, R. L. (2016). Misclassifications of Hispanics using Fordisc 3.1: Comparing cranial morphology in Asian and Hispanic populations. *Journal of Forensic Sciences*, *61*(5), 1311–1318.

Dupej, J., Krajíček, V., & Pelikán, J. (2015). Low-rank matrix approximations for Coherent point drift. *Pattern Recognition Letters*, *52*, 53–58.

Elliott, M., & Collard, M. (2009). FORDISC and the determination of ancestry from cranial measurements. *Biology Letters*, *5*(June), 849–852. https://doi.org/10.1098/rsbl.2009.0462

Emlen, T., & Oring, L. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science*.

Esan, T. A., Yengopal, V., & Schepartz, L. A. (2017). The Demirjian versus the Willems method for dental age estimation in different populations: A meta-analysis of published studies. *PloS One*, *12*(11), e0186682.

Falys, C. G., Schutkowski, H., & Weston, D. A. (2006). Auricular surface aging: worse than expected? A test of the revised method on a documented historic skeletal assemblage. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *130*(4), 508–513.

Ferembach, D., Schwidetzky, I., & Stloukal, M. (1979). Recommandations pour déterminer l ’ âge et le sexe sur le squelette. *Bulletins et Mémoires de La Société d’anthropologie de Paris*, *8*, 7–45.

Filipek, P. A., Richelme, C., Kennedy, D. N., & Caviness Jr., V. S. (1994). The young adult human brain: An MRI-based morphometric analysis. *Cerebral Cortex*, *4*(4), 344–360.

Flaherty, T. M., Byrnes, J. F., & Maddalena, A. (2023). Misgendering a transgender woman using FORDISC 3.1: A case study. *Forensic Science International: Synergy*, *7*, 100342.

Franklin, D., Freedman, L., Milne, N., & Oxnard, C. E. (2006). A geometric morphometric study of sexual dimorphism in the crania of indigenous southern Africans. *South African Journal of Science*, *102*(June), 229–238.

Franklin, Daniel. (2010). Forensic age estimation in human skeletal remains: Current concepts and future directions. *Legal Medicine*, *12*(1), 1–7. https://doi.org/10.1016/j.legalmed.2009.09.001

Franklin, Daniel. (2023). Estimation of Skeletal Sex. *Encyclopedia of Forensic Sciences: Volume 1-4, Third Edition*, *2*, 292–303. https://doi.org/10.1016/B978-0-12-823677-2.00098-2

Franklin, Daniel, & Blau, S. (2020). Physical and virtual sources of biological data in forensic anthropology: Considerations relative to practitioner and/or judicial requirements. In *Statistics and probability in forensic anthropology* (pp. 17–45). Elsevier.

Franklin, Daniel, Cardini, A., Flavel, A., & Kuliukas, A. (2013). Estimation of sex from cranial measurements in a Western Australian population. *Forensic Science International*, *229*(1–3), 158.e1-158.e8.

Franklin, Daniel, Cardini, A., Flavel, A., Kuliukas, A., Marks, M. K., Hart, R., … O’Higgins, P. (2013). Concordance of traditional osteometric and volume-rendered MSCT interlandmark cranial measurements. *International Journal of Legal Medicine*, *127*(2), 505–520. https://doi.org/10.1007/s00414-012-0772-9

Franklin, Daniel, & Flavel, A. (2019). Population specificity in the estimation of skeletal age and sex: case studies using a Western Australian population. *Australian Journal of Forensic Sciences*, *51*(sup1), S188–S192.

Franklin, Daniel, O'Higgins, P., Oxnard, C. E., & Dadour, I. (2008). Discriminant function sexing of the mandible of indigenous South Africans. *Forensic Science International*, *179*(1), 84-e1.

Franklin, Daniel, Swift, L., & Flavel, A. (2016). "Virtual anthropology" and radiographic imaging in the Forensic Medical Sciences. *Egyptian Journal of Forensic Sciences*, *6*(2), 31–43. https://doi.org/10.1016/j.ejfs.2016.05.011

Frayer, D. W., & Wolpoff, M. H. (1985). Sexual Dimorphism. *Annual Review of Anthropology*, Vol. 14, pp. 429–473. https://doi.org/10.1146/annurev.an.14.100185.002241

Gao, H., Geng, G., & Yang, W. (2018). Sex Determination of 3D Skull Based on a Novel Unsupervised Learning Method. *Computational and Mathematical Methods in Medicine*.

Gapert, R., Black, S., & Last, J. (2009). Sex determination from the foramen magnum: Discriminant function analysis in an eighteenth and nineteenth century British sample. *International Journal of Legal Medicine*, *123*, 25–33.

Garvin, H. M., & Klales, A. R. (2020). Adult skeletal sex estimation and global standardization. *Forensic Science and Humanitarian Action*, 199–209. https://doi.org/10.1002/9781119482062.ch14

Garvin, H. M., Passalacqua, N. V, Uhl, N. M., Gipson, D. R., Overbury, R. S., & Cabo, L. L. (2012). Developments in forensic anthropology: age-at-death estimation. *A Companion to Forensic Anthropology*, 202–223.

Garvin, H. M., & Ruff, C. B. (2012). Sexual dimorphism in skeletal browridge and chin morphologies determined using a new quantitative method. *American Journal of Physical Anthropology*, *147*(4), 661–670.

Garvin, H. M., Sholts, S. B., & Mosca, L. A. (2014). Sexual dimorphism in human cranial trait scores: Effects of population, age, and body size. *American Journal of Physical Anthropology*, *154*(October 2013), 259–269. https://doi.org/10.1002/ajpa.22502

Gasparini, A., Lugli, F., Silvestrini, S., Pietrobelli, A., Marchetta, I., Benazzi, S., & Belcastro, M. G. (2022). Biological sex VS. Archaeological Gender: Enamel peptide analysis of the horsemen of the Early Middle age necropolises of Campochiaro (Molise, Italy). *Journal of Archaeological Science: Reports*, *41*, 103337. https://doi.org/https://doi.org/10.1016/j.jasrep.2021.103337

Giannecchini, M., & Moggi-Cecchi, J. (2008). Stature in archeological samples from central Italy: Methodological issues and diachronic changes. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *135*(3), 284–292.

Giles, E., & Elliot, O. (1963). Sex determination by discriminant function analysis of crania. *American Journal of Physical Anthropology*, *21*(1), 53–68.

Gillet, C., Costa-Mendes, L., Rérolle, C., Telmon, N., Maret, D., & Savall, F. (2020). Sex estimation in the cranium and mandible: a multislice computed tomography (MSCT) study using anthropometric and geometric morphometry methods. *International Journal of Legal Medicine*, *134*, 823–832.

Gonzales, P. N., Bernal, V., & Perez, S. I. (2009). Analysis of sexual dimorphism of craniofacial traits using geometric morphometric techniques. *International Journal of Osteoarchaeology*. https://doi.org/10.1002/oa

Green, H., & Curnoe, D. (2009). Sexual dimorphism in Southeast Asian crania: A geometric morphometric approach. *HOMO- Journal of Comparative Human Biology*, *60*(6), 517–534. https://doi.org/10.1016/j.jchb.2009.09.001

Greening, V. A., Hernandez, E., Mongle, C. S., Billings, B. K., Mngomezulu, V., Wallace, I. J., & Grine, F. E. (2024). Variation, sexual dimorphism, and enlargement of the frontal sinus with age in adult South Africans. *American Journal of Biological Anthropology*.

Guglielmi, G., Giannatempo, G. M., Blunt, B. A., Grampp, S., Gler, C. C., Cammisa, M., & Genant, H. K. (1995). Spinal bone mineral density by quantitative CT in a normal Italian population. *European Radiology*, *5*(3). https://doi.org/10.1007/BF00185311

Guyomarc'h, P., & Brůžek, J. (2011). Accuracy and reliability in sex determination from skulls: A comparison of Fordisc® 3.0 and the discriminant function analysis. *Forensic Science International*. https://doi.org/10.1016/j.forsciint.2011.03.011

Habets, L., Bras, J., & Borgmeyer-Hoelen, A. (1988). Mandibular atrophy and metabolic bone loss: endocrinology, radiology and histomorphometry. *International Journal of Oral and Maxillofacial Surgery*, *17*(3), 208–211.

Hagelberg, E., & Clegg, J. B. (1991). Isolation and characterization of DNA from archaeological bone. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *244*(1309), 45–50.

Hagelberg, E., Sykes, B., & Hedges, R. (1989). Ancient bone DNA amplified. *Nature*, *342*(6249), 485.

Harvati, K., & Weaver, T. D. (2006). Human cranial anatomy and the differential preservation of population history and climate signatures. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology: An Official Publication of the American Association of Anatomists*, *288*(12), 1225–1233.

Hefner, J. T., Spradley, M. K., & Anderson, B. (2014). Ancestry assessment using random forest modeling. *Journal of Forensic Sciences*. https://doi.org/10.1111/1556-4029.12402

Hennessy, R. J., Mclearie, S., Kinsella, A., & Waddington, J. L. (2005). Facial surface analysis by 3D laser scanning and geometric morphometrics in relation to sexual dimorphism in cerebral – craniofacial morphogenesis and cognitive function. *Society*, 283–295.

Hohlweg-Majert, B., Schmelzeisen, R., Pfeiffer, B. M., & Schneider, E. (2006). Significance of osteoporosis in craniomaxillofacial surgery: A review of the literature. *Osteoporosis International*, *17*, 167–179.

Humphrey, L. T. (1998). Growth patterns in the modern human skeleton. *American Journal of Physical Anthropology*, *105*(1), 57–72.

Ibrová, A., Dupej, J., Stránská, P., Velemínský, P., Poláček, L., & Velemínská, J. (2017). Facial skeleton asymmetry and its relationship to mastication in the Early Medieval period (Great Moravian Empire, Mikulčice, 9th–10th century). *Archives of Oral Biology*, *84*(September), 64–73. https://doi.org/10.1016/j.archoralbio.2017.09.015

Imaizumi, K., Bermejo, E., Taniguchi, K., Ogawa, Y., Nagata, T., Kaga, K., … Shiotani, S. (2020). Development of a sex estimation method for skulls using machine learning on three-dimensional shapes of skulls and skull parts. *Forensic Imaging*, *22*, 200393.

Introna, F., Di Vella, G., & Campobasso, C. Pietro. (1998). Sex determination by discriminant analysis of patella measurements. *Forensic Science International*, *95*, 39–45. https://doi.org/10.1016/S0379-0738(98)00080-2

Isa, M., Michael, A., Blatt, S., & Flaherty, T. (2022). *Centering Gender Inclusive Strategies in Forensic Anthropology*.

Işcan, M. Y., Loth, S. R., King, C. A., Shihai, D., & Yoshino, M. (1998). Sexual dimorphism in the humerus: a comparative analysis of Chinese, Japanese and Thais. *Forensic Science International*, *98*, 17–29.

Iscan, M. Y., & Steyn, M. (2013). *The human skeleton in forensic medicine*. Charles C Thomas Publisher.

Jeanson, A. L., Santos, F., Villa, C., Dupej, J., Lynnerup, N., & Brůžek, J. (2017). Body mass estimation from the skeleton: An evaluation of 11 methods. *Forensic Science International*, *281*, 183-e1.

Johnson, R. A., & Wichern, D. W. (2007). Applied multivariate statistical analysis. 6th. *New Jersey, US: Pearson Prentice Hall.*

Kanchan, T., Gupta, A., & Krishan, K. (2013). Craniometric analysis of foramen magnum for estimation of sex. *Int J Med Health Biomed Pharm Eng*, *7*, 111–113.

Kanis, J. A., & Kanis, J. A. (1994). Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. *Osteoporosis International*, *4*, 368–381.

Kanis, J. A., Melton III, L. J., Christiansen, C., Johnston, C. C., & Khaltaev, N. (1994). The diagnosis of osteoporosis. *Journal of Bone and Mineral Research*, *9*(8), 1137–1141.

Karadayı, B., Afşin, H., Ozaslan, A., & Karadayı, Ş. (2014). Development of dental charts according to tooth development and eruption for Turkish children and young adults. *Imaging Science in Dentistry*, *44*(2), 103.

Kelley, S. R., & Tallman, S. D. (2022). Population-inclusive assigned-sex-at-birth estimation from skull computed tomography scans. *Forensic Sciences*, *2*(2), 321–348.

Kendall, D. G. (1977). The diffusion of shape. *Advances in Applied Probability*, *9*(3), 428–430.

Kenyhercz, M. W., Klales, A. R., Stull, K. E., McCormick, K. A., & Cole, S. J. (2017). Worldwide population variation in pelvic sexual dimorphism: A validation and recalibration of the Klales et al. method. *Forensic Science International*, *277*, 259-e1.

Kerley, E. R. (1965). The microscopic determination of age in human bone. *American Journal of Physical Anthropology*.

Kimmerle, E. H., Ross, A., & Slice, D. (2008). Sexual dimorphism in America: Geometric morphometric analysis of the craniofacial region. *Journal of Forensic Sciences*, *53*(1), 54–57.

Klales, A. R., Ousley, S. D., & Vollner, J. M. (2012). A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. *American Journal of Physical Anthropology*, *149*(May), 104–114. https://doi.org/10.1002/ajpa.22102

Klepinger, L. L., Katz, D., Micozzi, M. S., & Carroll, L. (1992). Evaluation of cast methods for estimating age from the os pubis. *Journal of Forensic Sciences*, *37*, 763–770.

Konigsberg, L. W., Algee-Hewitt, B. F. B., & Steadman, D. W. (2009). Estimation and evidence in forensic anthropology: Sex and race. *American Journal of Physical Anthropology*, *139*(August 2008), 77–90. https://doi.org/10.1002/ajpa.20934

Kotěrová, A., Štepanovský, M., Buk, Z., Brůžek, J., Techataweewan, N., & Velemínská, J. (2022). The computational age-at-death estimation from 3D surface models of the adult pubic symphysis using data mining methods. *Scientific Reports*, *12*(1), 10324.

Kotěrová, A., Velemínská, J., Dupej, J., Brzobohatá, H., Pilný, A., & Brůžek, J. (2017). Disregarding population specificity: its influence on the sex assessment methods from the tibia. *International Journal of Legal Medicine*, *131*(1), 251–261. https://doi.org/10.1007/s00414-016-1413-5

Kranioti, E. F., & Apostol, M. A. (2015). Sexual dimorphism of the tibia in contemporary Greeks, Italians, and Spanish: forensic implications. *International Journal of Legal Medicine*, *129*, 357–363.

Kranioti, E. F., İşcan, M. Y., & Michalodimitrakis, M. (2008). Craniometric analysis of the modern Cretan population. *Forensic Science International*, *180*(2–3), 110.e1-110.e5.

Krishan, K. (2008). Estimation of stature from cephalo-facial anthropometry in north Indian population. *Forensic Science International*, *181*, 1–6. https://doi.org/10.1016/j.forsciint.2008.08.001

Krishan, K., Chatterjee, P. M., Kanchan, T., Kaur, S., Baryah, N., & Singh, R. K. (2016). A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Science International*, *261*, 165.e1-165.e8.

Krishan, K., Kanchan, T., & Sharma, A. (2012). Multiplication factor versus regression analysis in stature estimation from hand and foot dimensions. *Journal of Forensic and Legal Medicine*, *19*(4), 211–214. https://doi.org/https://doi.org/10.1016/j.jflm.2011.12.024

L'Abbé, E. N., Van Rooyen, C., Nawrocki, S. P., & Becker, P. J. (2011). An evaluation of nonmetric cranial traits used to estimate ancestry in a South African sample. *Forensic Science International*. https://doi.org/10.1016/j.forsciint.2011.04.002

Lee, U., Jung, G., Choi, S., & Kim, Y. (2014). Anthropological age estimation with bone histomorphometry from the human clavicle. *Anthropologist*, *17*(3), 929–936.

Lentini, E., Kasahara, M., Arver, S., & Savic, I. (2013). *Sex Differences in the Human Brain and the Impact of Sex Chromosomes and Sex Hormones*. (October), 2322–2336. https://doi.org/10.1093/cercor/bhs222

Lesciotto, K. M. (2023). Daubert and the Effect on Biological Profile Research. *Forensic Anthropology (University of Florida)*, *6*(4).

Lestrel, P. E., Cesar, R. M., Takahashi, O., & Kanazawa, E. (2005). Sexual dimorphism in the Japanese cranial base: A fourier-wavelet representation. *American Journal of Physical Anthropology*, *128*(December 2003), 608–622. https://doi.org/10.1002/ajpa.20209

Lettre, G. (2009). Genetic regulation of adult stature. *Current Opinion in Pediatrics*, *21*(4), 515–522.

Lieberman, D. E., McBratney, B. M., & Krovitz, G. (2002). The evolution and development of cranial form in Homosapiens. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(3), 1134–1139.

Lieberman, D. E., Pearson, O. M., & Mowbray, K. M. (2000). Basicranial influence on overall cranial shape. *Journal of Human Evolution*, *38*(2), 291–315.

Lillie, E. M., Urban, J. E., Lynch, S. K., Weaver, A. A., & Stitzel, J. D. (2016). Evaluation of Skull Cortical Thickness Changes with Age and Sex from Computed Tomography Scans. *Journal of Bone and Mineral Research*, *31*(2), 299–307. https://doi.org/10.1002/jbmr.2613

Listi, G. A., & Elizabeth Bassett, H. (2006). Test of an alternative method for determining sex from the os coxae: applications for modern Americans. *Journal of Forensic Sciences*, *51*(2), 248–252.

Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., & Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proceedings. Biological Sciences / The Royal Society*, *269*(1496), 1095–1100.

Liversidge, H. M. (2015). Tooth eruption and timing. *A Companion to Dental Anthropology*, 159–171.

Lovejoy, C. (1988). Evolution of human walking. *Scientific American*, *259*(5), 118–125.

Lovell, N. C. (1989). Test of Phenice's technique for determining sex from the os pubis. *American Journal of Physical Anthropology*, *79*, 117–120. https://doi.org/10.1002/ajpa.1330790112

Lurdes, M. De, Maria, I., Emilia, C., David, R., Fernandes, C., & Ian, D. (2016). The role of sexually dimorphic skin colour and shape in attractiveness of male faces. *Evolution and Human Behavior*, *37*(2), 125–133.

MacLaughlin, S. M., & Bruce, M. F. (1990). The accuracy of sex identification in European skeletal remains using the Phenice characters. *Journal of Forensic Sciences*, *35*(6), 1384–1392.

Mahfouz, M., Badawi, A., Merkl, B., Fatah, E. E. A., Pritchard, E., Kesler, K., … Jantz, L. (2007). Patella sex determination by 3D statistical shape models and nonlinear classifiers. *Forensic Science International*, *173*, 161–170. https://doi.org/10.1016/j.forsciint.2007.02.024

Mahfouz, M. R., Merkl, B. C., Abdel Fatah, E. E., Booth Jr, R., & Argenson, J. N. (2007). Automatic methods for characterization of sexual dimorphism of adult femora: distal femur. *Computer Methods in Biomechanics and Biomedical Engineering*, *10*(6), 447–456.

Mall, G., Hubig, M., Büttner, A., Kuznik, J., Penning, R., & Graw, M. (2001). Sex determination and estimation of stature from the longbones of the arm. *Forensic Science International*, *117*, 23–30.

Mamabolo, B., Alblas, A., & Brits, D. (2020). Modern imaging modalities in forensic anthropology and the potential of low-dose X-rays. *Forensic Imaging*, *23*, 200406.

Manthey, L., & Jantz, R. L. (2020). Fordisc: Anthropological software for estimation of ancestry, sex, time period, and stature. In *Statistics and probability in Forensic Anthropology* (pp. 275–287). Elsevier.

Martin, R., & Saller, K. (1957). *Lehrbuch der Anthropologie* (G. Fischer, Ed.). Sttutgart.

Martínez-Abadías, N., Paschetta, C., de Azevedo, S., Esparza, M., & González-José, R. (2009). Developmental and genetic constraints on neurocranial globularity: Insights from analyses of deformed skulls and quantitative genetics. *Evolutionary Biology*, *36*(1), 37–56.

Meindl, R. S., & Lovejoy, C. O. (1985). Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *American Journal of Physical Anthropology*, *68*, 57–66. https://doi.org/10.1002/ajpa.1330680106

Meindl, R. S., Lovejoy, C. O., & Mensforth, R. P. (1985). *Accuracy and Direction of Error in the Sexing of the Skeleton : Implications for Paleodemography*. *85*.

Meinerová, T., Šutoová, D., Brukner Havelková, P., Velemínská, J., Dupej, J., & Bejdová, Š. (2023). How reliable is the application of the sex classifier based on exocranial surface (Musilová et al., 2016) for geographically and temporally distant skull series. *Forensic Science Inter*, *352*(September).

Meissner, C., & Ritz-Timme, S. (2010). Molecular pathology and age estimation. *Forensic Science International*, *203*(1–3), 34–43. https://doi.org/10.1016/j.forsciint.2010.07.010

Mendelson, B. C., Hartley, W., Scott, M., McNab, A., & Granzow, J. W. (2007). Age-related changes of the orbit and midcheek and the implications for facial rejuvenation. *Aesthetic Plastic Surgery*, *31*(5), 419–423.

Mendelson, B., & Wong, C.-H. (2012). Changes in the facial skeleton with aging: Implications and clinical applications in facial rejuvenation. *Aesthetic Plastic Surgery*, *36*(4), 753–760. https://doi.org/10.1007/s00266-012-9904-3

Mennatt, M. A., H. Attia, M., Tarek Farghaly, Y., Ahmed El-Sayed Abulnoor, B., & Curate, F. (2022). Performance of the supervised learning algorithms in sex estimation of the proximal femur: A comparative study in contemporary Egyptian and Turkish samples. *Science and Justice*, *62*(3), 288–309. https://doi.org/10.1016/j.scijus.2022.03.003

Merritt, C. E. (2015). The influence of body size on adult skeletal age estimation methods. *American Journal of Physical Anthropology*, *156*(1), 35–57.

Mikšík, I., Morvan, M., & Brůžek, J. (2023). Peptide analysis of tooth enamel–A sex estimation tool for archaeological, anthropological, or forensic research. *Journal of Separation Science*, *46*(15), 2300183.

Milella, M., Franklin, D., Belcastro, M. G., & Cardini, A. (2021). Sexual differences in human cranial morphology: Is one sex more variable or one region more dimorphic? *The Anatomical Record*, *304*(12), 2789–2810.

Mitteroecker, P., & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology*, *36*, 235–247.

Mittnik, A., Wang, C.-C., Svoboda, J., & Krause, J. (2016). A molecular approach to the sexing of the triple burial at the upper paleolithic site of Dolní Věstonice. *PloS One*, *11*(10), e0163019.

Mohanraj, K. G. (2020). Analysis of Sexual Dimorphic Traits Through Morphometric Features of Clavicle. *International Journal of Pharmaceutical Research (09752366)*, *12*(2).

Moore, F. R., Coetzee, V., Contreras-Garduno, J., Debruine, L. M., Kleisner, K., Krams, I., … Suzuki, T. N. (2013). Cross-cultural variation in women's preferences for cues to sex- and stress-hormones in the male face. *Biology Letters*, *9*(3), 10–13. https://doi.org/10.1098/rsbl.2013.0050

Moorrees, C. F. A. (1965). Normal variation in dental development determined with reference to tooth eruption status. *Journal of Dental Research*, *44*(1), 161–173.

Morriss-Kay, G. M., & Wilkie, A. O. M. (2005). Growth of the normal skull vault and its alteration in craniosynostosis: insights from human genetics and experimental studies. *Journal of Anatomy*, *207*(5), 637–653.

Moss, M. L. (1968). A theoretical analysis of the Functional matrix. *Acta Biotheoretica*, *18*(1–4), 195–202. https://doi.org/10.1007/BF01556727

Mulhern, D. M., & Jones, E. B. (2005). Test of revised method of age estimation from the auricular surface of the ilium. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *126*(1), 61–65.

Murphy, R. E., & Garvin, H. M. (2018). A Morphometric Outline Analysis of Ancestry and Sex Differences in Cranial Shape. *Journal of Forensic Sciences*, *63*(4), 1001–1009.

https://doi.org/10.1111/1556-4029.13699

Musilová, B., Dupej, J., Brůžek, J., Bejdová, Š., & Velemínská, J. (2019). Sex and ancestry related differences between two Central European populations determined using exocranial meshes. *Forensic Science International*, *297*, 364–369.

Musilová, B., Dupej, J., Velemínská, J., Chaumoitre, K., & Bruzek, J. (2016). Exocranial surfaces for sex assessment of the human cranium. *Forensic Science International*, *269*, 70–77. https://doi.org/10.1016/j.forsciint.2016.11.006

Navega, D., Coelho, C., & Vicente, R. (2015). AncesTrees: ancestry estimation with randomized decision trees. *International Journal of Legal Medicine*, *129*, 1145–1153.

Nikita, E. (2012). Age-associated Variation and Sexual Dimorphism in Adult Cranial Morphology: Implications in Anthropological Studies. *International Journal of Osteoarchaeology*, (November 2011). https://doi.org/10.1002/oa.2241

Nikita, P., Garoufi, N., Valakos, E., Constantinou, C., Nikita, E., & Chovalopoulou, M. (2024). Testing the accuracy of the SexEst software for sex estimation in a modern Greek sample. *International Journal of Osteoarchaeology*, e3283.

O'Connor, J. J. M., Fraccaro, P. J., Pisanski, K., Tigue, C. C., O'Donnell, T. J., & Feinberg, D. R. (2014). Social dialect and men's voice pitch influence women's mate preferences. *Evolution and Human Behavior*, *35*(5), 368–375. https://doi.org/10.1016/j.evolhumbehav.2014.05.001

Oddie, T. H., Meade, J. H., & Fisher, D. A. (1966). An Analysis of Published Data on Thyroxine Turnover in Human Subjects. *The Journal of Clinical Endocrinology & Metabolism*, *26*(4), 425–436. https://doi.org/10.1210/jcem-26-4-425

Ogawa, Y., Imaizumi, K., Miyasaka, S., & Yoshino, M. (2013). Discriminant functions for sex estimation of modern Japanese skulls. *Journal of Forensic and Legal Medicine*, *20*(4), 234–238.

Ohtani, S. (1995). Estimation of age from dentin by utilizing the racemization of aspartic acid: influence of pH. *Forensic Science International*, *75*(2–3), 181–187.

Omari, R., Hunt, C., Coumbaros, J., & Chapman, B. (2021). Virtual anthropology? Reliability of three-dimensional photogrammetry as a forensic anthropology measurement and documentation technique. *International Journal of Legal Medicine*, *135*, 939–950.

Ost, A. M. (2022). Age-at-death estimation from the auricular surface of the ilium: A test of a sex-specific component method. *Journal of Forensic Sciences*, *67*(3), 868–876.

Ousley, S., & Jantz, R. (2013). Fordisc 3: Third generation of computer-aided forensic anthropology. *Rechtsmedizin*, *23*, 97–99.

Ousley, Stephen, Jantz, R., & Freid, D. (2009). Understanding race and human variation: why forensic anthropologists are good at identifying race. *American Journal of Physical Anthropology*, *139*(1), 68–76.

Padmanabhan, A. S., & Sapna, S. (2023). Forensic Investigation Approaches of DNA Analysis and Criminal Investigation. *International Journal of Safety & Security Engineering*, *13*(1).

Pandey, N., Roshan, S., Kharate, R., & Sonawane, M. (2014). Prediction of stature based on foot length. *Journal of Nobel Medical College*, *3*(5), 66–70.

Pandya, A. M., Singel, T. C., Akbari, V. J., Dangar, K. P., Tank, K. C., & Patel, M. P. (2011). Sexual dimorphism of maximum femoral length. *National Journal of Medical Research*, *1*(02), 67–70.

Panuganti, P. K., Reddy, M. V., Poojari, M. S., Shrish, P., & Kumar, J. A. (2022). Morphometry and Sexual Dimorphism of the Human Clavicle in South Indian Population. *Int J Anat Res*, *10*(1), 8283–8290.

Papaloucas, C., Fiska, A., & Demetriou, T. (2008). Sexual dimorphism of the hip joint in Greeks. *Forensic Science International*, *179*(1), 17–19. https://doi.org/10.1016/j.forsciint.2008.03.007

Parker, G. J., Yip, J. M., Eerkens, J. W., Salemi, M., Durbin-Johnson, B., Kiesow, C., … Phinney, B. S. (2019). Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. *Journal of Archaeological Science*, *101*, 169–180. https://doi.org/https://doi.org/10.1016/j.jas.2018.08.011

Patel, C. R., Fernandez-Miranda, J. C., Wang, W.-H., & Wang, E. W. (2016). Skull base anatomy. *Otolaryngologic Clinics of North America*, *49*(1), 9–20.

Paulinus, S. O., Mba, E. E., Ukpong, E. V, Archibong, B. E., Udoh, B. E., Egom, A. E., … Igiri, A. O. (2019). Anthropometric study of the cranial parameters using computed tomography (CT) scan to establish cephalic index of a sampled population in Calabar, Nigeria. *Global Journal of Pure and Applied Sciences*, *25*(2), 153–159.

Paulson, M. M. (2023). *Comparison of Male and Female Rib Sternal Ends and the Effects of Aging using Elliptical Fourier Analysis*.

Peckmann, T. R., Orr, K., Meek, S., & Manolis, S. K. (2015). Sex determination from the calcaneus in a 20th century Greek population using discriminant function analysis. *Science and Justice*, *55*(6), 377–382. https://doi.org/10.1016/j.scijus.2015.04.009

Phenice, T. W. (1969). A newly developed visual method of sexing the os pubis. *American Journal of Physical Anthropology*, *30*(2), 297–301.

Pilli, E., Palamenghi, A., Marino, A., Staiti, N., Alladio, E., Morelli, S., … Cattaneo, C. (2023). Comparing Genetic and Physical Anthropological Analyses for the Biological Profile of Unidentified and Identified Bodies in Milan. *Genes*, *14*(5). https://doi.org/10.3390/genes14051064

Pilmann Kotěrová, A., Santos, F., Bejdová, Š., Rmoutilová, R., Attia, M. A. H., Habiba, A., … Brůžek, J. (2024). Prioritizing a high posterior probability threshold leading to low error rate over high classification accuracy: the validity of MorphoPASSE software for cranial morphological sex estimation in a contemporary population. *International Journal of Legal Medicine*, (0123456789). https://doi.org/10.1007/s00414-024-03215-1

Pretorius, E., Steyn, M., & Scholtz, Y. (2006). Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *129*(1), 64–70.

Profico, A., Piras, P., Buzi, C., Di Vincenzo, F., Lattarini, F., Melchionna, M., … Manzi, G. (2017). The evolution of cranial base and face in Cercopithecoidea and Hominoidea: Modularity and morphological integration. *American Journal of Primatology*, *79*(12), e22721.

Quintyn, C. B. (2010). *The existence or non-existence of race?: Forensic anthropology, population admixture, and the future of racial classification in the US*. Teneo Press.

Raff, J. A. (2019). Ancient DNA and bioarcheology. *A Companion to Anthropological Genetics*, 187–197.

Ramsthaler, F., Kreutz, K., & Verhoff, M. a. (2007). Accuracy of metric sex analysis of skeletal remains using Fordisc?? based on a recent skull collection. *International Journal of Legal Medicine*, *121*, 477–482. https://doi.org/10.1007/s00414-007-0199-x

Ramsthaler, Frank, Kettner, M., Gehl, A., & Verhoff, M. a. (2010). Digital forensic osteology: Morphological sexing of skeletal remains using volume-rendered cranial CT scans. *Forensic Science International*, *195*, 148–152.

Raxter, M. H., Auerbach, B. M., & Ruff, C. B. (2006). Revision of the Fully technique for estimating statures. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *130*(3), 374–384.

Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex-typicality and attractiveness: are supermale and superfemale faces super-attractive? *British Journal of Psychology (London, England : 1953)*, *91 (Pt 1)*, 125–140. https://doi.org/10.1348/000712600161718

Ritz, S., & Schütz, H. W. (1993). Aspartic acid racemization in intervertebral discs as an aid to postmortem estimation of age at death. *Journal of Forensic Sciences*, *38*(3), 633–640.

Rosas, A., & Bastir, M. (2002). Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. *American Journal of Physical Anthropology*, *117*(February 2001), 236–245.

Rosenberg, K. R., & Trevathan, W. R. (2002). Birth, obstetrics and human evolution. *International Journal of Obstetrics and Gynaecology*, *109*(November), 1199–1206.

Rosenberg, K., & Trevathan, W. (1996). Bipedalism and human birth: The obstetrical dilemma revisited. *Evolutionary Anthropology:Issues, News, and Reviews*, *4*(5), 161–168.

Ross, A. H., & Jantz, R. (1998). Cranial thickness in American females and males. *Journal of Forensic Sciences*, *43*(January), 267–272.

Ross, A. H., Slice, D. E., & Williams, S. E. (2010). *Geometric morphometric tools for the classification of human skulls*. Department of Justice.

Rowbotham, S. K., & Blau, S. (2020). The application of medical imaging to the anthropological estimation of sex. In *Sex estimation of the human skeleton* (pp. 351–369). Elsevier.

Ruengdit, S., Case, D. T., & Mahakkanukrauh, P. (2020). Cranial suture closure as an age indicator: A review. *Forensic Science International*, *307*, 110111.

Ruff, C. B. (1994). Morphological adaptation to climate in modern and fossil hominids. *American Journal of Physical Anthropology*, *37*(S19), 65–107.

Ruff, C. B. (2000). Body size, body shape, and long bone strength in modern humans. *Journal of Human Evolution*, *38*(2), 269–290.

Ruff, C. B., Holt, B. M., Niskanen, M., Sladek, V., Berner, M., Garofalo, E., … Niinimäki, S. (2012). Stature and body mass estimation from skeletal remains in the European Holocene. *American Journal of Physical Anthropology*, *148*(4), 601–617.

Ruff, C., Trinkaus, E., & Holliday, T. (1997). Body mass and encephalization in Pleistocene Homo. *Nature*, *387*, 173–176. Retrieved from http://www.hopkinsmedicine.org/FAE/CBRETTWH1997Nature.pdf

Saini, V., Srivastava, R., Rai, R. K., Shamal, S. N., Singh, T. B., & Tripathi, S. K. (2011). An Osteometric study of Northern Indian populations for sexual dimorphism in craniofacial region. *Journal of Forensic Sciences*, *56*(3), 700–705.

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, *94*(3), 441–448.

Santos, F., Guyomarc'h, P., & Brůžek, J. (2014). Statistical sex determination from craniometrics: Comparison of linear discriminant analysis, logistic regression, and support vector machines. *Forensic Science International*, *245*, 204.e1-204.e8.

Sauer, N. J. (1992). Forensic anthropology and the concept of race: if races don't exist, why are forensic anthropologists so good at identifying them? *Social Science & Medicine*, *34*(2), 107–111.

Schanandore, J. V, Wolden, M., & Smart, N. (2022). The accuracy and reliability of the Suchey–Brooks pubic symphysis age estimation method: Systematic review and meta-analysis. *Journal of Forensic Sciences*, *67*(1), 56–67.

Scheuer, L. (2002). Application of osteology to forensic medicine. *Clinical Anatomy*, Vol. 15, pp. 297–312. https://doi.org/10.1002/ca.10028

Scheuer, L., & Black, S. (2004). *The juvenile skeleton*. Elsevier.

Schulte-Geers, C., Obert, M., Schilling, R. L., Harth, S., Traupe, H., Gizewski, E. R., & Verhoff, M. A. (2011). Age and gender-dependent bone density changes of the human skull disclosed by high-resolution flat-panel computed tomography. *International Journal of Legal Medicine*, *125*(3), 417–425. https://doi.org/10.1007/s00414-010-0544-3

Sierp, I., & Henneberg, M. (2016). Reconstruction of body height from the skeleton: Testing a dozen different methods for consistency of their results. *Anthropologischer Anzeiger*, *73*(1).

Silva, R. H. A. da, Sales-Peres, A., Oliveira, R. N. de, Oliveira, F. T. de, & Sales-Peres, S. H. de C. (2007). Use of DNA technology in forensic dentistry. *Journal of Applied Oral Science*, *15*, 156–161.

Simmons-Ehrhardt, T. L., Ehrhardt, C. J., & Monson, K. L. (2019). Evaluation of the suitability of cranial measurements obtained from surface-rendered CT scans of living people for estimating sex and ancestry. *Journal of Forensic Radiology and Imaging*, *19*, 100338.

Sjøvold, T. (1990). Estimation of stature from long bones utilizing the line of organic correlation. *Human Evolution*, *5*, 431–447.

Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., … Jakobsson, M. (2012). Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science*, *336*(6080), 466–469.

Sládek, V., Macháček, J., Makajevová, E., Přichystalová, R., & Hora, M. (2018). Body mass estimation in skeletal samples using the hybrid approach: the effect of population-specific variations and sexual dimorphism. *Archaeological and Anthropological Sciences*, *10*(4), 833–847. https://doi.org/10.1007/s12520-016-0400-6

Slatkin, M. (1984). Ecological causes of sexual dimorphism. *Evolution*, *38*(3), 622–630. https://doi.org/10.2307/2408711

Slice, D. E. (2007). Geometric morphometrics. *Annu. Rev. Anthropol.*, *36*, 261–281.

Smay, D., & Armelagos, G. (2000). Galileo wept: A critical assessment of the use of race in forensic anthropology. *Transforming Anthropology*, *9*(2), 19–29.

Sorrentino, R., Belcastro, M. G., Figus, C., Stephens, N. B., Turley, K., Harcourt-Smith, W., … Benazzi, S. (2020). Exploring sexual dimorphism of the modern human talus through geometric morphometric methods. *PloS One*, *15*(2), e0229255.

Spradley, M. K. (2021). Use of craniometric data to facilitate migrant identifications at the United States/Mexico border. *American Journal of Physical Anthropology*, *175*(2), 486–496.

Spradley, M. K., Anderson, B. E., & Tise, M. L. (2015). Postcranial sex estimation criteria for Mexican Hispanics. *Journal of Forensic Sciences*, *60*, S27–S31.

Spradley, M. K., & Jantz, R. L. (2011). Sex estimation in forensic anthropology: Skull versus postcranial elements. *Journal of Forensic Sciences*, *56*(2), 289–296. https://doi.org/10.1111/j.1556-4029.2010.01635.x

Spradley, M. K., Jantz, R. L., Robinson, A., & Peccerelli, F. (2008). Demographic change and forensic identification: Problems in metric identification of hispanic skeletons. *Journal of Forensic Sciences*, *53*(1), 21–28. https://doi.org/10.1111/j.1556-4029.2007.00614.x

Spradley, M. K., & Stull, K. E. (2018). Advancements in Sex and Ancestry Estimation. *New*

Perspectives in Forensic Human Skeletal Identification, 13–21. https://doi.org/10.1016/B978-0-12-805429-1.00003-X

Steyn, M., & İşcan, M. Y. (2008). Metric sex determination from the pelvis in modern Greeks. *Forensic Science International*, *179*(1), 86-e1.

Stock, M. K., Stull, K. E., Garvin, H. M., & Klales, A. R. (2016). Development of modern human subadult age and sex estimation standards using multi-slice computed tomography images from medical examiner's offices. *Developments in X-Ray Tomography X*, *9967*, 48–61. SPIE.

Stone, A. C., Milner, G. R., Paäbo, S., & Stoneking, M. (1996). Sex determination of ancient human skeletons using DNA. *American Journal of Physical Anthropology*, *99*, 231–238. https://doi.org/10.1002/(SICI)1096-8644(199602)99:2<231::AID-AJPA1>3.0.CO;2-1

Susmaga, R. (2004). Confusion Matrix Visualization. In *Intelligent Information Processing and Web Mining* (pp. 107–116). https://doi.org/10.1007/978-3-540-39985-8_12

Tague, R. G., & Lovejoy, C. O. (1998). AL 288-1 — Lucy or Lucifer: gender confusion in the Pliocene. *Journal of Human Evolution*, *35*, 75–94.

Tallman, S. D., & Blanton, A. I. (2020). Distal humerus morphological variation and sex estimation in modern Thai individuals. *Journal of Forensic Sciences*, *65*(2), 361–371.

Tallman, S. D., Kincer, C. D., & Plemons, E. D. (2021). Centering transgender individuals in forensic anthropology and expanding binary sex estimation in casework and research. *Forensic Anthropology*, *5*(2), 161–180.

Thali, M. J., Braun, M., Buck, U., Aghayev, E., Jackowski, C., Vock, P., … Dirnhofer, R. (2005). VIRTOPSY - Scientific documentation, reconstruction and animation in forensic: Individual and real 3D data based geo-metric approach including optical body/object surface and radiological CT/MRI scanning. *Journal of Forensic Sciences*, *50*(2), 1–15.

Thomas, R. M. (2020). Sex determination using DNA and its impact on biological anthropology. In *Sex Estimation of the Human Skeleton* (pp. 343–350). Elsevier.

Thomas, R. M., Parks, C. L., & Richard, A. H. (2016). Accuracy Rates of Sex Estimation by Forensic Anthropologists through Comparison with DNA Typing Results in Forensic Casework. *Journal of Forensic Sciences*, *61*(5), 1307–1310. https://doi.org/10.1111/1556-4029.13137

Thomas, R., Parks, C., & Richard, A. (2017). Accuracy Rates of Ancestry Estimation by Forensic Anthropologists Using Identified Forensic Cases. *J*, *62*(4). https://doi.org/10.1111/1556-4029.13361

Thornhill, R., & Gangestad, S. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, *3*(12), 452–460.

Tiede, U., Bomans, M., Höhne, K. H., Pommert, A., Riemer, M., Schiemann, T., … Lierse, W. (1993). A computerized three-dimensional atlas of the human skull and brain. *American Journal of Neuroradiology*, *14*(3), 551–559.

Tierney, S. N., & Bird, J. M. (2015). Molecular sex identification of juvenile skeletal remains from an Irish medieval population using ancient DNA analysis. *Journal of Archaeological Science*, *62*, 27–38.

Toneva, D., Nikolova, S., Tasheva-Terzieva, E., Zlatareva, D., & Lazarov, N. (2022). A geometric morphometric study on sexual dimorphism in viscerocranium. *Biology*, *11*(9), 1333.

Torimitsu, S., Nishida, Y., Takano, T., & Koizumi, Y. (2014). Statistical analysis of biomechanical properties of the adult skull and age-related structural changes by sex in a Japanese forensic sample. *Forensic Science International*, *234*, 185.e1-185.e9.

Trevathan, W., & Rosenberg, K. (2000). The shoulders follow the head: postcranial constraints on human childbirth. *Journal of Human Evolution*, *39*(6), 583–586.

Trotter, M., & Gleser, G. C. (1952). Estimation of stature from long bones of American Whites and Negroes. *American Journal of Physical Anthropology*, *10*(4), 463–514.

Tsuji, A., Ishiko, A., Takasaki, T., & Ikeda, N. (2002). Estimating age of humans based on telomere shortening. *Forensic Science International*, *126*, 197–199.

Ubelaker, D. H., & DeGaglia, C. M. (2017). Population variation in skeletal sexual dimorphism. *Forensic Science International*, *278*, 407-e1.

Urbanová, P., Ross, A. H., Jurda, M., & Nogueira, M. I. (2014). Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample. *Legal Medicine*, *16*(5), 264–273. https://doi.org/10.1016/j.legalmed.2014.06.002

Uzun, Ö., Yegİnoglu, G., Ertemoğlu Öksüz, C., Kalkışım, Ş., & Zihni, N. (2019). Estimation of stature from upper extremity anthropometric measurements. *Journal of Clinical and Diagnostic Research*, *13*.

Velemínská, J., Bigoni, L., Krajíček, V., Borský, J., Šmahelová, D., Cagáňová, V., & Peterka, M. (2012). Surface facial modelling and allometry in relation to sexual dimorphism. *HOMO- Journal of Comparative Human Biology*, *63*(2), 81–93.

Velemínská, Jana, Fleischmannová, N., Suchá, B., Dupej, J., Bejdová, Š., Kotěrová, A., & Brůžek, J. (2021). Age-related differences in cranial sexual dimorphism in contemporary Europe. *International Journal of Legal Medicine*, *135*(5), 2033–2044. https://doi.org/10.1007/s00414-021-02547-6

Velemínská, Jana, Jaklová, L. K., Kočandrlová, K., Hoffmannová, E., Koudelová, J., Suchá, B., & Dupej, J. (2022). Three-dimensional analysis of modeled facial aging and sexual dimorphism from juvenile to elderly age. *Scientific Reports*, *12*(1), 1–12. https://doi.org/10.1038/s41598-022-26376-8

Vercellotti, G. (2012). *Assessment of inter and intra-population variation in stature and body proportions: a comparative study between living and bioarchaeological populations*. The Ohio State University.

Vercellotti, G., Piperata, B. A., Agnew, A. M., Wilson, W. M., Dufour, D. L., Reina, J. C., …

Stout, S. D. (2014). Exploring the multidimensionality of stature variation in the past through comparisons of archaeological and living populations. *American Journal of Physical Anthropology*, *155*(2), 229–242.

Verhoff, M. A., Ramsthaler, F., Krähahn, J., Deml, U., Gille, R. J., Grabherr, S., … Kreutz, K. (2008). Digital forensic osteology-Possibilities in cooperation with the Virtopsy® project. *Forensic Science International*, *174*, 152–156.

Verma, R., Krishan, K., Rani, D., Kumar, A., & Sharma, V. (2020). Stature estimation in forensic examinations using regression analysis: A likelihood ratio perspective. *Forensic Science International: Reports*, *2*, 100069. https://doi.org/https://doi.org/10.1016/j.fsir.2020.100069

Villa, C., Lynnerup, N., & Buckberry, J. (2019). Evaluating osteological ageing from digital data. *J*, *235*(2). https://doi.org/10.1111/joa.12544

Vystrčilová, M., & Novotný, V. (2000). Estimation of age at death using teeth. *Variability and Evolution*, *8*, 39–49. Retrieved from http://var-and-evo.biol.uni.torun.pl/03_8.pdf

Walker, P. L. (2008). Sexing skulls using discriminant function analysis of visually assessed traits. *American Journal of Physical Anthropology*, *136*(July 2007), 39–50.

Walker, R. A., & Lovejoy, C. O. (1985). Radiographic changes in the clavicle and proximal femur and their use in the determination of skeletal age at death. *American Journal of Physical Anthropology*, *68*, 67–78. https://doi.org/10.1002/ajpa.1330680107

Walrath, D. E., Turner, P., & Bruzek, J. (2004). Reliability test of the visual assessment of cranial traits for sex determination. *American Journal of Physical Anthropology*, *125*(August 2002), 132–137.

Weber, G. W., & Bookstein, F. L. (2011). *Virtual anthropology: a guide to a new interdisciplinary field*. Springer.

Weiss, E. (2024). No Bones About It: Sex Is Binary. *Archives of Sexual Behavior*, *53*(5), 1595–1608. https://doi.org/10.1007/s10508-024-02851-3

Wells, J. C. K., Treleaven, P., & Cole, T. J. (2007). BMI compared with 3-dimensional body shape: The UK National Sizing Survey. *American Journal of Clinical Nutrition*, *85*(2), 419–425.

Wilkinson, C., Liu, C. Y. J., Shrimpton, S., & Greenway, E. (2024). Craniofacial identification standards: a review of reliability, reproducibility, and implementation. *Forensic Science International*, 111993.

Willems, G., Ph, D., Olmen, A. Van, Spiessens, B., Carels, C., G, R. W., … Carels, C. (2001). Dental Age Estimation in Belgian Children : Demirjian ' s Technique Revisited. *J of Forensic Sciences*, (March 2000), 893–895.

Williams, B. a., & Rogers, T. L. (2006). Evaluating the accuracy and precision of cranial morphological traits for sex determination. *Journal of Forensic Sciences*, *51*(4), 729–735.

Wilson, L. A., MacLeod, N., & Humphrey, L. T. (2008). Morphometric criteria for sexing juvenile human skeletons using the ilium. *Journal of Forensic Sciences*, *53*(2), 269–278.

Wood, K., & Cunningham, C. (2000). Age Determination in the Juvenile. *Forensic Anthropology*.

Yoder, C., Ubelaker, D. H., & Powell, J. F. (2001). Examination of variation in sternal rib end morphology relevant to age assessment. *Journal of Forensic Sciences*, *46*(2), 223–227.

Zelditch, M., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric morphometrics for biologists: a primer*. academic press.

Zeman, T., & Beňuš, R. (2020). Stature estimation. In *Statistics and Probability in Forensic Anthropology* (pp. 249–256). Elsevier.

Zhang, M. (2024). The Application of Forensic Imaging to Sex Estimation: Focus on Skull and Pelvic Structures. *Perspectives in Legal and Forensic Sciences*, *1*(1), 10005–10005. https://doi.org/10.35534/plfs.2024.10005

Zhang, Q., Paz, A. R., Banner, J., Jørkov, M. L. S., & Villa, C. (2020). Stature estimation from postmortem CT femoral maximum length in contemporary Danish population. *Journal of Forensic Sciences*, *65*(3), 930–938.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3D	Three dimensional
aDNA	Ancient deoxyribonucleic acid
AI	Artificial intelligence
AUC	Area under the curve
BMI	Body mass index
CPD-DCA	Coherent point drift – dense correspondence analysis
CT	Computed tomography
CV	Cross-validation
CZE	Dataset collected in Czech Republic
DEN	Dataset collected in Denmark
DNA	Deoxyribonucleic acid
EGY	Dataset collected in Egypt
FBI	Federal Bureau of Investigation
FRA	Dataset collected in France
GM	Geometric morphometry
GPA	Generalised Procrustes analysis
LDA	Linear discriminant analysis
PC	Principal component
PCA	Principal component analysis
PP	Posterior probability
ROC	Receiver operating characteristic
SVK	Dataset collected in Slovakia
SVM	Support vector machine

APPENDICES

a. Submitted publications

Musilová, B., Dupej, J., Velemínská, J., Chaumoitre, K., & Brůžek, J. (2016). Exocranial surfaces for sex assessment of the human cranium. Forensic science international, 269, 70-77.

Musilová, B., Dupej, J., Brůžek, J., Bejdová, Š., & Velemínská, J. (2019). Sex and ancestry related differences between two Central European populations determined using exocranial meshes. Forensic science international, 297, 364-369.

Velemínská, J., Fleischmannová, N., **Suchá, B**., Dupej, J., Bejdová, Š., Kotěrová, A., & Brůžek, J. (2021). Age-related differences in cranial sexual dimorphism in contemporary Europe. International Journal of Legal Medicine, 135, 2033-2044.

Velemínská, J., Jaklová, L. K., Kočandrlová, K., Hoffmannová, E., Koudelová, J., **Suchá, B.**, & Dupej, J. (2022). Three-dimensional analysis of modeled facial aging and sexual dimorphism from juvenile to elderly age. Scientific Reports, 12(1), 21821.

b. Conferences

Musilová, B., Dupej, J., Chaumoitre, K., & Brůžek, J. (2015). Validity of the sex assessment method from skull using CT scans in European population. Presented at the 7th European Academy of Forensic Science Conference (EAFS 2015).

Musilová, B. (2017). Populační specificita a odhad pohlaví na základě lebky za využití metod virtuální antropologie. Oral presentation at the Anthropological Congress in Olomouc, Conference Antropologické dny 2017.