



Universiteit
Leiden
The Netherlands

Can the fibula determine sex? Osteometric analysis around the nutrient foramen

van Smeerdijk, Pien

Citation

Van Smeerdijk, P. (2024). *Can the fibula determine sex?: Osteometric analysis around the nutrient foramen.*

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/3800088>

Note: To cite this publication please use the final published version (if applicable).

Can the Fibula determine sex?

OSTEOMETRIC ANALYSIS AROUND THE NUTRIENT FORAMEN

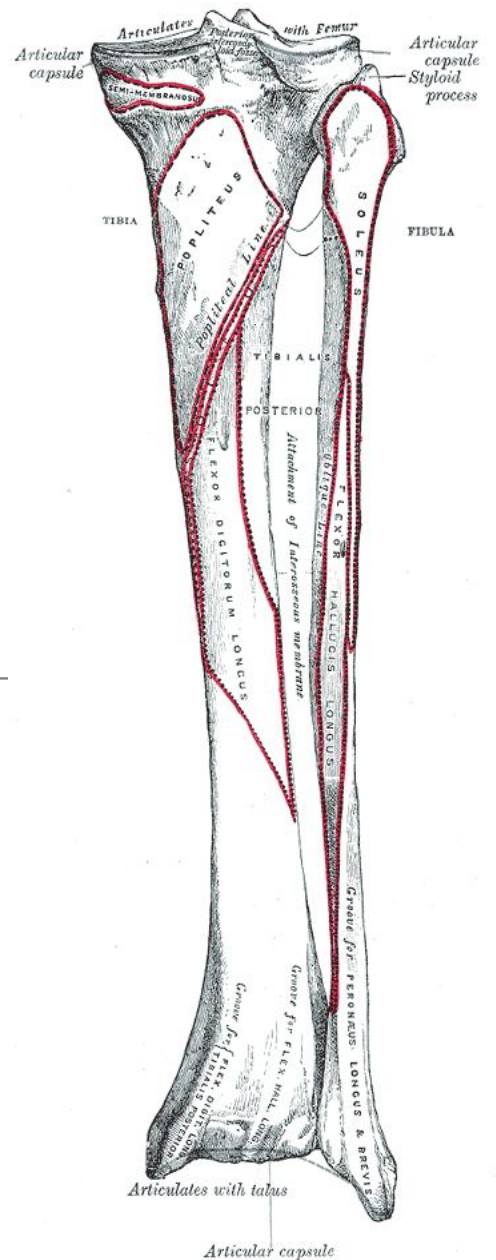


FIG. 259.—Bones of the right leg. Posterior surface.



Universiteit
Leiden
The Netherlands

Pien Willemijn van Smeerdijk

Front cover image: Anatomical picture of the lower leg from Gray's Anatomy.

Can the fibula determine sex?

Osteometric analysis around the nutrient foramen

Pien Willemijn van Smeerdijk (s3763838)

Master Thesis Archaeological Science, Human Osteoarchaeology (2023-2024)

Supervisor: Dr. R. Schats

Leiden University, Faculty of Archaeology

Leiden, June 14th, Final

Table of Contents

Acknowledgements	5
1. Introduction	6
1.1 <i>Aims and Research questions</i>	<i>8</i>
1.2 <i>Approach of this study</i>	<i>9</i>
1.3 <i>Ethical conduct.....</i>	<i>9</i>
1.4 <i>Thesis outline</i>	<i>10</i>
2. Theoretical background.....	11
2.1 <i>Sexual dimorphism and sex estimation.....</i>	<i>11</i>
2.1.1 <i>Biomolecular methods.....</i>	<i>11</i>
2.1.2 <i>Morphological methods.....</i>	<i>12</i>
2.1.3 <i>Metrics</i>	<i>13</i>
2.2 <i>Fibula</i>	<i>14</i>
2.3 <i>Nutrient foramen.....</i>	<i>15</i>
2.4 <i>Fasemore et al. method.....</i>	<i>15</i>
2.5 <i>Forensic relevance</i>	<i>16</i>
3. Materials and methodology	17
3.1 <i>Materials.....</i>	<i>17</i>
3.1.1 <i>Middenbeemster</i>	<i>17</i>
3.1.2 <i>Sample selection</i>	<i>18</i>
3.2 <i>Methodology.....</i>	<i>18</i>
3.2.1 <i>Fibular measurements</i>	<i>18</i>
3.2.2 <i>Descriptive statistics</i>	<i>21</i>
3.2.3 <i>Inferential statistics.....</i>	<i>21</i>

3.2.4 Discriminant analysis	22
4. Results	24
4.1 <i>Intra Observer Error</i>	24
4.2 <i>Descriptive statistics</i>	24
4.3 <i>Tests of normality</i>	25
4.4 <i>Left vs Right</i>	27
4.4.1 T-test	27
4.4.2 Mann-Whitney U-test	27
4.4.3 Normality Left and Right	27
4.5 <i>Male vs female</i>	28
4.5.1 T-test	28
4.5.2 Mann-Whitney U- test	29
4.6 <i>Discriminant function analysis</i>	29
4.6.1 Direct analysis Left	30
4.6.2 Stepwise analysis Left	31
4.6.3 Direct analysis combined	32
4.6.4 Created functions.....	33
5. Discussion	34
5.1 <i>General observations</i>	34
5.2 <i>Sexual dimorphism</i>	34
5.2.1 Proximal end to nutrient foramen	35
5.2.2 Circumference diameter	36
5.2.3 Anterior posterior diameter.....	37
5.2.4 Medial lateral diameter	37
5.3 <i>Comparison to Fasemore et al.</i>	38

5.4 Limitations	38
6. Conclusion	40
6.1 Summary of the research	40
6.2 Outcomes	40
6.3 Recommendations for future research	41
Abstract	42
References	43
List of Figures	50
List of Tables	51
Appendix I: Normality figures	52
Appendix II: Q-Q Plots	53
Appendix III: Raw data	56
Appendix IV: Intra observer error	59

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Rachel Schats for all the help and advice on the content of this thesis, and Dr. Sarah Schrader for stepping in to shape my thesis in the beginning.

I would also like to thank Marijke Langevoort, Rachael Hall, and Alex Tutweiler for their practical support in the lab and for helping to collect the individuals.

Lastly, my appreciation also goes out to my family and friends for their encouragement and support all through my studies. In particular, my mom for using her graphical skills to help create the figures for demonstrating the measurements.

1. Introduction

Osteoarchaeology is an important sub-discipline of archaeology and anthropology (Maggio, 2018, p. 1). It studies human skeletal remains from archaeological sites to reveal insights into past human populations. This field plays a vital role in reconstructing aspects of the life, health, diet, and social structures of ancient communities. Data such as age, sex, and population size can help gain an understanding of human evolution and population dynamics. Additionally, the examination of skeletal remains provides evidence of historical health conditions, dietary habits, physical activities, and cultural practices. Furthermore, osteoarchaeology helps to trace migration patterns and contributes to forensic anthropology by providing methods to identify human skeletal remains (Belcastro et al., 2022, p. 349). Building on the foundational work of osteoarchaeology, sex estimation plays a crucial role in the further analysis of human skeletal remains (Buikstra & Ubelaker, 1994, p. 272). Understanding the biological sex of individuals is instrumental in piecing together not only their identities but also their lifestyles and the environments they interact with. In an archaeological context, a population can be better understood by yielding this information. The identity of skeletal remains can also be important information in a forensic context. Estimating if the biological sex is male or female, is one of the initial steps in the analysis of an unknown person and linking a name to its skeletal remains. Knowing the biological sex of an individual greatly aids in identification efforts, since it reduces the number of people to search for by 50% (Leskovar et al., 2023, p. 2). However, keep in mind that there is a difference between gender and sex. Gender is an individual's phenotype how someone identifies and expresses themselves, whereas biological sex refers to the genes of an individual (Pryzgodna & Chrisler, 2000, p. 554). In doing research into a skeleton's biological identity, sex estimation is the first step since the results serve as a foundation for the application of the other identifying techniques (Klales, 2020, p. 31).

The estimation of sex can be done in three different ways. Namely, the visual assessment, metric analysis, or the use of DNA methods (Leskovar et al., 2023, p. 1). The distinction between a biological male or female can be seen in the DNA of the individual, whereas females have two X chromosomes, and males have a X and a Y chromosome. Although molecular techniques are the most reliable, they are time-consuming, complex, costly, and most importantly to archaeologists, invasive (Rösing et al., 2007, p. 78). Since extracting DNA is a destructive technique, the preference of archaeologists lays in the macroscopic methods. By the visual assessment, morphological traits and features that are different between the biological sexes, so called sexually dimorphic, are examined. Sex estimation can be achieved by recognizing the distinct physical characteristics that differentiate between male and female. These distinct skeletal characteristic enables the development of standardized methods to determine the sex (Buikstra & Ubelaker 1994; Maat & Mastwijk 2009; Phenice

1969; Rosing et al., 2007). The visual assessment method involves analyzing the cranium, pelvis, and mandible. Each morphological feature of these bones is rated using a numerical scoring system, which, in accordance with the WEA method used in the Netherlands, ranges from -2 to +2. (Workshop of European Anthropologists (WEA), 1980). Within this framework, a score of -2 indicates characteristics typically associated with females, while a score of +2 is indicative of traits commonly attributed to males. The internationally recognized method utilizes a categorical scale denoting M (male), PM (probable male), I (indeterminate), PF (probable female), and F (female). A trait is classified as indeterminate or zero when it does not distinctly align with male or female characteristics, rendering it ambiguous for sex determination. Human skeletons display sexual dimorphism, meaning there are measurable differences in the size and shape of bones between males and females. Metric analysis uses these disparities to estimate sex. Typically, male bones are generally larger and more robust, while female bones tend to be smaller and less robust (Sacragi & Ikeda, 1995, p. 139).

For estimating the sex of the individual, the pelvis and the skull are the bones that are used the most because the estimation based on these is the most precise, as these parts are the most sexually dimorphic (Mays, 2021, p. 55). However, those bones are not always preserved or are only found in fragments. Evaluating an individual's biological profile is constrained by the preservation quality and integrity of the available human skeletal remains. Osteologists often face challenges in discerning certain biological characteristics when the remains are in a deteriorated, cremated, or fragmented condition (Outram et al., 2005, p. 1700). For this reason, it is important to have additional sex estimation techniques that rely on other bones in the skeleton.

Long bones can be a good alternative for the estimation of sex when for example the sexual dimorphic skull and pelvis are not available (Knecht et al., 2023, p. 1887). A metric approach has been used the most in many recent studies on sex estimation on long bones, especially in forensic anthropology. Several authors have focused their efforts on deriving equations from measurements on the shaft and at the proximal and distal ends of long bones because these bones are often recovered with variable stages of fragmentation (Fasemore, et al., 2018, p. 222.e1). Metrics enables a detailed analysis of morphological characteristics that show variations in skeletons. Additionally, metric analysis decreases subjectivity and could improve inter- and intraobserver error, making it preferable to morphological approaches (Moore et al, 2016). Most of the long bones are already investigated intensively, but the fibula is one of the bones that is studied the least. A few attempts, like Sacragi and Ikeda (1995), Fasemore et al. (2018) and Pietrobelli et al. (2022), are made to study if the fibula can provide reliable information for estimating the sex of human bones, but still there is no clear answer of whether the bone can be used or not. Since the fibula's morphological variation between

males and females is expected to be too low to enable precise sex estimations, the fibula is frequently not employed. Nonetheless, it is possible that in archaeological or forensic contexts, the fibula is the only bone available (Maass & Friedling, 2020, p. 345). Thus, having a validated method for estimating the sex from only the fibula could be very useful and expands the opportunity to identify human remains.

Fasemore et al. (2018) proposed a method whereby the dimensions around the nutrient foramen, an opening in the bone shaft where an artery can run through, of the fibula were researched for the degree of sexual dimorphism. Because the nutrient foramen is fairly easy to recognise and may be preserved on the shaft of long bones, it is beneficial to use it in situations where long bones may be fractured or damaged. In the study of Fasemore and colleagues, 204 fibulae of South African whites and South African Africans were included in their sample. Results of this study show that the measurements around the nutrient foramen are sexually dimorphic; sex was correctly classified with an average accuracy rate of 69-77%. However, this method is only tested on anatomical African populations and hasn't been tested on an archaeological population. Since the variety of sexual dimorphism of long bones varies among populations, it is important to test it on a Dutch population to see if this method could be a potential method to use in the Netherlands (Iscan, 2005, p. 107).

1.1 Aims and Research questions

This study will investigate the sexual dimorphism of the fibula based on metrics in a Dutch archaeological population, and it will assess the accuracy of the method developed by Fasemore et al. (2018) to see if it can be recommended for future use.

Therefore, the main research question that this thesis aims to answer is as follows:

- How accurate are the dimensions of Fasemore et al. around the nutrient foramen of the fibula in estimating sex in a Dutch archaeological population?

Sub questions:

- To what extent can the left and right fibulae be combined?
- Which measurements have the greatest sexual dimorphism?
- How accurate are the developed formulae for determining the sex of fibulae from the Middenbeemster population?

1.2 Approach of this study

This thesis goal is to propose a method for Dutch populations to determine the biological sex of an individual based on only the fibula. The aim of this study is to replicate the sexual estimating method described by Fasemore et al. (2018) and test it on a Dutch population. This methodology is based on four different measurements around the nutrient foramen. One measurement includes the proximal end to the nutrient foramen, a circumference measurement, and two measurements regarding the posterior-anterior and the mediolateral diameters around the nutrient foramen. The values of these measurements obtained should then classify the individuals as male or female.

The selected skeletal collection for this study is Middenbeemster. Middenbeemster serves as an exemplary reference collection due to the availability of sex-specific archival data for its individuals. The prior knowledge of each individual's sex enables the validation of the measurement's effectiveness within the collection. By applying measurements for classification, one can straightforwardly compare these classifications against archival records. This comparison allows for an evaluation of the method's precision by assessing the correspondence between the assigned sex categories and the associated records. The idea is to create formulas based on the measurements of the Middenbeemster collection in order to distinguish between male and female.

1.3 Ethical conduct

This research carries a profound ethical responsibility. Human remains are not mere artifacts; they are the physical remnants of once-living people and are often considered sacred by their descendants and communities. Ethical conduct in this field is not just a professional obligation; it is a profound moral imperative. The essence of ethical engagement with human skeletal remains lies in the acknowledgment of the remains as vestiges of once-living beings, imbued with personal narratives and cultural significance. This respect is not restricted to the deceased alone but extends to encompass their descendants. Maintaining openness regarding research methodologies, discoveries, and interpretive frameworks is a staple of ethical practice. Those who handle human remains should adopt a stance of guardianship, prioritising the honourable treatment and respect of these remnants. Adherence to established protocols for the handling of human remains is essential to upholding these ethical standards (Belcastro et al., 2022, p. 354).

1.4 Thesis outline

This thesis consists of six chapters. After the introductory chapter, the second chapter will provide more background information for this thesis. Topics like sex estimation, sexual dimorphism, and the fibula will be discussed. The methods and materials used in order to answer the main research question are covered in chapter three. The methodology and the population sample for the skeletal remains will be addressed here. The fourth chapter covers the results of this research. In Chapter five, those results are discussed and compared with the literature. Finally, in the sixth chapter conclusions are drawn by answering the research questions and suggesting further research if necessary. Additional information like tables, and figures of raw data, can be seen in the appendices.

2. Theoretical background

In this chapter, background information related to this thesis is provided. Sexual dimorphism and the methods to estimate the sex will be explained in more detail. More information about the fibula, the bone at the centre of this research, and the nutrient foramen is presented in this chapter. Additionally, the forensic relevance of this analysis is mentioned.

2.1 Sexual dimorphism and sex estimation

Sexual dimorphism refers to the difference in morphological characteristics between females and males, primarily influenced by hormonal differences between the sexes. This dimorphism is rooted in biological sex, which is determined by their chromosomes. A female has two X chromosomes, whereas a male has one X and one Y chromosome. In accordance with the chromosomal sex, the foetus develops either the testes or the ovaries, including their associated sex-specific hormones. These hormones are crucial for the development of secondary sexual characteristics, which become pronounced during puberty. It is therefore not reliable to estimate the sex of a non-adult because it has not undergone puberty yet and the sexual dimorphism is not fully developed. In adults, however, males generally exhibit larger and more robust bones than females. Bone length is the most noticeable sex difference in the metrical properties of the long bone, including the fibula (Sacragi & Ikeda, 1995, p. 139). Sex estimation in an archaeological context can be done in three different ways. Looking at the biomolecular methods, morphological traits of a bone, or the metrics. Each of these methods has its strengths and limitations and may be employed depending on the condition of the skeletal remains and the specific requirements of the research. Effective sex estimation in archaeology often involves a combination of these approaches to increase accuracy and provide a more comprehensive understanding of the individuals being studied.

2.1.1 Biomolecular methods

Since chromosomes establish an individual's biological gender, DNA analysis provides a means of sex identification in situations where skeletal morphology alone cannot indicate an individual's gender. All bone cells, such as osteocytes, osteoblasts, and osteoclasts, as well as the cells found in the teeth's dentine and cementum, contain DNA (a few exceptions being red blood cells). Dental enamel is devoid of DNA because it lacks a cell structure (Mays, 2021, p. 292). Biomolecular methods for estimating the sex of an individual are destructive, which is something that archaeologists do not prefer to do. However, an advantage of DNA sequencing is that it can be used to determine the biological sex of young individuals who haven't reached puberty yet (Buonasera et al., 2020, p. 1). Human DNA was found to have survived in ancient skeletal remains and could be successfully recovered from them in the late 1980s (Hagelberg et al., 2014, p. 1). The petrous portion of the skull's temporal bone is

considered the best location for the extraction of human DNA, particularly in archaeological contexts where DNA preservation is typically poor. To obtain a sufficient quantity of DNA for analytical purposes, it is necessary to initially amplify the extracted DNA using the polymerase chain reaction (PCR) technique. This step is crucial due to the minimal amounts of DNA often recovered from archaeological specimens (Mays, 2021, p. 295).

An additional method of estimating sex in archaeological environments is proteomics. Recent proposals have suggested proteomic study of sex-specific amelogenin peptides in tooth enamel (Buonasera et al., 2020, p. 1). Since proteins' molecular structure is more conducive to preservation than DNA's, they can be valuable subjects for study in a variety of archaeological contexts (Wadsworth et al., 2017, p. 2). Furthermore, amelogenin peptides could be very stable and enduring over extended periods of time because the strongest and most robust material in the human body, tooth enamel, incorporates them into its mineral phase (Welker et al., 2019, p. 264). Amelogenin genes are essential for the manufacture of enamel and are found on the X and Y chromosomes in humans. AMELX_HUMAN (AMELX) and AMELY_HUMAN (AMELY) are two different isoforms of amelogenin proteins that these genes express. By identifying these proteins, it is possible to infer sex throughout archaeological time eras (Stamfelj, 2021, p. 1). This method was developed by Steward et al (2017), where peptides were identified by nanoflow liquid chromatography mass spectrometry (nanoLC-MS).

2.1.2 Morphological methods

When estimating an individual's biological sex with a morphological approach, specific sexually dimorphic traits are considered. By looking at the different traits, researchers can assign them a score based on how male or female their shape is. While the entire skeleton should be considered, most commonly only the pelvis, cranium and mandible are analysed. The skull and the pelvis differ the most between the sexes and are therefore the most accurate in estimating the sex. From those two, the pelvis is the most reliable since its distinctions observed are directly associated with functional disparities between males and females. Since the pelvis of a woman is typically wider than that of a man because giving birth requires a wider pelvis which is forced by the birth canal (Mays, 2021, p. 55). Numerous morphological traits contribute to the wider breadth of the female pelvis, and those dimorphic features are described and scored in the Phenice (1969) and Buikstra and Ubelaker (1969) methods. The Phenice (1969) method uses the ventral arc, subpubic concavity, and the medial

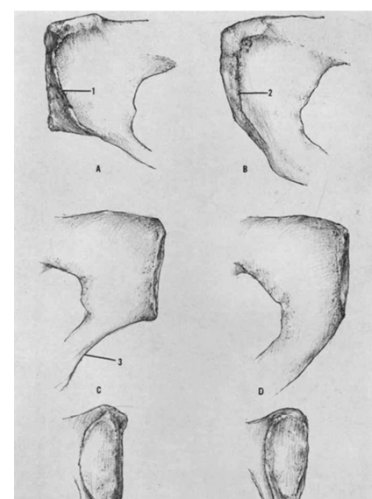


Figure 1. Phenice traits. (Phenice, 1969, p. 299, Figure 1.)

aspect of the ischio-pubis ramus to differentiate between male and female. In figure 1 are the different features illustrated. Where the left side is considered female and the right-side male. The ventral arc (1A) is only seen on females. The subpubic concavity is wider in females, and the ischio-pubis ramus has a ridge in females. The second accurate bone to estimate sex based on morphology is the cranium. The Buikstra and Ubelaker (1994) method uses traits like the mastoid process, supraorbital margin, nuchal crest, supraorbital ridge, and the glabella (Figure 2). For the mandible, traits like the angle and of the jaw, the flaring, and the robustness of the chin are considered.

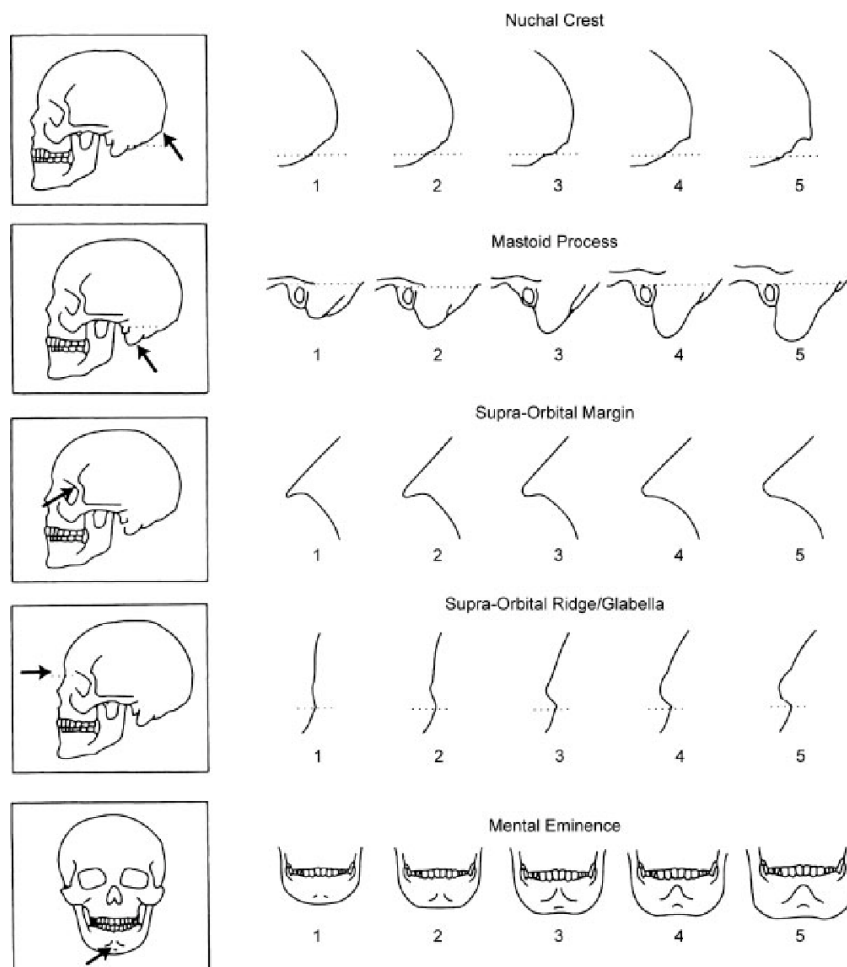


Figure 2. Traits of the Buikstra and Ubelaker method. (Buikstra & Ubelaker, 1994, p.272, Figure 2.

2.1.3 Metrics

The advantage of estimating sex based on metrics, is that it is less subjective, as it is not based on shape but on size. The metric assessment is highly repeatable due to its well-defined and

straightforward measurements. However, a population standard is needed since the variety of metrics is different between populations (Colman et al., 2018, p. 2). The requirement to first assign an unknown person to a particular group or population before a method can be used is a difficulty. Numerous characteristics, such as gender, age, birth year, race, ancestry, country, and continental origin, have been used to define membership in a group (Albanese et al., 2016, p. 59). Male bones tend to be larger and more robust than female bones. Based on this concept, bones are measured and classified as male or female according to their size (Stewart, 1979). The level of sexual dimorphism present and the association between specific metrics are usually taken into consideration when choosing measurements (İşcan and Steyn 2013, 168). Long bones, notably the humerus and femur, have been thoroughly measured, generating high accuracy in sex estimation (İşcan & Steyn, 1999, p. 83). In the Netherlands, metric analysis is done using a digital sliding calliper, a measure tape, and an osteometric board. The measure tape is mostly used for circumference measurements, and the osteometric board is used for total length measurements. A digital sliding calliper is very precise for smaller measurements and therefore useful for diameters.

In addition to sex estimation, metrics are also used for indices and stature. The metrics of a skull are typically used for the cranial index, which is part of indices to measure to categorize head shapes in populations. A recent study also uses the indicis for sexual dimorphism, but without promising results (Dhanwate & Ankushrao, 2017). In mathematical models for stature, a prediction formula is typically provided by long bone lengths regressed on stature. In these situations, a stature estimate is produced along with a confidence interval or prediction interval by entering measurements from a combination of bones or a measurement of the femur into a formula (Spradley, 2016, p. 396).

2.2 Fibula

The fibula together with the tibia forms the lower leg. It is a long bone in the lower limb that is positioned on the lateral side and connected to the tibia through the interosseous membrane. Compared to the tibia, the fibula is substantially thinner and smaller. The main purpose of the fibula, together with the tibia, is to stabilise the ankle joint. There are four components to the fibula's structure: the head (proximal end), neck, shaft, and distal end (lower part). Further features of the fibula are the, styloid process, the lateral malleolus, and the malleolar fossa (Figure 3).

Because the fibula is not directly connected to the knee joint, it does not support the weight of the body in the lower leg whenever the human body is standing or moving. As a result, the fibula has been considered a bone that, aside from length, may not show much sexual dimorphism (Sacragi & Ikeda, 1995, p. 142).

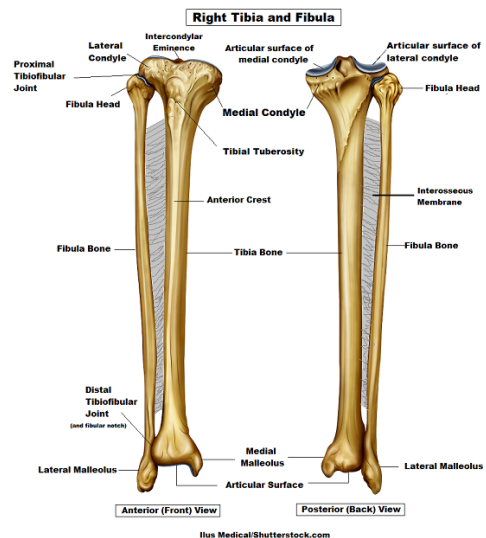


Figure 3. Tibia and fibula. (RegisteredNurseRn)

2.3 Nutrient foramen

On most of the fibula, a nutrient foramen can be seen. A nutrient foramen refers to an opening in the bone shaft. The purpose of this hole is that an artery can run through it and supply nutrients for the bone (Das et al., 2023, p. 1276). It provides the main blood supply for long bones, particularly during the early stages of ossification during the embryos and fetus' growth period (Ambekar & Sukre, 2016, p. 2685). As a result, they are crucial for the healing of fractures and the avoidance of avascular bone necrosis. (Dutta et al., 2020, p. 79). Humans are known to differ in the location and the direction of the fibula's nutritional foramina. In some cases, two or three nutrient foramina are present on the fibula. But often just one and the most common location of nutrient foramina in fibula is the middle 1/3rd of the shaft on the medial crest (Borthakur et al., 2023; Dutta et al., 2020). Due to the nutrient foramen's easy identification and potential preservation, the use of measurements surrounding it is beneficial in situations where long bones may be fractured and damaged. The nutrient foramen has greatly been used in archaeological context for the identification of human remains but also to differentiate between human and non-human (Corieri & Márquez-Grant, 2019, p. 257). Forensic experts have also employed measurements around the nutrient foramen of the fibula for sex discrimination, height interpretation, and human identification (Choudhary et al., 2023, pp. 1276-1277). Measurements of the fibula's nutrient foramen have not only been demonstrated to be sexually dimorphic, but they have also proven high accuracy in classifying sex (Mokoena et al., 2017, p. 7).

2.4 Fasemore et al. method

Fasemore et al. explore the use of measurements around the nutrient foramina of the tibia and fibula for estimating sex in forensic anthropology. This method is particularly beneficial

when dealing with fragmented or damaged long bones, as the nutrient foramen is often preserved and easily identifiable. The research focused on South African populations, utilising samples from the Raymond A. Dart Collection of Human Skeletons. It involved metric analysis of both the tibia and the fibula, attempting to develop specific osteometric standards tailored to these populations for more accurate sex determination. Four measurements were taken around the nutrient foramen of both the tibia and the fibula. These included the circumference, the distance from the nutrient foramen to the proximal end, the mediolateral diameter at the nutrient foramen, and the anteroposterior diameter at the nutrient foramen. The concordance correlation of reproducibility was used to test the measuring method's accuracy and precision. For this, a test sample of twenty bones was measured once and again two days apart. Additionally, the intra-observer error was also performed. The researchers applied discriminant function analysis to these measurements to create a model that can classify the sex of the individuals based on the dimensions related to the nutrient foramen. The accuracy of the discriminant functions was assessed by calculating their ability to correctly classify the sex of the individuals in the study. This involved comparing the predicted sex against the known sex from the skeletal records. Their two key findings include that the measurements around the tibia's nutrient foramen provided sex classification accuracy between 79-88%, with the circumference at the nutrient foramen level being the most reliable single predictor. The fibula's measurements were less accurate, with sex classification accuracy ranging from 69-77%. The study suggests that while these measurements can assist in sex determination, they should be used with caution due to their lower performance compared to tibial measurements (Fasemore et al., 2018).

2.5 Forensic relevance

There is a fine line between osteoarchaeology and forensic anthropology. Forensic analysis relies on an extensive understanding of research based on well-established generalised principles. One main aspect in forensic science which is like osteoarchaeology is the identification of an individual. Building this biological profile from skeletal remains consists of four main methods including stature, age, ancestry, and sex. In osteoarchaeology, the same methods are needed for the identification of an individual. However, a big difference between forensic and archaeological research is the context (Pinheiro e al., 2004, p. 138).

This research contributes to osteoarchaeology and forensic anthropology by refining and validating methods that can be applied particularly in cases where only partial skeletal remains are available, enhancing the accuracy of biological profiles in forensic contexts. The chances of finding only a part of the whole individual are higher in forensic cases. This is the main reason why it is important to establish identification methods based on only one bone, so that even when only some parts of the whole skeleton are available, a biological profile can still be made in some way.

3. Materials and methodology

This chapter introduces the materials and methods used in this research. It gives some background information about the used Middenbeemster collection, the used sample, and its preservation. Further, the methodology regarding the measurements and the statistics are discussed.

3.1 Materials

This first section provides an overview of the study's used materials, including a background on Middenbeemster and their population, and the sample selection.

3.1.1 Middenbeemster

For this research, a Dutch population from the Middenbeemster was used. Middenbeemster is a town in the Dutch province of North Holland (Figure 4). Located at the heart of the so-called Beemsterpolder, this area was formed after the Beemster lake was drained in the 17th century. A group of investors drained this the Beemster lake to obtain fertile land and created a grid of slots for cultivation. In 1999, UNESCO designated it as a World Heritage site due to its previously unknown and unusual design (Hoogland, 2011). The inhabitants primary source of income was dairy farming. They mostly consume eggs, potatoes, wheat, and dairy products. Barley, oats, beans, peas, were among the other crops (Bartholdy et al., 2019, p. 3). The original idea was to construct five churches in the Beemsterpolder, but only the one in Middenbeemster was completed. So, all residents of Beemster were laid to rest in the Middenbeemster cemetery. The cemetery has been operational since the early 17th century and remained active until 1866, when a new cemetery was established that is still operating today (Lemmers et al., 2013, p. 35).



Figure 4. *The location of Middenbeemster on a map of the Netherlands. (Bartholdy et al., 2018, p. 3 Figure 1)*

The Excavation

The former cemetery was excavated in 2011 by the University of Leiden's Laboratory for Human Osteoarchaeology in cooperation with the archaeological business Hollandia. The excavation of this cemetery took place from June 13th until August 5th and was necessary due to scheduled construction works adjacent to the church. The cemetery has graves dating from 1617 until 1866 and are located next to the Keyserkerk church of Middenbeemster, and

archival documents are available. The majority of the excavated remains date from the most recent period, between 1829 and 1866. In total, more than 500 individuals were excavated in just eight weeks. A big advantage from the Middenbeemster collection is that besides the individuals, also a cemetery map with all the names of the people buried in each plot was recovered (Hoogland, 2011).

3.1.2 Sample selection

As mentioned above, the Middenbeemster collection did have archival data on the buried individuals. This is of importance for this study as the retrieved data could be compared to the actual information about an individual being a biological male or female. To prevent biases, the individuals were blind-picked, whereas the sex of the individuals remained unknown until after measuring. Besides the known sex of the individuals, the preservation state and the age of the individual were factors in selecting or rejecting human remains for this study. A total of 300 individuals from the Middenbeemster collection were available at the faculty of Leiden for research. Due to preservation issues, only 151 individuals were suitable for examination. Concerning age criteria, individuals needed to be at least 18 years old, as this is the age at which the skeletal is fully developed, and to successfully show sexual dimorphism. After considering the requirements above, the sample used in this study consists of the left and right fibulae of 148 individuals.

3.2 Methodology

The aim of this study is to replicate the sexual estimating method described by Fasmore et al. (2018) and test it on a Dutch population. This following section provides an overview of the study's methodology, including measurements, equipment, and the statistical analysis used.

3.2.1 Fibular measurements

To take measurements of the fibula it was important to identify the nutrient foramen first. Hereby, bones with significant pathological damage were excluded since the nutrient foramen could not be identified and the bones would be too altered from their original size. Where possible, both the left and right fibula were measured, discriminating with an "L" and "R". The four measurements of the fibulae were done using a sliding calliper, measuring tape and an osteometric board. When both left and right fibulae were available from a single skeleton, both sides were examined for comparison, and where possible the left and right side were combined for statistical analysis and function construction.

1. PENF

Proximal end of fibula to nutrient foramen (penf): This is the length of the fibula measured linearly from its most proximal end to the nutrient foramen. This was taken from the inferior border of the opening of the foramen to the apex of the head of the fibula using an osteometry board. (Fasemore et al., 2018).



Figure 5. *PENF measurement.*
(Photograph: Pien van Smeerdijk)

2. CIRC

Circumference at nutrient foramen (circ): Using a normal measuring tape, the outlines of the bone were followed to take this measurement at the level of the fibula's nutrition foramen. (Fasemore et al., 2018).



Figure 6. *CIRC measurement.*
(Photograph: Pien van Smeerdijk)

3. APD

Antero-posterior diameter at nutrient foramen (apd): This is the linear measurement made using a digital sliding calliper at the level of the nutrient foramen, tracing the fibula's anterior border to posterior surface. (Fasemore et al., 2018).

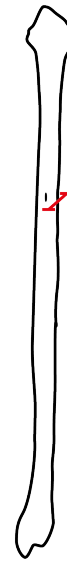


Figure 7. APD measurement.
(Photograph: Pien van Smeerdijk)

4. MLD

Mediolateral diameter at nutrient foramen (mld): At the level of the nutrient foramen, this is the distance measured with a digital sliding calliper between the medial and lateral margins of the fibula. (Fasemore et al., 2018).



Figure 8. MLD measurement
(Photograph: Pien van Smeerdijk).

3.2.2 Descriptive statistics

All the statistics were done using IBM SPSS (Statistical Package for the Social Sciences) software. The left side, and where possible, also the left and right sides combined, were used for statistical analysis.

Intra Observer Error

Intra-observer reproducibility was calculated on the basis of two sets of measurements by the same observer. This has been done by calculating the Intra class correlation coefficient. After measuring all the individuals, 10% of those measurements, have been measured again after two months. In total 14 left and right fibulae were re-measured 2 months later. The intraclass correlation coefficient ranges between zero and one. If the coefficient is 1.0 it indicates no variability between the measurements.

Descriptive statistics

Descriptive statistics are done to get an overview of the data. The means, standard deviations, and sample size were determined for both sexes as well as each measurement on the left and right fibula by descriptive statistics. Additionally, the sexual dimorphism index was calculated according to the following formula:

$$\text{Sexual dimorphism index} = \frac{Mm}{Mf} \times 100$$

The variables Mm and Mf represent the average measurements for males and females. The presence of significant sexual dimorphism within the sample was assessed using this formula. An index value exceeding 100 indicated that males exhibited larger dimensions relative to females.

3.2.3 Inferential statistics

Normal distribution / Shapiro-Wilk test

Before doing inferential statistical tests, it is important to know if the sample has a normal distribution. Testing the normal distribution of data can ensure the accuracy of your statistical analysis and accurate interpretations of the results. The normal distribution of the sample has been calculated by the Shapiro-Wilk test. The data is considered normally distributed if the significance value of the test is above 0.05. First, the left and right sides individually were tested for normality. After calculating if the left and right sides weren't statistically different, see (t-test), they were combined. The normality was then again tested for the combined measurements. To visualise the normal distribution, a histogram and the Mahalanobis distance Q-Q plots were used.

Histogram and Mahalanobis Q-Q plots

To visualise the normality, a histogram of all the measurements was made. If the samples on the histogram are bell curved, it has a normal distribution. In addition to visualising the normal distribution, outliers were detected in the Mahalanobis distance Q-Q plots. This plot gives an overview of the data and their distribution. On the Q-Q plot of the data, if the data are normally distributed, the points will roughly fall along a straight diagonal line; however, when the points depart from the straight line, it suggests that the data are outliers and are not distributed as expected by a normal distribution.

T-test/ Mann-Whitney U test

The independent t-test was employed to assess the presence of a statistically significant difference between male and female mean values for each measurement that has a normal distribution. If the significance is below 0.05 it means that the difference between the variables is statistically significant. If the data is not normally distributed, a t-test cannot be done. To determine if there is a significant difference between male and female with not normally distributed data, a Mann-Whitney U test is then recommended.

3.2.4 Discriminant analysis

Linear Discriminant Function Analysis (DFA) is a statistical method for differentiating between two or more object groupings or cases based on a set of variables that describe them. In this analysis, the grouping variables "Male" or "Female" serve as the dependent variables. The independent variables, which are the measurements, can be used to predict group membership by discovering which factors are most effective in discriminating between the groups. The DFA has been done in two different ways; a direct analysis and a stepwise method. A stepwise method has been used to identify the optimal variables for the discriminant function of the left side. For a direct analysis, all predictor variables are included in the model at once. Two direct analyses were done. One for the left side measurements, and one for the left and right combined. The outcome of a linear discriminant analysis is a function as follows: $Y=a+B_1X+B_2X+B_3X$. With "a" being the constant, and B1, B2, and B3 the created variables per measurement. A discriminant function analysis assumes that the predictor variables are normally distributed, thus normally distributed data is necessary.

Wilk's Lambda

The first element that is tested with a discriminant function analysis is the Wilk's lambda. Wilk's lambda is a powerful tool in discriminant function analysis to evaluate the effectiveness of the discriminant functions and to test whether the differences between the groups are statistically significant based on the set of independent variables. The Wilk's lambda indicates which element might be the best predictor for the discriminant function.

Log determinant and Box's M test

Before creating the discriminant functions, a Box's M test was done. A statistical test called Box's M test is used to determine whether covariance matrices for various groups are equal. It's particularly useful in multivariate statistics, when preparing to perform Discriminant function analysis, where an assumption needs to be made that the groups have equal covariance matrices. A non-significant result, a p-value less than 0.05, indicates that the covariance matrices of the groups are not significantly different, and thus, the assumption of homogeneity is met.

Standardized coefficients

In discriminant function analysis, standardised coefficients are used to understand the relative importance of each predictor variable in distinguishing between groups. These coefficients quantify how much the discriminant function score is affected by a unit change in a standardised predictor variable, holding other variables constant. It indicates the strength and direction of the relationship between that variable and the discriminant function. A larger absolute value of a standardised coefficient means that the variable has a stronger influence on the discriminant score. A positive coefficient suggests that higher values of the variable are associated with higher scores on the discriminant function, typically indicating a stronger presence in one group over another. Conversely, a negative coefficient implies that higher values are associated with lower discriminant scores.

Unstandardized coefficients

In discriminant function analysis, unstandardized coefficients, also known as raw coefficients, play a crucial role in defining the discriminant function equation. These coefficients are determined directly from the original scales of the variables involved, without any standardisation of the data. Unstandardized coefficients in discriminant function analysis are the weights assigned to each predictor variable in the discriminant function. These coefficients are derived from the covariance matrices of the variables and the groups being studied. The discriminant function can be expressed as a linear equation: $Y = a + B_1X_1 + B_2X_2 + B_3X_3$. Where B_1 , B_2 , and B_3 are the unstandardized coefficients.

Centroid

In discriminant function analysis, centroids are the mean values of the discriminant scores for each group or class. These centroids are critical for understanding the positioning of each group within the space defined by the discriminant functions and for classifying new cases into one of the predefined groups. When a new observation is analysed, it is assigned to the group whose centroid is closest to the observation's score on the discriminant function(s).

4. Results

This chapter presents the results of the collected data for this thesis. The results are divided into the descriptive statistics to give an overview of the data, the intraobserver error, the tests of normality, the T-test, the Mann-Whitney U, and the discriminant function analysis. The goal of this research is to investigate the sexual dimorphism of the fibula by testing the accuracy of the dimensions around the nutrient foramen for estimating the sex. The fibulae of 141 individuals were used to measure four different areas around the nutrient foramen. The proximal end of the fibula was frequently missing, which was adverse for the “PENF” measurement. In addition, the nutrient foramen could not be observed in some cases, due to taphonomic damage or simply the absence of it.

4.1 Intra Observer Error

To ensure the reliability and validity of the measurements, the intra class correlation coefficient (ICC) has been calculated. A test sample of 14 bones (10% of all the individuals) were measured and re-measured two months apart by the same observer. Not all measurements of the individuals were observable, so a total of 22 measurements were used for the intra class correlation coefficient. Both the left and right sides were measured and re-measured. The results of all variables are shown in Table 1. All measurements are statistically significant with a significance of less than 0.001. Additionally, they all have a correlation around 1.0, indicating of almost no variability between the multiple measurements. “PENF” has the highest correlation, but the lowest degrees of freedom. The rest of the measurements all have 21 degrees of freedom. “APD” has the lowest intraclass correlation of 0.960.

Table 1. Intra Class Correlation Coefficient

	Intraclass Correlation	95% interval		Value	df1	df2	Sig.
		Lower	Upper				
PENF	1.000	.996	1.000	5543.222	5	5	<.001*
CIRC	.990	.946	.997	159.618	21	21	<.001*
APD	.960	.905	.983	25.997	21	21	<.001*
MLD	.999	.997	.999	769.601	21	21	<.001*

*Statistical significant difference exists

4.2 Descriptive statistics

Descriptive statistics give an overview of the data. In Table 2, the descriptive statistics of the present study are shown. It pictures the distribution of the sample, showing the differences between female and male and left (_L) and right (_R) measurements. The sample consists out of 66 females and 75 males, with a total of 141 individuals. The number of individuals used

for each measurement is also shown. Evidently, the “PENF” left and right measurements have a remarkable lower sample size, since this measurement could often not be done due to the missing proximal end. A difference can be seen in all the fibula measurements between female and male, with males having higher overall higher averages. The standard deviation (SD) for both sexes was the highest for the proximal to nutrient foramen length. The smallest variation is observed for the medio-lateral diameter. Additionally, the sexual dimorphism index is illustrated in Table 2. All the values are above 100, meaning that the male measurements have greater dimensions compared to females.

Table 2. Descriptive statistics.

Measurement (mm)	Females (N=66)			Males (N=75)			Sexual dimorphism index
	N	Mean (mm)	SD	N	Mean (mm)	SD	
PENF_L	26	164.83	42.03	30	171.83	47.03	104.25
CIRC_L	58	39.28	6.34	57	43.05	8.94	109.60
APD_L	58	13.20	2.25	55	14.18	3.33	107.42
MLD_L	58	9.75	2.02	57	11.24	2.69	115.28
PENF_R	26	148.81	33.42	26	155.85	54.59	104.73
CIRC_R	51	39.82	6.46	57	43.74	9.16	109.84
APD_R	51	13.67	2.37	57	14.59	3.41	106.73
MLD_R	51	9.76	1.91	56	11.19	2.62	114.65

4.3 Tests of normality

Table 3 illustrates whether the various measurements that have been taken of the fibula are normally distributed. The Shapiro-Wilk test for normality on the left fibulae was non-significant for all measurements except for the proximal end to the nutrient foramen (PENF_L) because of the significance value below 0.05. Moreover, the statistics of “APD”, “MLD” and “CIRC” are the closest to 1.0, which also indicates the normality. The medio-lateral measurement “MLD_L” has the highest normality, and “PENF_L” the lowest. The normality of the measurements of the right fibulae is distributed differently. On the right side, the “PENF” and “APD” measurements have a significance above 0.05, meaning that these measurements are normally distributed. The circumference and medio-lateral measures are not normally distributed. The proximal end to the nutrient foramen has the highest significance of the right “PENF_R” fibulae. “CIRC_R” has the lowest value. For the measurements that significantly deviate from normality, non-parametric testing has to be performed in subsequent analyses.

Table 3. Results of the Shapiro-Wilk test. Df=degree of freedom

Measurement	Statistic	df	Significance
PENF_L	.940	58	.007*
CIRC_L	.980	121	.071
APD_L	.990	119	.530
MLD_L	.992	121	.740
PENF_R	.974	53	.307
CIRC_R	.971	113	.015*
APD_R	.978	113	.054
MLD_R	.974	112	.030*

*Statistical significant difference exists

Furthermore, the normality was visualised, and outliers were detected in the Mahalanobis distance Q-Q plots. The “PENF_L”, “CIRC_R” and “MLD_R” measurements were proven to be not normally distributed by the Shapiro-Wilk test. If the data are normally distributed, the points will fall approximately along a straight diagonal line on the Q-Q plot of the data, but when points fall far from the straight line, it suggests outliers and indicates that they deviate from what is expected in a normal distribution. On the histogram, a bell curve will appear for normally distributed data. A right skewed histogram can be observed in Figure 9, meaning that “PENF_L” is not normally distributed. On the Q-Q plot of PENF_L, almost none of the measurements lay around the line of distribution (Figure 10).

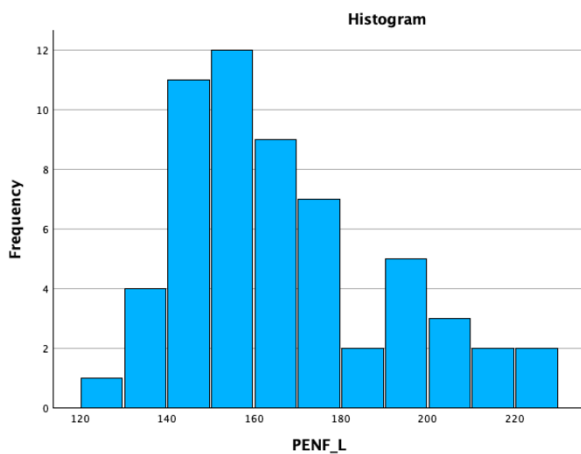


Figure 9. Histogram of PENF_L.

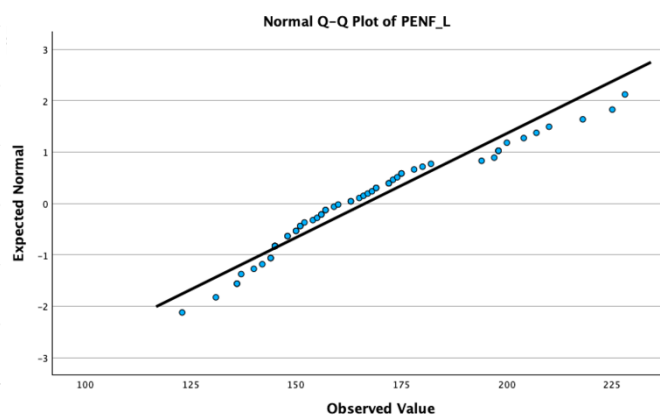


Figure 10. Mahalanobis distance Q-Q Plot of PENF_L.

4.4 Left vs Right

In order to make the sample size bigger and combine the left and right measurements, there needs to be no statistical difference between left and right. A t-test and a Mann-Whitney U-test have been done to determine this.

4.4.1 T-test

Left and right APD had a normal distribution. For this reason, the t-test is used to determine if left and right have a significant difference. As seen in Table 4, the two-tailed significance value is above 0.05 for the “APD” measurement, meaning that there is no statistically significant difference between left and right. Thus, the left and right measurements for “APD” are allowed to be combined.

Table 4. Results of the t-test. Significance difference between left and right.

Measurement	t	df	Sig. (2-tailed)	Mean difference	Std. Error difference	95% interval	
						Lower	Upper
APD	-1.876	230	.062	-.45954	.24493	-.94213	.02306

4.4.2 Mann-Whitney U-test

The rest of the measurements did not have a normal distribution (see Table 3), so the non-metric Mann-Whitney U test is used instead. It can be seen in Table 5 that all the three measurements do not have a statistically significant difference between left and right ($p > 0.05$). Meaning that the left and right measurements can be combined.

Table 5. Results of the Mann-Whitney U-test. Significance difference between left and right.

Measurement	Null Hypothesis	Test	Sig.	Decision
PENF	The distribution of PENF is the same across L and R.	Independent-Samples Mann-Whitney U Test	.191	Retain the null hypothesis.
CIRC	The distribution of PENF is the same across L and R.	Independent-Samples Mann-Whitney U Test	.265	Retain the null hypothesis.
MLD	The distribution of PENF is the same across L and R.	Independent-Samples Mann-Whitney U Test	.733	Retain the null hypothesis.

4.4.3 Normality Left and Right

Because left and right showed no significant difference in Tables 4 and 5, they can be combined to expand the sample size and possibly improve the normality. However, a second test of normality needs to be done before the following statistical analysis to know what test

needs to be used. Table 6 shows the normality of the measurements when the left and right sides are combined. It can be seen that only the anterior-posterior and the medio-lateral measurements have a normal distribution, according to the significance above 0.05.

Table 6. Normality left and right combined.

Measurement	Statistic	df	Significance
PENF	.958	111	.001*
CIRC	.981	234	.003*
APD	.990	232	.101
MLD	.992	233	.276

*Statistical significant difference exists

4.5 Male vs female

An independent sample t-test and Mann-Whitney U test were conducted to evaluate whether statistically significant differences exist between males and females across the four measurements. Since there are no statistical differences between left and right, only the left side measurements were used for the t-test and Mann-Whitney U test of male versus female. A t-test can only be done if the samples are normally distributed. For this reason, "PENF" was extracted for the t-test and the difference between male and female for "PENF" will be done using a non-parametric Mann-Whitney U test in the chapter below (4.5.2).

4.5.1 T-test

The findings are shown in Table 7. Examination of the data reveals that for all measurements, there are significant discrepancies, indicated by significance values less than 0.05. Notably, the mean difference and standard error difference for the circumference measurement are comparatively higher. Additionally, it is worth noting that the presence of a significant difference between males and females is evidenced by the exclusion of zero within the 95% confidence interval bounds for the difference.

Table 7. Results of the T-test. Calculated if there was a difference between male and female.

Measurement	t	df	Sig. (2-tailed)	Mean difference	Std. Error difference	95% interval	
						Lower	Upper
CIRC_L	6.563	113	<.001*	4.70659	.71712	3.28585	6.12733
APD_L	3.981	111	<.001*	1.30037	.32666	.65307	1.94767
MLD_L	3.533	113	<.001*	1.36575	.38653	.59997	2.13153

*Statistical significant difference exists

4.5.2 Mann-Whitney U- test

For the proximal end to nutrient foramen measurement (PENF_L) the Mann-Whitney U test was done to determine if there is a difference between males and females. The results are shown in Table 8. The resulting significance is below 0.05, meaning that we can reject the null hypothesis. Thus, there is a difference between the male and female groups.

Table 8. Mann-Whitney U-test

Measurement	Null Hypothesis	Test	Sig.	Decision
PENF_L	The distribution of PENF_L is the same across categories of the Sex	Independent-Samples Mann-Whitney U Test	.021*	Reject the null hypothesis

*Statistical significant difference exists

4.6 Discriminant function analysis

Discriminant Function Analysis (DFA) is used to distinguish between males and females based on the set of measurements that describe them. The methods to perform DFA can vary, with direct and stepwise discriminant function analysis being two distinct approaches. First a direct analysis and a stepwise analysis were done using the measurements from the left side. Additionally, a direct function analysis was done to create a function based on the left and right combined measurements.

Wilk's Lambda was acquired for the three left side measurements. A low value of Wilk's lambda reflects high significance and the good discriminating power of the model. As seen in Table 9, "CIRC_L" has the lowest value, and "APD_L" the highest. This indicates that circumference might be the best predictor for the discriminant function, and the anterior-posterior diameter the worst. Further, all the measurements are statistically significant.

Table 9. Tests of Equality of Group Means.

Measurements	Wilk's Lambda	F	df1	df2	Sig.
CIRC_L	.719	42.903	1	110	<.001*
APD_L	.879	15.172	1	110	<.001*
MLD_L	.854	18.781	1	110	<.001*

*Statistical significant difference exists

Table 10 illustrates the correlation between the measurements. Circumference has the best correlation with the anterior-posterior diameter (.455) and vice versa (.687). The medial-lateral diameter also correlates the best with circumference. Between the anterior-posterior and the medial-lateral measurements, a negative correlation exists, which means that they

have an inverse relationship. As one variable increases, the other tends to decrease. Overall, the best correlation is between “CIRC_L” and “APD_L” with the matrices closest to 1.0.

Table 10. Pooled Within-Groups Correlation Matrices.

	CIRC_L	APD_L	MLD_L
CIRC_L	1.00	.687	.455
APD_L	.687	1.00	-0.30
MLD_L	.455	-0.30	1.00

4.6.1 Direct analysis Left

First, a direct discriminant analysis was done. In direct discriminant function analysis, all predictor variables are included in the model at once.

Log determinant and Box’s M test

The log determinants of male, female and pooled are all close together. This implies that there is homogeneity in the covariance matrices across the female and male groups (Table 11). The significance of Box’s M test is 0.007 which is significant because it is above 0.001. Thus, all the matrices are equal, and the assumption for proceeding with the discriminant technique is met.

Table 11. Log determinant.

Sex	Rank	Log determinant
Male	3	4.438
Female	3	3.103
Pooled within groups	3	3.924

Table 12. Box’s M.

Box’s M		18.274
F	Approx.	2.955
	df1	6
	df2	87358.261
	Sig.	.007

The standardised canonical discriminant function coefficients are each predictor's weight in the linear combination that is the discriminant function. In other terms, Table 13 summarises which variable is the best predictor. The circumference is the best and the anterior-posterior diameter is the worst predictor of sex.

Table 13. Standardized coefficients

	Function 1
CIRC_L	.860
APD_L	-.005
MLD_L	.261

Discriminant function coefficients

The discriminant function coefficient, centroids, and average accuracy of the direct discriminant function are shown in Table 14. For function one, all the available variables on the left side were used. For function 1, the unstandardized coefficient of the constant is -10.652. The coefficients of the measurements are -0.003, 0.222, and 0.138. The female centroid is -0.626 whereas the male centroid is 0.648. So, in this function, a negative score of -0.626 would be classified as female, and a positive score of 0.648 would be classified as male. 71.4% of the original group cases are correctly classified, and 70.5 of cross-validated grouped cases are correctly classified.

Table 14. Discriminant function coefficients. Unstandardized coefficient, centroids and accuracy.

Function	Variables	Unstandardized Coefficient	Centroids	Average accuracy %	
				Original	Cross
1	CIRC_L	.222	F= -.626 M= .648	71.4	70.5
	APD_L	-.003			
	MLD_L	.138			
	Constant	-10.652			

4.6.2 Stepwise analysis Left

Besides the direct analysis of the left side variables, a stepwise analysis is also done. The standardised canonical discriminant function coefficient of the circumference is shown in Table 15.

Table 15. Standardized coefficients

	Function 1
CIRC_L	1.000

Log determinant and Box's M test

In Table 16 the log determinants of male, female and pooled are all around 3.0. This implies that there is homogeneity in the covariance matrices across the female and male groups. The significance of the Box's M test (table 17) is 0.570, which is significant because it is above 0.001. This signifies that the matrices are equal, and the assumption for proceeding with the discriminant technique is met.

Table 16. Log determinant.

Sex	Rank	Log determinant
Male	1	2.782
Female	1	2.629
Pooled within groups	1	2.707

Table 17. Box's M.

Box's M		.325
F	Approx.	.322
	df1	1
	df2	36268.011
	Sig.	.570

Stepwise analysis of all variables selected a discriminant function of only the circumference measurement. The discriminant function coefficient, centroids, and average accuracy of the function are shown in Table 18. For function two, the unstandardized coefficient of "CIRC_L" is 0.258 and the corresponding constant is -10.759. The centroid for female is -0.608 and 0.630 for males. the original classification accuracy for this function is 71.3% and for the cross-validation, it is also 71.1%.

Table 18. Discriminant function coefficients.

Function	Variables	Unstandardized Coefficient	Centroids	Average accuracy %	
				Original	Cross
2	CIRC_L	.258	F= -.608	71.1	71.1
	Constant	-10.759	M= .630		

4.6.3 Direct analysis combined

For two of the four variables (APD and MLD), the left and right sides could be combined (table 6). So, additionally direct analysis of the combined "APD" and "MLD" is done.

Log determinant and Box's M test

The log determinant of male is 2.325, female is 1.476, and pooled is 1.973 (Table 19). The significance of the Box's M test is 0.002 which is significant because it is above 0.001. Thus, all the matrices are equal, and the assumption for proceeding with the discriminant technique is met.

Table 19. Log determinant.

Sex	Rank	Log determinant
Male	2	2.324
Female	2	1.476
Pooled within groups	2	1.973

Table 20. Box's M.

Box's M		15.037
F	Approx.	2.955
	df1	4.962
	df2	8645535.028
	Sig.	.002

In Table 21, the standardised coefficients are shown. The medial-lateral diameter is the best and the anterior-posterior diameter is the worst predictor of sex in this discriminant function.

Table 21. Standardized coefficients

	Function 2
APD	.495
MLD	.828

Discriminant function coefficients

For function three, a direct analysis of the normal distributed left and right combined variables is used (Table 22). The unstandardized coefficient of “APD” is 0.279, and “MLD” is 0.546. A positive score of 0.643 would be classified as male, and a negative score of -0.655 as female. The accuracy of this function is 73.6%.

Table 22. Discriminant function coefficients. Unstandardized coefficient, centroids and accuracy.

Function	Variables	Unstandardized Coefficient	Centroids	Average accuracy %	
				Original	Cross
3	APD	.279	F= -.655 M= .643	73.6	73.6
	MLD	.546			
	Constant	-9.724			

4.6.4 Created functions

After containing all the constants and unstandardized coefficients, the functions depicted in Table 23 were created. A total of three functions were created after discriminant analysis. One function using the stepwise method, and two direct analyses. The stepwise method only used the variable circumference with a canonical coefficient of 0.258 and a constant of -10.759. Function one of the direct analyses is created using all the available left side variables with a constant of -10.652. Function three, which is a direct analysis of the left and right variables combined, has a constant of -9.724.

Table 23. Created functions. One stepwise and two direct discriminant functions.

Function	
1-Direct	$Y = -10.652 + \text{CIRC}(0.222) + \text{APD}(-.003) + \text{MDL}(.138)$
2-Stepwise	$Y = -10.759 + \text{CIRC}(0.258)$
3-Direct	$Y = -9.724 + \text{APD}(0.279) + \text{MLD}(0.546)$

5. Discussion

The goal of this thesis is to test the accuracy of the dimensions around the nutrient foramen of the fibula in estimating the sex of a Dutch archaeological population. In this chapter, the interpretation of the results is discussed. The following topics are discussed: General observations, sexual dimorphism, comparison to Fasemore et al., and future research.

5.1 General observations

The observed fragmentation of the fibulae, encountered during metric analysis, corroborates existing scholarly discourse (Novak, 2016). Frequently, the epiphyses of the bones were absent, with the proximal extremity notably absent in a majority of the examined fibulae. Consequently, this resulted in considerable challenges in obtaining accurate measurements of the "PENF" (proximal end to nutrient foramen), making it a difficult metric to ascertain reliably. Observable was the difference in location of the nutrient foramen in some cases. This could be because there may be two nutrient foramina present on the fibula (Borthakur et al., 2023, p. 263). In addition to the fragmentation of bones, the preservation state of the bones significantly impeded the sample size and precise measurement of the. In some cases, a nutrient foramen could not be found. Numerous specimens exhibited taphonomic alterations or were subject to waterlogged conditions, which considerably obstructed the identification of the nutrient foramen. Another possibility is that some of the nutrient foramina may have been overlooked because of the variation between fibulae (Kocabiyik et al., 2006). Such factors also raise concerns regarding the potential inaccuracy of the metric assessments conducted.

5.2 Sexual dimorphism

The sexual dimorphism index of this study indicated that the bones originating of biological males tend to be larger than that of females. The mediolateral measurement had the greatest difference between sexes, and the proximal end to nutrient foramen the least. Different populations exhibit different levels of sexual dimorphism and this is influenced by genetics, culture, and the environment (Charisi & Jiménez-Brobeil, 2016). The fibula is known to play a minimal role in weight-bearing, a function primarily undertaken by the tibia (Bae et al., 2006, p. 86). Given this, it is reasonable to attribute that observed sexual dimorphic differences between the sexes in different population groups to genetic variations rather than to activity patterns or environmental influences (Barrier & L'abbé, 2008, p. 85.e5). This understanding underpins the recommendation by forensic and physical anthropologists for the development of population-specific discriminant function and logistic regression equations. Consequently, there is a need to establish osteometric standards for sex estimation that utilize various dimensions from different bones of the human skeleton. The findings of Pietrobelli et al. 2022

underscore the significance of adopting a population-specific strategy for sex determination using the human fibula. They studied sexual dimorphism of the fibula on two different populations and observed that sexual dimorphism could only be identified in the Italian population. Which implies that the extent of sexual dimorphism may vary among populations with different genetic backgrounds and is not necessarily linked to size differences alone. Within the population, there also is an anatomical variation of the nutrient foramen on bones (Aggarwal et al., 2022). Aggarwal et al. (2022) studied the location, number, and the position of the nutrient foramen in the femur, tibia, and fibula. The findings suggest that 94 percent of the fibula had one single nutrient foramen, and that they were mostly found on the peroneal crest.

5.2.1 Proximal end to nutrient foramen

According to the sexual dimorphism index, the proximal end to nutrient foramen (PENF) measurement came out to be the least sexually dimorphic, with an index of 104. However, this does not mean that there is no statistical difference between males and females. The Mann-Whitney U test suggests that there is indeed a statistical significance between male and female for the PENF measurement. These results, however, did show that the difference between male and female for proximal end to the nutrient foramen was the lowest. The results of the Mann-Whitney U test that the proximal end to nutrient foramen measurement differs between sexes match the findings presented by Pietrobelli et al. (2022), which suggest that the proximal extremity of the fibula can determine sex with an accuracy ranging from 80 to 93 percent. Their study revealed that the primary distinctions between males and females involve a larger and more robust epiphysis in males compared to females. The findings may point to a pattern in knee and ankle or foot posture that is unique to a given sex, highlighting unique morpho-functional characteristics of the lower limbs associated with sex. In the literature, measurements of the proximal end, without considering the nutrient foramen, seem to be accurate in estimating the sex of the individual. The study by Kieser et al. (1992) found significant sexual dimorphism in all measurements taken of the proximal end of the tibia, where the biarticular breadth was the most discriminate. The result of this study contradicts the research of Atamtürk et al. (2010), which demonstrated that the proximal rather than the middle or distal portion of the humerus has higher diagnostic accuracy.

The examination for normality revealed that the left side measurements did not conform to a normal distribution, but the right side does. It is recognized within the literature that the likelihood of normal distribution increases with sample size literature (Krithikadatta, 2014). However, the left side has the same number of samples than the right so this cannot be a cause of non-normality in this case. A remarkable difference between left and right is the mean value. The right side has a mean around 15 cm while the mean of the left measurements

is around 17 cm. Which is right in between the mean distance of 15.91 cm of the foramina from the proximal end was for another study (Borthakur et al., 2023, p.262). The study of Gualdi-Russo (1998) shows that there is variation between left and right long bones. He argues that noticeable asymmetries in long bones are likely due to the prominent role of a left or right limb in work-related activities (Gualdi-Russo, 1998, p. 294). This could be the cause of the measurement discrepancies seen in this study. When measuring the proximal end to nutrient foramen a variation in location was observed. This variation in location can give quite some impact on the distance of the proximal end to the nutrient foramen. Which, together with the low sample size, can also be a possible reason why these measurements were not normally distributed.

5.2.2 Circumference diameter

The t-test of the circumference diameter (CIRC) around the nutrient foramen was proven to be sexually dimorphic, with a statistical significance of 0.001. This is in concordance with other studies. Safont et al. (2000) confirm the potential of long bone circumference to assist in the determination of the sex of individuals. Their research demonstrated that functions derived from a single circumference measurement yielded accuracy rates exceeding 80%. However, the authors also state that functions based on the arm circumference were found to be more effective than those based on leg circumferences. In the study of Nagaoka and Hirata (2009) the accuracy of sex classification exceeds 80% when using discriminant functions that incorporate a single variable, and can reach up to 90% with functions that combine multiple variables. They argue that functions that utilise multivariate analysis tend to yield superior outcomes compared to those that employ only univariate analysis. The sexual dimorphism index indicated that the circumference has one of the biggest differences between males and females among the measurements taken. This aligns with Sacragi et al. (1993), which demonstrated that males had a significantly larger fibula circumference compared to females. A comparable study that analysed the tibia reported a low Wilk's Lambda value of 0.365 for its circumference, also displaying the importance of the circumference (Steyn & Iscan, 1997, p. 114).

The stepwise discriminant analysis selected only the circumference measurement as a good predictor for the function. These findings are consistent with previous research which demonstrated that circumference and width measurements contribute more significantly to distinguishing between sexes than length measurements do. This could be the result of distinct bone remodelling in men and women, which causes men to build their cortical bones more during adolescence and to maintain this development into adulthood. Measurements of circumference and width are principally impacted by this uneven cortical bone growth (Slaus et al, 2013, p. 302.e3). With the stepwise analysis, using the best discriminating variable

for a function, you would suspect that the accuracy of this function is higher than that of a direct analysis where all the available variables are used. But this is not the case. The original accuracy of the direct function is 71.4% and the stepwise accuracy is 71.1%. However, this does compliment the research of Asala et al. (2004) where the combined variables were more useful than the use of variables individually. However, the accuracy percentage is still lower than for other bones like the tibia. In a Hispanic population, Safont et al. (2000) found a greater average accuracy of 91% when utilising the tibia's circumference at the level of the nutrient foramen.

5.2.3 Anterior posterior diameter

The anterior posterior diameter (APD) measurement had a statistical significance of 0.001 for the t-test between male and female, meaning that this measurement is sexually dimorphic and a statistical difference between male and female exists. The difference in normality between the left and right anterior posterior diameters was quite large, which also confirms the research of Gualdi-Russo (1998) which shows variation between the left and right long bones and suggests that the reason could be due to work-related activities. The left and right combined did however have a normal distribution, and these measurements were used for the discriminant function analysis. In the study of Iscan and Miller-Shaivitz (1984, p. 54), the anteroposterior dimension was selected for the stepwise discriminant analysis. This is contractionary with the current study, where the anteroposterior diameter was chosen to be the least discriminator (Table 9). Iscan and Miller-Shaivitz however used the tibia instead of the fibula, which may be the cause of this difference.

5.2.4 Medial lateral diameter

The highest value of the sexual dimorphism index was seen in the medial lateral diameter, which exhibits a difference of around 115%. Slaus et al (2013) also discovered that the diameter had the highest percentage difference in sex. The t-test of the medial lateral diameter (MLD) revealed a statistical significance of 0.001 between male and female, which confirms the sexual dimorphism of the measurement. The medial lateral diameter for the right fibulae was proven to be not normally distributed, while the left side was. The sample size for the left is comparable to the right-side measurements, suggesting that factors other than sample size may contribute to the lack of a normal distribution. A possible contributor to the non-normal distribution could be outliers. However, it can also be an indicator that there are significant differences between left and right. The medial lateral diameter in the discriminant function analysis of the left and right combined measures, was proved to be a better predictor than the anterior posterior diameter. According to McGeer (1990) and O'Connor and Kuo (2009), foot placement is generally thought to be passively controlled in the anteroposterior direction, but primarily regulated by active muscular control in the

mediolateral direction (Roelker et al., 2019, p. 1). This primary muscle use in the medio-lateral direction could be a possible contributor.

5.3 Comparison to Fasemore et al.

In their 2018 study, Fasemore et al. analysed 204 fibulae from two distinct African populations, South African African (SAA) and South African White (SAW). With each population consisting of approximately 50 males and 50 females. In this study, 66 females and 75 males are utilised. They developed discriminant functions separately for each population group and also for the combined populations. Their findings indicated that the discriminant functions derived from the combined populations exhibited lower accuracy percentages. Furthermore, the measurement from the proximal end to the nutrient foramen showed the lowest significance in their analysis. Fasemore et al. (2018) also noted that white South Africans display osteometric dimensions that differ from other European nationalities, likely due to admixture with other native population groups, affecting their genetic makeup. This is complemented by the differences between the two population groups. With an average classification rate of 74%, (penf b) and (circ b) together provided the best discriminant measurements for sex determination in the SAA population, while in the SAW population a single variable (circ b) with an average accuracy of 77%, gave the best results. Regarding sexual dimorphism, the study found it evident in all measurements except for the proximal end to the nutrient foramen distance. Interestingly, even though this measurement showed the least difference between males and females based on the Mann-Whitney U test in the current study, the difference was still statistically significant. The stepwise analysis of Fasemore selected a combination of (mld b) and (circ b), while in the current study, only the circumference was selected as the best predictor of sex. Both Fasemore et al. and the current thesis evaluate the reliability of their measuring techniques. Comparing the results of the intra-observer error, Fasemore et al. reported coefficients ranging from 0.981 to 0.999, while the present study had slightly lower coefficients which ranged from 0.960 to 1.00. Additionally, Fasemore et al. conducted an extra test to assess the coefficient of reproducibility. The sample size used by Fasemore et al., which included 20 bones, was comparable to the 22 bones examined in this thesis. In terms of classification accuracy, Fasemore et al. (2018) reported that sex was correctly classified with an average accuracy ranging from 69% to 77%. These percentages are comparable to his present study, which achieved an accuracy between 71% and 73%.

5.4 Limitations

It is important to note that the Shapiro-Wilk test, like other statistical tests, has limitations and may be influenced by sample size. Very large samples may result in significant p-values for very small deviations from normality, which may not be meaningful in practice. Conversely, with very small samples, the test might not have enough power to detect significant

deviations from normality. Therefore, the Shapiro-Wilk test's results should be considered in the context of the data and alongside other normality tests and visual assessments like Q-Q plots. Within each group or class, discriminant function analysis, particularly linear discriminant analysis (LDA), presumes a normal distribution of the predictor variables. When the normality assumption is violated, the efficiency and accuracy of LDA might be compromised (Rausch & Kelley, 2009). Non-normality can affect the estimation of group means and covariances, which are central to calculating the discriminant functions. Which can lead to misclassification and biased results. Additionally, a minimum of $n = 20$ is required for each variable and each set of dependent variables in order for the DFA results to be stable (Antonogeorgos et al. 2009, 3). An essential factor that should be remembered is that the discriminant function formulas for sex are population specific and cannot be applied to other populations; this is something that needs to be kept in mind.

The biggest limitation in this study was the lack of normality. Because of this, different tests (metric and non-metric) needed to be done. In the best-case scenario, one test for all the measurements is used. Because of the lack of normality, not all the variables could be considered in creating the discriminant functions. This can be seen as a limitation, since it differs from the Fasemore et al method, and therefore cannot be fully compared.

6. Conclusion

This final chapter highlights the key findings and interpretations of this study. It concludes with a brief summary of the research, the main outcomes, and recommendations for future research. It answers the main question of whether the methodology of Fasmore et al. (2018) is suitable for the archaeological collection of Middenbeemster and with what accuracy.

6.1 Summary of the research

The goal of this thesis is to test the accuracy of the Fasmore et al. method in estimating the sex of a Dutch archaeological population based on measurements of the fibula. This has been done by examining four measurements from both the left and right fibulae around the nutrient foramen of 141 individuals, using a digital sliding calliper, a tape measure, and an osteometric board. The following measurements were collected: the proximal end of the fibula to the nutrient foramen (PENF), the circumference at the nutrient foramen (CIRC), the antero-posterior diameter at the nutrient foramen (APD), and the medio-lateral diameter at the nutrient foramen (MLD). First, normality was tested on these measurements, and it was examined whether the left and right measures could be combined. To examine whether there was a significant difference between males and females, a T-test and Mann-Whitney U test were done. Discriminant function analysis was used to classify the fibula as male or female, where two direct functions, and one stepwise function were created.

6.2 Outcomes

The main goal in this thesis was to test the accuracy of the dimensions around the nutrient foramen of the fibula in estimating sex in a Dutch archaeological population. The test for the intra-observer error showed a correlation around 1.0 for all the measurements taken, suggesting that there is essentially little variation among the several metrics. The intraclass correlations were varied between 0.960 and 1.0, with "PENF" having the highest correlation and "APD" the lowest. Observations revealed some discrepancies between the left (PENF_L) and right (PENF_R) side measurements. Notably, PENF_L exhibited a normal distribution, whereas PENF_R did not. Nonetheless, statistical tests (t-test and Mann-Whitney U-test) indicated that the differences between the left and right measurements were not statistically significant, allowing for their combination in further analysis. However, after combining the left and right measurements, "PENF" and "CIRC" were not normally distributed. The results of the independent t-test and Mann-Whitney U test on the left side demonstrated that, there were statistically significant differences between males and females across all recorded metrics. The proximal end to nutrient foramen was the least statistically significant and sexually dimorphic. These findings are supported by the sexual dimorphism index, which results in males having larger dimensions than females. The significance levels from the t-

tests of the left measurements were consistent, suggesting a uniform degree of sexual dimorphism across these measurements. In this study, three discriminant functions were developed. One direct and one stepwise discriminant function utilising left-side measurements, and one direct function using combined left and right measurements of “APD” and “MLD”. The function that integrated both left and right measurements demonstrated greater accuracy (73.6%), whereas the created function of only the left side had an accuracy of 71.1% (direct) and 71.4 (stepwise). Among the metrics analysed, the circumference measurement emerged as the most effective predictor for distinguishing between male and female within the discriminant function analysis. This comprehensive approach highlights the potential for using fibular measurements in archaeological and forensic applications to reliably estimate sex.

6.3 Recommendations for future research

The main issue in this research was the lack of normality in some of the measurements. It is recognised within the literature that the likelihood of a normal distribution increases with sample size (Krithikidatta, 2014, p. 96). So, increasing the sample size might be a feasible strategy to mitigate the effects of non-normality in some cases. To employ discriminant function analysis as a statistical approach, a number of data-related assumptions must be made. One of the most significant of which is related to sample size and normal distribution. For the DFA results to be stable, there must be a minimum of $n = 20$ for each variable and each group of dependent variables (Antonogeorgos et al., 2009, p. 3). Conversely, no presumptions regarding the data's distribution are necessary when using logistic regression (Pohar et al., 2004, p. 144). For further analysis with non-normally distributed data, quadratic discriminant function analysis can be considered. Unlike linear discriminant analysis, quadratic discriminant function analysis does not assume equal covariance matrices across groups, which makes it slightly more flexible when dealing with non-normal data since it allows each group to have its own covariance matrix. Future studies should consider including a broader range of archaeological populations from different geographical locations and time periods to validate and potentially refine the developed discriminant functions. To make sure that the Fasmore et al. (2018) method can be used on Dutch populations, it would be beneficial to apply a similar approach to other Dutch populations. The results can be compared by repeating this study and putting the formulas to the test on a representative test sample. Also, testing the developed discriminant functions on different populations would help in understanding the applicability and limitations of these measurements across diverse genetic backgrounds. While sex prediction accuracy varies depending on the measurement set, it often falls between the middle to upper 80th and low 90th percentile ranges (Slaus & Tomicic, 2005, p. 148). In this case the accuracy of the created functions is lower, around 70%. For this reason, the created formulas should be used with caution.

Abstract

The estimation of sex is an essential component of osteoarchaeology. Comprehending the biological sex of individuals is crucial for piecing together their lives, interactions with environments, and identities. Sex estimation is the initial stage in determining the biological identification of a skeleton since the results form the basis for applying subsequent identifying techniques. Three methods are available for estimating sex. Specifically, the use of DNA techniques, metric analysis, or visual evaluation. The purpose of this thesis is to suggest a technique for Dutch populations to ascertain an individual's biological sex just from the dimensions of the fibula. This study aims to evaluate the sexual estimation approach in a Dutch population by reproducing the method of Fasmore et al. (2018). Four distinct measures around the nutrient foramen serve as the foundation for this research: the proximal end to the nutrient foramen, the circumference at the nutrient foramen, the medio lateral and the postero-anterior diameters around the nutrient foramen. In all measurements, there were statistically significant differences between male and female, according to the findings of the independent t-test and the Mann-Whitney U test. The proximal end to the nutrient foramen was the least sexually dimorphic, since all the measurements had a significance of 0.001, while PENF_L had a value of 0.21. The three functions created had accuracies ranging between 71.1% and 73.6%. The left and right combined measurements resulted in the most accurate discriminant function.

References

- Aggarwal, P., Roy, R., & Medda, S. (2022). Variations in nutrient foramen in the long bones of lower limb. *Indian Journal of Clinical Anatomy and Physiology*, 9(1), 35-41.
- Albanese, J., Tuck, A., Gomes, J., Cardoso, H. F. V. (2016). *Forensic Science International*, 259, 59-68. <https://doi.org/10.1016/j.forsciint.2015.12.011>
- Ambekar, S., & Sukre, S. (2016). Diaphyseal nutrient foramen of lower limb long bones: variations and importance. *International Journal of Anatomy and Research*, 2684-2688.
- Antonogeorgos, G., Panagiotakos, D. B., Priftis, K. N., & Tzonou, A. (2009). Logistic regression and linear discriminant analyses in evaluating factors associated with asthma prevalence among 10- to 12-years-old children: divergence and similarity of the two statistical methods. *International Journal of Pediatrics*, 1-6.
- Asala, S. A., Bidmos, M. A., & Dayal, M. R. (2004). Discriminant function sexing of fragmentary femur of South African blacks. *Forensic Science International*, 145(1), 25-29.
- Atamtürk, D., Akif Akçal, M., Duyar, I., Mas, N. (2010). Sex estimation from the radiographic measurements of the humerus. *Eurarian Journal of Anthropology*, 1(2), 99-108.
- Bae, D. S., Peter, M., Waters, M. D. (2006). Free Vascularized Fibula Grafting: Principles, Techniques, and Applications in Pediatric Orthopaedics, *The Orthopaedic Journal at Harvard Medical School*, 86-89.
- Barrier, I. L. O., L'Abbé, E. N. (2008). Sex determination from the radius and ulna in a modern South African sample. *Forensic Science International*, 179, 85.e1-85.e7.
- Bartholdy, B., Hoogland, M., & Water-Rist, A. (2019). How old are you now? A new aging method for nonadults based on dental wear. *International Journal of Osteoarchaeology*, 3.
- Belcastro, M. G., Pietrobelli, A., Nicolosi, T., Millela, M., & Mariotti, V. (2022). Scientific and Ethical Aspects of Identified Skeletal Series: The Case of the Documented Human Osteological Collections of the University of Bologna (Northern Italy). *Forensic science*, 2(2), 349-361.
- Borthakur, D., Kumar, R., & Dada, R. (2023). Clinical relevance of nutrient foramina of fibula: A morphological study. *Clin Ter*, 174(3), 261-265.

- Buikstra, J., & Ubelaker, D. (1994). Standards for data collection from human skeletal remains. *American Journal of Human Biology*, 272.
- Buonasera, T., Eerkens, J., De Flamingh, A., Engbring, L., Yip, J., Li, H., Haas, R., DiGiuseppe, D., Grant, D., Salemi, M., Nijmeh, C., Arellano, M., Leventhal, A., Phinney, B., Byrd, B. F., Malhi, R. S., Parker, G. (2020). A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation. *Scientific reports*, 1-15. <https://doi.org/10.1038/s41598-020-68550-w>
- Charisi, D., & Jiménez-Brobeil, S. A. (2016). Sexual dimorphism in two mediaeval Muslim populations from Spain. *Journal of Comparative Human Biology*, 397-408.
- Choudhary, C., Mishra, M., & Das, S. (2023). Sexual dimorphism of the nutrient foramen of the fibula and its importance in vascularised bone crafting: a descriptive cross-sectional study. *Indian Journal of Orthopaedics*, 1276-1282.
- Colman, K. L., Janssen, M. C., 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*, 268.e1-268.e8.
- Corieri, B., & Márquez-Grant, N. (2019). Using nutrient foramina to differentiate human from non-human long bone fragments in bioarchaeology and forensic anthropology. *Homo*, 255-268.
- Das, S. S., Choudhary, C., & Mishra, M. (2023). Sexual dimorphism of the nutrient foramen of the fibula and its importance in vascularised bone grafting: a descriptive cross-sectional study. *Indian Journal of Orthopaedics*, 1276-1282.
- Dutta, M., Mahanta, P., Baro, B., Hazarika, D. (2020). Study of nutrient foramen in fibula. *Journal of Indian Academy of Forensic Medicine*, 42(2), 137-139.
- Fasemore, M., Bidmos, M., Mokoena, P., Imam, A., Billings, B., & Mazenganya, P. (2018). Dimensions around the nutrient foramina of the tibia and fibula in the estimation of sex. *Forensic Science International*, 222.e1-222.e7.
- Gualdi-Russo, E. (1998). Study on long bones: variation in angular traits with sex, age, and laterality. *Anthropologischer Anzeiger*, 56(4), 289-299.
- Hagelberg, E., Hofreiter, M., Keyser, C. (2015). Ancient DNA: the first three decades. *Philosophical transactions of the Royal Society B*, 370(1660), 1-6. <http://dx.doi.org/10.1098/rstb.2013.0371>

- Hoogland, M. (2011). *The Middenbeemster Excavation 2011*. Retrieved from univesiteitleiden: <https://www.universiteitleiden.nl/en/research/research-projects/archaeology/the-middenbeemster-excavation-2011>
- Iscan, M. Y. (2005). Forensic anthropology of sex and body size. *Forensic Science International*, 107-112.
- Iscan, M., & Loth, S. (1997). The scope of forensic anthropology. In *Introduction to Forensic Sciences* (pp. 343-369).
- Iscan, M., & Miller-Shaivitz, P. (1984). Determination of sex from the tibia. *American Journal of Physical Anthropology*, 64, 53-57
- Iscan, M., & Steyn, M. (1999). Osteometric variation in the humerus: sexual dimorphism in South Africans. *Forensic Science International*, 77-85.
- Işcan, M.Y. and M. Steyn, 2013. *The Human Skeleton in Forensic Medicine*. 3rd edition. Springfield, Illinois (IL): Charles C. Thomas.
- Kieser, J. A., Moggi-Cecchi, J., Groeneveld, H.T. (1992). Sex allocation of skeletal material by analysis of the proximal tibia. *Forensic Science International*, 56(1), 29-36.
- Klales, A. (2020). Introduction to sex estimation and this volume. In *Sex Estimation of the Human Skeleton: History, Methods, and Emerging Techniques* (p. 31). United Kingdom: Elsevier.
- Knecht, S., Santos, F., Ardagna, Y., Alunni, V., Adalian, P., & Nogueira, L. (2023). Sex estimation from long bones: a machine learning approach. *International Journal of Legal Medicine*, 1887-1895.
- Kocabiyik, N., Yalçın, B., & Ozan, H. (2009). Variations of the nutrient artery of the fibula. *Clinical Anatomy*, 20(4), 440-443. <https://doi.org/10.1002/ca.20442>
- Krithikadatta, J. (2014). Normal distribution. *Journal of Conservative dentistry*, 96-97.
- Lemmers, S., Schats, R Hoogland, M., Waters-Rist, A. (2013). *Fysisch antropologische analyse Middenbeemster* (464). Hollandia. Retrieved on May 27 2024, https://archisarchief.cultureelerfgoed.nl/Archis2/Archeorapporten/33/AR30344/464.%20Middenbeemster%20kerkhof%20Keyserkerk%20AO%20DEF_upload.pdf
- Leskovar, T., Mlinsek, T., Pocivavsek, T., & Pajnic, I. Z. (2023). Comparison of Morphological Sex Assessment and Genetic Sex Determination on Adult and Sub-Adult 17th–19th Century Skeletal Remains. *Genes*, 14(8), 1-13.

- Maass, P., & Friedling, L. J. (2020). Assessment of sex and ancestry variation of the morphology of the fibula. *Anthropologischer Anzeiger*, 77(4), 345-354.
- Maggio, A. (2018). Interdisciplinarity and Archaeology. In C. Smith, *Encyclopedia of Global Archaeology* (pp. 1-7). Springer.
- Mays, S. (2021). DNA analysis. In S. Mays, *The archaeology of human bones* (p. 292). New York: Routledge.
- Mays, S. (2021). Sex indicators in the adult skeleton. In S. Mays, *The archaeology of human bones* (p. 55). New York: Routledge.
- McGeer, T. (1990). Passive dynamic walking. *The International Journal of Robotics Research*, 9(2), 62-82. <https://doi.org/10.1177/0278364990000900206>
- Mokoena, P., Billings, B., Bidmos, M., & Mazenganya, P. (2017). Sex estimation using dimensions around the nutrient foramen of the long bones of the arm and forearm in South Africans. *Forensic Science International*, 404.e1-404.e5.
- Moore, M. K., DiGangi, E. A., Davila, O. J., Medina, C. S., & Ruíz, F. P. (2016). metric sex estimation from the postcranial skeleton from the Colombian population. *Forensic Science International*, 286.e1-286.e8.
- Nagaoka, T., & Hirata, K. (2009). Reliability of metric determination of sex based on long-bone circumferences: perspectives from Yuigahama-minami, Japan. *Anatomical Science International*, 7-16.
- Novak, M. (2016). Sex assessment using the femur and tibia in medieval skeletal remains from Ireland: discriminant function analysis. *Collegium Antropologicum*, 17-22.
- O'Connor, S. M., Kuo, A. D. (2009). Direction-dependent control of balance during walking and standing. *Journal of Neurophysiology*, 102, 1411-1419. <https://doi.org/10.1152/jn.00131.2009>
- Outram, A. K., Knüsel, C. J., Knight, S., & Harding, A. F. (2005). Understanding complex fragmented assemblages of human and animal remains: a fully integrated approach. *Journal of Archaeological Science*, 1699-1710.
- Phenice, T. (1969). A newly developed visual method of sexing the os pubis. *American Journal of Biological Anthropology*, 297-301.
- Pietrobelli, A., Sorrentino, R., Durante, S., Marchi, D., Benazzi, S., & Belcastro, M. G. (2022). Sexual Dimorphism in the Fibular Extremities of Italians and South Africans of

- Identified Modern Human Skeletal Collections: A Geometric Morphometric Approach. *Biology*, 11(7), 1079.
- Pinheiro, J., Cunha, E., Cordeiro, C., Vieira, D. N. (2004). Bridging the gap between Forensic Anthropology and Osteoarchaeology – a case of vascular pathology. *International Journal of Osteoarchaeology*, 137-144. DOI: 10.1002/oa.718
- Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodološki Zvezki* 1(1): 143-161.
- Pyzgodna, J., Chrisler, J. C. (2000). Definitions of gender and sex: the subtleties of meaning. *Sex Roles*, 43, 553-569.
- Rausch, J. R., Kelley, K. (2009). A comparison of linear and mixture models for discriminant analysis under nonnormality. *Behavior Research Methods*, 41(1), 85-98.
doi:10.3758/BRM.41.1.85
- Roelker, S. A., Kautz, S. A., Neptune, R. R. (2019). Muscle contributions to mediolateral and anteroposterior foot placement during walking. *Journal of Biomechanics*, 109310, 1-7. <https://doi.org/10.1016/j.jbiomech.2019.08.004>
- Rösing, F., Graw, M., Marré, B., Ritz-Timme, S., Rotschild, M., Röttscher, K., . . . Geserick, G. (2007). Recommendations for the forensic diagnosis of sex and age from skeletons. *Journal of Comparative Human Biology*, 58(1), 75-89.
- Sacragi, A., & Ikeda, T. (1995). Sex identification from the distal fibula. *International Journal of Osteoarchaeology*, 139-143.
- Sacragi, A., Ikeda, T., & Terada, H. (1993). Fibulo-tibial weight index - a new criterion for sex identification based on the lower leg bones. *Acta Anatomica*, 193-196.
- Safont, S., Malgosa, A., & Subira, E. (2000). Sex assesment on the basis of long bone circumference. *American Journal of Physical Anthropology*, 317-328.
- Slaus, M., Bedic, Z., Strinovic, D., Petrovecki, V. (2013). Sex determination by discriminant function analysis of the tibia for contemporary Croats. *Forensic Science International*, 226, 302.e1-302.e4.
- Slaus, M., & Tomicic, Z. (2005). Discriminant function sexing of fragmentary and complete tibiae from medieval Croatian sites. *Forensic Science International*, 147, 147-152.
<https://doi.org/10.1016/j.forsciint.2004.09.073>
- Spradley, M. K. (2016). Metric methods for the biological profile. *Academic Forensic Pathology*, 6(3), 391-399.

- Stamfelj, I. (2021). Sex estimation based on the analysis of enamel peptides: False assignments due to AMELY deletion. *Journal of archaeological science*, 130, 1-4. <https://doi.org/10.1016/j.jas.2021.105345>
- Stewart, T. (1979). *Essentials of Forensic Anthropology*. Springfield: Charles C Thomas.
- Stewart, N. A., Gerlach, R. F., Gowland, R. L., Montgomery, J. (2017). Sexes determination of human remains from peptides in tooth enamel. *Anthropology*, 114(2), 13649-13654. <https://doi.org/10.1073/pnas.1714926115>
- Steyn, M., & Iscan, M. Y. (1997). Sex determination from the femur and tibia in South African Whites. *Forensic Science International*, 111-119.
- White, T. D., & Folkens, P. A. (2005). Determination of sex. In T. D. White, & P. A. Folkens, *The human bone manual* (p. 385). Oxford: Elsevier Academic Press.
- Wadsworth, C., Procopio, N., Anderung, C., Carretero, J., Iriarte, E., Valdiosera, C., Elburg, R., Penkman, K., Buckley, M. (2017). Comparing ancient DNA survival and proteome content in 69 archaeological cattle tooth and bone samples from multiple European sites. *Journal of Proteomics*, 158, 1-8.
- Welker, F., Ramos-Madrugal, J., Kuhlwilm, M., Liao, W., Gutenbrunner, P., de Manuel, M., Samodova, D., Mackie, M., Allentoft, M. E., Bacon, A., Collins, M. J., Cox, J., Lalueña-Fox, C., Olsen, J. V., Demeter, F., Wang, W., Margues-Bonet, T., Cappelini, E. Enamel proteome shows that *Gigantopithecus* was an early diverging pongine. *Nature* 576, 262–265 (2019). <https://doi.org/10.1038/s41586-019-1728-8>
- Workshop of European Anthropologists (WEA). (1980). Recommendations for age and sex diagnosis of skeletons. *Journal of Human Evolution*, 9, 517-549.

Images

Barthody, B. P. (2019). Map of The Netherlands showing the location of Middenbeemster [Photograph]. Researchgate. https://www.researchgate.net/figure/Map-of-The-Netherlands-showing-the-location-of-Middenbeemster_fig7_332648472

Buikstra and Ubelaker (1994). Standards for scoring cranial traits [Photograph]. Scirp. <https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1446825>

RegisteredNurse (n.d.) Tibia and Fibula bone anatomy [Photograph]. RegisteredNursern. <https://www.registerednursern.com/tibia-and-fibula-anatomy/>

Phenice, T. (1969). A Newly developed visual method of sexing the os pubis [Photograph]. Wiley. <https://onlinelibrary.wiley.com/doi/abs/10.1002/ajpa.1330300214>

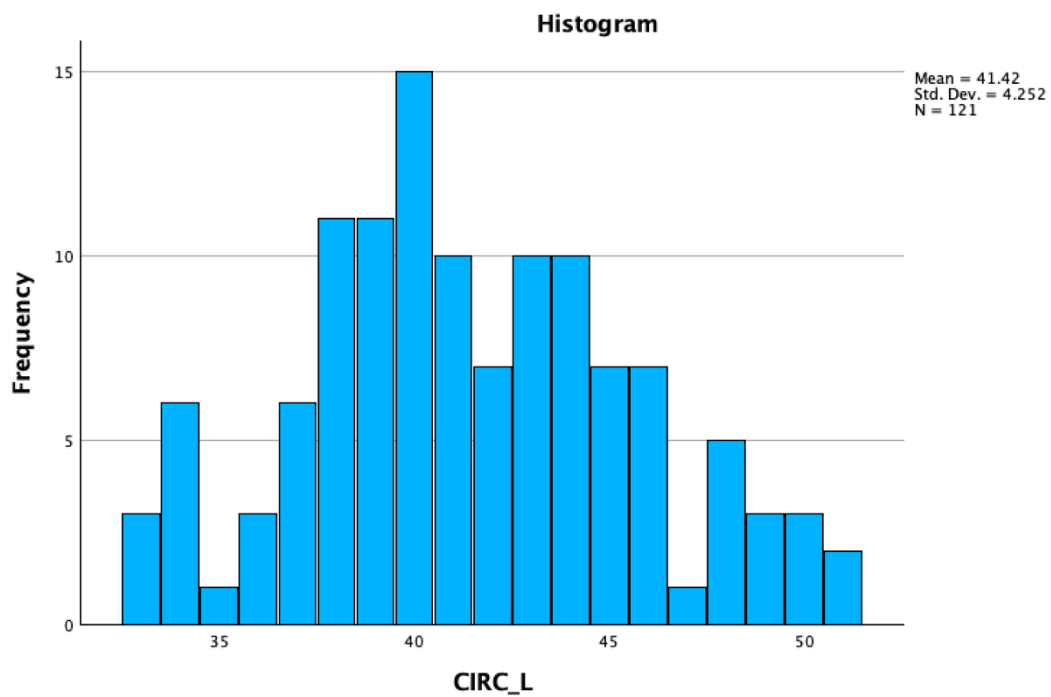
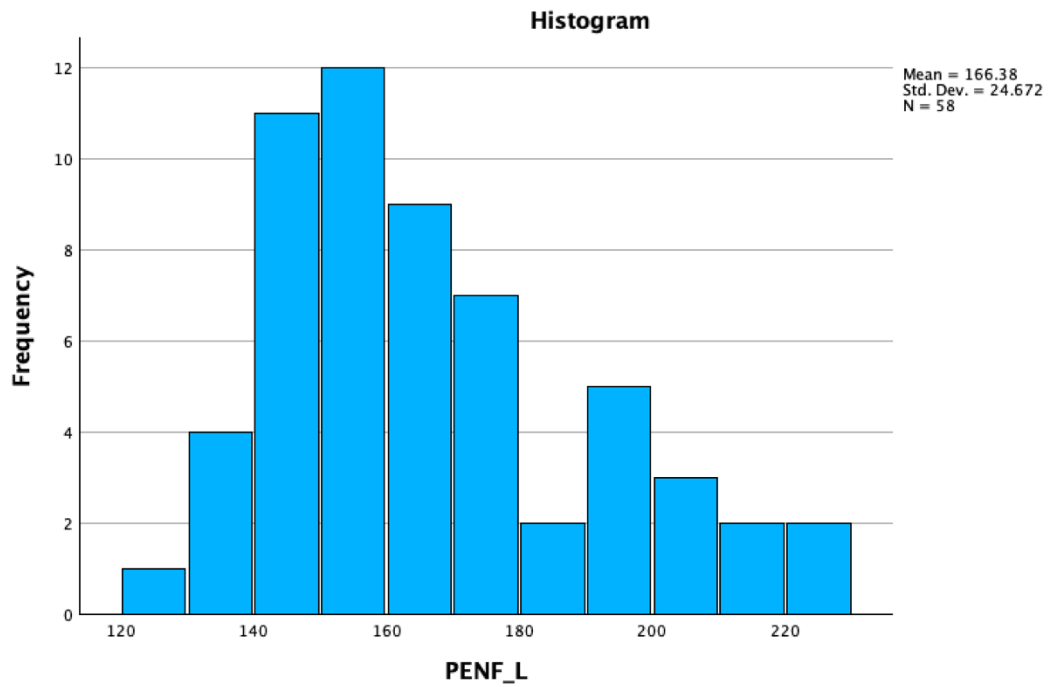
List of Figures

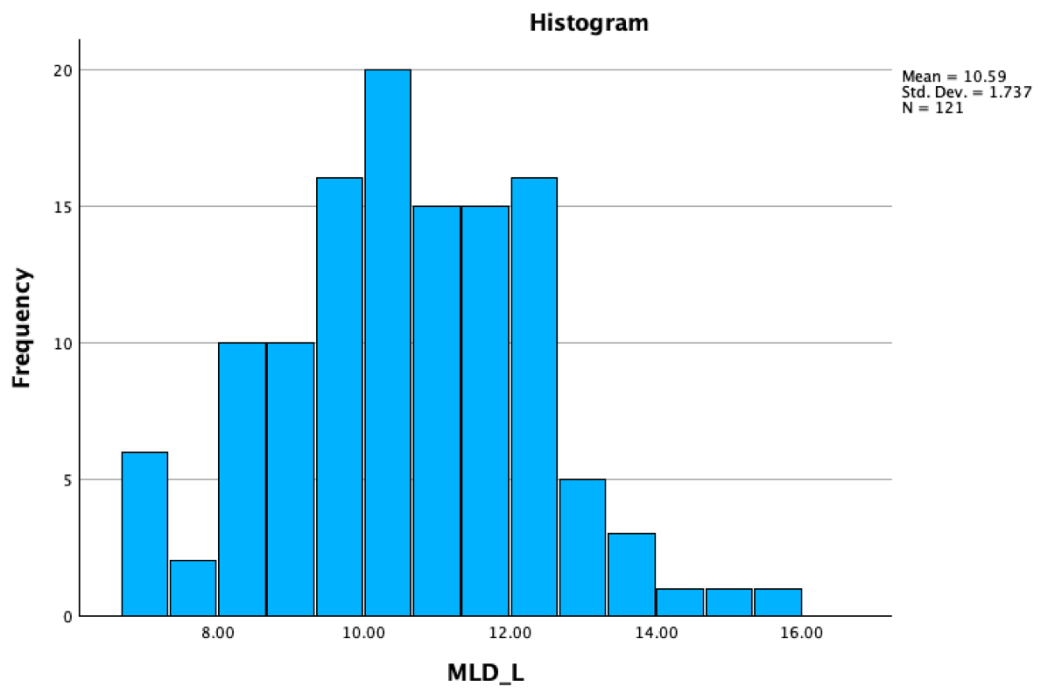
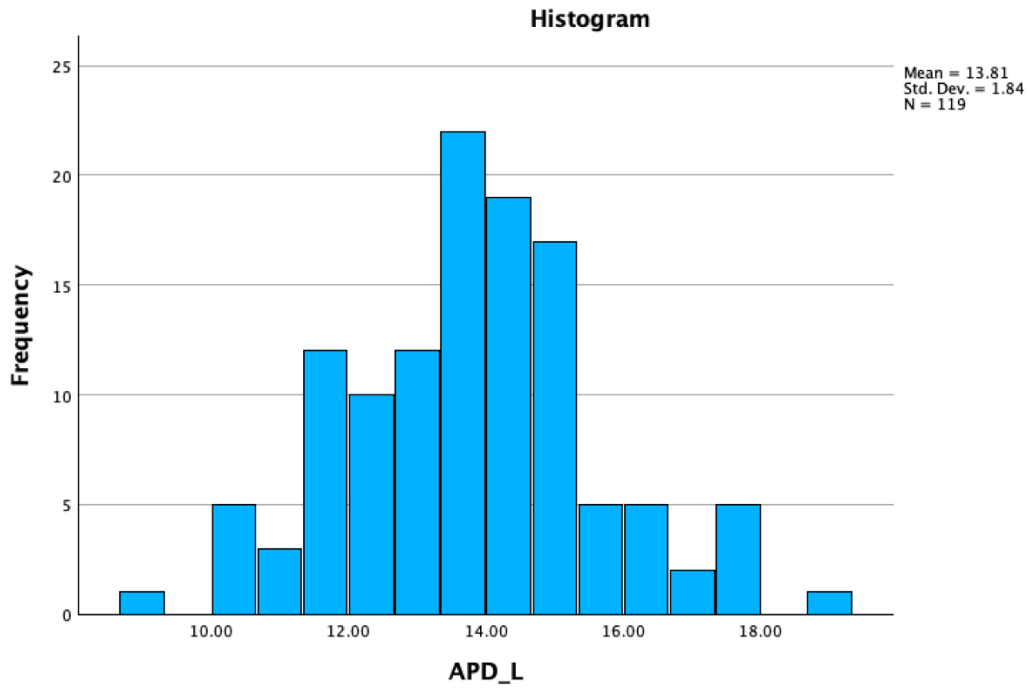
Figure 1: Phenice traits.	12
Figure 2: Traits of the Buikstra and Ubelaker method.	13
Figure 3: Tibia and fibula.	15
Figure 4: The location of Middenbeemster on a map of the Netherlands.	17
Figure 5: PENF measurement.	19
Figure 6: CIRC measurement.	19
Figure 7: APD measurement.	20
Figure 8: MLD measurement.	20
Figure 9: Histogram of PENF_L.	26
Figure 10: Mahalanobis distance Q-Q Plot of PENF_L.	26

List of Tables

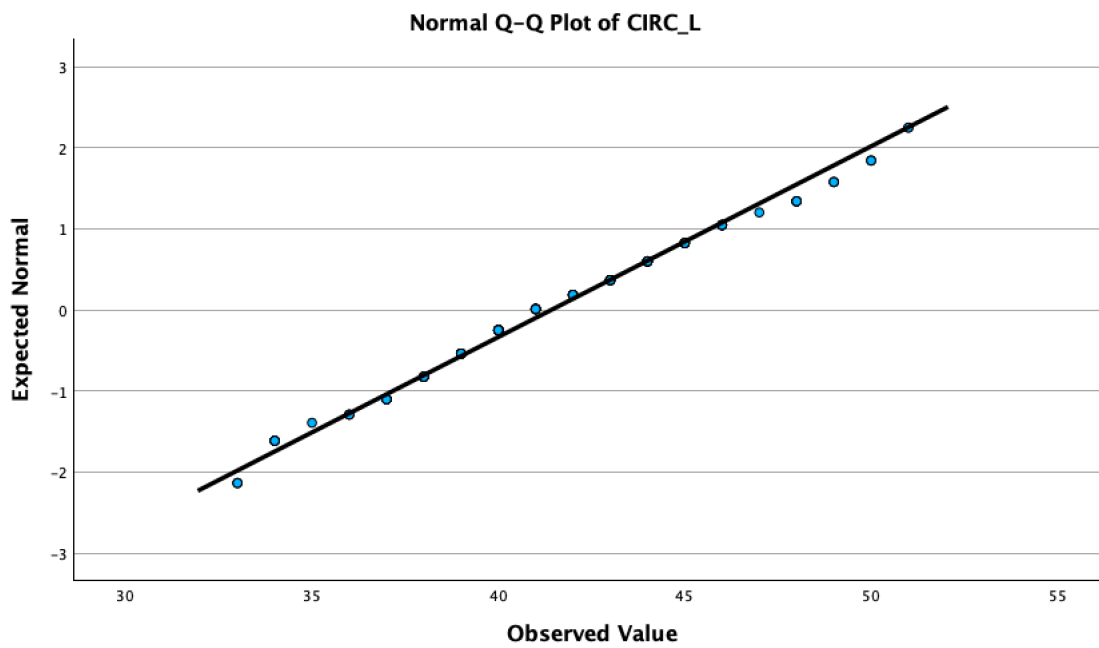
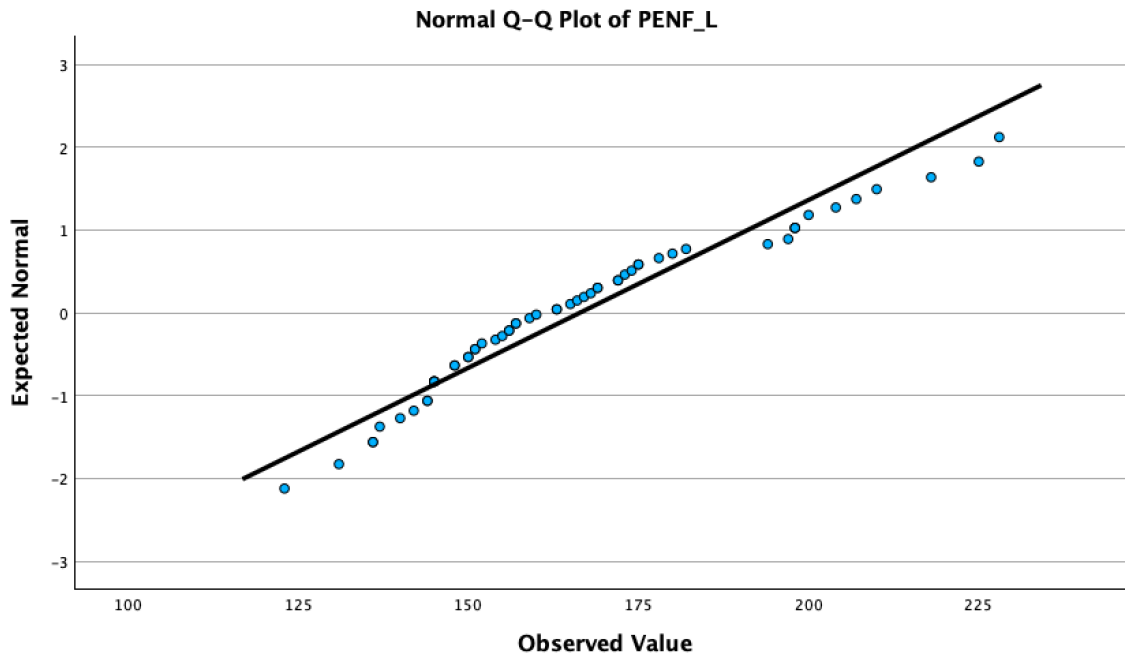
Table 1: Intra Class Correlation Coefficient.	24
Table 2: Descriptive statistics.	25
Table 3: Results of the Shapiro-Wilk test.	26
Table 4: Results of the T-test L vs R.	27
Table 5: Mann-Whitney U test L vs R.	27
Table 6: Normality left and right combined.	28
Table 7: Results of the T-test Male vs Female.	28
Table 8: Mann-Whitney U test Male vs Female.	29
Table 9: Tests of Equality of Group Means.	29
Table 10: Pooled Within-Groups Correlation Matrices.	30
Table 11: Log determinant.	30
Table 12: Box's M.	30
Table 13: Standardized coefficients.	30
Table 14: Discriminant function coefficients.	31
Table 15: Standardized coefficients.	31
Table 16: Log determinant.	31
Table 17: Box's M.	31
Table 18: Discriminant function coefficients.	32
Table 29: Log determinant.	32
Table 20: Box's M.	32
Table 21: Standardized coefficients.	32
Table 22: Discriminant function coefficients.	33
Table 23: Created functions.	33

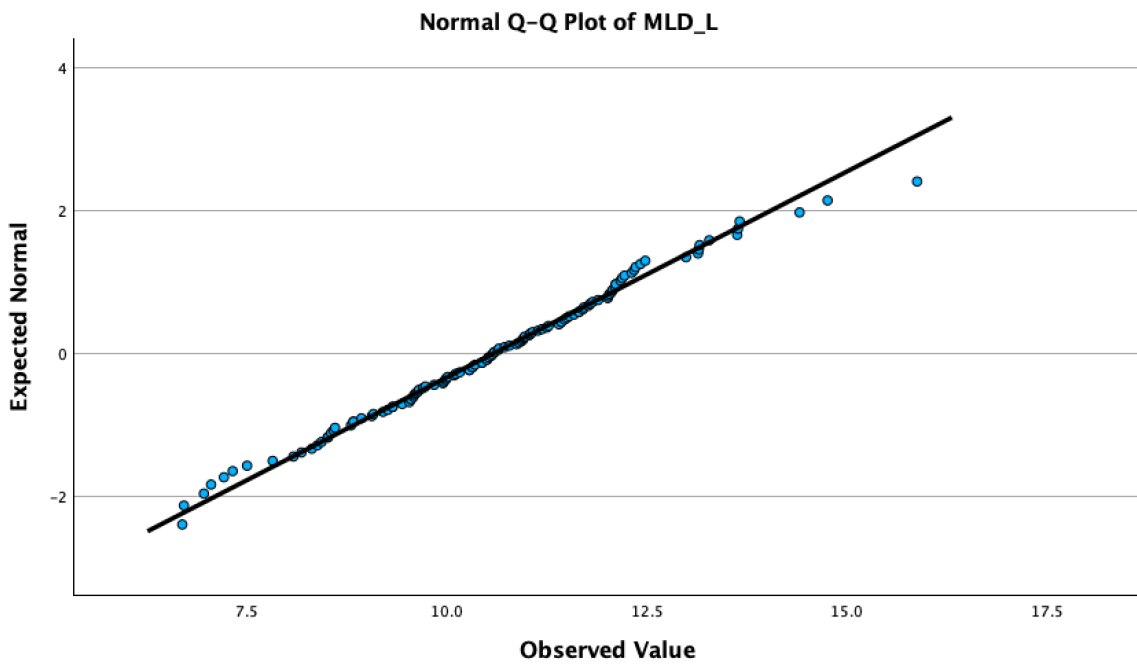
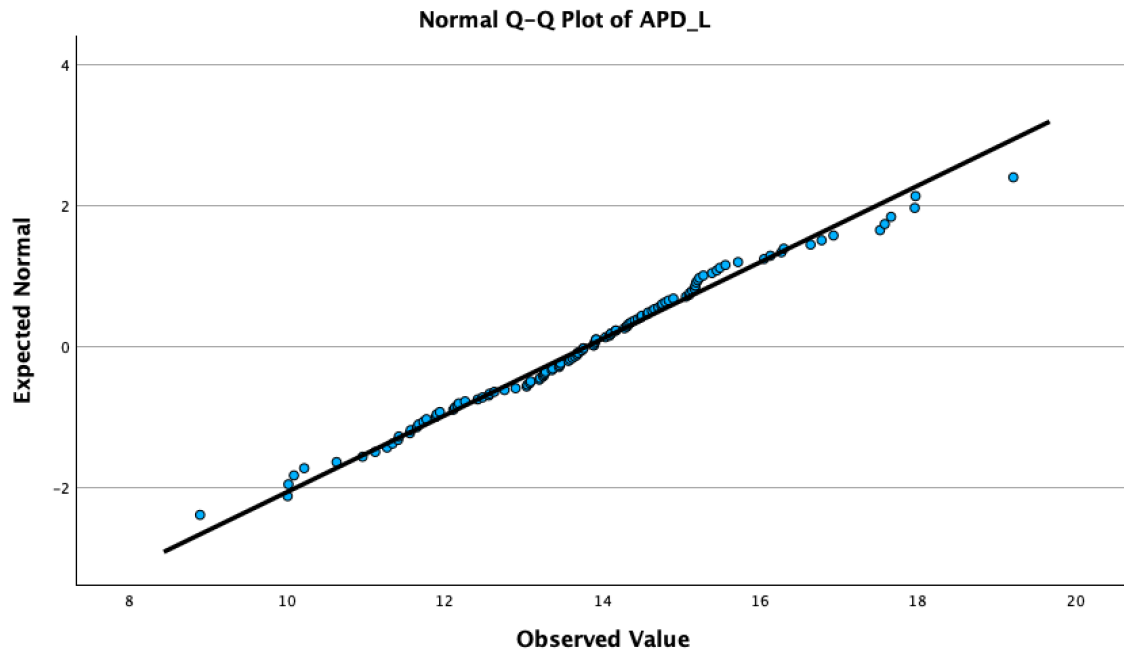
Appendix I: Normality figures





Appendix II: Q-Q Plots





Appendix III: Raw data

	D	E	F	G	H	I	J	K	L
1		Left	in mm			Right	in mm		
2	Sex	PENF_L	CIRC_L	APD_L	MLD_L	PENF_R	CIRC_R	APD_R	MLD_R
3	M	X	X	X	X	202	41	12,58	10,95
4	F		159	34	12,13	9,08	X	X	X
5	PF		160	38	13,57	8,58	156	42	14,25
6	F	X		41	14,17	9,84	X	42	14,31
7	F	X		34	11,89	7,21	X	34	12,79
8	PM		178	43	12,63	11,71	X	X	X
9	M	X		40	13,69	8,51	X	41	9,83
10	F	X		33	10,63	7,82	X	35	12,19
11	F	X		48	13,89	11,65	X	48	14,14
12	F	X		33	11,94	6,96	X	X	X
13	F	X		40	14,17	8,6	145	39	14,27
14	M	X		43	11,27	11,47	X	45	12,01
15	F	X		37	12,76	9,32	X	39	13,67
16	M	X	X	X	X	151	47	14,66	11,73
17	PF	X		40	13,28	11,25	X	40	12,95
18		X		40	14,41	10,78	X	41	15,39
19	F	X		38	12,11	10,01	X	39	12,8
20	M	X		37	12,16	9,98	X	38	12,12
21	F	X		36	12,42	8,51	143	36	12,37
22	F	X		39	12,18	11,04	168	36	11,05
23	M	X		40	11,9	11,4	NF	NF	NF
24	M	X	X	X	X	147	42	11,58	12,08
25	F	X		40	13,45	9,2	NF	NF	NF
26	M		145	46	15,39	12,17	X	54	17,27
27	F	X		43	13,92	10,35	X	40	13,06
28	PF		140	45	14,66	10,55	150	43	14,16
29	M		175	44	13,08	11,59	188	44	15,44
30	F		157	33	13,04	6,69	NF	NF	NF
31	M		198	45	13,92	13,64	163	43	11,79
32	PM		172	50	19,21	10,87	A	A	A
33	PF	X		43	15,18	11,65	X	NF	NF
34	PF		148	34	10,96	8,08	NF	NF	NF
35	PM		198	43	14,37	11,19	X	43	13,31
36	F	X		41	13,36	12,02	X	38	12,72
37	F		151	38	11,73	8,93	156	40	13,41
38	PF	X		36	11,41	9,55	X	X	X
39	M	X		45	13,90	13,14	164	44	15,3
40	PM	NF	NF	NF	NF	X		44	16,05
41	M	NF	NF	NF	NF		190	51	18,65
42	PF	X	X	X	X	X		39	14,85
43	PM	X		46	14,75	12,19	X	X	X
44	M		175	52	17,52	13,66	164	50	18,88
45	M		131	41	14,32	9,96	162	43	14,81
46	M	NF	NF	NF	NF		154	41	14,67
47	F	X		38	13,89	8,43	NF	NF	NF
48	M	X		41	13,47	11,43	X	X	X
49	M		200	48	15,17	12,42	214	45	14,9
50	M		169	43	14,9	10,11	NF	NF	NF
51	M	X		47	16,13	10,59	X	44	14,78

	D	E	F	G	H	I	J	K	L
52	PM	155	39	13,62	10,44	A	A	A	A
53	M	198	41	14,31	9,73	NF	NF	NF	NF
54	PM	X	49	16,27	12,11	X	45	15,96	10,35
55	M	X	X	X	X	X	43	15,05	8,79
56	PM	X	38	13,59	9,61	A	A	A	A
57	M	A	A	A	A	134	37	11,23	9,90
58	F	X	37	12,57	8,83	X	38	14,49	8,79
59	M	173	44	14,76	10,44	194	46	15,82	11,75
60	F	151	42	14,71	10,09	X	43	15,05	9,91
61	M	X	NF	NF	NF	X	49	17,68	10,67
62	M	156	49	14,28	9,32	NF	NF	NF	NF
63	M	X	X	X	X	X	49	15,38	13,8
64	F	X	X	X	X	152	42	14,73	8,41
65	PM	X	42	15,21	8,55	X	41	15,72	8,85
66	M	166	45	16,3	10,57	155	45	15,67	9,45
67	M	169	52	17,66	12,31	X	52	17,74	1,62
68	M	145	53	17,58	12,34	188	54	17,84	1,47
69	F	150	37	11,34	8,38	146	39	13,13	9,94
70	F	X	X	X	X	135	42	13,78	10,74
71	PF	123	37	11,42	10,32	NF	NF	NF	NF
72	F	X	40	14,45	9,26	X	36	12,62	9,74
73	PF	145	40	14,57	9,53	136	40	13,94	9,43
74	F	X	40	13,25	7,32	X	40	14,31	8,27
75	M	163	44	15,45	11,53	X	42	13,96	11,03
76	M	NF	NF	NF	NF	151	50	17,36	12,76
77	F	150	39	14,09	8,18	156	41	13,75	8,78
78		137	36	10,01	10,5	NF	NF	NF	NF
79	M	NF	NF	NF	NF	153	45	17,22	9,60
80	M	136	48	15,49	12,06	X	53	15,29	13,75
81	M	204	40	13,74	12,01	149	44	16,59	10,53
82	M	218	43	14,49	12,48	NF	NF	NF	NF
83	M	Nf	NF	NF	NF	166	45	15,67	13,95
84	M	A	A	A	A	161	39	11,58	10,88
85	F	174	44	14,58	12,1	X	46	16	13,18
86	PF	X	49	16,64	11,82	133	43	14,36	11,16
87	F	210	41	15,06	10,93	189	43	15,15	10,97
88	F	X	34	10,22	10,28	X	34	11,41	9,13
89	M	X	41	13,91	10,96	X	42	13,10	10,25
90	I	154	40	14,34	11,07	NF	NF	NF	NF
91	F	144	35	13,09	10,64	NF	NF	NF	NF
92	M	165	44	14,04	12,11	180	46	14,56	15,18
93	M	NF	NF	NF	NF	X	37	12,2	9,67
94	M	X	41	12,56	11,79	A	A	A	A
95	PF	X	46	11,56	9,99	NF	NF	NF	NF
96	M	X	44	14,11	10,97	X	X	X	X
97		X	38	13,37	9,55	X	37	12,04	9,56
98	F	X	34	12,48	6,71	X	35	12,11	7,93
99	M	X	40	13,27	8,8	158	44	16,02	9,79
100	M	X	55	17,97	14,76	X	51	17,69	14,17
101	F	X	39	11,65	11,5	X	39	13,29	10,06
102	M	X	45	16,93	11,03	X	48	18,22	10,27
103	F	142	39	13,26	9,06	158	38	13,13	10,09

	D	E	F	G	H	I	J	K	L	
104	M	X		37	8,9	13,16	NF	NF	NF	NF
105	F		182	38	13,69	9,65	154	40	14,1	11,84
106	M		168	46	16,05	13,15	X	47	16,64	11,46
107	F	X	X	X	X	X		40	14,87	8,72
108	PF		145	40	10,09	13,28	133	39	11,9	10,13
109	F	X		34	11,77	7,5	X	35	11,35	8,57
110		X		45	15,19	12,99	X	41	12,31	10,43
111	F		228	38	14,5	7,05	153	40	15,71	7,43
112	F		136	39	13,65	8,83	140	38	14,1	8,47
113		NF		NF	NF	NF	X	37	14,38	8,21
114	PM	X		42	X	10,52	X	38	14,36	9,84
115	PM		157	43	14,63	10,32	X	44	14	11,17
116	F	NF		NF	NF	NF	X	51	18,66	10,67
117	M		172	45	13,46	12,07	NF	NF	NF	NF
118	F	NF		NF	NF	NF	119	37	12,32	9,47
119	M	X		46	10,02	17,88	X	43	14,09	11,66
120	M	X		48	17,96	9,44	X	48	18,4	9,51
121	PF	X		42	14,8	9,59	X	44	15,18	10,58
122	M	X		44	15,18	11,14	NF	NF	NF	NF
123	F		225	38	12,9	9,95	180	39	14,48	9,56
124	M		163	46	16,78	1,43	187	50	16,06	12,96
125	M		152	40	X	1,72	X	46	13,57	14,15
126	F	X		39	13,2	9,64	X	X	X	X
127	PF	X		44	13,21	12,36	NF	NF	NF	NF
128	F		197	39	13,05	11,70	X	39	12,48	11,60
129	PF	Nf		NF	NF	NF	173	42	13,04	11,94
130	PM	A		A	A	A	X	43	13,83	12,25
131	M	X		39	11,12	10,95	X	41	13,47	10,21
132	F		156	44	15,11	10,9	A	A	A	A
133	M	X		46	14,84	14,41	18	45	15,68	12,26
134	F		194	42	15,14	11,27	136	38	12,79	8,37
135	F		144	38	11,57	10,65	135	38	13,11	9,58
136	M		167	43	15,10	11,89	177	44	15,43	11,62
137	F	X		39	13,68	10,28	NF	NF	NF	NF
138		X		38	13,47	10,59	169	42	12,01	11,72
139	M		180	43	13,76	12,03	X	X	X	X
140	F		145	41	15,72	10,16	139	41	16,15	9,08
141	M		207	48	15,56	13,63	X	50	15,06	14,14
142	M	X		40	13,76	9,7	X	40	12,44	10,79
143	PM	X		X	X	X	170	37	13,31	X
144	M	X		39	11,67	10,5	X	41	12,51	11,97
145	PM	NF		NF	NF	NF	X	43	14,49	10,90
146	M	X		41	15,23	9,58	NF	NF	NF	NF
147	F		148	44	14,11	11,77	X	41	12,91	10,95
148	F	X		42	12,26	12,22	126	40	12,48	10,02
149	PF	X		42	15,28	8,31	X	42	14,96	9,17
150	F	A		A	A	A	158	38	12,12	10,52

Appendix IV: Intra observer error

Sex	Left				Right			
	PENF_L	CIRC_L	APD_L	MLD_L	PENF_R	CIRC_R	APD_R	MLD_R
M	X	X	X	X	203	40	12,8	10,62
F	160	34	11,68	9,14	X	X	X	X
PF	160	37	13,66	8,52	156	41	14,25	10,12
F	X	41	14,26	9,87	X	41	14,74	9,9
F	X	33	11,82	7,23	X	34	12,82	6,34
PM	X	X	X	X	X	X	X	X
M	X	39	13,81	8,46	X	41	9,86	13,53
F	X	32	10,09	7,72	X	35	12,06	8,48
F	X	45	14,54	11,58	X	47	14,73	13,11
F	X	32	12,18	6,93	X	X	X	X
F	X	40	14,18	8,58	146	38	14,26	8,65
M	X	43	11,26	11,50	X	45	14,27	12,78
F	X	37	12,76	9,32	X	39	13,60	9,18
M	X	X	X	X	151	47	14,62	11,72
M	X	X	X	X	202	41	12,58	10,95
F	159	34	12,13	9,08	X	X	X	X
PF	160	38	13,57	8,58	156	42	14,25	9,73
F	X	41	14,17	9,84	X	42	14,31	9,9
F	X	34	11,89	7,21	X	34	12,79	6,46
PM	178	43	12,63	11,71	X	X	X	X
M	X	40	13,69	8,51	X	41	9,83	13,2
F	X	33	10,63	7,82	X	35	12,19	8,51
F	X	48	13,89	11,65	X	48	14,14	13,14
F	X	33	11,94	6,96	X	X	X	X
F	X	40	14,17	8,6	145	39	14,27	8,59
M	X	43	11,27	11,47	X	45	12,01	12,85
F	X	37	12,76	9,32	X	39	13,67	9,27
M	X	X	X	X	151	47	14,66	11,73