Making the Invisible Visible

Test of an osteological population-specific non-adult sexing approach using permanent odontometrics on a post-medieval Dutch skeletal collection

Veronica Jackson

Crania of Middenbeemster (MB11) individuals S167V0270 (left) and S396V0877 (right) (photograph taken by the author).

Making the invisible visible:
Test of an osteological population-specific non-adult sexing approach using permanent odontometrics on a post-medieval Dutch skeletal collection

Veronica Jackson S2075261

MSc Thesis Archaeology 4ARX-0910ARCH

Dr. S.A. Schrader and Dr. C.L. Burrell

Human Osteology and Funerary Archaeology

University of Leiden, Faculty of Archaeology

Final Version

Leiden 15-06-2019

Table of Contents

Acknowledgements

I would like to thank Dr. Carla Burrell for her incredibly helpful and prompt feedback, editing and advice throughout the writing process. I would like to thank Dr. Sarah Schrader and Bjørn Peare Bartholdy for encouraging me to pursue this topic and providing the practical foundations on which this study was based. Special thanks goes to Bjørn for his invaluable guidance through the intimidating world of statistical analyses as well as his thorough and detailed editing. Thank-you again to Dr. Schrader for editing the first draught of my thesis. Thanks to Dr. Menno Hoogland for permitting the use of the Middenbeemster collection. And finally, a very special thank-you to my perennial editor and cornerstone, Jill Jackson, without whom I would not have had the opportunity to conduct this research, let alone see it through to completion.

1. Introduction

1.1 Non-adult sex estimation within osteoarchaeology

Sex estimation of non-adult skeletal remains has long been regarded as a problematic or even an unattainable objective within physical anthropology and forensic science (Aris *et al.* 2018, 672; Baker *et al.* 2005, 3, 10; Black 1978, 77; Cardoso 2008, 158; Cunningham *et al.* 2016, 17-18; Klales and Burns 2017, 747; Saunders 2008, 117; Wilson and Humphrey 2017, 33). In adult skeletal remains, sex estimation is largely based on the differences in morphology between males and females in the pelvis and skull (Olivares and Aguilera 2016, 1623; Stull *et al.* 2017, 64; Wilson *et al.* 2016, 255; Wilson and Humphrey 2017, 34). This systematic difference in shape and size between males and females of the same species is referred to as sexual dimorphism; the pelvis and skull are universally considered to be the most sexually dimorphic aspects of the human adult skeleton (Buikstra and Ubelaker 1994, 16; Schutkowski 1993, 200; Shankar *et al.* 2013, 753). Consequently, consistently high rates of correct sex allocation can be obtained for adult skeletal remains using widely-accepted and well-established morphological methods. Therefore, nascent methodologies in this field are generally expected to have accuracy rates of at least 85% (Klales and Burns 2017, 750).

Unfortunately, the same sexually dimorphic features that allow such accurate classification in adult remains do not develop until puberty and so cannot be reliably analysed in the remains of individuals under the age of about 10 years in girls and 12 years in boys (Aris *et al.* 2018, 672; Baker *et al.* 2005, 10; Cardoso 2008, 159; Cunningham *et al.* 2016, 17; Hassett 2011, 486; Olivares and Aguilera 2016, 1623; Shankar *et al.* 2013, 753; Stull *et al.* 2017, 64; Viciano *et al.* 2011, 97; Wilson and Humphrey 2017, 34). In recent decades, many methods have been developed that attempt to identify and categorise early indicators of sexual dimorphism in the non-adult pelvis and skull (e.g. Klales and Burns 2017, Luna *et al.* 2017, Molleson *et al.* 1998, Schutkowski 1993, Weaver 1980). However, few of these methods were able to match the accuracy rates of methods designed for adult remains and those that did were found to have significantly lower accuracy rates when tested on a population other than the one on which the method was originally developed (Baker *et al.* 2005, 10; Cardoso 2008, 159; Lewis 2006, 50-54; Olivares and Aguilera 2016, 1624; Wilson and Humphrey 2017, 34). In addition, several studies have found that the reliability of many of the traits used in morphological methods varies with age and so may be more closely linked with growth than with sex (Cardoso 2008, 159; Cardoso and Saunders 2008, 28; Lewis 2006, 51; Wilson *et al.* 2016, 263-264; Wilson and Humphrey 2017, 36). It has been proposed an accuracy rate of 75% should be considered acceptable for the estimation of sex in non-adults in acknowledgement of the immense challenge posed by the variability of and overlap in the expression of sexually dimorphic traits in pre-pubescent individuals (Klales and Burns 2017, 750; Olivares and Aguilera 2016, 1623). Even so, few advances have been made towards the establishment of a dependable means of sexing non-adults with an accuracy rate better than 75% (Klales and Burns 2017, 750). As such, the Scientific Working Group for Forensic Anthropology (SWGFA) lists "Sub-Adult Sex Assessment" on any individual less than 12 years of age as an unacceptable practice due to its inconsistency (SWGFA 2010, 3).

A definitive and reliable technique to estimate sex in non-adult osteological remains would contribute greatly to the field of osteoarchaeology and to archaeology as a whole. Not only would accurate sex estimates allow for the refinement of current osteological procedures such as age estimation and growth studies, but it would also allow for more perceptive interpretations of the social, economic, or environmental implications of osteological evidence including palaeodemographic studies that could illuminate differential morbidity and mortality rates (Baker *et al.* 2005, 4; Cardoso 2008, 158, 167; Cunningham *et al.* 2016, 5; Luna *et al.* 2017, 898; Viciano *et al.* 2011, 105). Accurate sexing is needed in order to elucidate gendered patterns of behaviour and experience as related to biological sex and to properly contextualise an individual historically (Aris *et al.* 2018, 672; Lewis 2006, 187; Luna *et al.* 2017, 898; Olivares and Aguilera 2016, 1623). This can relate to treatment and care of younger individuals, the impact of disease, access to resources, and gendered activity patterns or division of labour (Baker *et al.* 2005, 4-5; Cardoso 2008, 158-159, 167; Lewis 2006, 185; Saunders 2008, 122; Wilson *et al.* 2016, 263). These can be observed through signs of nutritional stress and growth retardation, and through the distribution of activity markers and pathological conditions including trauma (Baker *et al.* 2005, 4-5; Cardoso 2008, 159, 167; Lewis 2006, 187; Viciano *et al.* 2011, 105). Differences in funerary practices between girls and boys could be examined as well, such as the age at which they were given the same burial rites as adults (Baker *et al.* 2005, 4; Lewis 2006, 186; Saunders 2008, 118).

Women and children are often excluded from written sources and are therefore essentially invisible in the historical record. The ability to draw all possible information from non-adult osteoarchaeological remains is therefore essential to a more comprehensive understanding of the past lives of children and of women, whose lives were habitually closely intertwined with those of children.

1.2 Sex, growth and development of the non-adult skeleton

One of the many problems caused by the inability to reliably estimate sex in non-adult skeletal remains is that growth itself is sexually dimorphic (Cunningham *et al.* 2016, 17; Hillson 2005, 210). This causes difficulties in the extrapolation of chronological age from biological age. Biological age is based on biological changes in the skeleton related to growth and development and subsequently to degeneration (Uhl 2013, 63). Chronological age, also known as calendar age, is a representation of the length of an individual's life after birth in calendric terms, which advances at a uniform rate from birth to death (Uhl 2013, 63). Though the two are related, chronological age is a social construct universally applied to individuals of varying physical development, whereas biological age can be greatly affected by extrinsic factors such as nutrition, diet, disease exposure, and mechanical stress (Uhl 2013, 63). Girls are about 10% more developed physically than their male peers from 20 weeks *in utero* to puberty, meaning that their biological development is more advanced in relation to chronological age (Lewis 2006, 48; Saunders 2008, 123). Without knowledge of the sex of the individual, the potential chronological age range of any one stage of development is wider than necessary in order to encompass the associated chronological ages of both sexes (Baker *et al.* 2005, 3; Cardoso 2008, 159; Cunningham *et al.* 2016, 17; Lewis 2006, 48, 186; Olivares and Aguilera 2016, 1623; Saunders 2008, 122, 125). Despite this, non-adult age-at-death estimates are much more precise than the wide age categories routinely used in adult age-at-death estimation (Baker *et al.* 2005, 3; Olivares and Aguilera 2016, 1623; Saunders 2008, 117; Uhl 2013, 63). Adult ageing methods rely on observations of the systematic degeneration of the body that begins once full maturation has been attained. In contrast to this, non-adult ageing methods rely on the progress of rapid, genetically determined changes that occur in the non-adult skeleton during growth and development (Lewis 2006, 184). As these changes occur at different times in males and females, accurate sex estimations would also allow for the refinement of growth patterning studies and the onset of puberty within various populations (Lewis 2006, 186; Saunders 2008, 134). Age estimation and growth studies have so dominated the field of juvenile osteology that most texts dedicated to non-adult osteology ignore the issue of sex estimation entirely and concentrate instead on element identification and ageing techniques (Baker *et al.* 2005, 5, 10).

1.3 Non-adult skeletal collections

1.3.1 Retrieval and preservation

A major obstacle to the development and testing of any method designed for use on non-adult remains is the paucity of large, well-preserved skeletal samples of known sex and age-at-death (Baker *et al.* 2005, 3; Cardoso and Saunders 2008, 29; Cunningham *et al.* 2016, 15, 18; Klales and Burns 2017, 747; Olivares and Aguilera 2016, 1624; Schutkowski 1993, 204; Wilson and Humphrey 2017, 34). This is in part due to the susceptibility of immature skeletal remains to taphonomic processes, but over time the problem has been exacerbated by negligence during excavation (Baker *et al.* 2005, 3; Buikstra and Ubelaker 1994, 39; Cunningham *et al.* 2016, 16; Lewis 2006, 186; Saunders 2008, 117, 118; Viciano *et al.* 2011, 97). The non-adult skeleton consists of significantly more bones than the adult skeleton and epiphyses or small ossification centres are often missed in the field due to inadequate excavation techniques (Baker *et al.* 2005, 2; Lewis 2006, 186). As they are generally physically smaller than adults, non-adult inhumations create smaller features that can be more easily missed or inadvertently destroyed during excavation (Baker *et al.* 2005, 11; Saunders 2008, 118). Non-adults were also frequently buried in shallower graves than their adult counterparts, increasing the likelihood of their accidental destruction during the initial stages of excavation (Baker *et al.* 2005, 11; Lewis 2006, 186). Insufficient training through the 20th century repeatedly led to the misidentification of non-adult remains, especially those of infants and perinates, as animal remains (Baker *et al.* 2005, 2). Even when these skeletal remains were recognised as human, on occasion they were purposefully not collected because it was thought that they could provide little information (Baker *et al.* 2005, 11).

Differential funerary treatment of non-adults may also have played a role in their underrepresentation in the archaeological record (Cunningham *et al.* 2016, 16; Lewis 2006, 186; Saunders 2008, 118). Culturally, many past societies tended to inter infants and children apart from the main burial area of the adult members of the society, some grouped together in designated areas and some in more private spaces such as within the walls or doorways of homes (Lewis 2006, 186; Saunders 2008, 118-119). As most excavations are concentrated on one area that was previously found to contain archaeological material, the individuals who were excluded from the adult burial grounds have a higher chance of going unnoticed (Baker *et al.* 2005, 11; Saunders 2008, 118). Practically, child mortality rates were much higher in the past and many families simply may not have been able to afford recurrent funerary costs, which would also lead to exclusion of non-adults from the adult cemetery (Cunningham *et al.* 2016, 16).

Poor preservation is frequently cited as a major factor in the lack of non-adult skeletal material in archaeological samples (Aris *et al.* 2018, 673; Baker *et al.* 2005, 3, 11; Saunders 2008, 117; Viciano *et al.* 2011, 97). Although non-adult bones are indeed not as robust or as well mineralized as those of adults, recent taphonomic research has shown that under the same conditions that allow for good preservation in adult bone, non-adult remains will fair just as well (Baker *et al.* 2005, 11; Lewis 2006, 185). This being said, because of their size and fragility, non-adult remains will deteriorate more quickly than those of adults in unfavourable soil conditions, which would indeed lead to a bias in the sample (Lewis 2006, 185; Saunders 2008, 119). So, although non-adult skeletal material is more susceptible to diagenesis, large excavations with good overall preservation have the potential to yield enough non-adult individuals to allow for detailed analysis at both the individual and population levels (Lewis 2006, 185). However, as non-adult bones are more easily affected by taphonomic processes, it is more likely that the skeletal features used in morphological and metric methods of sex estimation may be obscured (Shankar *et al.* 2013, 753; Tuttösí and Cardoso 2015, 307).

1.3.2 Documented collections

As rare as non-adult remains are within skeletal collections, it is indeed far more rare for these collections to have verifiable documentation pertaining to the individuals that constitute the mortuary sample (Cunningham *et al.* 2016, 15; Klales and Burns 2017, 747; Moore 2013, 107; Olivares and Aguilera 2016, 1624; Wilson and Humphrey 2017, 34). Documented collections are invaluable to the development and testing of new methods, as they provide conclusive evidence of sex and age-at-death independent of osteological assessments (Cunningham *et al.* 2016, 18). Due to the scarcity of adequate collections, however, many methods have been developed and tested on samples of unknown age and sex, which completely undermines the ability to independently ascertain the efficacy and validity of the method (Baker *et al.* 2005, 10; Schutkowski 1993, 204; i.e. Viciano et al. 2011, Tuttösí and Cardoso 2015).

1.3.3 Population specificity

Even with the few documented collections that contain non-adult remains in sufficient quantities, differences in body size and the degree and patterning of sexual dimorphism between populations hinders the ability to create a universal method (Buikstra and Ubelaker 1994, 16; Cunningham *et al.* 2016, 18; Hillson 2005, 257; Lewis 2006, 49; Shankar *et al.* 2013, 753; Tuttösí and Cardoso 2015, 306; Wilson and Humphrey 2017, 34, 35). These differences are influenced by genetic (intrinsic) and environmental (extrinsic) factors and can be observed temporally as well as geographically (Aris *et al*. 2018, 677-678; Bosman *et al.* 2017, 331; Cardoso 2008, 159; Inskip *et al.* 2018, 682; Shankar *et al.* 2013, 755; Wilson and Humphrey 2017, 34, 35). Predisposition to disease, hormone levels, tendency towards congenital abnormalities and the genetically-dictated aspect of body size are examples of intrinsic factors. Extrinsic factors include a wide variety of external influences to the body such as nutrition, activity patterns, diet, and illness (Moore 2013, 94). Population variation is the main reason that many of the extant non-adult methods cannot be reliably applied across multiple populations, an issue that will be further explored in Chapter 2.

1.3.4 Selective mortality and the heterogeneity in risks

A deceptively obvious fact that could have significant repercussions on any osteological study is that the individuals that constitute the skeletal sample are deceased. This is particularly poignant for non-adult studies, as these individuals represent the members of a living population that did not survive into adulthood (Aris *et al*. 2018, 673; Cardoso 2008, 166; Wood *et al*. 1992, 344). As such, a high proportion of the sample population will have pathological skeletal changes that increase the risk of juvenile mortality (Wood *et al.* 1992, 344). For example, since children are continually undergoing growth and development, nutritional stress can affect them more strongly than adults and so the non-adult skeletal sample is likely to be highly selective for individuals displaying lesions associated with malnutrition (Saunders 2008, 133, 134-136; Wood *et al.* 1992, 344).

The specific cause of death is inscrutable in most archaeological circumstances. As such, the premature death of the non-adults in question may have been caused by any number of acute or chronic illnesses as well as a variety of fatal injuries (Cunningham *et al.* 2016, 16; Lewis 2006, 187; Saunders 2008, 133; Wilson and Humphrey 2017, 35). This variation leads to problems in interpreting the degree to which the deceased children, in contrast to those that survived to adulthood, may have been differentially affected in their nutrition, exposure to disease, activity patterns, or intrinsic frailty. Some of the non-adult skeletal sample may represent individuals who suffered from malnutrition or chronic illness, which would have affected their growth and development (Cardoso 2008, 166; Moore 2013, 94; Saunders 2008, 125-126; Wilson and Humphrey 2017, 35). Others may have perished of a sudden fever or serious accident (Aris *et al*. 2018, 673; Saunders 2008, 133; Wilson and Humphrey 2017, 35). Looking only at non-traumatic causes of death, susceptibility to disease varies by individual in relation to genetic causes, socioeconomic status, and microenvironmental context (Wood *et al.* 1992, 345). Due to the general uncertainty surrounding cause of death, it cannot be assumed that all non-adult skeletal material represents the members of a population with the greatest underlying frailty (Aris *et al*. 2018, 673; Saunders 2008, 133; Wilson and Humphrey 2017, 35; Wood *et al.* 1992, 345).

It is therefore possible that some, but not necessarily all, of the non-adult skeletal sample represents individuals who experienced long-term physical stress, such as prolonged malnutrition or chronic illness. Extrinsic stressors such as these can cause marked changes to the non-adult skeleton along its growth and development trajectory (Hammerl 2013, 263; Moore 2013, 94; Saunders 2008, 125-126; Wilson and Humphrey 2017, 35; Uhl 2013, 63). Malnutrition, signs of which are often found in the skeletal record, in particular can reduce observable sexual dimorphism in body size (Lewis 2006, 49-50). Males are more strongly affected by malnutrition than females (Cardoso 2008, 159; Moore 2013, 93). Females, including non-adults, have a defence against nutritional stress in the form of the high level of fat in their bodies. Females can therefore survive longer periods of poor nutrition than their male peers without sacrificing much of the energy normally allotted to growth and development (Moore 2013, 93, 94; Stull *et al.* 2017, 65). This results in a reduction in sexual dimorphism of size, because from the age of 3 years until puberty, boys undergoing normal growth and development have more muscle mass than girls. As bone responds and remodels to mechanical stress caused by muscles, the long-bones of well-nourished boys are generally larger in breadth than their female counterparts of the same age (Stull *et al.* 2017, 65). If growth retardation is experienced, this sexual dimorphism is less readily apparent as the skeletons of malnourished boys would not be able to achieve the same size as they conserve energy (Moore 2013, 94; Stull *et al.* 2017, 65). In such a way, skeletal elements of non-adults can be greatly affected by extrinsic factors, a fact which in turn serves to obscure the genetically-determined sexual dimorphism in size and morphology. As discussed above, it is difficult to determine whether or not and to what extent a non-adult experienced physical stress from the skeletal record. Sex estimation methods based on skeletal elements are therefore problematic as varying levels of sexual dimorphism can be displayed even at the individual level within a single population.

The permanent dentition, on the other hand, is strongly regulated by genetic and hormonal influences and is less influenced by environmental factors than is dynamic bone (Cunningham *et al.* 2016, 13; Hammerl 2013, 263; Hillson 2005, 210; Moore 2013, 93-94, 107; Uhl 2013, 68). In addition, when the individual has undergone a period of prolonged stress, this will be evidenced by enamel hypoplastic defects on the tooth surface (Hillson 2005, 211). In the present study, the dimensions of the adult and non-adult permanent dentitions will be compared to one another using independent-samples *t*-tests to assess whether the phenomena of selective mortality or underlying frailty have significantly affected the dimensions of the permanent dentition of the non-adult sample. As this is a method intended for use on archaeological samples for the purpose of estimating sex, the issue of whether the skeletal sample is an accurate representation of the living population is immaterial as long as no significant differences in dental dimensions are observed between the adult and non-adult permanent dentitions.

1.4 A population-specific odontometric approach

Decades of research into the permanent dentition have consistently found sufficient sexual dimorphism in the dimensions of the tooth crown and cervix to allow for statistical discrimination between male and female individuals (Black 1978, 81; Cardoso 2008, 159; Cunningham *et al.* 2016, 18; Saunders 2008, 124; Schwartz and Dean 2005, 312; Shankar *et al.* 2013, 753; Viciano *et al.* 2013, 31). Many recent studies have achieved accuracy rates comparable to those seen in adult sexing methods (Aris *et al.* 2018, 676, 680; Cardoso 2008, 163, 164-165; Viciano *et al.* 2011, 105; Viciano *et al.* 2013, 36). Teeth are especially useful for the study of non-adults in archaeological contexts for a number of reasons. Firstly, enamel is the hardest substance in the human body and is therefore significantly more resistant to taphonomic processes than bone, which consists of a much high percentage of degradable organic material (Baker *et al.* 2005, 53; Cunningham *et al.* 2016, 13; Hammerl 2013, 263; Hillson 2005, 158-159; Hillson 2014, 70, 110). As such, teeth are frequently recovered from archaeological sites where unfavourable soil conditions lead to poor preservation of skeletal material (Cunningham *et al.* 2016, 13). Secondly, the formation and mineralization of teeth are genetically dictated and are consequently less affected by environmental factors such as nutrition and chronic illness than bone growth or tooth eruption (Cunningham *et al.* 2016, 13; Hillson 2005, 257). The third advantage relates to the fact that bone is a dynamic material and continuously remodels throughout life in response to such influences as systemic pathological conditions and trauma. This remodelling process is particularly active in non-adults as their bodies grow and develop (Hillson 2005, 207). This means that the non-adult skeleton is constantly changing and morphological features thought to reflect sexual dimorphism may in fact be more closely linked to growth and development (Cardoso 2008, 159; Lewis 2006, 51; Wilson *et al.* 2016, 263-264; Wilson and Humphrey 2017, 36). Conversely, dental enamel develops and mineralises during childhood and remains largely unchanged throughout life, apart from tooth-wear and pathological changes such as caries (Cardoso 2008, 159; Hillson 2005, 207-208, 257; Hillson 2014, 110). It is therefore possible to make direct comparisons between the permanent dentition of adults and the permanent dentition of non-adults, which is not true of any other skeletal element (Aris *et al.* 2018, 673; Cardoso 2008, 159; Hillson 2005, 257; Viciano *et al.* 2011, 98; Viciano *et al.* 2013, 31). Any sexual dimorphism observable in the adult permanent dentition should also be observable within the non-adult permanent dentition of the same population (Hassett 2011, 486; Cardoso 2008, 159). Archaeological collections often do not contain enough non-adult individuals to constitute a sample large enough to carry out compelling statistical analysis. This can be circumvented by developing discriminatory functions using the larger adult population and subsequently applying those parameters to the non-adult individuals (Hassett 2011, 486). This also negates the issue of population specificity, which is especially problematic for metric techniques as these depend more heavily on the general body size of the population (Buikstra and Ubelaker 1994, 16; Inskip *et al.* 2018, 681; Tuttösí and Cardoso 2015, 306). A great advantage of population-specific methods is that they work within the range of sexual dimorphism of the population rather than imposing the patterning and magnitude of sexual dimorphism of an unrelated population, which may be considerably different (Cardoso 2008, 166, 167; Tuttösí and Cardoso 2015, 306).

1.5 Aims and research questions

This study will use the post-medieval documented collection of Middenbeemster, the Netherlands, to examine the potential of odontometrics of the permanent canines and upper first molars to classify sex in adults. If successful, the same criteria will then be applied to the documented non-adult sample from the same collection. In order to determine the accuracy of the method, the odontrometrically estimated sex will be compared to the documented sex of the individual.

This thesis contains six chapters: Introduction (1), Non-adult sex estimation in osteoarchaeology (2), Materials and methods (3), Results (4), Discussion (5), and Conclusion (6). Following this introduction to the issue of non-adult sex estimation within osteoarchaeology, a review of skeletal and dental methods currently used for sex estimation in non-adult remains will be presented in Chapter 2. Particular focus will be given to the accuracy rates of these methods and their ability to be reliably reproduced. This background chapter will culminate in a detailed description of the recent odontometric studies on which the underlying theory of this thesis is based. Chapter 3 will describe the ways in which these odontometric methods have been adapted for the Middenbeemster collection and the means by which data was collected. This chapter will also expand upon the statistical methods employed in this analysis. Chapter 4 will present the results of the statistical analyses and the logistic regression models and equations developed on the adult sample and applied to the non-adult sample. Interpretation of the raw data presented in Chapter 4 will be undertaken in Chapter 5. Confounding factors to the study and the method will be contemplated and the overall viability of population-specific odontometric sex estimation methods based on the permanent canines and upper first molars will be assessed. This chapter will attempt to provide comprehensive answers the following research questions:

1. Can sexual dimorphism be observed in the dimensions of the permanent canines or maxillary first molars in the documented adult population of Middenbeemster?
    a. Which tooth or measurement, if any, displays the greatest degree of sexual dimorphism and therefore the greatest discriminating power?
    b. Were any measurements not sexually dimorphic or too problematic?
2. Can logistic regression analysis be applied to accurately categorize the adult population?
    a. Do any accuracy rates meet the 85% threshold of acceptability?
    b. How does this method compare to skeletal estimations based on the methods listed in Buikstra and Ubelaker (1994)?
        i. Would the classification parameters meet acceptable levels of accuracy if documented sex was unavailable and the functions were developed based on osteological estimations of sex?
3. Can this method be successfully applied to the permanent dentition of the documented non-adults from the Middenbeemster collection?
    a. Does the accuracy meet the 75% acceptability threshold for non-adult sexing methods? Does it meet the 85% considered acceptable in adult sexing?
    b. From what age can this method be reliably applied? What practical considerations, such as the embedment of the tooth in the jaw bone, may impede the application of this methodology?

    c. Is this method, previously applied to archaeological English and Italian populations and a modern Spanish sample, applicable to a post-medieval Dutch skeletal collection?

4. Do any extrinsic factors influence tooth dimensions, such as age or health?
    a. Is selective mortality discernable in the non-adult sample?

A summary of the research undertaken and its importance to the field of archaeology will be presented in Chapter 6, the Conclusion. Here, suggestions for future research will be discussed, based on the issues encountered during the preparation of this thesis. The aim of this study is to test and promote a reliable and accurate population-specific method to estimate sex in non-adults. This would contribute enormously to the field of osteoarchaeology and would help to illuminate the lived experiences of girls and boys in the past, which are currently indiscernible due to the poor resolution of current morphometric methods for sexing non-adults.

2. Non-adult sex-estimation in osteoarchaeology

The value of any method is assessed through its accuracy and reliability or its ability to be reproduced as compared to other available methods. It is therefore necessary to undertake a brief but comprehensive survey of the most promising methods currently available for sex estimation in non-adults. In this chapter, methods concerning the skeleton and the dentition will be considered separately, followed by a more detailed explanation of the odontometric studies on which the underlying theory of this thesis is based.

2.1 Genetic methods

2.1.1 Contamination

Great progress has been made in the field of DNA analysis since the discipline first began to gain ground within physical anthropology, although the financial cost and time needed are often too great for many research institutions (Cunningham *et al.* 2016, 18; Lewis 2006, 54-55, 187; Moore 2013, 109; Shankar *et al.* 2013, 753; Wilson and Humphrey 2017, 33). Though provisions exist to prevent and identify contamination and to successfully amplify degraded DNA sequences, the multi-tiered process of confirming the validity of results is still time-consuming and expensive (Cabana *et al.* 2013, 468, 474; Lewis 2006, 55; Tierney and Bird 2015, 34). Current methods of DNA extraction can isolate human DNA, but cannot discriminate between the DNA of two or more individuals (Cabana *et al.* 2013, 468). This means that researchers must take all measures to avoid any and all contact with the sample both in the field and in the laboratory (Cabana *et al.*, 468; Tierney and Bird 2015, 28). To monitor for possible contamination after extraction, it is recommended that samples from each individual be analysed in at least two separate facilities by separate technicians and that each of these analyses be repeated at least once (Cooper and Poinar 2000, 1139; Tierney and Bird 2015, 28, 32). The results of each assay are then compared to those of other samples from the same individual to determine whether any inconsistencies can be observed between laboratories or repetitions (Cooper and Poinar 2000, 1139; Tierney and Bird 2015, 28, 32, 35). In addition to this test of replication, it is recommended that each laboratory create a database of the DNA of all those who work in the facility. The results of each analysis can then be compared to the database in order to identify any match or partial match between sequences (Tierney and Bird 2015, 32). Although each of these precautions is cumbersome, they have proven highly successful in isolating DNA without contamination or at least in identifying when contamination has taken place. Nevertheless, these provisions do not control for the possibility of contamination during excavation and post-excavation processing.

2.1.2 Degradation

A more concerning issue for the use of DNA analysis within archaeology is that of degradation (Cabana *et al.* 2013, 468; Cunningham *et al.* 2016, 18; Moore 2013, 109; Olivares and Aguilera 2016, 1625; Tierney and Bird 2015, 34). The preservation of DNA is influenced by a number of factors including environmental conditions, the type of skeletal element sampled and its biochemical integrity (Cabana *et al.* 2013, 468). Over time, DNA will degrade to the point where it is not possible to amplify, or copy, targeted areas unless degradation is delayed or prevented by environmental conditions (Cabana *et al.* 2013, 467-468; Tierney and Bird 2015, 28). Degradation is particularly problematic in sex estimation, because there are substantially more copies of the X-chromosome than the Y-chromosome in each cell during life and the Y-chromosome degrades at a significantly faster rate than the X-chromosome (Cabana *et al.* 2013, 458, 468, 474; Lewis 2006, 54; Tierney and Bird 2015, 28). If the Y- and X-chromosomes have degraded at different rates within a sample, the X-chromosome may amplify successfully while the Y-chromosome does not, resulting in readings showing only X-chromosomes and a false identification as female (Cabana *et al.* 2013, 474; Lewis 2006, 54; Moore 2013, 109; Tierney and Bird 2015, 28). For this reason, two guidelines are generally followed: that results be deemed inconclusive if the sample is not sufficiently well-preserved as to allow for multiple high-resolution assays and that the detection of a Y-chromosome can lead to a male identification, but its absence cannot result in a conclusive female identification (Cabana *et al*. 2013, 474; Tierney and Bird 2015, 32, 34, 35). It has been noted that the hormone fluctuations experienced by pubescent adolescents may differentially affect the presentation of sexually dimorphic genes and result in a false positive identification of males as female (Lewis 2006, 55). A reliable means of sexing non-adults skeletally would therefore provide a way to sex those individuals whose DNA was not sufficiently preserved.

2.1.3 A case study in recent genetic sex estimation (Tierney and Bird 2015)

An excellent example of the tremendous promise of sex estimation through DNA analysis, as well as the limitations still plaguing the field, is a 2015 study on a medieval population from Ireland by Tierney and Bird. After first following a series of protocols to limit and detect contamination as outlined above, the researchers tested the reliability of their methods of extraction, amplification, and analysis on a small sample of adult individuals whose sex had been estimated morphologically. These techniques were then applied to a sample of non-adults from the same population. Thirty-eight adults and 19 non-adults were chosen for sampling based on their excellent state of preservation (Tierney and Bird 2015, 28). However, due to the magnitude of degradation, conclusive sex determinations were only possible for 20 adult individuals and 4 non-adult individuals (Tierney and Bird 2015, 32, 35). Poor amplification due to partial degradation meant that the DNA of 10 adults and 11 non-adults could not be analysed repeatedly at both facilities and so were assigned a probable sex based on the obtainable results (Tierney and Bird 2015, 32, 34, 35). The results obtained were nonetheless compelling: 14 of the 15 non-adult sex estimations were either male or probable male, meaning that the Y-chromosome was observable in all 14 samples (Tierney and Bird 2015, 36). Though the results of 11 of these samples were not confirmed through duplication at separate locations, the presence of the Y-chromosome is fairly conclusive, especially given the many precautions taken against contamination. However, it is not encouraging that so many samples did not yield any DNA or were too poorly preserved to allow for multiple analyses and therefore conclusive results. In order to better understand the demographic patterns emerging from these preliminary results, another method for sex estimation in non-adults is needed.

2.2 Skeletal morphometric methods

It is debated whether morphological or metric methods provide the most reliable results overall and some studies choose to use a combination of the two (Cardoso and Saunders 2008, 28; Klales and Burns 2017, 747; Wilson and Humphrey 2017, 33). Morphological methods rely on the visual assessments of non-metric skeletal traits, that is, traits whose expressions can be seen in the shape or form of the skeletal element rather than its size (Wilson *et al.* 2016, 255). Metric methods, as the name suggests, quantify the size and potentially the shape of a trait using measurements, angles, and ratios (Wilson *et al.* 2016, 255). Proponents of morphological methods emphasize their simplicity and intuitive nature, whereas supporters of metric methods emphasize the objective nature of quantitative methods and greater efficiency, accuracy and accessibility (Cardoso 2008, 159, 167; Cardoso and Saunders 2008, 28; Olivares and Aguilera 2016, 1625; Wilson *et al.* 2016, 255-256; Wilson and Humphrey 2017, 33-35). Though many of the following methods showed great promise upon their development, subsequent application of these methods to separate populations resulted in a significant reduction in accuracy, limiting their practicality within osteoarchaeology.

2.2.1 Morphology of the ilium and mandible

The morphological method that is most commonly applied to archaeological non-adult human skeletal remains is that of Schutkowski (1993) (Olivares and Aguilera 2016, 1624). This method was developed on the Named Spitalfields collection in London. The sex and age of many individuals in this collection is known from coffin plates with their names and years of birth and death (Schutkowski 1993, 199-200). Schutkowski identified three sexually dimorphic morphological features in the lower jaws (hereafter called mandibles) and four such features in the largest of the hip bones (the ilium) of 61 non-adult individuals (Schutkowski 1993, 200-201). Descriptions of each of these features and their differential expressions would prove burdensome for this brief review of methods, but the features identified largely correspond to features that show pronounced sexual dimorphism in adult individuals (Buikstra and Ubelaker 1994, 16; Schutkowski 1993, 200-201). The author concluded that these features were consistently and appreciatively different in boys and girls and that the accuracy of the observation of these features was comparable to that of methods for sex estimation in adults (Schutkowski 1993, 204). He even declared that children could now be included in palaeodemographic analysis (Schutkowski 1993, 205). However, this declaration was premature as several subsequent studies found that population-specific differences in the expression and patterning of not only sexual dimorphism but also growth and development meant that the method could not be reliably applied to a population other than Spitalfields (Cardoso and Saunders 2008, 28; Loth and Henneberg 2001, 180; Olivares and Aguilera 2016, 1628; Wilson *et al.* 2016, 256). These studies also consistently reported high rates of inter- and intraobserver error, tested by having both the original observer and at least one other osteologist observe and repeatedly score the pertinent traits of a subsample of individuals after an interval of time and comparing the results (Cardoso and Saunders 2008, 25; Olivares and Aguilera 2016, 1628-1629). This led many authors to conclude that the dimorphic expressions of the traits were not well-defined and could therefore be interpreted subjectively by each researcher, making the results of each test highly variable (Cardoso and Saunders 2008, 25; Loth and Henneberg 2001, 180; Olivares and Aguilera 2016, 1629). Tests of Schutkowski's study (1993) have invariably reached the conclusion that the method is not effective (Cardoso and Saunders 2008, 28; Loth and Henneberg 2001, 180; Olivares and Aguilera 2016, 1631).

Loth (1996) tested the mandibular features as defined by Schutkowski (1993) and obtained significantly lower accuracy rates (Loth and Henneberg 2001, 180). After these disappointing results, Loth and Henneberg (2001) developed a new method for non-adult sex estimation using the shape of the mandible (Loth and Henneberg 2001, 180). In this study the authors proposed that the overall shape of the mandible is curved in females and angular in males and that the shape of the chin region is rounded in females and squared in males (Loth and Henneberg 2001, 181-182). The original study documented good interobserver agreement and easy identification of traits based on the findings of three observers (Loth and Henneberg 2001, 183). The accuracy rate of this method, averaged between the three observers, was 81% (Loth and Henneberg 2001, 183). A year later, Scheuer (2002) published a blind test of this method on 36 individuals from the Spitalfields collection and achieved a predictive accuracy of only 64% (Scheuer 2002, 189, 191). Another study of the same year also tested the ability of morphological features of the mandible to classify sex as outlined by Loth and Henneberg (2001), attaining 62.5% accuracy in the female sample of 40 girls and 41.6% accuracy in the male sample of 36 boys (Coqueugniot *et al.* 2002, 135, 136). Though the authors of both studies were able to identify the characteristics described by Loth and Henneberg (2001), it was found that many individuals displayed contradictory expressions, meaning one female trait and one male trait (Coqueugniot *et al.* 2002, 136; Scheuer 2002, 191). It is unclear whether sample-specificity or inter-observer error influenced these results, but certainly this method has not proven reliable.

2.2.2 Morphology of the *os coxa*

More recent work has also focussed on aspects of the pelvic girdle (the *os coxa*). In 2017, Klales and Burns (2017) attempted to apply Phenice (1969), a morphological method commonly used in adults, to non-adults from birth to 20.5 years of age. This method has achieved accuracy rates of up to 95% using three sexually dimorphic traits of the *os cox*a (Klales and Burns 2017, 747). Klales and Burns divided their sample into six age groups and presented the accuracy rates achieved according to these groupings. The results are presented in Table 1. Accuracy rates increased with age, with accuracy rates above 85% only achieved for those individuals above the age of 12.6 years (Klales and Burns 2017, 750). For younger individuals the results were much less promising; those under 12.5 years of age were all classified with accuracies less than 75% (Klales and Burns 2017, 750). It can therefore be safely concluded that the traits established on the adult skeleton by Phenice 1969 are not applicable to the non-adult skeleton before puberty.

Table 1. Accuracy rates by age group as achieved by Klales and Burns (2017) (after Klales and Burns 2017, 750).

Age range (years)	Accuracy rate
1-3.5	53.9%
3.6-6.5	59.1%
6.6-9.5	64.5%
9.6-12.5	71.7%
12.6-15.5	85.3%
15.6-20.5	97.2%

Luna and colleagues (2017) used macroscopic observation along with quantitative measurements in order to discriminate for sex based on the shape of one aspect of the ilium (the auricular surface) (Luna *et al.* 2017, 899). Using 34 individuals aged 7 to 18 years from the Identified Skeletons Collection of the University of Coimbra, Portugal (1887-1934) the authors achieved accuracy rates of 82.35% and 88.23%, depending on the statistical method employed (Luna *et al.* 2017, 903-904). A similar method had previously been carried out on the Spitalfields collection by Wilson and colleagues (2008) with an 84% accuracy rate. However, a subsequent study by the same authors published in 2011 reported high inter-observer error and an overall accuracy of only 65.2% (Luna *et al.* 2017, 899-900; Wilson *et al.* 2011, 39-40). Thus, although this method does show potential, it has had varying results and reports of high inter-observer error due to the ambiguity of the landmarks on which the measurements are based (Wilson *et al.* 2011, 37).

2.2.3 Morphometric methods using 3D scanning

It has been suggested that metric approaches to non-adult sex estimation are less reliable than morphological ones due to the smaller degree of and greater overlap in size dimorphism among non-adults as compared to adults (Moore 2013, 107). This overlap impairs any attempt to definitively discriminate between male and female based on size and the comparatively small amount of difference between the sexes increases the likelihood that slight errors in measurement could result in an incorrect estimation (Moore 2013, 107). These issues are being addressed through technological advances that allow for more precise and complex measurements, such as digital 3D scanning (Moore 2013, 107). This recent innovation allows researchers to manipulate a digital replica of a skeletal element and to precisely pinpoint certain landmarks from which to take measurements (Wilson *et al.* 2016, 256). Wilson and colleagues (2016) applied this procedure to three skeletal collections from London dating to the 16th and 17th centuries to quantify different elements of the curvature of the ilium (Wilson *et al.* 2016, 257). Although the study was able to achieve 86.8% accuracy for one series of measurements, the researchers advised caution in using this method due to the differential expressions of the traits analysed at different stages of growth and development (Wilson *et al.* 2016, 262, 264). Two of the same researchers subsequently undertook an investigation into the development of the ilium throughout childhood using 3D images (Wilson and Humphrey 2017, 33). The authors concluded that although certain traits showed sufficient sexual dimorphism to allow for discrimination at certain ages, which traits were useful and how each expressed dimorphism varied significantly (Wilson and Humphrey 2017, 36-37). Though 3D imaging may contribute greatly to the field in the future, more detailed studies of how sexually dimorphic traits are expressed in relation to growth and development are needed before these can be widely adopted.

2.2.4 Long-bone diaphyseal dimensions

It has recently been proposed that although the length of non-adult long bones can be too greatly affected by such factors as growth retardation and the sexual dimorphism of growth to be an effective means of differentiating between the sexes, the breadth of long bones may be more closely linked to sex (Cardoso 2008, 159; Moore 2013, 94; Stull *et al.* 2017, 65). Stull and colleagues (2017) used radiographs of 1310 modern South African children of known sex to take 18 measurements from six long bones (Stull *et al.* 2017, 65-66). Several statistical analyses were performed on these measurements (Stull *et al.* 2017, 66-70). The only results to show any promise were those obtained from multivariate equations based on the measurements of several bones (Stull *et al.* 2017, 69). With these equations, accuracy rates ranged from 70% to 93% (Stull *et al.* 2017). Though these results may seem encouraging, the multivariate analysis requires the complete, or near complete, preservation of multiple long bones in excellent condition in order to obtain the necessary measurements, a scenario rarely encountered within archaeology. In addition, as outlined in Chapter 1, male non-adults may be differentially affected by malnutrition, reducing the level of sexual dimorphism in both the length and breadth of long-bones.

2.3 Dental methods

In archaeological contexts, bone can be too poorly preserved to allow for thorough morphological or metric sex estimation (Tuttösí and Cardoso 2015, 307). Under similar conditions, teeth are extremely resilient to post-depositional changes (Cunningham *et al.* 2016, 13, 149). Sexual dimorphism in the size and shape of teeth has been observed in various populations, with males consistently having larger teeth (Cunningham *et al.* 2016, 18; Hammerl 2013, 268; Hillson 2005, 257; Shankar *et al*. 2013, 753). The level of sexual dimorphism can vary within species and so comparisons between populations must be done with caution (Hillson 2005, 257). Several studies have sought to quantify this sexual dimorphism in either the deciduous or permanent dentitions and to develop odontometric methods to estimate the sex of non-adult skeletal remains. Some of these studies made provisions for population-level differences by developing a discriminant method on the permanent dentitions of the adults of known or estimated sex and applying this method to the non-adult individuals of the same population. In order to understand the below methods, a basic knowledge of the anatomical terminology used to describe various aspects of the dentition is necessary. Please refer to figure 1 for a visual representation of these terms. The mesial surface of the tooth is the surface closest to the point where the central incisors meet,

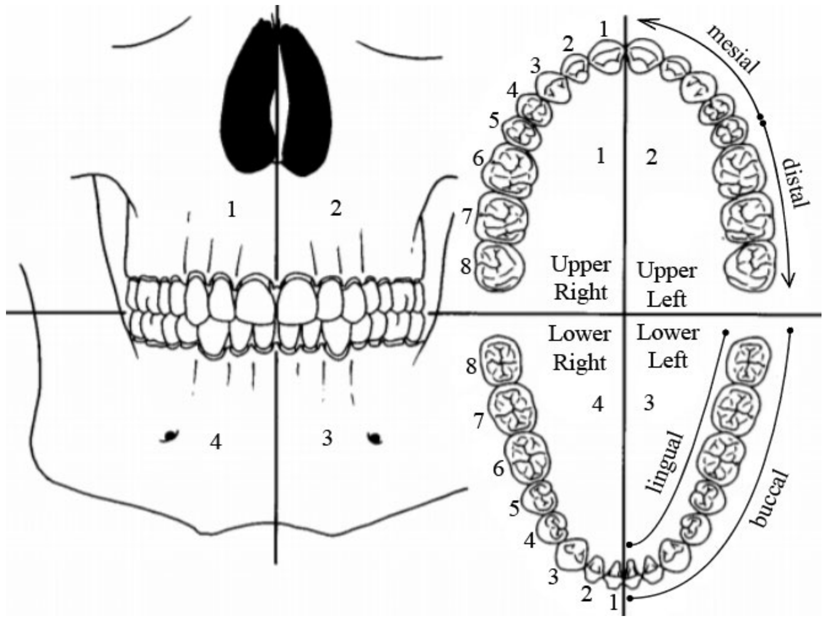


Figure 1. Illustration showing dental directions and the FDI notation system (after Hillson 1996, 7).

or generally towards the front of the mouth (White and Folkens 2005, 128). The distal surface is opposite the mesial surface, generally facing the back of the mouth (White and Folkens 2005, 128). The lingual surface is that closest to the tongue (White and Folkens 2005, 128). Within this study, the term buccal will be used to refer to the surface opposite to the lingual surface, nearest the cheeks or lips. The term crown refers to the part of the tooth encased in enamel visible in life above the gumline (White and Folkens 2005, 129). The cervix is located at the point where the enamel of the crown meets the root (Cunningham *et al.* 2016, 151; White and Folkens 2005, 129). The chewing surface of the tooth is referred to as the occlusal surface (White and Folkens 2005, 128). The teeth of the upper jaw are referred to as the maxillary dentition, while the teeth of the lower jaw are referred to as the mandibular dentition (Hillson 2005, 10). This thesis will use the FDI method of dental notation to denote each specific tooth. This method divides the mouth into quadrants numbered (1) upper right, (2) upper left, (3) lower left and (4) lower right (Hillson 1996, 9). The teeth within each quadrant are also numbered consecutively from the most mesial to most distal. In this way, each tooth is assigned a two-digit label indicating its position in the mouth (Hillson 1996, 9). For example, the most central tooth in the upper right quadrant would be referred to as 1.1. Again, please see figure 1 for visual clarification.

2.3.1 Odontometric methods using the deciduous dentition

Several studies have shown that sexual dimorphism is discernible in the size of the deciduous dentition, but to a lesser extent than that observable in the permanent dentition (Black 1978, 81; Cunningham *et al.* 2016, 18; Hammerl 2013, 268; Lewis 2006, 48; Viciano *et al.* 2011, 97; Żądzińska *et al.* 2008, 179, 185). The first major foray into using the deciduous dentition for sex estimation was made by Black (1978). Black used the mesio-distal (MD) and bucco-lingual (BL) crown diameters, measured on the casts of 133 children from the University of Michigan, to conduct discriminant function analysis and determine the utility of deciduous odontometrics for sex estimation (Black 1978, 77). This procedure achieved 75% accuracy, but the method is reliant on 10 MD and 10 BL measurements (Black 1978, 77-78, 81). The sample used in this study was selected for their complete dentitions, allowing for full resolution in all the discriminant equations developed (Black 1978, 77). Despite its accuracy of 75%, chances are slim that the entire dentition is preserved in an archaeological context and slimmer still that all the deciduous crowns are fully formed, unworn and undamaged by caries or post-mortem deterioration, thus limiting the usefulness of this method in an archaeological context.

Żądzińska and colleagues (2008) studied the deciduous dentition of 133 non-adult skeletonised individuals from a medieval archaeological sample from Poland (Żądzińska *et al.* 2008, 177). The researchers applied multiple regression statistical analysis to the MD and BL crown diameters and compared these results to those obtained through DNA analysis (Żądzińska *et al.* 2008, 177). The DNA analysis was conclusive for 101 of the 133 individuals (Żądzińska *et al.* 2008, 179). Using multiple regression, the researchers were able to correctly estimate sex in 69% of males and 88% of females, or about 78% overall (Żądzińska *et al.* 2008, 184). The authors note that the level of sexual dimorphism in the deciduous dentition is relatively low and varies between populations and so only advise the application of the multiple regression equation developed to other central European medieval populations (Żądzińska *et al.* 2008, 186).

Taking a similar approach, Shankar and colleagues (2013) used stepwise discriminant function analysis on the crown dimensions of the deciduous canines and molars to estimate sex in a sample of 183 modern children from India between 5 and 13 years of age (Shankar *et al.* 2013, 753, 755). The individuals selected all had fully erupted canines and molars with no caries, damage, or orthodontic alterations so that all 20 measurements per individual could be recorded (Shankar *et al.* 2013, 753). As was the case with Black (1978), although this method obtained accuracy rates of 87.2% to 88% its applicability within archaeology is hindered by the necessity of undamaged teeth that have not been lost prior to analysis.

2.3.2 Odontometric methods using the permanent dentition

The permanent dentition has consistently shown higher levels of sexual dimorphism than the deciduous dentition (Black 1978, 81; Cardoso 2008, 159; Cunningham *et al.* 2016, 18; Hammerl 2013, 268; Lewis 2006, 48; Viciano *et al.* 2011, 97; Saunders 2008, 124). It should therefore follow that statistical analyses based on the permanent dentition should find more success in correctly classifying an individual as male or female. In addition to this advantage, the permanent odontometrics of adults and non-adults within the same population have been shown to be comparable, so methods can be developed on the adult sample and applied to the non-adult sample (Cardoso 2008, 159; Hassett 2011, 486; Hillson 2005, 257). This is advantageous for two reasons: firstly, there are usually far more adults than non-adults within an archaeological sample due to a variety of reasons outlined in Chapter 1. Secondly, if the sample is of unknown sex, osteological methods of sex estimation for adults are significantly more accurate and reliable than those currently available for non-adult sex estimation. It is therefore possible to develop an odontometric method that is relatively reliable based on an adult sample of skeletally estimated sex.

2.3.3 Sample-specific sexual dimorphism in the permanent dentition and its application to non-adult sexing (Cardoso 2008)

The issues of population-level differences encountered by many of the studies described above led to the recognition of the need for sample-specific methods that could nonetheless be conducted on separate populations (Cardoso 2008, 159). One such method is that proposed by Cardoso (2008), in which the dental dimensions of an adult sample of known sex, whether through documentation or skeletal analysis, are used to develop a statistical method of sex estimation that can in turn be applied to the non-adults of the same population (Cardoso

2008, 159). This method could also hold potential for estimating sex in adult individuals whose sexually dimorphic skeletal elements are not preserved for observation (Cardoso 2008, 159; Tuttösí and Cardoso 2015, 306-307). This method requires a large adult sample either with documented evidence of sex or reliable skeletal estimations of sex, which require good preservation (Cardoso 2008, 159). However, these conditions are frequently met by archaeological collections and so this method could prove useful in elucidating the demographics of past populations.

Table 2. Allocation accuracies achieved by Cardoso (2008) using logistic regression on the non-adult subsample (after Cardoso 2008, 163-164).

Tooth	**Measurement(s) used**	**Maxillary**		**Mandibular**	
		N	**Accuracy (%)**	***N***	**Accuracy (%)**
1st incisor	MD crown	22	54.5	25	60.0
	BL crown	23	52.2	24	70.8
	MD crown BL crown	22	54.5	11	55.0
2nd incisor	MD crown	21	57.1	24	45.8
	BL crown	23	69.6	23	60.9
	MD crown BL crown	21	47.6	22	59.1
Canine	MD crown	17	58.8	21	71.4
	BL crown	17	**100.0**	21	**85.7**
	MD crown BL crown	17	**88.2**	21	**85.7**
1st premolar	MD crown	22	68.2	20	70.0
	BL crown	22	68.2	20	**75.0**
	MD crown BL crown	22	63.6	20	70.0
2nd premolar	MD crown	18	50.0	18	61.1
	BL crown	18	72.2	18	55.6
	MD crown BL crown	18	72.2	18	66.7
1st molar	MD crown	40	55.0	38	50.0
	BL crown	37	62.2	37	59.5
	MD crown BL crown	36	61.1	35	54.3
2nd molar	MD crown	17	58.8	19	68.4
	BL crown	16	60.0	18	**77.8**
	MD crown BL crown	16	62.5	18	**77.8**

N number of individuals. MD mesio-distal. BL bucco-lingual. Accuracies equal to or above 75% are marked in bold.

The study used logistic regression to analyse the maximum MD and BL crown dimensions of 107 adults and 49 non-adults aged 1.7 to 15 years, from the documented skeletal collection housed in the National Museum of Natural History in Lisbon, Portugal (Cardoso 2008, 160). As described above, the statistical method was derived from the adult sample and subsequently applied to the non-adult sample as if sex were unknown (Cardoso 2008, 161). The allocation accuracies of this process are presented in table 2. The upper and lower canines were the only teeth to achieve acceptable accuracy rates among the non-adults when only a single measurement was used. The BL crown dimension of the maxillary canine discriminated for sex correctly in 100% of the 17 non-adult individuals who retained this tooth (Cardoso 2008, 163). The same dimension in the mandibular canine resulted in an accuracy rate of 85.7% on 18 non-adults (Cardoso 2008, 163). The allocation accuracies of the single-tooth logistic regression analyses are presented in table 2. These results did not improve significantly when measurements from several teeth were analysed in combination. Among these, the highest accuracy rate (88.2%) was achieved by two combinations: the BL mandibular canine with the BL mandibular second molar and the BL mandibular canine with the BL mandibular first premolar (Cardoso 2008, 164). All methods using canine dimensions showed great promise with accuracies of over 85% (Cardoso 2008, 164). This is a level of accuracy comparable to methods of sex estimation in adults, inspiring several research teams to test the method.

2.3.4 Test of a sample-specific odontometric approach to sex estimation on an undocumented archaeological population (Herculaneum) (Viciano *et al.* 2011)

One issue that is apparent in the Cardoso (2008) study is that many of the measurements were unobservable in certain individuals, significantly reducing the sample size for each measurement. Though 49 non-adult individuals were included in the analysis, the number of individuals used for each discriminant method ranged from 9 to 22 (Cardoso 2008, 165). In order to compensate for this, Viciano and colleagues (2011) included several alternative measurements as proposed by Hillson and colleagues (2005). In addition to the maximum MD and BL crown diameters, the crowns of the molars were also measured diagonally (Viciano *et al.* 2011, 99). This allows for measurements to be taken on teeth whose mesial or distal surfaces are blocked by neighbouring teeth (Hillson *et al.* 2005, 415; Shankar *et al.*

2013, 755). Maximum cervical diameters were also recorded for all observable teeth (Viciano *et al.* 2011, 99). These two measurements are taken along the line of the cervix and, like crown diameters, run between the most mesial and distal points and the most buccal and lingual points of the cervix respectively (Hillson *et al.* 2005, 416; Viciano *et al.* 2011, 99). All these alternative measurements provide the researcher with more options and therefore allow for at least some measurements to be taken on teeth that have been damaged or altered by pathological changes such as caries or by post-mortem alterations (Hassett 2011, 486; Hillson *et al.* 2005, 425; Tuttösí and Cardoso 2015, 307, 310).

In order to develop a statistical procedure for sex estimation, Viciano and colleagues (2011) used the dental dimensions of 87 adult individuals excavated from Herculaneum and curated in the Museum of Biomedical Sciences in Chieti, Italy (Viciano *et al.* 2011, 98). The sex of these individuals was estimated according to the morphological method of Ferembach and colleagues (1980) (Viciano *et al.* 2011, 98). By correlating the osteologically estimated sex with the dental dimensions, discriminant function analysis was performed and the resulting equations were applied to the non-adult sample (Viciano *et al.* 2011, 100). The sex of the non-adult individuals was independently estimated using Schutkowski's (1993) method in order to allow some degree of authentication for the odontometric results (Viciano *et al.* 2011, 98). As previously discussed, the method of Schutkowski (1993) has been proven unreliable when applied to populations temporally or geographically separated from Spitalfields, London. The accuracy rates reported in this study are therefore inherently flawed, as they are ultimately based on a flawed skeletal method.

2.3.5 Test of a sample-specific odontometric approach on a documented modern collection (Granada osteological collection) (Viciano *et al.* 2013)

In order to better assess the accuracy of this method, some of the researchers from the Herculaneum study went on to apply a similar technique to a modern, documented osteological collection from Granada, Spain (Viciano *et al*. 2013, 32). As this collection contains hundreds of non-adult individuals, it was not necessary to develop the statistical criteria on a larger adult population. Instead, the researchers used a subsample of 221 non-adult individuals to create several logistic regression equations and then blind-tested these equations on another subsample of 48 non-adults (Viciano *et al.* 2013, 34). The study reported relatively high levels of inter- and intra-observer error in the crown dimensions of the molars, but the difference between measurements was still within the acceptable limits as proposed by Hillson and colleagues (2005) (Viciano *et al.* 2013, 41). The researchers also found that in 26 of the 46 measurements, the cervix displayed a significantly higher level of sexual dimorphism in size than the crown, underlining the utility of cervical measurements for sex estimation (Viciano *et al*. 2013, 41). The researchers analysed the degree of sexual dimorphism of each measurement by comparing the male and female means through the Student's *t*-test (Viciano *et al.* 2013, 33). The results of these tests are given in table 3. Any results equal to or less than 0.05 are considered statistically significant and therefore potentially useful in estimating sex. Overall, the logistic

Table 3. Viciano and colleagues' (2013) *P*-value results listed by tooth (after Viciano *et al*. 2013, 38).

Tooth	**Measurement**	**Maxillary**	**Mandibular**
		P	*P*
1st incisor	MD crown	1.000	0.236
	BL crown	0.023*	0.367
	MD cervical	0.141	0.211
	BL cervical	0.453	0.823
2nd incisor	MD crown	0.038*	0.286
	BL crown	0.188	0.047*
	MD cervical	0.083	0.372
	BL cervical	0.000**	0.001**
Canine	MD crown	0.004**	0.018*
	BL crown	0.000**	0.001**
	MD cervical	0.000**	0.000**
	BL cervical	0.000**	0.001**
1st premolar	MD crown	0.312	0.003**
	BL crown	0.934	0.084
	MD cervical	0.025*	0.990
	BL cervical	0.048*	0.038*
2nd premolar	MD crown	0.314	0.001**
	BL crown	0.067	0.010**
	MD cervical	0.005**	0.000**
	BL cervical	0.012*	0.057
1st molar	MD crown	0.420	0.001**
	BL crown	0.000**	0.151
	MD cervical	0.749	0.013*
	BL cervical	0.000**	0.002**
2nd molar	MD crown	0.487	0.089
	BL crown	0.059	0.006**
	MD cervical	0.052	0.041*
	BL cervical	0.005**	0.000**
3rd molar	MD crown	0.413	0.256
	BL crown	0.165	0.023*
	MD cervical	0.042*	1.000
	BL cervical	0.827	0.014*

A single asterisk (*) indicates statistical significance at the $P \leq 0.05$ level. A double asterisk (**) indicates statistical significance at $P \leq 0.01$ (after Viciano *et al.* 2013, 38).

equations developed on the permanent dentition correctly estimated sex in 80.0% to 87.5% individuals depending on the measurements used (Viciano *et al.* 2013, 40). The results of the aforementioned blind test are given in table 4.

Table 4. Viciano and colleagues' (2013) allocation accuracy results for univariate and multivariate logistic regression equations. Only those using the permanent dentition are presented here (after Viciano *et al*. 2013, 40).

Tooth/teeth	Measurement(s)	*N*	Accuracy (%)
Maxillary canine	BLcrn and BLcerv	14	85.7
	MDcerv	12	83.3
	MDcerv and BLcerv	12	83.3
	BLcerv	15	80.0
Mandibular canine	MDcerv	16	87.5
Mandibular canine and mandibular first premolar	BLcerv and MDcrn	15	80.0
	BLcerv and BLcrn	16	87.5
Mandibular second premolar	MDcerv	15	80.0
Mandibular second molar	BLcrn and MLDBcrn	15	86.7
	MLDBcrn	16	81.3

BL bucco-lingual. MD mesio-distal. MLBD diagonal measurement from mesio-lingual corner to bucco-distal corner. crn crown. cerv cervical.

2.3.6 Test of a sample-specific odontometric approach on the maxillary first molars of a known archaeological population (Spitalfields) (Aris *et al.* 2018)

While Viciano and colleagues (2011) used an archaeological collection of unknown sex and Viciano and colleagues (2013) used a modern documented collection, the research carried out by Aris and colleagues (2018) made use of the documented archaeological sample from Spitalfields, London (Aris *et al.* 2018, 673). This study only included the maxillary first molar, which is usually the first permanent tooth to develop (Aris *et al.* 2018, 673). Although the canine has proven to be the most sexually dimorphic tooth, first molars also display high levels of sexual dimorphism and are more frequently preserved in the archaeological record since as multi-rooted teeth, they are more likely to be retained in the bone than their single-rooted counterparts (Aris *et al.* 2018, 673; Cardoso 2008, 165; Cunningham *et al.* 18; Hammerl 2013, 268). Statistical functions were first developed on the crown and cusp dimensions of 37 adult individuals and subsequently applied to 22 non-adult individuals with a resultant accuracy rate of 94.6% using multivariate linear discriminant analysis (Aris *et al.* 2018, 673-674, 676). Using binary logistic regression, accuracy rates above 90% were achieved using only the MD crown diameter (Aris *et al*. 2018, 677). The authors then further tested the method by applying the functions derived from the adult Spitalfields sample to the morphologically-assessed adult sample of medieval remains from Black Gate cemetery in Newcastle-upon Tyne, England (Aris *et al.* 2018, 673). This cross-application yielded comparatively poor results ranging from 57.7% to 83.8% depending on the formula used (Aris *et al.* 2018, 677). The authors attributed this to population-level differences in tooth size and the expression of sexual dimorphism in the dimensions of the maxillary first molar and to the highly specific nature of the formulae developed on the Spitalfields sample (Aris *et al.* 2018, 677-678). The two sites are relatively close geographically, but are separated temporally by several centuries (Aris *et al.* 2018, 673). This indicates that unless two populations are extremely close to one another in both space and time, the statistical procedures developed on one should not be applied to the other, rather that new, sample-specific formulae should be developed for each population. In this vein, Aris and colleagues (2018) also developed new formulae on the Black Gate population as had been done for Spitalfields (Aris *et al.* 2018, 678). Within the adult sample, the resultant sex estimations matched the morphologically estimated sex in 83.3% of cases (Aris *et al.* 2018, 678). The accuracy of the method as applied to non-adults in the Black Gate population was incalculable, because no documentation exists for the collection and there is no other accessible and reliable method to independently estimate sex (Aris *et al.* 2018, 679). Among the Black Gate population, the study found a marked variation in size between the adult and non-adult samples, emphasizing the need to check for such discrepancies that may point to a mortality bias within the sample (Aris *et al*. 2018, 680; Cardoso 2008, 166).

2.4 Summary

Most of the results obtained through the analysis of the permanent dentition via population-specific odontometric methods are comparable in accuracy to adult sex estimation methods and generally outperform other morphometric skeletal methods for non-adults. Unfortunately, these methods cannot be used on individuals whose permanent tooth crowns have not yet formed or are embedded within the bone (Cardoso 2008, 167). Further research is needed in order to better assess the utility of this method on diverse populations as well as to identify the most efficient and effective method of estimating sex in non-adult archaeological remains using these techniques. Efforts must be taken to maximise the number of measurements attainable per tooth and to closely regulate the measurement of molars, which consistently show comparatively high rates of inter- and intraobserver error. As demonstrated by the analysis of odontometrics in the Black Gate population, it is necessary to compare the mean dimensions of adult and non-adult permanent teeth in order to detect any signs of mortality bias and to assure that the two sets of dentition are indeed comparable. If these factors are taken into careful consideration, population-specific odontometric approaches hold exceptional potential to reliably estimate the sex of non-adult skeletal remains at an acceptable level of accuracy. The remainder of this thesis will be devoted to a test of a population-specific odontometric method for sex estimation in non-adults following an adaptation of Cardoso's (2008) method with the inclusion of cervical measurements, as described by Hillson and colleagues (2005) and as used by Aris and colleagues (2018) and Viciano and colleagues (2011) and (2013). The following chapter will outline the parameters of this research.

3. Materials and Methods

This thesis will attempt to use the dimensions of the permanent canines and maxillary first molars of a documented adult skeletal sample from post-medieval Holland to develop logistic regression formulae that will allow the sex of the documented non-adult sample from the same population to be accurately estimated. Following an introduction to the Middenbeemster skeletal collection, this chapter will present a detailed account of how this study was conducted.

3.1 The Middenbeemster collection

The Middenbeemster collection, housed at the Laboratory for Human Osteoarchaeology at Leiden University**,** the Netherlands, consists of the remains of over 450 individuals that are on the whole in very good condition (Carroll *et al*. 2016, 54; Inskip *et al*. 2018, 676; Veselka *et al*. 2015, 668; Waters-Rist and Hoogland 2013, 244). The collection was excavated from Keyserkerk, Middenbeemster, in the summer of 2011 by Leiden University in cooperation with Hollandia Archeologen (Waters-Rist and Hoogland 2013, 244). Though the church and cemetery had been in use since the early 17th century, most of the burials date from a comparatively narrow period of time from 1829 to 1866 (Waters-Rist and Hoogland 2013, 244). About one quarter of the individuals excavated have been identified in the cemetery ledger and so are of known age and sex (Inskip *et al*. 2018, 676; Waters-Rist and Hoogland 2013, 244). Military records also exist for most of the identified males, listing their height and any noteworthy medical remarks (Blom *et al.* 2018, 1393; Waters-Rist and Hoogland 2013, 244).

Middenbeemster became the central village of the Beemster polder, which was created through the draining of Beemster Lake in the early 17th century (Carroll *et al.* 2016, 54; Veselka *et al*. 2015, 667; Vikatou *et al*. 2017, 54). Beemster is located in the province of North Holland, north of the city of Amsterdam as shown in figure 2. As such, the Keyserkerk served the needs of the entire Beemster population, mainly composed of farmers focused primarily on raising cattle and dairy production (Bosman *et al*. 2017, 339; Carroll *et al.* 2016, 54; Chilcote 2018, 36, 39; Palmer *et al.* 2016, 79; Veselka *et al*. 2018, 70, 73; Vikatou *et al*. 2017, 54). Traditionally, labour in the rural dairy industry began at a young age, typically by the age of six, and the assigned roles were highly dependent upon sex (Chilcote 2018, 40; Palmer *et al.* 2016, 79; Veselka *et al*. 2015, 672, 673; Veselka *et al*. 2018, 73). Though in times of need adult women may have assisted their male relatives in the fields, women and girls were by and large tasked with indoor activities such as the production or maintenance of clothing, churning butter, making cheese and taking care of young children (Veselka *et al.* 2015, 672; Veselka *et al.* 2018, 73). Meanwhile, boys from about six years of age were expected to help in the more rigorous outdoor tasks such as herding and milking the cattle (Palmer *et al.* 2016, 79; Veselka *et al*. 2015, 672; Veselka *et al.* 2018, 73; Vikatou *et al*. 2018, 60). It has been hypothesized, based on the adult and documented non-adult individuals in the collection, that the higher prevalence of residual rickets in adult females as opposed to adult males is primarily caused by these gender-based activity patterns (Veselka *et al*. 2018, 73). If it were possible to estimate the sex of the undocumented non-adult portion of the sample it would be possible to study wider patterns of morbidity, trauma, mortality, nutritional deficiencies, and gender-based division of labour.

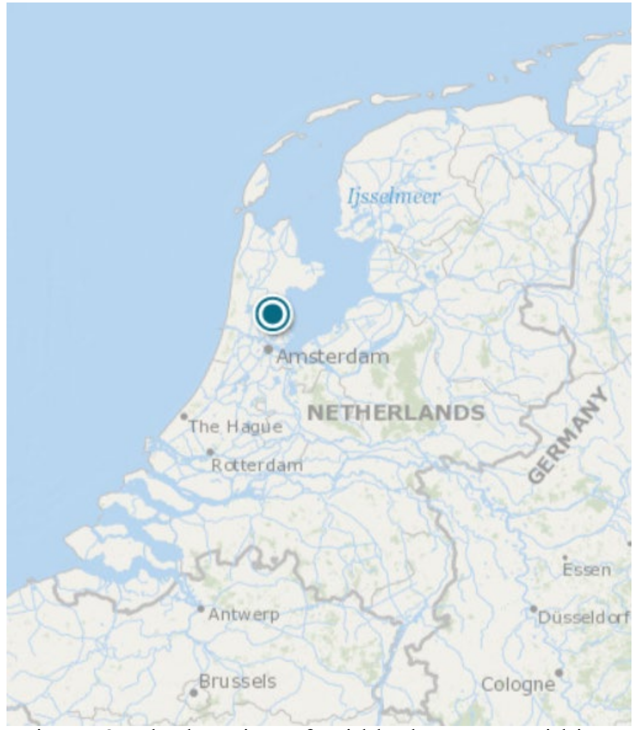


Figure 2. The location of Middenbeemster within the Netherlands, indicated by the blue circle (mapmaker.nationalgeographic.org).

As the oldest area of reclaimed land in the Netherlands, the Beemster Polder was made a UNESCO world heritage site in 1999 in recognition of this early feat of engineering and landscape architecture (Vikatou *et al*. 2017, 54; UNESCO 2019). Though the region flourished in the 17th and 18th centuries, an onslaught of diseases, which affected crops as well as the inhabitants and livestock, from the mid-18th century meant that by the mid-19th century, many members of the Beemster population suffered from malnutrition and a wide variety of infectious diseases (Carroll *et al.* 2016, 54; Chilcote 2018, 37, 39; Veselka *et al*. 2015, 667-668).

One hundred and eighteen adult individuals and 33 non-adult individuals aged three to 17 years were examined for this study, drawn from a list of individuals of known sex compiled by a person not involved in the collection of data. Those that did not retain at least one canine or maxillary first molar with observable measurements were excluded from the study. Teeth whose crowns were not yet fully formed were also excluded. The documented sex of the adult and non-adult individuals was not known to the researcher at the time of data collection. Osteological sex was estimated using the methods set out by Buikstra and Ubelaker (1994) and the WEA (1980). Dietary reconstruction has indicated that the Beemster population consumed large amounts of soft foods such as bread, potatoes and the dairy products they manufactured, which may reduce the severity of dental abrasion in the collection (Bosman *et al*. 2017, 330; Vikatou *et al*. 2017, 59). At the same time, the common occurrence of pipe notches has been noted in the collection (Carroll *et al*. 2016, 58). Pipe notches are the result of habitual tobacco smoking through clay pipes, which were very popular in the region from the 17th century (Carroll *et al*. 2016, 58). The continual pressure caused by biting on the pipe shaft while smoking abrades the affected teeth in a very characteristic fashion, which results in a circular void at the intersection between two or four teeth, with a semi-circular edge on each tooth as shown in figure 3 (Carroll *et al.* 2016, 58). As this is commonly found in

Figure 3. Individual MB11-S325-V0676 displaying several pipe notches, the most prominent of which is encircled in red (photograph taken by the author).

the Middenbeemster collection, many individuals may not have fully intact tooth crowns, obscuring maximum crown measurements. This issue will be addressed through the use of the alternative measurements proposed by Hillson and colleagues (2005) and used by Viciano and colleagues (2011) and (2013) as well as Aris and colleagues (2018).

3.2 Selection of teeth

Following the results of Cardoso (2008), Viciano and colleagues (2013), and Aris and colleagues (2018) the permanent canines and maxillary first molars have been selected as the most appropriate teeth for odontometric sex estimation in non-adults. The permanent canine is known to be the most sexually dimorphic tooth in terms of size, weight, and relative composition and therefore represents the tooth with the most potential to correctly classify sex (Cardoso 2008, 162, 164, 165, 166; Cunningham *et al.* 2016, 18; Lewis 2006, 49; Saunders *et al.* 2007, 735; Schwartz and Dean 2005, 314; Tuttösí and Cardoso 2015, 311; Viciano *et al.* 2011, 102, 105; Viciano *et al.* 2013, 36, 37). Yet the crowns of the permanent canines are not fully formed until between five and six years of age and do not emerge from the jawbone until the individual is approximately 10-11 years old (Cunningham *et al.* 2016, 168; Hillson 2014, 46, 64). This effectively means that even if proven successful, a method based solely on the canine dimensions could not be applied to any children under the age of about five years. Between the ages of about five and 10, such a method could only be applied to those whose jaws had been sufficiently damaged post-mortem to allow for extraction of the unerupted tooth crown. It is for these reasons that the first upper molar will also be examined in all documented individuals from the Middenbeemster collection.

Though results regarding the sexual dimorphism of the upper first molar are more variable than those pertaining to canines, sexual dimorphism has been consistently observed at varying degrees and was used exclusively by Aris and colleagues (2018) to develop discriminant function and logistic regression formulae that correctly classified the biological sex of 90.9% of 22 non-adult individuals from an English collection of known age and sex (Aris *et al.* 2018, 677; Cardoso 2008, 165; Viciano *et al.* 2013, 37). The first molar is the earliest permanent tooth to develop and is often recovered in archaeological contexts (Aris *et al.* 2018, 673; Cunningham *et al.* 2016, 159; Hillson 2014, 45). Mineralization of the upper first molar begins *in utero*; the crown is fully formed at 2.37 years on average and emerges from about six years of age (Cunningham *et al.* 2016, 159, 168; Hillson 2014, 45, 64). The inclusion of the upper first molar in this study will hopefully allow for younger individuals and those whose canines were lost ante- or post-mortem to be incorporated into the study, as well as provide a means by which to compare results with previous studies that analysed the maxillary first molar. All documented individuals in the Middenbeemster collection aged three years and above were examined for this study.

3.3 Measurements

When possible, mesio-distal (MD) and bucco-lingual (BL) cervical and crown dimensions (fig. 4) were taken on canines and maxillary first molars from both sides of the dental arcade in order to maximize the amount of attainable data from each individual. By collecting data from both antimeres, there is twice the potential to obtain any measurements that were unilaterally obscured by pathological or taphonomic changes, tooth-wear, deposits of calculus, or their embedded position in the maxilla or mandible. Before any substitutions are made, the degree of asymmetry between antimeres will be calculated to ensure that there are no significant differences in size. The inclusion of cervical measurements, as proposed and tested by Hillson and colleagues (2005), will allow for the measurement of teeth whose maximum crown diameters have been affected by wear or ante- and post-mortem damage (Hillson *et al*. 2005, 415; Tuttösí and Cardoso 2015, 310). Due to this, cervical measurements are especially beneficial in archaeological contexts, as moderate-to-severe tooth-wear was ubiquitous in many past populations (Hassett 2011, 486; Hillson *et al.* 2005, 416). Following the recommendations of Hillson and colleagues (2005), the measurements taken were aimed at recording a simple maximum diameter without prioritizing making the MD and BL measurements perpendicular to one another (Hillson *et al*. 2005, 415). All measurements were taken using a digital calliper with a precision of 0.01mm. The measurements are described below according to the definitions of Hillson and colleagues (2005) and are pictured in figure 4. As the perpendicular relationship between MD and BL dimensions is difficult to conclusively establish while measuring a tooth with simple callipers, Hillson considers it more practical to define measurements based on consistent and readily identifiable landmarks (Hillson 2005, 260; Hillson *et al*. 2005, 415).

Maximum mesio-distal crown diameter (MDcrn). The distance between two parallel planes aligned with the most mesial and most distal points of the crown (Hillson *et al.* 2005, 418).

Maximum bucco-lingual crown diameter (BLcrn). The distance between two parallel planes aligned with the most buccal and most lingual points of the crown. These points are easily identified on the incisors, canines and premolars as each has only one convexity on the buccal surface. Molars, on the other hand, may prove more problematic as there are usually two or more bulges of varying sizes on the buccal surface of the crown. In this case, Hillson and colleagues (2005) suggest following the recommendation of Tobias (1967) in measuring the distance between the maximum bulge on the buccal surface and the maximum bulge on the lingual surface. This may result in the measurement being taken at an angle relative to the tooth's BL plane, because the two bulges are not necessarily directly opposite one another (Hillson *et al.* 2005, 418).

Mesio-distal cervical diameter (MDcerv). The distance between a point at the central aspect of the cervix on the mesial side and a point at the central aspect of the cervix on the distal side. In canines, the mesial point is more occlusal than the distal point and so the measurement is taken at an angle rather than parallel to the occlusal surface. This is less of an issue in molars, though the axis of the measurement is not used as a defining characteristic here either (Hillson *et al.* 2005, 418).

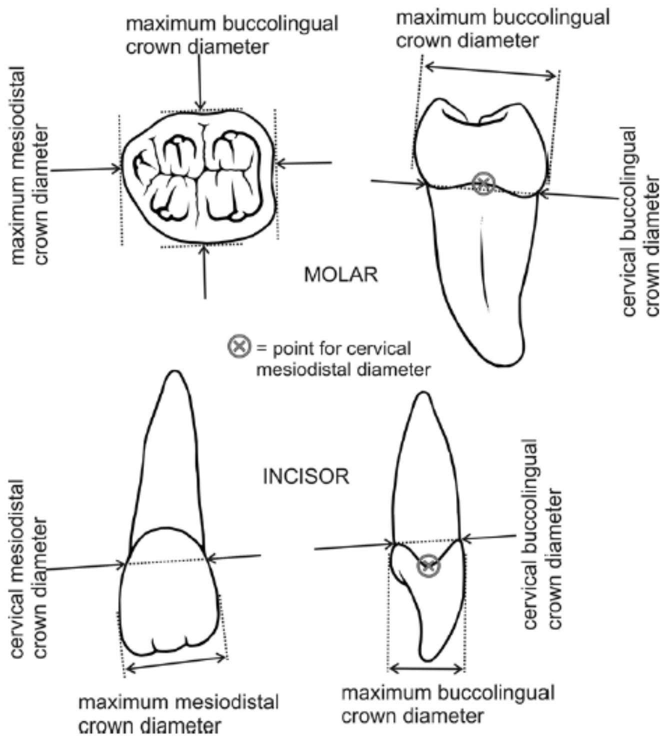


Figure 4. Locations of the maximum MD and BL crown diameters and cervical dimensions for molars and incisors. The criteria for incisors also apply to canines (Hillson 2005, 261).

Bucco-lingual cervical diameter (BLcerv). The distance between a point at the central aspect of the cervix on the buccal side and a point at the central aspect of the cervix on the lingual side. Again, the landmarks may not be directly opposite one another, so its angle relative to the occlusal and MD planes is not strictly defined (Hillson *et al*. 2005, 418).

This study therefore sought to obtain a maximum number of 24 measurements per individual (4 canines and 2 maxillary molars per individual, 4 measurements per tooth). In order to avoid possible biases, all measurements were taken prior to consulting documented sex. Any pathological or taphonomic changes were noted; some measurements were immediately excluded due to severe changes, while others were taken but noted as potentially problematic, such as those displaying enamel hypoplasia. If these measurements do indeed prove problematic, this procedure will help to identify potential confounding factors in future studies. Apart from damage to or loss of a particular tooth, many measurements were unattainable if the tooth in question was still embedded in its socket. In this scenario, cervical measurements were generally more difficult to obtain, but in cases where neighbouring teeth were also retained both the MDcrn and MDcerv measurements were unobservable.

The teeth of 13 randomly selected individuals (208 measurements) from the sample were re-measured by the author after an interval of 6 to 8 weeks in order to calculate intraobserver error. A second observer trained in osteology measured a subsample of 12 individuals (140 measurements) in order to assess interobserver error. This will be particularly valuable for the assessment of maxillary molar measurements, as their landmarks are more difficult to identify and measure consistently than those of canines (Hillson *et al.* 2005, 418; Tuttösí and Cardoso 2015, 310; Viciano *et al.* 2011, 101; Viciano *et al.* 2013, 35). As with Hillson and colleagues (2005), while collecting data the author often had to decide between several possible values (Hillson *et al*. 2005, 424). This was done by repeating the measurement several times and noting the most frequently obtained result.

3.4 Statistical analyses

All statistical analyses were conducted using the software program SPSS 21.0. Intra- and interobserver error were evaluated using the intraclass correlation coefficient (ICC), which serves to quantify the amount of agreement or disagreement between repetitions of the same measurement. The ICC will produce a value between 0 and 1. A high level of similarity will result in a high ICC of close to 1.

The adult sample was tested for normalcy using the Kolmogorov-Smirnov one-sample test. Descriptive statistics were performed for the adult sample as a whole and then by sex. The same was done for the non-adult sample. This procedure calculated the sample size, mean, and standard deviation for each measurement in each subsample.

The mean measurements of antimeres were then compared using a paired-samples *t*-test to evaluate asymmetry. This procedure is used to quantify the extent of variability between samples for any given variable (Cardoso 2008, 160). If no significant asymmetry is detected, it will be possible to maximize the sample size for each measurement through the substitution of measurements from the right dentition when the left is unobservable. An independent-samples *t*-test was used to compare the means of each measurement in the whole adult sample to those of the non-adults. Any variance between the two groups may indicate dimensions that change with age or the presence of the phenomena selective mortality and differential frailty introduced in Chapter 1. The independent-samples *t*-test was lastly used to compare the mean values of each measurement between the male and female adult subsamples. This was done in order to confirm the existence of a significant relationship between sex and dental dimensions before proceeding to regression analysis. The level of sexual dimorphism for each dimension was calculated as follows (Hillson 2005, 267):

$$\% \; of \; sexual \; dimorphism = \frac{\text{mean difference between males and females}}{\text{female mean}} \text{ x } 100$$

Binomial logistic regression is a statistical tool that computes the probability of a binary dependent variable, such as sex, from a series of independent variables, such as dental dimensions. Other classification methods, such as discriminant function analysis, do not give results as probabilities, meaning that logistic regression allows for a more nuanced interpretation of the results (Cardoso 2008, 161; Olivares and Aguilera 2016, 1624; Tuttösí and Cardoso 2015, 308). Unlike discriminant function analysis, logistic regression does not necessitate a normal distribution of the independent variables and equality of variance-covariance matrices in the male and female subsamples, making it a more robust statistical approach (Aris *et al.* 2018, 675, 680; Cardoso 2008, 161; Tuttösí and Cardoso 2015, 308; Viciano *et al*. 2013, 33-34). Univariate logistic regression equations to distinguish between male and female were created for each measurement based on the data collected from the documented adult population. Multivariate logistic regression equations were created for each type of tooth. Additional combinations of measurements were chosen based on which measurements showed the most significant correlations with sex.

These logistic regression formulae were then applied to the non-adult data to obtain sex estimations in the form of probabilities of being either male or female. These results were subsequently tested against the documented sex to determine their accuracy.

The results of these tests will be discussed in the following chapter.

4. Results

4.1 Sample

Of the 118 adult individuals of documented sex, 76 preserved at least one of the teeth under study, allowing for a total of 761 measurements to be taken. Of these 76 adult individuals, 43 were female and 33 male.

Though documentation is available for 33 non-adult individuals between the ages of three and 17 years, only 16 individuals retained at least one accessible canine or maxillary first molar with fully formed crowns (11 females and 5 males). A total of 165 measurements were taken on the non-adult sample.

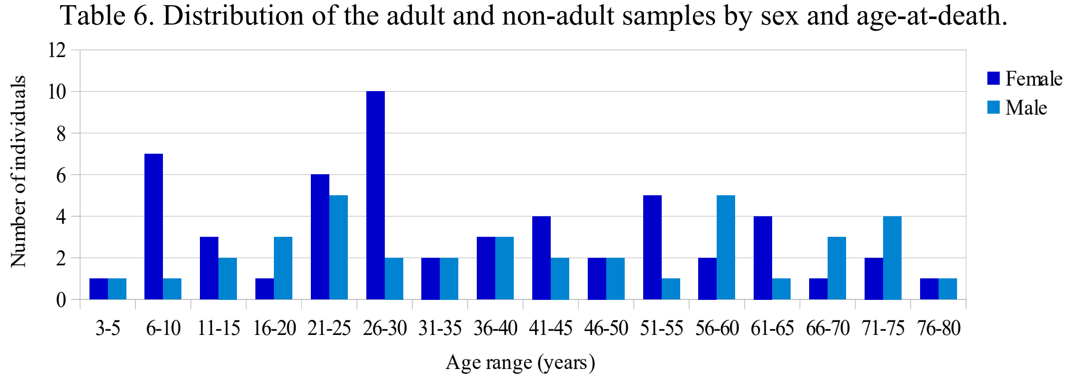
Table 5. Distribution of the non-adult sample by sex and age-at-death in years.

Individual	Age	Sex
S089V0091	3	F
S140V207	5	M
S018V0102	7	F
S365V0773	8	F
S384V0839	8	M
S396V0877	8	F
S248V0393	9	F
S334V0716	9	F
S471V1020	9	F
S286V0469	10	F
S367V803	11	F
S167V0270	12	F
S196V0437	13	M
S522V1127	13	F
S465V1001	15	M
S462V0987	16	M

The distribution of the non-adults sample by age-at-death and sex is given in table 5. Table 6 shows the distribution of both the adult and non-adult samples by sex and age-at-death. The age-at-death among the non-adult sample ranges from three to 16 years, while the adult sample includes individuals aged 19 to 80 years.

All measurements were found to be normally distributed within the adult sample through the Kolmogorov-Smirnov one-sample test. The results of these tests can be found in appendix 1.

Table 6. Distribution of the adult and non-adult samples by sex and age-at-death.



4.2 Intra- and interobserver error

As outlined in the previous chapter, 13 individuals were re-examined by the original observer after a period of approximately 6 to 8 weeks in order to assess intraobserver error. For the assessment of interobserver error, 9 individuals were examined by a separate observer trained in osteology but not trained in odontometrics. This resulted in comparisons between 392 measurements for intraobserver error and 258 measurements for interobserver error. When compared to the original measurements obtained by the author, all measurements aside from the mesio-distal measurements of the upper first molars showed excellent concordance (tables 7 and 8).

Table 7. Assessment of intraobserver error: intraclass correlation coefficients.

Measurement	*N*	ICC
1.3 MDcrn	10	0.976
1.3 BLcrn	12	0.997
1.3 MDcerv	12	0.990
1.3 BLcerv	12	0.996
2.3 MDcrn	9	0.991
2.3 BLcrn	9	0.999
2.3 MDcerv	9	0.966
2.3 BLcerv	9	0.918
3.3 MDcrn	6	0.995
3.3 BLcrn	7	0.999
3.3 MDcerv	6	0.985
3.3 BLcerv	7	0.995
4.3 MDcrn	8	0.631
4.3 BLcrn	10	0.999
4.3 MDcerv	12	0.948
4.3 BLcerv	12	0.995
1.6 MDcrn	4	0.985
1.6 BLcrn	11	0.990
1.6 MDcerv	3	**0.011**
1.6 BLcerv	7	0.987
2.6 MDcrn	5	0.874
2.6 BLcrn	8	0.991
2.6 MDcerv	3	**0.580**
2.6 BLcerv	5	0.988

N number of teeth. An ICC of 0.6 is considered acceptable. ICCs below this threshold are marked in bold.

Table 8. Assessment of interobserver error: intraclass correlation coefficients.

Measurement	*N*	ICC
1.3 MDcrn	6	0.868
1.3 BLcrn	8	0.986
1.3 MDcerv	8	0.994
1.3 BLcerv	8	0.974
2.3 MDcrn	5	0.995
2.3 BLcrn	5	0.941
2.3 MDcerv	5	0.997
2.3 BLcerv	5	0.972
3.3 MDcrn	5	0.959
3.3 BLcrn	5	0.984
3.3 MDcerv	5	0.967
3.3 BLcerv	4	0.977
4.3 MDcrn	7	0.998
4.3 BLcrn	7	0.963
4.3 MDcerv	8	0.944
4.3 BLcerv	7	0.972
1.6 MDcrn	3	**-0.241**
1.6 BLcrn	7	0.976
1.6 MDcerv	3	0.777
1.6 BLcerv	4	0.789
2.6 MDcrn	4	0.852
2.6 BLcrn	5	0.943
2.6 MDcerv	2	0.998
2.6 BLcerv	3	0.980

N number of teeth. ICCs below 0.6 are marked in bold.

4.3 Comparisons between antimeres

The measurements of antimere pairs in the adult and non-adult samples were compared separately using the paired-sample *t*-test. The results of these tests are presented in tables 9 and 10. None of the measurements differed to a statistically significant degree ($P \leq 0.05$), meaning that any measurements that were present on one side of the dental arcade but not the other could be substituted for the missing or unobservable element from the same individual, maximizing the number of observations for each measurement. Measurements from the right side of the dental arcade, quadrants 1 and 4, were used preferentially, with antimeric measurements from quadrants 2 and 3 used when necessary and possible. All subsequent statistical analyses were performed on the basis of measurement-type from this combined dataset.

Table 9. Paired-sample *t*-tests of antimere odontometrics in the adult sample.

Tooth Type	Paired measurements	*N*	Mean	SD	*t*	Sig. (2-tailed)
Upper Canine	1.3 MDcrn	17	7.411	0.352	-0.819	0.425
	2.3 MDcrn	17	7.434	0.334		
	1.3 BLcrn	27	8.1267	0.643	1.68	0.105
	2.3 BLcrn	27	8.0793	0.649		
	1.3 MDcerv	19	5.260	0.486	-1.103	0.285
	2.3 MDcerv	19	5.312	0.524		
	1.3 BLcerv	23	7.7513	0.78288	1.874	0.074
	2.3 BLcerv	23	7.6761	0.77156		
Lower Canine	3.3 MDcrn	20	6.588	0.423	2.056	0.54
	4.3 MDcrn	20	6.518	0.366		
	3.3 BLcrn	40	7.613	0.661	1.012	0.318
	4.3 BLcrn	40	7.585	0.638		
	3.3 MDcerv	34	5.084	0.511	0.800	0.429
	4.3 MDcerv	34	5.069	0.495		
	3.3 BLcerv	37	7.479	0.726	1.656	0.106
	4.3 BLcerv	37	7.408	0.667		
Upper Molar	1.6 MDcrn	1a	11.060	n/a	n/a	n/a
	2.6 MDcrn	1a	11.330			
	1.6 BLcrn	19	11.266	0.805	-1.387	0.0182
	2.6 BLcrn	19	11.302	0.829		
	1.6 MDcerv	2	9.010	1.301	5.000	0.126
	2.6 MDcerv	2	8.985	1.308		
	1.6 BLcerv	8	10.339	0.849	-1.219	0.262
	2.6 BLcerv	8	10.469	0.787		

N number of observations. SD standard deviation of the mean. *t* Student's *t*-test values statistically significant at the $P \leq 0.05$ level.

Table 10. Paired-sample *t*-tests of antimere odontometrics in the non-adult sample.

Tooth Type	Paired measurements	*N*	Mean	SD	*t*	Sig. (2-tailed)
Upper Canine	1.3 MDcrn 2.3 MDcrn	5 5	7.340 7.354	0.626 0.677	-0.305	0.776
	1.3 BLcrn 2.3 BLcrn	5 5	8.150 8.070	0.901 0.961	2.401	0.74
	1.3 MDcerv 2.3 MDcerv	5 5	5.262 5.312	0.327 0.336	-1.327	0.255
	1.3 BLcerv 2.3 BLcerv	5 5	7.754 7.740	0.803 0.820	0.263	0.805
Lower Canine	3.3 MDcrn 4.3 MDcrn	7 7	6.133 6.090	0.302 0.321	0.989	0.361
	3.3 BLcrn 4.3 BLcrn	8 8	7.288 7.309	0.673 0.612	-0.362	0.728
	3.3 MDcerv 4.3 MDcerv	8 8	4.909 4.9063	0.483 0.556	0.046	0.965
	3.3 BLcerv 4.3 BLcerv	7 7	7.1929 7.1643	0.783 0.652	0.413	0.694
Upper Molar	1.6 MDcrn 2.6 MDcrn	3 3	10.223 10.110	0.268 0.429	1.190	0.356
	1.6 BLcrn 2.6 BLcrn	4 4	11.115 11.038	0.296 0.345	0.935	0.419
	1.6 MDcerv 2.6 MDcerv	3 3	7.790 7.540	0.270 0.104	1.916	0.195
	1.6 BLcerv 2.6 BLcerv	4 4	10.005 9.945	0.496 0.447	1.492	0.232

N number of observations. SD standard deviation of the mean. *t* Student's *t*-test values statistically significant at the $P \leq 0.05$ level.

4.4 Comparison between the adult and non-adult samples

Independent-samples *t*-test comparisons between the means of each measurement showed no significant differences between the adult and non-adult samples (tab. 11). This indicates that the odontometrics of the non-adult sample were not significantly affected by selective mortality as discussed in Chapter 1 and that any statistical models developed on the adult sample can be reliably applied to the non-adult sample. The descriptive statistics for each measurement type within each sample are also listed in table 11.

Table 11. Descriptive statistics and independent sample *t*-tests comparing the adult and non-adult samples for each measurement.

Measurement	Adult sample			Non-adult sample			*t*	Sig. (2-tailed)
	N	Mean	SD	*N*	Mean	SD		
UC MDcrn	35	7.421	0.397	7	7.229	0.545	-1.096	0.279
UC BLcrn	46	8.030	0.631	7	7.893	0.857	-0.509	0.613
UC MDcerv	45	5.339	0.511	7	5.317	0.307	-0.108	0.914
UC BLcerv	50	7.583	0.765	7	7.473	0.817	-0.354	0.725
LC MDcrn	36	6.459	0.347	9	6.316	0.527	-0.992	0.327
LC BLcrn	64	7.536	0.608	9	7.366	0.597	-0.789	0.433
LC MDcerv	58	5.042	0.504	9	4.901	0.521	-0.777	0.440
LC BLcerv	59	7.338	0.665	9	7.163	0.588	-0.743	0.460
UM MDcrn	26	10.369	0.513	7	10.506	0.426	0.645	0.524
UM BLcrn	39	11.163	0.619	11	10.992	0.568	-0.824	0.414
UM MDcerv	21	8.01	0.825	8	8.266	1.121	0.687	0.498
UM BLcerv	34	10.293	0.654	11	10.026	0.889	-1.079	0.287

UC upper canine. LC lower canine. UM upper molar. *N* number of observations. SD standard deviation of the mean. *t* independent-sample t-test absolute values. No differences are significant at the ≤ 0.05 level.

4.5 Univariate analysis of sexual dimorphism in the adult sample

The female and male means of each measurement type were compared using independent-samples *t*-tests to determine the degree of significance of any sex-linked differences. The results are presented in table 12 along with descriptive statistics grouped by sex.

The level of sexual dimorphism, given as a percentage, quantifiably illustrates which measurements show the greatest amount of sexual dimorphism. The cervical measurements of both the upper and lower canines display the greatest level of sexual dimorphism at 13-14%, while the upper molar MD crown diameter is the least sexually dimorphic at 1.34%. The absolute values of the *t*-test results give an indication of the distance between groups and the degree of overlap between distributions (Tuttösí and Cardoso 2015, 308). Accordingly, the odontometric with the most distance and least overlap between the sexes is the lower canine BL cervical diameter. The significance levels for both the independent sample *t*-tests and Pearson's correlation coefficients show that there is a highly significant relationship between sex and tooth size in all measurements except the MD measurements of the upper molars. All canine measurements were found to be significant at the ≤ 0.01 level.

Table 12. Descriptive statistics and independent sample *t*-tests comparing the male and female adult subsamples.

Measurement	Female			Male			Level of sex. dim.	*t*	*r*	*d*	Sig. (2-tailed)
	N	Mean	SD	*N*	Mean	SD					
UC MDcrn	25	7.308	0.367	10	7.701	0.339	5.37%	-2.917	0.453	1.110879	0.006**
UC BLcrn	29	7.736	0.495	17	8.531	0.517	10.27%	-5.171	0.615	1.570453	0.000**
UC MDcerv	27	5.507	0.408	18	5.743	0.364	13.29%	-5.661	0.653	1.742493	0.000**
UC BLcerv	26	7.112	0.605	24	8.093	0.571	13.80%	-5.887	0.648	1.668329	0.000**
LC MDcrn	30	6.392	0.307	6	6.790	0.374	6.22%	-2.799	0.433	1.163268	0.008**
LC BLcrn	35	7.197	0.451	29	7.945	0.518	10.39%	-6.171	0.617	1.539127	0.000**
LC MDcerv	32	4.754	0.379	26	5.397	0.405	13.53%	-6.235	0.640	1.639975	0.000**
LC BLcerv	29	6.860	0.462	30	7.799	0.481	13.69%	-7.651	0.712	1.993194	0.000**
UM MDcrn	15	10.311	0.516	11	10.449	0.521	1.34%	-0.673	0.136	0.266859	0.507
UM BLcrn	22	10.913	0.513	17	11.487	0.606	5.25%	-3.195	0.465	1.020286	0.003**
UM MDcerv	11	7.744	0.649	10	8.298	0.930	7.16%	-1.597	0.344	0.691300	0.127
UM BLcerv	20	10.060	0.577	14	10.627	0.629	5.64%	-2.720	0.433	0.940157	0.010**

UC upper canine. LC lower canine. UM upper molar. *N* number of observations. SD standard deviation of the mean. Level of sex. dim. Level of sexual dimorphism. *t* independent-sample t-test absolute values. *r* Pearson's correlation coefficient. *d* Cohen's effect size. * significant at the ≤ 0.05 level. ** significant at the ≤ 0.01 level.

4.6 Logistic regression analysis

4.6.1 Logistic regression models

Twenty-five univariate and multivariate logistic regression models were developed using the adult data and are presented in tables 13a and 13b. The selection of measurement combinations was guided by the above results as well as previous research conducted by Aris and colleagues (2018), Cardoso (2008), Tuttösí and Cardoso (2015), and Viciano and colleagues (2013). The practicalities of working with an archaeological collection were also considered. Models 1 to 3 use the full complement of measurements for each tooth type. Models 4, 8, and 11 employ only the crown measurements of each type of tooth, while models 5, 9, and 12 use cervical measurements by tooth. Models 6, 7, 10, and 13 were developed based on the results of backwards logistic regression, which serves to identify the most effective combinations of variables by progressively eliminating those that do not have a consequential effect on classification. Models 14 through 25 each rely on a single measurement. Sixteen of the 25 models surpassed the 75% correct classification threshold considered acceptable in methods for non-adult sex estimation. Half of these sixteen models also surpass the 85% correct classification rates considered acceptable in adult sex estimation. All models developed on both the upper and lower canines were found to significantly improve classification accuracy over the base model, though the upper canine

MDcrn (model 14) failed to meet the 75% threshold. The predicted values, on which the above classification percentages are based, and standardized residuals of each model are listed by individual in appendix 2. The large standardized residuals of individual S521V1150, a 69-year-old male, indicate that he is an outlier within the sample.

Table 13a. Summaries of logistic regression models 1 through 10.

Model number	Equation	Measurement(s) used in model	*N*		% correct			Model summary		Omnibus tests of model coefficients		
			Females	Males	Females	Males	Total	AIC	Nagelkerke R square	Chi-square	df	Sig.
1	L_1	UC MDcrn UC BLcrn UC MDcerv UC BLcerv	19	10	94.7	90.0	**93.1**	18.69	0.867	28.673	4	0.000**
2	L_2	LC MDcrn LC BLcrn LC MDcerv LC MDcrn	21	6	95.2	66.7	**88.9**	16.742	0.849	21.862	4	0.000**
3	L_3	UM MDcrn UM BLcrn UM MDcerv UM BLcerv	10	9	80.0	66.7	73.7	29.408	0.405	6.879	4	0.142
4		UC MDcrn UC BLcrn	22	10	90.9	60.0	**81.3**	30.917	0.522	14.832	2	0.001**
5		UC MDcerv UC BLcerv	24	17	79.2	70.6	**75.6**	37.504	0.599	24.133	2	0.000**
6		UC MDcrn UC BLcrn UC MDcerv	19	10	94.7	90.0	**93.1**	14.743	0.866	28.62	3	0.000**
7	L_7	UC BLcrn UC MDcerv	19	9	89.5	90.0	**89.7**	16.583	0.832	26.779	2	0.000**
8	L_8	LC MDcrn LC BLcrn	28	6	100	83.3	**97.1**	17.169	0.747	20.519	2	0.000**
9	L_9	LC MDcerv LC BLcerv	27	26	88.9	84.6	**86.8**	42.455	0.670	37.000	2	0.000**
10		LC MDcrn LC BLcrn LC BLcerv	21	6	95.2	66.7	**88.9**	14.998	0.843	21.606	3	0.000**

UC upper canine. LC lower canine. UM upper molar. *N* number of individuals used in the development of the model. AIC Akaike information criterion. Total percentages correct above 75% are marked in bold. * model significant at the ≤ 0.05 level. ** significant at the ≤ 0.01 level.

Table 13b. Summaries of logistic regression models 11 through 25

Model number	Equation	Measurement(s) used in model	*N*		% correct			Model summary		Omnibus tests of model coefficients		
			Females	Males	Females	Males	Total	AIC	Nagelkerke R square	Chi-square	df	Sig.
11		UM MDcrn UM BLcrn	15	11	86.7	27.3	61.5	39.251	0.108	2.175	2	0.337
12		UM MDcerv UM BLcerv	11	10	63.6	70.0	66.7	28.606	0.353	6.458	2	0.040*
13		UM BLcrn UM BLcerv	18	12	83.3	41.7	66.7	41.564	0.201	4.817	2	0.090
14		UC MDcrn	25	10	84.0	40.0	71.4	38.052	0.287	7.827	1	0.005**
15	L_{15}	UC BLcrn	29	17	89.7	70.6	**82.6**	43.480	0.503	21.122	1	0.000**
16	L_{16}	UC MDcerv	27	18	81.5	72.2	**77.8**	41.293	0.546	23.278	1	0.000**
17	L_{17}	UC BLcerv	26	24	80.8	70.8	**76.0**	47.347	0.539	25.887	1	0.000**
18	L_{18}	LC MDcrn	30	6	100.0	50.0	**91.7**	28.725	0.325	7.716	1	0.005**
19	L_{19}	LC BLcrn	35	29	82.9	75.9	**79.7**	62.248	0.499	29.912	1	0.000**
20	L_{20}	LC MDcerv	32	26	81.3	73.1	**77.6**	54.466	0.531	29.317	1	0.000**
21	L_{21}	LC BLcerv	29	30	82.8	83.3	**83.1**	45.935	0.655	39.839	1	0.000**
22		UM MDcrn	15	11	80.0	9.1	50.0	38.942	0.025	0.484	1	0.487
23	L_{23}	UM BLcrn	22	17	86.4	58.8	74.4	47.283	0.307	10.14	1	0.001**
24		UM MDcerv	11	10	81.8	50.0	66.7	30.443	0.157	2.621	1	0.105
25		UM BLcerv	20	14	80.0	57.1	70.6	42.962	0.254	7.107	1	0.008**

UC upper canine. LC lower canine. UM upper molar. *N* number of individuals used in the development of the model. AIC Akaike information criterion. Total percentages correct above 75% marked in bold. * model significant at the ≤ 0.05 level. ** significant at the ≤ 0.01 level.

Table 14a. Variables in equations 1, 2, 3, 7 and 8.

Equation number	Measurement(s) used	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
L_1	UC MDcrn	-8.243	8.309	0.984	1	0.321	0.000	0.000	3112.361
	UC BLcrn	10.422	11.207	0.865	1	0.352	33597.405	0.000	1.163E+14
	UC MDcerv	15.697	10.111	2.410	1	0.121	6562507.754	0.016	2.653E+15
	UC BLcerv	1.536	6.623	0.054	1	0.817	4.647	0.000	2017568.132
	Constant	-125.319	80.134	2.446	1	0.118	0.000		
L_2	LC MDcrn	-21.015	23.077	0.829	1	0.362	0.000	0.000	32863939099.527
	LC BLcrn	23.531	25.900	0.825	1	0.364	16571133669.113	0.000	1.843E+32
	LC MDcerv	2.043	4.537	0.203	1	0.653	7.71 0	0.001	56130.733
	LC BLcerv	25.727	30.603	0.707	1	0.401	1.490E+11	0.000	1.670E+037
	Constant	-245.346	285.127	0.740	1	0.390	0.000		
L_3	UM MDcrn	-1.095	1.477	0.550	1	0.459	0.334	0.018	6.053
	UM BLcrn	3.427	2.347	2.132	1	0.144	30.781	0.309	3061.779
	UM MDcerv	1.198	0.772	2.410	1	0.121	3.313	0.730	15.031
	UM BLcerv	0.167	1.893	0.008	1	0.930	1.182	0.029	48.319
	Constant	-38.402	23.976	2.565	1	0.109	0.000		
L_7	UC BLcrn	4.619	2.752	2.818	1	0.093	101.399	0.461	22300.943
	UC MDcerv	9.826	5.054	3.779	1	0.052	18500.282	0.922	371128576.2
	Constant	-93.817	47.364	3.923	1	0.048	0.000		
L_8	LC MDcrn	-5.830	4.863	1.437	1	0.231	0.003	0.000	40.481
	LC BLcrn	13.913	7.290	3.642	1	0.056	1102240.288	0.687	1.770E+12
	Constant	-70.186	37.450	3.512	1	0.061	0.000		

UC upper canine. LC lower canine. UM upper molar. B intercept coefficient. S.E. standard error of B. Wald Wald Chi-square test. df degrees of freedom for the Wald Chi-square test. Sig. Significance of Wald Chi-square test over null hypothesis that the constant equals 0. Exp (B) exponentiation of B coefficient. 95% C.I. for Exp (B) Confidence interval for the exponentiation of the coefficient at the 95% level.

Table 14b. Variables in equations 9, 15-21 and 23.

Equation number	Measurement(s) used	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
L_9	LC MDcerv	1.168	1.334	0.766	1	0.381	3.215	0.235	43.923
	LC BLcerv	3.602	1.461	6.081	1	0.014	36.683	2.094	642.583
	Constant	-32.419	8.591	14.240	1	0.000	0.000		
L_{15}	UC BLcrn	3.157	0.946	11.125	1	0.001	23.491	3.676	150.131
	Constant	-26.279	7.802	11.344	1	0.001	0.000		
L_{16}	UC MDcerv	4.006	1.151	12.482	1	0.000	58.344	6.113	556.822
	Constant	-22.391	6.282	12.74	1	0.000	0.000		
L_{17}	UC BLcerv	2.758	0.767	12.923	1	0.000	15.764	3.505	70.903
	Constant	-21.112	5.891	12.846	1	0.000	0.000		
L_{18}	LC MDcrn	4.381	1.850	5.606	1	0.018	79.937	2.127	3004.595
	Constant	-30.509	12.395	6.058	1	0.014	0.000		
L_{19}	LC BLcrn	3.251	0.823	15.618	1	0.000	25.811	5.148	129.414
	Constant	-24.846	6.281	15.647	1	0.000	0.000		
L_{20}	LC MDcerv	4.073	1.022	15.883	1	0.000	58.704	7.922	435.012
	Constant	-20.866	5.210	16.039	1	0.000	0.000		
L_{21}	LC BLcerv	4.300	1.107	15.088	1	0.000	73.673	8.416	644.945
	Constant	-31.466	8.105	15.073	1	0.000	0.000		
L_{23}	UM BLcrn	2.263	0.896	6.385	1	0.012	9.611	1.662	55.598
	Constant	-25.553	10.022	6.501	1	0.011	0.000		

UC upper canine. LC lower canine. UM upper molar. B intercept coefficient. S.E. standard error of B. Wald Wald Chi-square test. df degrees of freedom for the Wald Chi-square test. Sig. Significance of Wald Chi-square test over null hypothesis that the constant equals 0. Exp (B) exponentiation of B coefficient. 95% C.I. for Exp (B) Confidence interval for the exponentiation of the coefficient at the 95% level.

4.6.2 Logit equations

Based on a critical evaluation of the above models in conjunction with a consideration of which combinations of measurements would be most practical within archaeological investigations, fourteen models were selected to create logit equations. The variables for the equations are listed in tables 14a and 14b and the equations themselves are presented in table 15.

Table 15. Logit equations with variables.

L_1 = -125.319 – 8.243(UC MDcrn) + 10.422(UC BLcrn) + 15.697(UC MDcerv) + 1.536(UC BLcerv)
L_2 = -245.346 – 21.015(LC MDcrn) + 23.531(LC BLcrn) + 2.043(LC MDcerv) + 25.727(LC BLcerv)
L_3 = -38.402 – 1.095(UM MDcrn) +3.427(UM BLcrn) + 1.198(UM MDcerv) + 0.167(UM BLcerv)
L_7 = -93.817 + 4.619(UC BLcrn) + 9.826(UC MDcerv)
L_8 = -70.186 – 5.830(LC MDcrn) + 13.913(LC BLcrn)
L_9 = -32.419 + 1.168(LC MDcerv) + 3.602(LC BLcerv)
L_{15} = -26.279 + 3.157(UC BLcrn)
L_{16} = -22.391 + 4.006(UC MDcerv)
L_{17} = -21.112 + 2.758(UC BLcerv)
L_{18} = -30.509 + 4.381(LC MDcrn)
L_{19} = -24.846 + 3.251(LC BLcrn)
L_{20} = -20.866 + 4.073(LC MDcerv)
L_{21} = -31.466 + 4.300(LC BLcerv)
L_{23} = -25.553 + 2.263(UM BLcrn)

4.6.3 Estimated regression equations

Through algebraic transformation, the logit equations in which L = log(odds) can be transformed to calculate the probability of group membership ($\hat{p}$) for each individual. The formula for this transformation is given below.

$$\hat{p} = \frac{e^{L_x}}{1 + e^{L_x}}$$

In this study, the probability score predicts group membership in the male category, i.e. the higher the probability, the more likely it is that the individual in question is male. The probably of the individual being female can be calculated as 1-$\hat{p}$.

4.6.4 Application to non-adult sample

The estimated regression equations were applied to the non-adult sample and group membership was estimated based on the probability values, sectioned at 0.5 (male > 0.5 > female). A probability value of exactly 0.5 would be classified as indeterminate.

4.7 Results on the non-adult sample

4.7.1 Classification accuracies by model

In order to validate the method, the results of each estimated regression equation were compared to the individual's documented sex. A summary of this validation can be found in table 16.

Nine of the equations tested produced accuracies above 75%. Seven of these nine produced accuracies above 85%. The most successful equation was L_{20}, based on model 20 using the lower canine mesio-distal cervical diameter. This estimated regression equation proved to be correct in all nine cases. However, the small sample size cautions against putting too much emphasis this flawless assessment, as more extensive repetition may lead to misclassifications. Furthermore, model 20 returned a correct classification of 77.6% in the adult sample (tab. 13b), showing that the model is not always perfect. No regression equation using the upper molars reached a correct classification rate of 75% or above, nor did models 15, 17, or 18, all of which were based on single canine measurements.

Table 16. Summaries and correct classification rates of the estimated regression equations as applied to the non-adult sample.

Logit equation	Measurement(s) used	*N*	Females		Males		Total
			n	**%**	***n***	**%**	
L_1	UC MDcrn UC BLcrn UC MDcerv UC BLcerv	7	4	100.0	3	66.7	**85.7**%
L_2	LC MDcrn LC BLcrn LC MDcerv LC BLcerv	9	6	100.0	3	66.7	**88.9**%
L_3	UM MDcrn UM BLcrn UM MDcerv UM BLcerv	4	1	100.0	3	33.3	50.0%
L_7	UC BLcrn UC MDcerv	7	4	100.0	3	66.7	**85.7**%
L_8	LC MDcrn LC Blcrn	9	6	100.0	3	66.7	**88.9**%
L_9	LC MDcerv LC BLcerv	9	6	83.3	3	100.0	**88.9**%
L_{15}	UC BLcrn	7	4	75.0	3	66.7	71.4%
L_{16}	UC MDcerv	7	4	100.0	3	66.7	**85.7**%
L_{17}	UC BLcerv	7	4	75.0	3	66.7	71.4%
L_{18}	LC MDcrn	9	6	83.3	3	33.3	66.7%
L_{19}	LC BLcrn	9	6	83.3	3	66.7	**77.8**%
L_{20}	LC MDcerv	9	6	100.0	3	100.0	**100.0**%
L_{21}	LC BLcerv	9	6	83.3	3	66.7	**77.8**%
L_{23}	UM BLcrn	11	8	62.5	3	33.3	54.6%

UC upper canine. LC lower canine. UM upper molar. *N* total number of individuals. *n* number of individuals by sex. % percentage of correct classifications by sex. Total total percentage of correct classifications. Correct classification rates above 75% marked in bold.

4.7.2 Classification accuracies by individual

Table 17 presents a summary of which models correctly classified sex for each non-adult separately. Of the sixteen individuals from whom at least one applicable measurement could be obtained, sex estimations could not be made for one, S384V0839. The only measurements available from this individual were both the crown and cervical mesio-distal diameters of a first maxillary molar. These measurements were not found to be significantly related to sex in the univariate and logistic regression analyses performed on the adult sample and as such, no estimated regression equations were developed solely on these measurements. L_{23} was the only equation that was able to be applied to four individuals. The equation classified three of those four correctly, though overall the equation correctly classified only 54.6% of the non-adult sample. The percentage correct presented in table 17 is only useful in identifying individuals as potential outliers, as it is dependent on the total number of equations applied to the individual, which is purely a function of differential preservation of and access to pertinent teeth.

Table 17. Classification accuracies per individual for each regression equation.

Individual	Age	Sex	L_1	L_2	L_3	L_7	L_8	L_9	L_{15}	L_{16}	L_{17}	L_{18}	L_{19}	L_{20}	L_{21}	L_{23}	% correct
S089V0091	3	F	-	-	-	-	-	-	-	-	-	-	-	-	-	N	0.00
S140V207	5	M	-	-	Y	-	-	-	-	-	-	-	-	-	-	Y	100.00
S018V0102	7	F	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	100.00
S365V0773	8	F	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	100.00
S396V0877	8	F	-	-	Y	-	-	-	-	-	-	-	-	-	-	Y	100.00
S248V0393	9	F	-	Y	-	-	Y	Y	-	-	-	Y	Y	Y	Y	-	100.00
S334V0716	9	F	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	100.00
S471V1020	9	F		-	-	-	-	-	-	-	-	-	-	-	-	Y	100.00
S286V0469	10	F	Y	Y	-	Y	Y	N	N	Y	N	N	N	Y	N	N	**46.15**
S367V803	11	F	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	100.00
S167V0270	12	F	-	Y	-	-	Y	Y	-	-	-	Y	Y	Y	Y	Y	100.00
S196V0437	13	M	N	N	N	N	N	Y	N	Y	N	N	N	Y	N	N	**21.43**
S522V1127	13	F	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	92.31
S465V1001	15	M	Y	Y	-	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	-	91.67
S462V0987	16	M	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	78.57

Age is given in years. F documented female. M documented male. Y yes sex match. N no sex match. % correct overall percentage of correct estimations per individual. % correct under 50% marked in bold.

On the whole, the number of equations applicable to each individual increased with age. Two individuals were classified incorrectly in more than half the trials. One of these individuals, a 13-year-old male (S196V0437), was classified correctly by only three of fourteen applicable equations. The other, a 10-year-old female (S286V0469), was classified correctly in five of thirteen applicable equations. Both skeletons were re-examined for signs of systemic pathological changes that may have affected their tooth dimensions. No indications of such a systemic disturbance were observed in individual S286V469, but individual S196V0437 was clearly affected by microdontia of the maxillary second incisors (1.2 and 2.2). This condition will be further discussed in Chapter 5. As microdontia is readily visible and may affect the dimensions of other teeth as well, the classification accuracies of each equation were recalculated with S196V0437 removed from the sample (Cantekin and Celikoglu 2015, 188). These are given in table 18.

Table 18. Summaries and correct classification rates of the estimated regression equations as applied to the non-adult sample with S196V0437 removed.

Logit equation	**Measurement(s) used**	***N***	**Females**		**Males**		**Total**
			n	**%**	***n***	**%**	
L_1	UC MDcrn UC BLcrn UC MDcerv UC BLcerv	6	4	100.0	2	100.0	**100.0%**
L_2	LC MDcrn LC BLcrn LC MDcerv LC BLcerv	8	6	100.0	2	100.0	**100.0%**
L_3	UM MDcrn UM BLcrn UM MDcerv UM BLcerv	3	1	100.0	2	50.0	66.7%
L_7	UC BLcrn UC MDcerv	6	4	100.0	2	100.0	**100.0%**
L_8	LC MDcrn LC BLcrn	8	6	100.0	2	100.0	**100.0%**
L_9	LC MDcerv LC BLcerv	8	6	83.3	2	100.0	**87.5%**
L_{15}	UC BLcrn	6	4	75.0	2	100.0	**83.3%**
L_{16}	UC MDcerv	6	4	100.0	2	50.0	**83.3%**
L_{17}	UC BLcerv	6	4	75.0	2	100.0	**83.3%**
L_{18}	LC MDcrn	8	6	83.3	2	50.0	**75.0%**
L_{19}	LC BLcrn	8	6	83.3	2	100.0	**87.5%**
L_{20}	LC MDcerv	8	6	100.0	2	100.0	**100.0%**
L_{21}	LC BLcerv	8	6	83.3	2	100.0	**87.5%**
L_{23}	UM BLcrn	10	8	62.5	2	50.0	60.0%

N total number of individuals. *n* number of individuals by sex. % percentage of correct classifications by sex. Total total percentage of correct classifications. Correct classification rates above 75% marked in bold.

4.7.3 Classification accuracies with S196V0437 removed

With S196V0437 removed, all models based on canine measurements returned correct classification rates above 75% (tab. 18). Only those using measurements of the upper molar, models 3 and 23, did not meet this threshold. The correct classification rates of the male subsample increased substantially, but this improvement is accentuated by the very small sample size (n = 2).

4.8 Summary

Overall, the logistic regression equations developed on the adult odontometrics proved highly successful on the non-adult sample from the same population. The equations were generally better at classifying females than males, but this is likely due to the extremely small male sample of 3 individuals. These results will be discussed in further detail in the following chapter.

5. Discussion

In this chapter, the promising results of this research will be interpreted and put into context with current methods in non-adult sex estimation. Limits to the study and to the method will be discussed and those individuals who did not fit the models will be presented in further detail.

5.1 Interpreting results

5.1.1 Intra- and interobserver error

In the evaluation of both intra- and interobserver error, the intraclass correlation coefficients (ICC) indicate near-perfect reproducibility for all canine measurements, with only the ICC of the interobserver comparison of 1.3 MDcrn (0.868) falling below 0.9. The results pertaining to the maxillary first molars are more variable, a trend that has also been observed in previous studies (Hillson *et al*. 2005, 418; Tuttösí and Cardoso 2015, 310; Viciano *et al*. 2011, 101; Viciano *et al*. 2013, 41). The intraobserver comparison of the MDcerv measurements of both maxillary molars showed poor concordance with ICCs of below 0.6. However, when taken by the second observer, this measurement produced ICCs well above the 0.6 threshold in both 1.6 and 2.6. Similarly, though in the evaluation of intraobserver error the ICCs of the maxillary molar MDcrn for both 1.6 and 2.6 were above 0.98, in interobserver error these ICCs varied greatly from -0.241 to 0.852. These discrepancies highlight the difficulties in obtaining consistent measurements from molars, whose morphologies do not present clear landmarks from which to take measurements (Hillson *et al*. 2005, 424; Viciano *et al*. 2013, 35, 41). As pointed out by Hillson and colleagues (2005), even small rotations during measurement can result in variations of 1mm or more and an element of personal judgement is required (Hillson *et al*. 2005, 414). That being said, all other molar measurements showed good-to-excellent concordance.

These results are in accordance with those obtained by Tuttösí and Cardoso (2015), Viciano *et al.* (2011), and Viciano and colleagues (2013), in which the canines were invariably found to produce more consistent measurements than molars (Tuttösí and Cardoso 2015, 310; Viciano *et al*. 2011, 101; Viciano *et al*. 2013, 35, 41). Aris and colleagues' (2018) study, which dealt only with the maxillary first molar, did not evaluate interobserver error, but did attain an extremely small level of intraobserver error for all measurements (Aris *et al*. 2018, 675-676). The consistency achieved by Aris and colleagues (2018) in molar measurements can be ascribed to the primary researcher's high level of experience with molar odontometrics, especially considering the second round of observations was undertaken only 7 days after the initial collection of measurements from the entire adult and non-adult sample, consisting of 149 individuals yielding 1490 measurements (Aris *et al*. 2018, 674).

The fact that an observer with no previous training in odontometrics was able to nearly duplicate most of the original measurements bodes well for future applications of the method presented in this thesis. Unlike morphological methods of sex estimation, which are improved by experience with both the method and the range and level of sexual dimorphism within the skeletal collection under study, this analysis shows that odontometric methods such as this are readily accessible to researchers trained in osteology but unfamiliar with the method or the collection. All of the morphological methods of non-adult sex estimation described in Chapter 2 reported high levels of interobserver error when tested by researchers not associated with the creation of the method (Cardoso and Saunders 2008, 25; Luna *et al.* 2017, 899-900; Olivares and Aguilera 2016, 1628-1629; Wilson *et al.* 2011, 39-40). These were in large part attributed to the subjectivity inherent in interpreting non-metric traits whose full range of expression was not adequately described by the methods' developers (Cardoso and Saunders 2008, 25; Loth and Henneberg 2001, 180; Olivares and Aguilera 2016, 1629). Metric methods, on the other hand, have been found to be more objective and easier to learn and apply (Cardoso 2008, 167). The high levels of intra- and interobserver agreement in the current study and in past studies indicate that, especially pertaining to the canines, odontometric methods can be easily and reliably employed by any osteologist (Aris *et al*. 2018, 675-676; Hassett 2011, 487; Hillson 2005, 420; Tuttösí and Cardoso 2015, 308; Viciano *et al*. 2011, 101; Viciano *et al*. 2013, 34-35).

5.1.2 Comparisons of antimeres

A comparison of antimeric measurements in the adult sample showed no significant difference, which allowed for the maximisation of the number of observations for each measurement through the substitution of the left measurement when the right was absent or unobservable. The same procedure was carried out on the non-adult sample in order to ensure that no bias was introduced through any asymmetry potentially caused by systemic stress affecting growth and development that may not have been present in the individuals that survived to adulthood. Since no significant differences were observed, antimeric measurements within both samples were pooled. It should be noted that no other study using similar odontometric methods of sex estimation has found there to be a prohibitive difference in size between the antimeres of canines or maxillary first molars (Aris *et al*. 2018, 676; Cardoso 2008, 160; Hassett 2011, 487; Tuttösí and Cardoso 2015, 308-309;).

Due to the wide variety of ante- and post-mortem changes that can obscure or damage the landmarks needed to obtain measurements, the ability to substitute antimeric measurements without a loss of accuracy greatly enhances the potential of odontometrics for sex estimation in osteological collections. By increasing the number of data points available for each measurement, a more effective statistical analysis is achieved through this process.

5.1.3 Comparison between the adult and non-adult samples

As explained in Chapter 1, the non-adult individuals of any skeletal collection are representative of the living members of a population that did not survive to adulthood and may therefore display a high prevalence of pathological changes associated with an increased risk of mortality in juveniles (Wood *et al.* 1992, 344). Growth retardation may be especially prevalent, as non-adults are more strongly affected by malnutrition than adults and conserve energy by halting or slowing growth and development (Cardoso 2008, 166; Moore 2013, 94; Saunders 2008, 125-126, 133, 134-136; Wilson and Humphrey 2017, 35; Wood *et al.* 1992, 344). Though malnutrition is known to affect the size of skeletal elements, dental dimensions are under more strict genetic control and are therefore less likely to be affected by extrinsic factors such as nutrition (Hillson 2005, 210). Even so, it was necessary to investigate the possibility that the mortality bias of the sample affected the odontometrics of the non-adult sample to confirm that the permanent dentitions of the Middenbeemster adult and non-adult samples are indeed comparable. No evidence of selective mortality was detected through a comparison of the mean adult and non-adult measurements, none of which differed to a significant degree. This confirmation is integral to the underlying theory of the method, which seeks to impose statistical models developed on an adult sample to a non-adult sample. Other studies based on the same principle also found no significant differences between adult and non-adult permanent dentitions in the crown or cervical dimensions, supporting the premise that nutritional and physiological stress have a negligible effect on the dentition as compared to the skeleton (Aris *et al*. 2018, 680; Cardoso 2008, 161; Hillson 2005, 210; Tuttösí and Cardoso 2015, 309).

5.1.4 Univariate analysis of sexual dimorphism in the adult sample

In all measurements included in this study, male mean values were found to be greater than female mean values. All measurements showed a very strong relationship with sex except for the MD measurements of the maxillary first molar. While both the crown and cervical MD dimensions in the maxillary first molars were found to be poorly correlated with sex, the BL measurements were related to sex, though not as strongly as most of the canine measurements. The maxillary first molar MDcerv was found to have the largest standard deviations in both the male and female adult subsamples. This may reflect a greater variation in the morphology of the maxillary first molar, but given the intraobserver error observed for this measurement it is possible that the large standard deviation, particularly in the male sample, is related to the difficulties encountered when measuring molars. The MD molar measurements were also found to have the lowest absolute t values of all measurements under study. This is indicative of a greater amount of overlap and less distance between the sexes in the MD cervical and crown diameters of the maxillary first molars than in any other measurement under study. In accordance with previous studies, this analysis demonstrates that all canine measurements hold greater potential for the development of methods of sex estimation than do those of the maxillary first molars, particularly the MDcrn and MDcerv (Cardoso 2008, 162; Tuttösí and Cardoso 2015, 311; Viciano *et al*. 2011, 101-102; Viciano *et al*. 2013, 36).

All eight canine measurements were found to be sexually dimorphic to a highly significant degree, with the level of sexual dimorphism of all but the MDcrn measurements exceeding 10%. This is in accordance with previous studies, which have consistently found the canine to be the most sexually dimorphic tooth and which therefore holds the most discriminating power (Cardoso 2008, 164, 165; Hassett 2011, 486; Tuttösí and Cardoso 2015, 311; Viciano *et al*. 2011, 105; Viciano *et al*. 2013, 37). The level of sexual dimorphism and the most sexually dimorphic measurement of the canines do, however, differ between studies (Cardoso 2008, 165; Hassett 2011, 486; Tuttösí and Cardoso 2015, 311; Viciano *et al*. 2011, 102-103; Viciano *et al*. 39-40). This may be a result of varying sample sizes, but is likely a manifestation of the differences in the range and magnitude of sexual dimorphism in the various populations studied, emphasizing the need to use population-specific data to develop accurate discriminatory models.

In this study, the maxillary canine BLcerv was found to have the highest level of sexual dimorphism at 13.80%, closely followed by the mandibular canine BLcerv at 13.69%. The level of sexual dimorphism of the MDcerv measurements of both the maxillary and mandibular canines also exceeded 13%, indicating that in this population canine cervical dimensions were more sexually dimorphic than crown dimensions. Though second in level of sexual dimorphism, the mandibular canine BL cervical diameter displayed the greatest absolute *t* value, suggesting good discriminating potential. These results are also consistent with previous studies that have demonstrated the utility of cervical measurements for sex estimation, particularly those of the canines (Hassett 2011, 489; Hillson *et al*. 2005, 425; Tuttösí and Cardoso 2015, 311; Viciano *et al*. 2011, 105; Viciano *et al*. 2013, 38, 41).

5.1.5 Developing logistic regression models and selecting equations

The measurements used in each logistic regression model were selected to explore the efficacy of varying combinations of measurements while attempting to create effective models suitable for archaeological application. Since individual dentitions are often variously preserved within archaeological collections, only models employing one tooth were developed so that, if successful, only a single tooth would be needed to estimate the sex of a non-adult individual. Similarly, a model was created for each measurement individually in order to identify any models with the potential to accurately classify sex from a single measurement. Pairings of crown and of cervical measurements were developed to see if the combinations would provide more accurate results than single measurements. These could then hypothetically be used in cases where the crown and cervix were differentially preserved. In addition, these models allow for comparisons with the studies of Cardoso (2008) and Viciano and colleagues (2013), which both tested logistic regression models based on two crown or two cervical measurements. Backwards logistic regression, conducted on the measurements of each tooth separately found that in this population the BLcerv measurement did not significantly contribute to the maxillary canine model, as the same correct classification rate (93.1%) was achieved by models 1 and 6. Likewise, backwards logistic regression conducted on the mandibular canine resulted in the creation of model 10, which returned the same correct classification rate (88.9%) as model 2 with the MDcerv measurement removed. Backwards logistic regression also identified useful combinations of crown and cervical measurements, represented by models 7 and 13.

Due to the wide variety of factors that can obscure dental measurements in non-adult individuals, it is both possible and probable that few measurements will be available for any non-adult individual in an archaeological collection. Therefore, multiple univariate and multivariate models within single teeth were developed on the much larger adult sample. This was done in order to recognize and create as many useful equations as possible, thereby increasing the likelihood that at least one equation could be applied to each non-adult individual. Of the 25 logistic regression models developed, 16 achieved correct classification rates above 75% and 14 were selected for the creation of logit equations to be subsequently tested on the non-adult sample. Equations were created for the first three models, based on the full complement of measurements for each tooth type under investigation. Models 1 and 2 were chosen over models 6 and 10 despite the identical correct classification rates, because no individual within the Middenbeemster non-adult sample lacked the measurements eliminated by backwards logistic regression and therefore the development of both equations would not have broadened applicability.

Among the canine models, only multivariate models that improved the correct classification accuracies as compared to single measurements were developed into equations. For example, model 7 using the maxillary canine BLcrn and MDcerv measurements achieved a correct classification rate of 89.7% in the adult sample. The maxillary canine BLcrn and MDcerv univariate models, models 15 and 16, achieved accuracy rates of 82.6% and 77.8%

respectively. Conversely, other multivariate maxillary canine models, models 4 and 5, achieved correct classifications of 81.3% and 75.6%. Therefore, a logit equation was developed for model 7 but not for models 4 or 5, because they did not achieve higher classification accuracies than their univariate counterparts. Past studies have similarly concluded that univariate statistical models developed on canine measurements are not improved by the inclusion of additional measurements, whether from the same or different teeth (Cardoso 2008, 163, 165; Viciano *et al*. 2011, 102-103; Viciano *et al*. 2013, 40). The results of these studies support the decision to eliminate multivariate models that did not improve univariate correct classification rates in the adult sample.

Each univariate model that achieved total accuracy rates above 75% was developed into an equation, resulting in univariate equations for every canine measurement except for the maxillary canine MDcrn, which correctly classified only 71.4% of the adult population. Despite accuracy rates below 75%, in order to include the maxillary first molar in the testing of the method, equations were developed based on models 3 and 23. These two models were very near the 75% threshold at 73.7% and 74.4% respectively and so it was considered worthwhile to test these equations on the non-adult sample in order to better assess the benefits and drawbacks of non-adult sex estimation using molar odontometrics.

5.1.6 Application to the non-adult sample

Of the 16 non-adults with at least one applicable measurement preserved, at least one sex estimation was possible for 15 individuals, demonstrating good applicability within an archaeological sample. Because several equations were developed based on single measurements, most non-adult individuals included in this study retained at least one of these measurements, allowing for a sex estimation to be reached. Only the MD maxillary first molar measurements were observable in the individual for whom no sex estimation was possible, an 8-year-old male (S384V0839). Though all four canines were present, none were erupted and since both the mandible and maxillae were intact, it was not possible to obtain any canine measurements. As the maxillary first molar MD crown and cervical measurements were deemed unreliable on their own, no equations were developed for these measurements. Therefore, none of the resultant equations were applicable to S384V0839.

When applied to the entirety of the non-adult sample, nine of the 14 equations exceeded the 75% threshold for non-adult sex estimation and seven of these nine exceeded the 85% threshold for adult sex estimation. After identifying and removing outlier S196V0437 from the evaluation, the number of equations with correct classification rates above 75% increased to 12 of 14, with eight of these 12 exceeding the level considered acceptable in adult sex estimation standards as well. Neither of the two equations developed on maxillary first molar measurements were able to correctly estimate sex in more than 75% of the non-adult sample. This is in contrast to the results obtained by Aris and colleagues (2018), whose multivariate approach led to accuracy rates of up to 94.6% (Aris *et al*., 2018, 676-679). However, all of the functions developed in the 2018 study included at least 1 and as many as 4 cusp dimensions, which were beyond the scope of the present research. It can therefore be concluded that the cusp measurements used by Aris and colleagues (2018) are necessary in order to achieve acceptable levels of accuracy in non-adult sex estimation with maxillary first molar odontometrics. This is supported by the findings of previous work using only the maximum crown and cervical diameters, none of which identified the dimensions of the maxillary first molar as reliable indicators of sex (Cardoso 2008, 165; Garn 1977, 697; Viciano *et al*. 2011, 102; Viciano *et al*. 2013, 39-40). The high rates of intraobserver and interobserver error as well as large standard deviations of the means seen in the MD maxillary first molar measurements underscore the problematic nature of maxillary first molars within sex estimation. If the measurements cannot be reliably reproduced, this could have negatively affected the initial formulation of the molar models as well as the results when applied to the second sample, obviously diminishing the accuracy of these models.

All 12 of the equations developed on canine measurements achieved correct classification rates above 75%. Five of the 12 equations correctly estimated sex in 100% of cases, albeit on very small samples of at most eight individuals. Though these results are likely inflated by the small sample size, they are nonetheless very promising. Contrary to the implications of the univariate analysis for sexual dimorphism in the adult sample, the univariate logit equation with the highest correct classification rate in the non-adult sample was the mandibular canine MDcerv rather than the BLcerv of either maxillary or mandibular canines. This demonstrates the need to develop multiple useful logistic regression models based on several measurements that have been proven to be significantly related to sex, as the specific levels of sexual dimorphism in the adult sample may not be exactly reflected in the non-adult sample. This also creates the opportunity of applying several regression equations to a single individual. This can be used to more confidently estimate sex in undocumented populations, in which multiple predictions in agreement can convey a greater degree of certainty. The logit equation based on the mandibular canine MDcerv also outperformed all other equations based on canine or molar measurements when blind tested by Viciano and colleagues (2013) (Viciano *et al*. 2013, 39-40).

Unlike many morphometric methods of non-adult sex estimation, the classification accuracies listed by individual in order of age illustrate that this method performs equally well on individuals from the age of five to 16 years (Klales and Burns 2017, 750; Wilson *e al*. 2016, 262, 264; Wilson and Humphrey 2017, 36-37). It is also evident that the number of applicable equations per individual generally increased with age. This trend was clearly caused by the greater accessibility to the teeth of older individuals whose permanent canines have partially or fully erupted. Although this method can only be applied to those with canines with fully-formed and accessible crowns, it has nevertheless been demonstrated that from the age of 5-7 years, this method can be reliably applied to non-adult individuals.

5.2 Limitations and confounding factors

5.2.1 Utility of permanent maxillary first molar odontometrics

This study has demonstrated that for a number of reasons, the permanent maxillary first molar is not ideally suited to non-adult sex estimation. Firstly, the only measurements with poor inter- and intraobserver agreement were the MD measurements of the maxillary first molars. Whether this is due to a greater variance in morphology or to greater difficulty in consistently identifying the landmarks needed for the measurements, any model based on unreliable and unrepeatable measurements is problematic. Secondly, the univariate analysis of sexual dimorphism in the adult sample revealed that only the BL crown and cervical measurements of the maxillary first molars were significantly related to sex. Thirdly and most conclusively, none of the logistic regression models developed on the dimensions of the maxillary first molars reached a 75% correct classification rate within the adult or non-adult samples. It is therefore recommended that any future research into non-adult sex estimation using the odontometrics of the maxillary first molar include cusp measurements as described by Aris and colleagues (2018) (Aris *et al*. 2018, 675).

5.2.2 Impediments to measurements

Many measurements were obscured by a variety of factors in both the adult and non-adult samples. Both samples were affected by tooth-wear, enamel hypoplasia, caries, and the accumulation of calculus. In the adult population, pipe notches were indeed found to be very common and frequently obscured the MDcrn of both maxillary and mandibular canines. This can be observed in the adult sample sizes for each measurement; the male sample sizes for the MDcrn of both canine types are substantially lower than any other canine measurement, with only 6 observations possible on the mandibular MDcrn as compared to 29 on the BLcrn. The MDcrn diameter is more susceptible to abrasion than the BLcrn because its axis is located nearer the occlusal surface and is therefore more quickly obfuscated by tooth-wear. In the non-adult sample, measurements were frequently unattainable because the tooth in question was unerupted and inaccessible within the alveolar bone.

A great number of both the adult and non-adult samples displayed enamel hypoplasia, very slight caries, or very slight calculus deposits on both the crown and the cervix. This was observed to such an extent that to exclude all of these cases would have greatly reduced the sample size and therefore the statistical power of the method. Measurements were therefore taken and though these changes were noted during data collection as potentially problematic, they do not seem to have affected the accuracy of this method. This increases the method's value within archaeological contexts, where enamel hypoplasia, caries, and deposits of calculus are very common. The inclusion of cervical measurements in the investigation as well as the substitution of antimeric measurements also greatly aided in circumventing these issues and maximizing the number of observations available for each measurement.

5.2.3 Age range of non-adult sample

One of the reasons the permanent maxillary first molar was included in this study was because it is the first tooth to develop and therefore holds the potential of estimating sex in non-adults as young as three years of age at death. Sex estimations using the maxillary first molar were possible for the two youngest individuals analysed, one female aged 3 years (S089V0091) and one male aged 5 years (S140V207). However, these were done using models 3 and 23, which proved to be unreliable overall. For this reason, the sexes of these two individuals are not considered to have been reliably estimated.

Accordingly, although this study examined all documented individuals aged 3 years and above, reliable sex estimations through the canine models was only possible for those aged 7 years and above. The canines of one documented five-year-old individual in the collection (S336V0709) were visible within the alveolar bone, but were not extractable. This indicates that under certain circumstances, such as the poor preservation of the mandible or maxilla, it may be possible to achieve reliable sex estimations for individuals as young as five years of age at death. The only two documented six year-old individuals in the collection did not retain any of the teeth under research in this thesis. Therefore, the youngest individual for whom a reliable sex estimate was obtained was S018V0102, a seven-year-old female.

Of the 33 non-adult individuals examined, only 16 yielded any of the measurements with which this study is concerned. This low proportion can be attributed to the decision to look at all individuals from the age of three years, when typically the crown of the permanent maxillary first molar has fully formed but has not yet begun to erupt (Cunningham *et al*. 2016, 159, 168; Hillson 2014, 45, 64). Now that the logit equations developed on the maxillary molar measurements have proven ineffective, it is recommended that future applications concentrate on individuals aged 5 years and above. Had the current study been concerned only with those aged 5 years and above, a total of 19 non-adult individuals would have been examined with only 4 of the 19 unable to be included in the analysis. These 4 include the five-year-old and two six-year-olds mentioned above as well as one 17-year-old male whose skull was not preserved. This is indicative of the approach's utility and applicability within archaeological contexts.

5.2.4 Non-adult sample size

Some degree of caution is advised in interpreting the high levels of correct classification rates achieved by the canine models when applied to the non-adult sample due to the small sample sizes for each equation. Though 15 individuals were included in the final analysis, the number of individuals on which each model was tested ranged from three to 10 and within the male non-adult subsample all equations were tested on only two individuals. These low numbers make the correct classification percentages highly sensitive, as evidenced by the increase in accuracies when outlier S196V0437 was excluded. However, the consistency with which the canine models achieved accuracy rates well above the 75% threshold in combination with the fact that the same models correctly classified 7 of the 8 non-adults on which they were tested in 90-100% of cases (tab. 19) indicates that this is a reliable means of estimating sex in non-adult skeletal remains.

Table 19. Classification accuracies per individual for only the canine regression equations.

Individual	Age	Sex	L1	L2	L7	L8	L9	L15	L16	L17	L18	L19	L20	L21	% Correct
S018V0102	7	F	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100.00%
S248V0393	9	F	-	Y	-	Y	Y	-	-	-	Y	Y	Y	Y	100.00%
S286V0469	10	F	Y	Y	Y	Y	N	N	Y	N	N	N	Y	N	50.00%
S367V803	11	F	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100.00%
S167V0270	12	F	-	Y	-	Y	Y	-	-	-	Y	Y	Y	Y	100.00%
S522V1127	13	F	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100.00%
S465V1001	15	M	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	91.67%
S462V0987	16	M	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	91.67%

Age is given in years. F documented female. M documented male. Y yes sex match. N no sex match.
% correct overall percentage of correct estimations per individual.

Aside from being the earliest developing permanent tooth, the maxillary first molar was chosen for this study because it is often recovered in archaeological contexts. The veracity of this statement is borne out in the fact that four of the 12 non-adult individuals aged seven years and above did not retain a single canine measurement. Consequently, of the 19 documented non-adults in the Middenbeemster collection aged five years and above, the logit equations based on canine odontometrics were only applicable to eight. All methods of sex estimation in adults and non-adults alike rely on the preservation of whichever skeletal elements are necessary for analysis. Though not ideal, the small number of non-adults on whom this method could be tested yielded promising and consistent results. A reliable method by which the sexes of even part of a non-adult skeletal sample could be estimated would be a substantial addition to osteoarchaeology and would provide insight into the lives of past children, who are often neglected by historical sources.

5.2.5 Outliers

5.2.5.1 S521V1150

Individual S521V1150, a 69 year-old male, preserved all measurements of the maxillary and mandibular canines except MDcrn and was therefore included in the development of models 4, 5, 7, 9, 15-17, and 19-21. The high standardized residuals of this individual (appendix 1) indicate that he is an outlier within the adult sample. During data collection it was noted that this individual had abnormally flat buccal crown surfaces and severe horizontal wear, but otherwise no pathological changes to the dentition or to the rest of the skeleton were macroscopically apparent. However, this individual does not appear to have significantly affected the models developed on the adult sample, as demonstrated by the high rates of correct classification achieved by these models. Because there were no readily apparent indications that this individual would be an outlier during data collection, it was decided not to remove S521V1150 from the logistic regression models.

5.2.5.2 S2860469

In the non-adult sample, the sexes of two individuals were incorrectly estimated in more than half of the equations applied to them. A 10 year-old female (S286V0469) was correctly classified in six of 13 application equations tested and a 13 year-old male (S196V0437) was correctly classified in only three of 14 equations. Upon re-examination of S286V0469, no evidence of systemic changes or illnesses was visible apart from healed *cribra orbitalia*, which has been associated with nutritional deficiency. This condition was also observed in other adult and non-adult individuals in the sample and did not seem to have any effect on their dental dimensions. It was decided not to remove this individual from the test results as there were no visible indications that could have allowed for the identification of this individual as an outlier in an undocumented population.

5.2.5.3 S196V0437

Figure 5. Anterior view of the mandible of S196V0437 displaying severe resorption and new bone growth (photograph taken by the author).

Individual S196V0437, on the other hand, was immediately recognizable as atypical. Most strikingly, this 13-year-old boy lost his permanent mandibular second premolars, first molars, and second molars ante-mortem. All of the corresponding sockets display differing stages of healing and resorption, from initial new bone growth in the premolar region to full closure and pronounced resorption in the molar region. These changes can be seen in figures 5, 6, and 7. There is widespread bilateral new bone growth and porosity on the mandibular body and rami (fig. 7). The area surrounding the mandibular notch is especially thickened and smooth, indicating that this individual likely had a disorder of the temporomandibular joints (figures 6 and 7). The maxillae were unfortunately not preserved and so it is impossible to tell whether any maxillary teeth were lost ante-mortem. However, the presence of all four permanent maxillary incisors, one canine (2.3), all four premolars, and both first molars implies that any maxillary ante-mortem loss was less severe than the observed mandibular loss. No severe enamel hypoplasia was present, though slight linear enamel hypoplasia was visible on the mandibular canines.

Figure 6. Superior view of the mandible of S196V0437 displaying ante-mortem tooth loss with differing stages of healing and thickened mandibular notches (photograph taken by the author).

For this study, the most relevant pathological change is that the permanent maxillary second incisors (1.2 and 2.2) have been affected by microdontia (fig. 8). Microdontia is a condition resulting in abnormally small teeth, which can also be abnormally shaped, and has been associated with several congenital conditions such as other dental anomalies, heart disease, Down's syndrome, and cleft palate (Cantekin and Celikoglu 2015, 186, 188; Ortner 2003, 597; Sajnani and King 2014, 209, 211). Microdontia has been genetically linked to many other dental anomalies and has been observed in conjunction with a general reduction in tooth size, which would clearly have an effect on the method under study. The maxillary second incisors of S196V0437 are indeed abnormally small and rounded or peg-shaped (fig. 7) and upon re-examination of this individual, it was evident that the shape of the canines had also been affected, especially that of the maxillary canine (2.3), which was more round than is typical. Only three of the 14 (21.43%) equations tested on individual S196V0437 correctly estimated his sex, a rate significantly below that of any other non-adult individual analysed. This suggests that the present method is not suitable for individuals affected by microdontia. When preservation allows, microdontia can be readily detected in the skeletal record through the observation of

Figure 7. Right lateral view of the mandible of S196V0437 displaying ante-mortem tooth loss of the second premolars, first molars and second molars and dense new bone growth on the right mandibular ramus (photograph taken by the author).

Figure 8. Anterior maxillary teeth of S196V0437 displaying microdontia of the second incisors (1.2 and 2.2). From left to right: 1.2, 1.1, 2.1, 2.2, 2.3.

abnormally small teeth or tooth sockets. Consequently, it is recommended that in future studies of this sort, any individuals displaying microdontia be removed from analysis as was done here.

5.2.6 Requirements of the adult sample

5.2.6.1 Temporally and geographically restricted

The development of population-specific approaches to sex estimation has several advantages. These methods work within the range of body size and sexual dimorphism of the population studied. This is especially important when using metrics as these can be more easily confounded by population-level differences in body size (Buikstra and Ubelaker 1994, 16; Cardoso 2008, 166, 167; Inskip *et al*. 2018, 681; Tuttösí and Cardoso 2015, 306). Cross-population methods have often been found to be less accurate when applied to a second population, because they impose the patterning and magnitude of sexual dimorphism from an unrelated population (Buikstra and Ubelaker 1994, 16; Cardoso 2008, 166, 167; Cunningham *et al*. 2016, 18; Hillson 2005, 257; Lewis 2006, 49; Shankar *et al*. 2013, 753; Tuttösí and Cardoso 2015, 306; Wilson and Humphrey 2017, 34, 35). Consequently, in order to properly develop such a method, one must be certain that the skeletal collection under study represents a population that is relatively constrained both geographically and temporally. If the skeletal sample is composed of individuals from different populations, the accuracy of any metric and statistical approaches would be compromised by population-level disparities.

Aris and colleagues (2018) attempted to evaluate the effects of chronological separation between populations by applying discriminant functions developed on the permanent dentition of adult individuals from the Spitalfields collection (London, 17^{th}-19^{th} centuries) to a geographically similar but temporally distinct collection, that of Black Gate (Newcastle-upon Tyne, 8^{th}-12^{th} centuries) (Aris *et al*. 2018, 673). When tested on the Black Gate adults of morphologically estimated sex, the Spitalfields function with the best results decreased in accuracy from 87.3% to 83.3%, a loss of only 4.8% (Aris *et al.* 2018, 680). Though the authors describe these two sites as geographically similar but chronologically distinct, the sites are in fact separated by just under 400 km (Aris *et al*. 2018, 680; www.distancefromto.net). Though both are located within the United Kingdom, the genetic and extrinsic factors affecting the two populations are likely to differ substantially between northern and southern England and between the early Middle Ages and the modern era. Aris and colleagues (2018) found that although stepwise methods of statistical analysis increased accuracy on the Spitalfields collection, they resulted in more drastically decreased accuracies when applied to the Black Gate sample (Aris *et al*. 2018, 680-681). The authors therefore concluded that when developing an equation for cross-population application, it is best not to over-fit the data to the developing population through stepwise analysis and exclusion of variables. This will result in less accuracy within the original population, but will also generate greater applicability across populations that are somewhat geographically and temporally distinct.

The other studies using the permanent odontometrics discussed in this thesis have been developed and tested on skeletal assemblages formed over the course of as little as one day (Herculaneum) and as much as approximately 150 years (Granada) (Viciano *et al*. 2011, 98; Viciano *et al*. 2013, 31-32). The present study of Middenbeemster includes documented individuals from a single municipality interred from 1829 to 1866, a span of just under 40 years, and is therefore restricted both temporally and geographically. Accordingly, issues pertaining to the population-specificity of the equations developed on the adult sample were not explicitly considered. More work is necessary in order to ascertain the extent to which geographic and chronological differences between populations affect the odontometrics of the permanent dentition and to what extent reasonably accurate and reliable functions can be applied to distinct populations. The work of Aris and colleagues (2018) holds promising results for methods that are at least regionally applicable.

5.2.6.2 Adult sample size and preservation

Due to the need to develop equations for each population studied, a reasonably large adult population of known or confidently estimated sex is required. Though it is not necessary to have an equal ratio of males and females in logistic regression, both sexes should be well-represented within the sample in order to adequately represent the range of sexual dimorphism within the population. If no documentation is available, relatively complete adult individuals are necessary in order to morphologically estimate sex with any degree of certainty (Cardoso 2008, 159). Features of the pelvis and skull are especially important to adult osteological sex estimation and so good preservation of these elements in a substantial portion of the adult population is necessary. In the Middenbeemster collection, the methods of Buikstra and Ubelaker (1994) and the WEA (1980) were used to estimate the sex of adult individuals with sufficiently preserved skeletons. The reliability of these adult sex estimation methods is borne out by the fact that the morphologically estimated sex of every documented individual used in the development of this method matched documented sex. It can therefore be concluded that the present analysis would have performed equally well had the population not been documented, though testing of the method would not have been possible.

5.3 Comparisons with other current methods in non-adult sex estimation

5.3.1 Metric approach

Metric techniques have been found to be more objective than morphological ones, as they require less familiarity with the method and with the range of expression of sexual dimorphic traits within a population (Cardoso 2008, 159, 167; Cardoso and Saunders 2008, 28; Wilson *et al*. 2016, 255-256; Wilson and Humphrey 2017, 33-35). This is exemplified by the fact that many morphological methods greatly decrease in accuracy in both correct classification rates and intra- and interobserver error when tested on distinct populations by researchers unassociated with the development of the method, including those developed by Schutkowski (1993) and Loth and Henneberg (2001) (Cardoso and Saunders 2008, 25, 28; Coqueugniot *et al.* 2002, 135, 136; Loth and Henneberg 2001, 180; Olivares and Aguilera 2016, 1628-1629; Scheuer 2002, 191; Wilson *et al.* 2016, 256). Morphometric methods such as that of Wilson and colleagues (2008) have also suffered from poor intra- and interobserver agreement due to difficulties in identifying landmarks from which to measure (Luna *et al.* 2017, 899-900; Wilson *et al.* 2011, 37, 39-40). The current study found near perfect intra- and interobserver agreement for all canine measurements. That the second observer was untrained in odontometrics and unfamiliar with both the method and the collection under study suggests that this approach is readily accessible to osteologists. Similar results from previous studies support the argument that techniques based on the odontometrics of the permanent canine can be reliably reproduced by multiple researchers (Tuttösí and Cardoso 2015, 310; Viciano *et al.* 2011, 34-35; Viciano *et al.* 2013, 101).

Due to the small range of and large overlap in the size dimorphism of non-adults experiencing normal growth, small errors in measurement can result in an incorrect estimation (Moore 2013, 107). This situation is exacerbated by the potential inclusion of individuals who suffered from growth retardation, which would result in a reduction of size dimorphism between non-adult males and females (Lewis 2006, 49-50). More so than morphological methods, metric methods are sensitive to differences in overall body size and in the patterning and level of sexual dimorphism between populations. It is therefore difficult to develop a reliable method that is universally applicable.

5.3.2 Population specificity

Many of the methods described in Chapter 2 proved unsuccessful when tested on a population geographically and temporally distinct from that on which it was developed (Cardoso and Saunders 2008, 25, 28; Coqueugniot *et al.* 2002, 135, 136; Loth and Henneberg 2001, 180; Olivares and Aguilera 2016, 1628-1629; Scheuer 2002, 191; Wilson *et al*. 2011, 37; Wilson *et al*. 2016, 256). Following previous research, by using the adult sample to develop discriminant equations for application on a non-adult sample within the same population, this study sought to circumvent many of the issues associated with universal methods. This technique reflects the degree and patterning of sexual dimorphism within the population under study and does not impose any extraneous assumptions. Though tentative, the work of Aris and colleagues (2018) shows that such discriminant equations may be applicable to regionally similar populations over a relatively large time span if the equations are not overly fitted to the original population (Aris *et al*. 2018, 681).

5.3.3 Mortality bias and growth retardation

The efficacy of skeletal metric methods of non-adult sex estimation can be greatly affected by the potential reduction in size dimorphism caused by the differential effects of nutritional stress on boys as opposed to girls (Cardoso 2008, 159; Lewis 2006, 49-50; Moore 2013, 93; Stull *et al.* 2017, 65). Some studies have also found that certain morphometric traits often identified as dimorphic are in fact more closely related to growth and development than to sex (Wilson *et al.* 2016, 262, 264; Wilson and Humphrey 2017, 36-37). Consequently, growth retardation can significantly alter the results obtained from morphometric skeletal methods. The dimensions of the permanent dentition are less affected by dietary stress than are skeletal elements and, as such, present consistent levels of adult-level size dimorphism from the time of formation through adulthood (Cunningham *et al.* 2016, 13; Hammerl 2013, 263; Hillson 2005, 210, 257; Moore 2013, 93-94, 107; Uhl 2013, 68). Selective mortality and the heterogeneity of risks may differentially affect the permanent odontometrics of the non-adults who did not survive to adulthood. This can, however, be easily checked by comparing the means of the adult and non-adult samples using independent-samples *t*-tests, a practice that should be universally adopted when using this approach. In this way, statistical analyses of permanent odontometrics can function equally well on non-adults who did and did not experience growth retardation. In the present study, this can be demonstrated through a comparison of the documented ages of the non-adult sample with their osteologically estimated ages. Table 20 shows that the biological ages of the older non-adult individuals included in this study are retarded in comparison to their documented calendar age. Despite this, the sexes of all these individuals were correctly estimated by all or nearly all of the equations tested on them.

Table 20. The documented and osteologically estimated ages of all individuals for whom sex could be estimated based on canine odontometrics. No estimated age was available for individual S522V1127.

Individual	Documented age	Estimated age
S018V0102	7 yrs	7.5 yrs ± 12 mo
S248V0393	9 yrs	8.5 yrs ± 12 mo
S286V0469	10 yrs	10 yrs ± 3.5 yrs
S367V803	11 yrs	11 yrs ± 30 mo
S167V0270	12 yrs	**10.5 yrs ± 12 mo**
S465V1001	15 yrs	**12 yrs ± 12 mo**
S462V0987	16 yrs	**14 yrs ± 2 yrs**

Yrs years. Mo months. Estimated ages lower than the documented age marked in bold.

5.3.4 Requirements of the non-adult sample

Unlike methods developed on the deciduous dentition such as those of Black (1978), Żądzińska and colleagues (2008) and Shankar and colleagues (2013), the current method does not require a large homogenous non-adult sample on which to develop effective statistical inquiries. Since all statistical models are developed on the adult dentition, non-adult sex estimations can be made even if there are only a few non-adult individuals in the collection. Additionally, the abovementioned methods require maximal MD and BL measurements from the entire deciduous dentition in order to estimate sex (Black 1978, 77; Shankar *et al.* 2013, 753). In contrast, the method presented in this thesis can achieve high levels of accuracy with as little as one measurement of the permanent canine (tab. 18). Although the retention and accessibility of a permanent canine is not by any means guaranteed, it is a much more likely occurrence than the survival of a complete set of undamaged deciduous teeth. Similarly, some metric methods of non-adult sex estimation require the complete or near-complete preservation of several skeletal elements in order to attain the necessary measurements. For example, the method of Stull and colleagues (2017) was developed based on 18 measurements obtained from six long bones. Morphological methods usually require preservation of a sufficient quality to recognize and assess dimorphic traits (Tuttösí and Cardoso 2015, 307). These conditions are often not met by archaeological collections, especially in relation to non-adult skeletal remains which deteriorate more readily in unfavourable soil conditions (Lewis 2006, 185; Saunders 2008, 119). Teeth, on the other hand, are extremely resistant to diagenetic processes and are frequently recovered in archaeological contexts (Aris *et al*. 2018, 673; Baker *et al*. 2005, 53; Cardoso 2008, 159; Cunningham *et al*. 2016, 13; Hammerl 2013, 263; Hillson 2005, 158-159; Hillson 2014, 70, 110; Shankar *et al*. 2013, 752; Tuttösí and Cardoso 2015, 306, 307; Viciano *et al*. 2011, 97; Viciano *et al.* 2013, 31). Odontometric approaches are therefore ideally suited to archaeological collections in which non-adult remains may be more poorly preserved than those of adult individuals.

This method can also be applied to all individuals above the age of five years, given that at least one permanent canine crown is preserved. This is in contrast to other methods that have had varying results that have proven more effective on certain age groups than others (Klales and Burns 2017, 750; Wilson and Humphrey 2017, 36-37). Since the permanent dentition forms in childhood and, apart from damage and pathological changes, remains unchanged thereafter, this method can be applied not only to all non-adults from the age of five years, but also to adult individuals of the same population whose sex was indeterminable due to poor preservation or because their features were not sufficiently dimorphic as to allow for sex estimation.

5.3.5 Logistic regression

Studies comparing the permanent odontometrics of adult and non-adult individuals in order to estimate sex have used a variety of statistical approaches to analyse their data. These include logistic regression, discriminant function analysis, and sectioning point procedure. Though all three approaches have attained similar results, both sectioning point procedure and discriminant function analysis entail certain requirements of the sample used for their development (Aris *et al*. 2018, 680; Cardoso 2008, 165; Tuttösí and Cardoso 2015, 311; Viciano *et al.* 2013, 40). Sectioning point procedure requires a relatively large sample with a broadly even ratio of males to females (Cardoso 2008, 159-160). Discriminant function analysis requires a normal distribution of the independent variables and equality of variance-covariance matrices in the male and female subsamples (Aris *et al*. 2018, 675, 680; Cardoso 2008, 161; Tuttösí and Cardoso 2015, 308; Viciano *et al.* 2013, 33-34). Conversely, logistic regression makes none of these assumptions of the data and is better equipped to offset any issues arising from differing independent variable sample sizes (Cardoso 2008, 160, 161). This is particularly useful in archaeological contexts. In addition, logistic regression is the only one of these three to automatically provide probability scores of binary group membership, in this case male or female, allowing for a more nuanced interpretation of results (Cardoso 2008, 161; Olivares and Aguilera 2016, 1624; Tuttösí and Cardoso 2015, 308). Both univariate and multivariate models and equations can be created relatively easily, allowing multiple combinations of measurements to be assayed in order to identify the most effective discriminating measurement within the population. It is beneficial to develop many equations, because this increases the likelihood that at least one equation will be applicable to each non-adult within the sample. The results of multiple equations can also be compared and used to bolster any sex estimate made on undocumented non-adult skeletal remains.

5.3.6 Consistently high accuracy rates

The accuracy rates obtained by the current study using the odontometrics of permanent canines are comparable to those considered acceptable within adult sex estimation. Previous studies developing similar methods have also consistently found that non-adult sex estimation could be made with confidence levels above 85% based on univariate and multivariate equations using only the permanent canine (Cardoso 2008, 163; Viciano *et al*. 2013, 40). Though accuracy rates will never be able to compare to those attainable through genetic analyses, these odontometric methods are much more accessible to the average osteologist, less expensive, and less fraught with complications such as degradation and contamination. That multiple studies have now found the approach presented in this thesis to be reliable and accurate across several modern and archaeological populations is indicative of the promise it holds for the future of non-adult sex estimation, a feat that has been long sought after and is currently prohibited by the SWGFA (SWGFA 2010, 3).

5.4 Summary

Permanent canine odontometrics have shown tremendous promise for the estimation of sex in non-adult skeletal remains. Though this approach requires a large sexed adult population, few conditions of the non-adult sample are required. Unlike many other methods of non-adult sex estimation, this approach does not require a large non-adult sample nor the good preservation of several skeletal elements. If accurate logistic regression models can be developed on the adult sample, as little as one measurement can be used to estimate sex with an acceptable level of confidence. This is especially advantageous within archaeological contexts, in which the bones of non-adults may deteriorate more readily than their adult counterparts, but teeth are frequently recovered. At least one accessible canine with a fully formed crown is, however, necessary and for that reason this approach is not applicable to individuals under five years of age at death. As long as all statistical analyses are undertaken to ensure that asymmetry between antimeres and mortality bias have not affected the odontometric of the adult and non-adult samples, the permanent canine odontometrics have consistently shown tremendous discriminating potential through logistic regression. By allowing reliable sex estimation in non-adult skeletal remains, this method makes it possible to gain insights into the past lives of at least a portion of the non-adult population, who all too often seem invisible in archaeological and historical narratives.

6. Conclusion

The consistency with which population-specific approaches to non-adult sex estimation using the permanent canine odontometrics have achieved accuracy rates above 75% is all the more notable considering the irregularity with which current methods correctly estimate non-adult sex across populations and age groups. To date, no method of non-adult sex estimation has been widely accepted by the field of osteology (Aris *et al.* 2018, 672; Baker *et al.* 2005, 3, 10; Black 1978, 77; Cardoso 2008, 158; Cunningham *et al*. 2016, 17, 18; Klales and Burns 2017, 747; Saunders 2008, 117; SWGFA 2010, 3; Wilson and Humphrey 2017, 33). For this reason, non-adults are routinely omitted from archaeological palaeodemographic studies exploring the relationship between biological sex and gendered patterns of behaviour and experience through the interpretation of morbidity and mortality (Aris *et al.* 2018, 672; Baker *et al.* 2005, 4; Cardoso 2008, 158, 167; Cunningham *et al.* 2016, 5; Lewis 2006, 187; Luna *et al.* 2017, 898; Olivares and Aguilera 2016, 1623; Viciano *et al.* 2011, 105). Children are also frequently omitted from historical accounts and so there is currently little information on the minutiae of life as a child in the past. Archaeology is regularly used to corroborate, nuance, or supplement historical records, which are frequently more concerned with affairs of public adult life rather than the private spheres to which women and children were so often restricted. Though it is not possible to create new contemporary written descriptions of the mundane, physical remains hold a wealth of potential information about the lived realities of those in the past. Additionally, since growth and development is sexually dimorphic, non-adult sex estimations would allow for the refinement of age estimates and growth studies (Lewis 2006, 186; Saunders 2008, 134; Uhl2013, 65). If it were possible to assess the sex of non-adult individuals in an archaeological assemblage, a more detailed, nuanced, and thorough understanding of the society under study could be achieved.

6.1 The approach

Accordingly, this thesis set out to test one of the most promising approaches to non-adult sex estimation currently available by developing logistic regression equations based on the permanent canine and maxillary first molar odontometrics of the documented adult sample of Middenbeemster and subsequently applying these equations to the non-adult sample. This generates a population-specific discriminant process that reflects the overall body size as well as the patterning and magnitude of sexual dimorphism of the Middenbeemster population. Both the permanent canines and the maxillary first molars were used based on previous work that showed the permanent canine to be the most sexually dimorphic tooth and the maxillary first molar to be the earliest-developing permanent tooth and the one most likely to be recovered in archaeological excavations (Aris *et al*. 2018, 673, 681; Cardoso 2008, 165; Cunningham *et al*. 2016, 19, 159; Hammerl 2013, 268; Hassett 2011, 486; Hillson 2014, 45, 64; Tuttösí and Cardoso 2015, 311;Viciano *et al*. 2011, 102; Viciano *et al*. 2013, 37). In addition to traditional maximum mesio-distal and bucco-lingual crown diameters, this study made use of the alternative cervical measurements proposed by Hillson and colleagues (2005) (Hillson *et al*. 2005, 416). The inclusion of both crown and cervical measurements in the data collection process allowed for measurements to be collected on teeth that displayed extensive tooth-wear, carious activity, calculus deposits, or damage to either the crown or the cervical region. A large adult sample of confidently documented or estimated sex is necessary in order to attain enough data points for robust statistical analyses.

These measurements were then used in various statistical analyses to assess intra- and interobserver error, the differential effects of mortality bias on the non-adult sample in comparison to the adult sample, and the level of sexual dimorphism for each measurement. Multiple logistic regression models were developed and based on these, the work of previous researchers, the results of the univariate analysis for sexual dimorphism, and the practical realities of archaeological collections, certain models were selected to be developed into logit equations and applied to the non-adult sample. The resultant sex estimates were then compared to documented sex in order to assess accuracy. Logistic regression is highly suitable for archaeological applications as it necessitates fewer requirements of the data and provides the probability of group (sex) membership for each individual, allowing for a more nuanced interpretation of results. Logistic regression is also fairly accessible and several multivariate and univariate models can be created and reviewed relatively quickly. This facilitates the creation of multiple models and equations that will maximize the number of sex estimations possible in the non-adult sample.

6.2 Results

The logistic regression equations developed on the maxillary and mandibular canines of the adults sample were demonstrated to have great potential in correctly estimating sex in non-adult individuals from the same population. All 12 equations based on canine odontometrics correctly classified the sex at rates above 75% and eight of these 12 attained correct classification rates between 85% and 100%. Intra- and interobserver errors were found to be minimal in the canine measurements, but were less consistent in the molar measurements. The maxillary first molars were also shown to be less sexual dimorphic in the univariate analysis, which was confirmed in the poor performance of the two logit equations developed on molar measurements, neither of which reached or surpassed 75%. As past studies have reported similar results, it was determined that the maxillary first molar MD and BL crown and cervical diameters were not sufficiently reproducible and sexually dimorphic to allow for reliable sex estimation in the adult or non-adult samples.

With the maxillary first molar equations removed, the number of documented non-adults on whom sex estimation was possible significantly decreased. This indicates that the requirement of at least one accessible canine with a fully formed crown is frequently unmet in archaeological assemblages. In spite of this, the accuracy and consistency with which the canine logit equations estimated sex demonstrate the utility, if not the applicability, of this approach. Within the documented Middenbeemster population, the youngest individual for whom a reliable sex estimation was possible was seven years of age at death, though in theory it should be possible to attain estimates for individuals as young as five years. It is important that no mortality bias was detectable through a comparison of the adult and non-adult odontometrics, as such a bias would have undermined the underlying theory of this approach.

6.3 Research questions

1. Can sexual dimorphism be observed in the dimensions of the permanent canines or maxillary first molars in the documented adult population of Middenbeemster?
    a. Which tooth or measurement, if any, displays the greatest degree of sexual dimorphism and therefore the greatest discriminating power?

b. Were any measurements not sexually dimorphic or too problematic?

All of the crown and cervical dimensions of both the maxillary and mandibular canines were significantly related to sex, while in the maxillary first molars only the BL crown and cervical measurements were found to be significantly correlated with sex. In this population, the maxillary canine BLcerv was found to be the most dimorphic. However, the logistic regression model that displayed the greatest discriminating power in the adult sample was based on a combination of the lower canine MD and BL crown diameters. In the non-adult sample several multivariate equations based on the canine attained correct classification rates of 100%. The only univariate equation to do so was based on the lower canine MDcerv. This shows that the measurement found to have the greatest level of sexual dimorphism is not necessarily the most effective measurement for sex estimation in the non-adult sample. This highlights the need to develop several logistic regression equations on the adult sample, a process that also serves to maximise the number of equations that are potentially applicable to the non-adult sample.

In addition to showing low correlation with biological sex, the MD crown and cervical maxillary first molar measurements failed to meet acceptable levels of intra- and interobserver error. Consequently, these measurements were deemed too problematic to produce reliable sex estimations. This, in combination with the poor performance of the logistic regression models based on the maxillary first molar, has led this study to conclude that its MD and BL crown and cervical diameters are not adequately reproducible or sexually dimorphic to reliably estimate sex within the Middenbeemster population.

2. Can logistic regression analysis be applied to accurately categorize the adult population?
    a. Do any accuracy rates meet the 85% threshold of acceptability?
    b. How does this method compare to skeletal estimations based on the methods listed in Buikstra and Ubelaker (1994)?
        i. Would the classification parameters meet acceptable levels of accuracy if documented sex was unavailable and the functions were developed based on osteological estimations of sex?

Eight of the logistic regression models developed correctly estimated sex in more than 85% of the adult population. These equations, all based on canine odontometrics, therefore surpass the level of accuracy considered to be acceptable within adult osteological sex estimation. This signifies that this method could also be employed to estimate sex in other adult individuals in Middenbeemster collection whose state of preservation does not allow for morphometric sex estimation or whose morphological sex estimate was indeterminate. In this study, the morphologically estimated sexes of all adult individuals were consistent with documented sex, meaning that had documentation not been available for this collection, the approach would have performed equally well. As there is currently no accepted method of estimating sex in non-adults, in such a case validation would not be possible for non-adult skeletal samples. Nevertheless, the results of this study indicate that a similar approach based on the odontometrics of undocumented adults of morphologically estimated sex could produce reliable non-adult sex estimations.

3. Can this method be successfully applied to the permanent dentition of the documented non-adults from the Middenbeemster collection?
    a. Does the accuracy meet the 75% acceptability threshold for non-adult sexing methods? Does it meet the 85% considered acceptable in adult sexing?
    b. From what age can this method be reliably applied? What practical considerations, such as the embedment of the tooth in the jaw bone, may impede the application of this methodology?
    c. Is this method, previously applied to archaeological English and Italian populations and a modern Spanish sample, applicable to a post-medieval Dutch skeletal collection?

All of the 12 logit equations developed on canine odontometrics met the 75% threshold considered acceptable in non-adult sexing methods and eight surpassed the 85% adult threshold, including five that attained 100% accuracy in the non-adult sample. This demonstrates that this method is applicable to a post-medieval Dutch assemblage. Due to the need of at least one canine with a fully formed crown, this method is broadly applicable to any individual above the age of five years. In this study, however, no reliable sex estimate was possible for any individual under the age of seven years, because the canines of many individuals both below and above the age of seven years were lost, damaged, or unerupted, impeding the ability to take measurements.

4. Do any extrinsic factors influence tooth dimensions, such as age or health?
    a. Is selective mortality discernable in the non-adult sample?

No differential mortality bias was discernable in the non-adult sample when compared to the adult sample, allowing for the accurate application of logistic regression models developed on the adult sample to be applied to the non-adult sample. Age did not affect the accuracy of the method, but microdontia was shown to be a confounding factor as this condition can alter dental dimensions. No other pathological changes were found to alter the efficacy of this approach.

6.4 Future research

The promising results attained by this and previous studies using this approach suggest that this method can be applied to undocumented skeletal assemblages with some confidence. Of course, future assays on larger documented archaeological collections are desirable and could contribute further information as to the reliability and applicability of the method, but these collections are rare. As this and previous studies have achieved consistently high accuracy rates using the canine odontometrics, there is substantial justification for the application of this approach to other archaeological assemblages. As all tests of this approach have been conducted on European populations, future assessments on populations outside of Europe would help to establish whether this approach is universally applicable within populations. Further research is needed into the effects of temporal and geographic separation between populations in order to ascertain whether regionally-applicable equations could be derived from the canine odontometrics as was done in England for the maxillary first molars by Aris and colleagues (2018).

The use of specialised dental callipers with thin measuring points, such as those developed and described by Hillson and colleagues (2005), would allow for more measurements to be taken on teeth that remain embedded in the alveolar bone. These callipers are especially useful for MD measurements, which were often obscured by neighbouring teeth in this study.

There is also the possibility that this type of calliper would be able to measure some of the unerupted canine crowns observed in the non-adult sample, allowing for the sex estimation of individuals who were excluded from the current research.

This thesis has shown that the MD and BL crown and cervical diameters of maxillary first molars are not well-suited for sex estimation in non-adult skeletal remains. Future attempts to estimate sex using the maxillary first molar should include the cusp measurements described by Aris and colleagues (2018) (Aris *et al*. 2018, 675). Overall, equations based on canine odontometrics have consistently proven to be more accurate and accessible than those based on the maxillary first molar. It is therefore recommended that canine odontometrics be used whenever possible.

6.5 Significance within archaeology

As no widely-accepted and reliable means of estimating sex in non-adult skeletal remains currently exist, non-adult individuals are habitually excluded from osteological and palaeodemographic studies concerned with biological sex. Though a plethora of approaches based on a wide array of potentially dimorphic pre-pubescent traits have been attempted, few have been able to attain the accuracy rates considered acceptable in adult or non-adult sex estimation and those that did subsequently failed to achieve acceptable accuracy rates when tested on a population other than the one on which it was originally developed. Additionally, research has shown that many of the traits used in morphological methods of non-adult sex estimation vary with age and so may be more closely tied to growth than to sex. The approach tested in this thesis not only successfully circumvents many of these issues, but has also been shown to provide consistently high rates of accurate sex estimations in non-adult individuals across several populations.

The results of the present research, in combination with previous studies, indicates that this method of osteological sex estimation can be reliably applied to any archaeological sample that includes a sufficient number of adult individuals for whom confident morphological sex estimation is possible. In addition to non-adult sex estimation, this approach can be used to estimate the sex of adult individuals for whom a decisive morphological sex estimate is not possible. As indicated by the consistently high rates of interobserver agreement in canine odontometrics, these measurements can be accurately collected by multiple separate researchers of varying levels of experience. The creation and application of binomial logistic regression equations would be relatively quick and simple if canine odontometrics were collected during the initial osteological analysis of an archaeological skeletal collection. Widespread adoption of this method would unearth a wealth of previously inaccessible information.

The ability to estimate sex in non-adult individuals would greatly improve the precision and accuracy of established osteological practices such as non-adult age estimation and growth studies. Sex estimates are necessary in order to properly contextualise an individual historically and to illuminate gendered patterns of behaviour and experience as they relate to biological sex. Palaeodemographic studies are hindered by the fact that, generally, sex and age estimates are possible for only a portion of any skeletal assemblage. As sex estimation in any individual under the age of 12 years at death is currently considered an unacceptable practice within forensic anthropology, these individuals are excluded from investigations into any gendered social, economic, or environmental implications of osteological evidence. Though the application of the approach outlined in this thesis would not allow for every individual within a collection to be sexed, it could significantly increase the number of individuals included in such studies, creating a better, more comprehensive picture of past populations. Through osteological indicators of nutritional stress, growth retardation, and the distribution of activity markers and of pathological conditions including trauma, it would be possible to investigate potentially sex-linked trends such as differential treatment, impact of and susceptibility to disease, mortality rates, access to resources, activity patterns, or division of labour among non-adults.

The experiences of children in the past are often inscrutable in both historical and archaeological sources, rendering these individuals virtually invisible to modern researchers. The ability to attain all possible information from non-adult osteoarchaeological material is therefore imperative in order to reach a more comprehensive and accurate understanding of the activities of children in the past, their treatment, and their place within society. Through the application of the method developed in this thesis, children can at last be made visible in archaeological and historical reconstructions of the past.

Abstract

Sex estimation of non-adult skeletal remains has long been regarded as a problematic or even an unattainable objective within physical anthropology and forensic science. Few extant methods have been able to match the accuracy rates of methods designed for adult remains and those that have failed to achieve similarly acceptable rates when tested on a population other than the one on which the method was originally developed. Due to this, children are habitually excluded from archaeological investigations since a major component of their biological profiles is considered inaccessible. A definitive and reliable technique to estimate sex in non-adult osteological remains would contribute greatly to the field of osteoarchaeology, allowing for the refinement of osteological age estimation and growth studies as well as more perceptive interpretations of the social, economic, or environmental implications of osteological evidence.

In this thesis, a population-specific statistical approach to non-adult sex estimation based on the crown and cervical dimensions of the permanent canines and maxillary first molars was tested on the documented post-medieval skeletal collection of Middenbeemster, the Netherlands. The odontometrics of the adult component of the population (n = 76) were used to develop 14 binomial logistic regression formulae, which were subsequently applied to the non-adult individuals of the same population (n = 15). Though the two formulae based on the maxillary first molar odontometrics performed little better than chance, all 12 of the formulae based on the permanent canines achieved accuracy rates above 75%, with eight surpassing 85% and five achieving 100% accuracy. It was demonstrated that as little as one dimension of the permanent maxillary or mandibular canine can be used to estimate sex with an acceptable level of confidence. Due to the necessity of a permanent canine, this method is only applicable to individuals aged five years and above at the time of death, including adult individuals whose state of preservation does not allow for morphometric sex estimation or whose morphological sex estimate was indeterminate. By allowing reliable sex estimation in non-adult skeletal remains, this method makes it possible to gain insights into the past lives of non-adult individuals, who all too often seem invisible in archaeological and historical narratives.

Internet Pages

https://mapmaker.nationalgeographic.org/r9aIhZahSgEwjJbQpDFk7//?edit=hWp9tAYq33ZqcBJLLNQyiH, accessed on 2 April 2019.

https://www.distancefromto.net/distance-from-newcastle-upon-tyne-gb-to-london-gb, accessed on 29 April 2019.

Bibliography

Afacan, G.O., T. Onal, G. Akansel and A.S. Arslan, 2017. Is the lateral angle of the internal acoustic canal sexually dimorphic in non-adults? An investigation by routine cranial magnetic resonance imaging. *Journal of Comparative Human Biology* 68 (5), 393-397.

Aris, C., P. Nystrom and E. Craig-Atkins, 2018. A new multivariate method for determining sex of immature human remains using the maxillary first molar. *American Journal of Physical Anthropology* 167 (3), 672-683.

Baker, B.J., M.W. Tocheri and T.L. Dupras, 2005. *The Osteology of Infants and Children*. College Station: Texas A & M University Press (Texas A & M University Anthropology Series 12).

Black, T.K., 1978. Sexual dimorphism in the tooth-crown diameters of the deciduous teeth. *American Journal of Physical Anthropology* 48 (1), 77-82.

Blom, A.A., S.A. Inskip, W.A. Baesten and M.L.P. Hoogland, 2018. Testing the sternal clavicle ageing method on a post-medieval Dutch skeletal collection. *Archaeometry* 60 (6), 1391-1402.

Bosman, A.M., S.R. Moisik, D. Dediu and A. Waters-Rist, 2017. Talking heads: morphological variation in the human mandible over the last 500 years in the Netherlands. *Journal of Comparative Human Biology* 68, 329-342.

Buikstra, J.E. and D. H. Ubelaker (eds), 1994. *Standards for Data Collection from Human Skeletal Remains*, Fayetteville (AR): Arkansas Archeological Survey.

Cabana, G.S., B.I. Hulsey and F.L. Pack, 2013. Molecular methods, in E.A. DiGiangi and M.K. Moore (eds), *Research Methods in Human Skeletal Biology*. Oxford (UK): Academic Press, 449-480.

Cantekin, K. and M. Celikoglu, 2015. Comparison of the dental anomaly frequency in patients with and without mandibular second premolar agenesis. *Journal of Dental Sciences* 10 (2), 185-189.

Cardoso, H.F.V., 2008. Sample-specific (universal) metric approaches for determining the sex of immature human skeletal remains using permanent tooth dimensions. *Journal of Archaeological Science* 35, 158-168.

Cardoso, H.F.V. and S.R. Saunders, 2008. Two arch criteria of the ilium for sex determination of immature skeletal remains: A test of their accuracy and an assessment of intra- and inter-observer error. *Forensic Science International* 178 (1), 24-29.

Carroll, G.M.A., A. Waters-Rist and S.A. Inskip, 2016. A suspected case of intranasal inverted Schneiderian papilloma in an adult male from post-Medieval Holland. *International Journal of Paleopathology* 12, 53-62.

Chilcote, C., 2018. *Social identity over the lifecourse at historic Middenbeemster: a biocultural approach*. Berkeley (PhD thesis University of California, Berkeley).

Cooper, A. and H.N. Poinar, 2000. Ancient DNA: do it right or not at all. *Science* 289 (5482), 1139.

Coqueugniot, H., G. Giacobini and G. Malerba, 2002. L'utilisation de caractères morphologiques dans la diagnose sexuelle des mandibules d'enfants: application à la collection ostéologique de Turin (Italie). *Bulletins et Mémoires de la Société d'Anthropologie de Paris* 14 (1-2), 131-139.

Cunningham, C., L. Scheuer and S. Black, 2016. *Developmental Juvenile Osteology*. Amsterdam: Academic Press.

Hammerl, E., 2013. Dental anthropology, in E.A. DiGiangi and M.K. Moore (eds), *Research Methods in Human Skeletal Biology*. Oxford (UK): Academic Press, 263-291.

Hassett, B., 2011. Technical note: estimating sex using cervical canine odontometrics: a test using a known sample. *American Journal of Physical Anthropology* 146 (3), 486-489.

Hillson, S., 1996. *Dental Anthropology*. Cambridge: Cambridge University Press.

Hillson, S., 2005. *Teeth*. Cambridge: Cambridge University Press (Cambridge Manuals in Archaeology).

Hillson, S., 2014. *Tooth Development in Human Evolution and Bioarchaeology*. Cambridge: Cambridge University Press.

Hillson, S., C. FitzGerald and H. Flinn, 2005. Alternative dental measurements: proposals and relationships with other measurements. *American Journal of Physical Anthropology* 126 (4), 413-426.

Hua, F.; H. He, P. Ngan and W. Bouzid, 2013. Prevalence of peg-shaped maxillary permanent lateral incisors: a meta-analysis. *American Journal of Orthodontics and Dentofacial Orthopedics* 144(1), 97-109.

Inskip, S., M. Constantinescu, A. Brinkman, M. Hoogland and J. Sofaer, 2018. The effect of population variation on the accuracy of sex estimates derived from basal occipital discriminant functions. *Archaeological and Anthropological Sciences* 10 (3), 675-683.

Kemkes-Grottenthaler, A., F. Lobig and F. Stock, 2002. Mandibular ramus flexure and gonial eversion as morphologic indicators of sex. *Journal of Comparative Human Biology* 53 (2), 97-111.

Klales, A.R. and T.L. Burns, 2017. Adapting and applying the Phenice (1969) adult morphological sex estimation technique to subadults. *Journal of Forensic Sciences* 62 (3), 747-752.

Loth, S.R. and M. Henneberg, 2001. Sexually dimorphic mandibular morphology in the first few years of life. *American Journal of Physical Anthropology* 115 (2), 179-186.

Lewis, M.E., 2006. *The Bioarchaeology of Children: Perspectives from Biological and Forensic Anthropology*. Cambridge: Cambridge University Press.

Luna, L.H., C.M. Aranda and A.L. Santos, 2017. New method for sex prediction using the human non-adult auricular surface of the ilium in the Collection of Identified Skeletons of the University of Coimbra. *International Journal of Osteoarchaeology* 27 (5), 98-911.

Molleson, T., K. Cruse and S. Mays, 1998. Some sexually dimorphic features of the human juvenile skull and their value in sex determination in immature skeletal remains. *Journal of Archaeological Science* 25 (8), 719-728.

Moore, M.K., 2013. Sex estimation and assessment, in E.A. DiGiangi and M.K. Moore (eds), *Research Methods in Human Skeletal Biology*. Oxford (UK): Academic Press, 91-116.

Olivares, J.I. and I.A. Aguilera, 2016. Validation of the sex estimation method elaborated by Schutkowski in the Granada Osteological Collection of identified infant and young children: analysis of the controversy between the different way of analysing and interpreting the results. *International Journal of Legal Medicine* 130 (6), 1623-1632.

Ortner, D.J., 2003. *Identification of Pathological Conditions in Human Skeletal Remains*. Oxford: Academic Press.

Palmer, J.L.A., M. Hoogland and A.L. Waters-Rist, 2016. Activity reconstruction of post-medieval Dutch rural villagers from upper limb osteoarthritis and entheseal changes. *International Journal of Osteoarchaeology* 26 (1), 78-92.

Sajnani, A.K. and N.M. King, 2014. Dental anomalies associated with buccally- and palatally-impacted maxillary canines. *Journal of Investigative and Clinical Dentistry* 5 (3), 208-213.

Saunders, S., A. Chan, B. Kahlon, H. Kluge and C. FitzGerald, 2007. Sexual dimorphism of the dental tissues in human permanent mandibular canines and third premolars. *American Journal of Physical Anthropology* 133 (1), 735-740.

Saunders, S., 2008. Juvenile skeletons and growth-related studies, in M.A. Katzenberg and S. Saunders (eds), *Biological Anthropology of the Human Skeleton*. Hoboken (NJ): John Wiley and Sons, 117-148.

Scheuer, L., 2002. Brief communication: a blind test of mandibular morphology for sexing mandibles in the first few years of life. *American Journal of Physical Anthropology* 119 (2), 189-191.

Scheuer, L. and S.M. Black, 2004. *The Juvenile Skeleton*. Boston (MA): Academic Press.

Schutkowski, H., 1993. Sex determination of infant and juvenile skeletons: morphognostic features. *American Journal of Physical Anthropology* 90 (2), 199-205.

Schwartz, G.T. and M.C. Dean, 2005. Sexual dimorphism in modern human permanent teeth. *American Journal of Physical Anthropology* 128 (2), 213-217.

Scientific Working Group for Forensic Anthropology, 2010. *Sex Assessment*. New York: Scientific Working Group for Forensic Anthropology. https://www.nist.gov/sites/default/files/documents/2018/0 3 /13/swganth_sex_assessment.pdf, accessed on 1 October 2018.

Shankar, S., K. Anuthama, M. Kruthika, V. Suresh Kumar, K. Ramesh, A. Jaheerdeen and M. Mohamed Yasin, 2013. Identifying sexual dimorphism in a paediatric South Indian population using stepwise discriminant function analysis. *Journal of Forensic and Legal Medicine* 20 (6), 752-756.

Stull, K., E. L'Abbé and S. Ousley, 2017. Subadult sex estimation from diaphyseal dimensions. *American Journal of Physical Anthropology* 163 (1), 64-74.

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

Tuttösí, P. and H.F.V. Cardoso, 2015. An assessment of sexual dimorphism and sex estimation using cervical dental measurements in a Northwest Coast archeological sample. *Journal of Archaeological Science: Reports 3*, 306-312.

Uhl, N.M., 2013. Age-at-death estimation, in E.A. DiGiangi and M.K. Moore (eds), *Research Methods in Human Skeletal Biology*. Oxford (UK): Academic Press, 63-90.

UNESCO, 2019. *World Heritage List: Droogmakerij de Beemster (Beemster Polder)*. https://whc.unes co.org/en/list/899, accessed on 9 March 2019.

Veselka, B., M.L.P. Hoogland and A.L. Waters-Rist, 2015. Rural rickets: vitamin D deficiency in a post-medieval farming community from the Netherlands. *International Journal of Osteoarchaeology* 25 (5), 665-675.

Veselka, B., A.E. van der Merwe, M.L.P. Hoogland and A.L. Waters-Rist, 2018. Gender-related vitamin D deficiency in a Dutch 19th century farming community. *International Journal of Paleopathology* 23, 69-75.

Vikatou, I., M.L.P. Hoogland and A.L. Waters-Rist, 2017. Osteochondritis dissecans of skeletal elements of the foot in a 19th century rural farming community from the Netherlands. *International Journal of Paleopathology* 19, 53-63.

Viciano, J., I. Alemán, R. D'Anastasio, L. Capasso and M.C. Botella, 2011. Odontometric sex discrimination in the Herculaneum sample (79 AD, Naples, Italy), with application to juveniles. *American Journal of Physical Anthropology* 145 (1), 97-106.

Viciano, J., S. López-Lázaro and I. Alemán, 2013. Sex estimation based on deciduous and permanent dentition in a contemporary Spanish population. *American Journal of Physical Anthropology* 152 (1), 31-43.

Waters-Rist, A.L. and M.L.P. Hoogland, 2013. Osteological evidence of short-limbed dwarfism in a nineteenth century Dutch family: achondroplasia or hypochondroplasia. *International Journal of Paleopathology* 3, 243-256.

White, T.D. and P.A. Folkens, 2005. *The Human Bone Manual*, San Diego (CA): Elsevier Academic Press.

Wilson, L.A.B., H.F.V. Cardoso and L.T. Humphrey, 2011. On the reliability of a geometric morphometric approach to sex determination: a blind test of six criteria of the juvenile ilium. *Forensic Science International* 206 (1), 35-42.

Wilson, L.A.B., R. Ives and L.T. Humphrey, 2016. Quantification of 3D curvature in the iliac crest: ontogeny and implications for sex determination in juveniles. *Americal Journal of Physical Anthropology* 162 (2), 255-266.

Wilson, L.A.B and L.T. Humphrey, 2017. Voyaging in to the third dimension: a perspective on virtual methods and their application to studies of juvenile sex estimation and the ontogeny of sexual dimorphism. *Forensic Science International* 278, 32-34.

Wood, J.W., G.R. Milner, H.C. Harpending, K.M. Weis, M.N. Cohen, L.E. Eisenberg, D.L. Hutchinson, R. Jankauskas, G. Česnys, M. A. Katzenberg, J. R. Lukacs, J. W. McGrath, E. A. Roth, D. H. Ubelaker and R. G. Wilkinson, 1992. The osteological paradox: problems of inferring prehistoric health from skeletal samples. *Current Anthropology* 33 (40), 343-370.

Żądzińska, E., M. Karasińska, K. Jedrychowska-Dańska, C. Watalac and H.W. Witasb, 2008. Sex diagnosis of subadult specimens from medieval Polish archaeological sites: metric analysis of deciduous dentition. *Journal of Comparative Human Biology* 59 (3), 175-187.

List of Figures

List of Tables

List of Appendices

Appendix 1

Results of the Kolmogorov-Smirnov one-sample test for normalcy on the adult sample.

Measurement	*N*	Mean	SD	*Z*	Sig. (2-tailed)
1.3 MDcrn	27	7.4181	0.39632	0.655	0.784
1.3 BLcrn	37	8.0697	0.62631	0.518	0.951
1.3 MDcerv	32	5.3516	0.48661	0.477	0.977
1.3 BLcerv	37	7.6267	0.73097	0.584	0.885
2.3 MDcrn	25	7.4324	0.35758	0.431	0.992
2.3 BLcrn	36	8.0292	0.65415	0.498	0.965
2.3 MDcerv	32	5.3100	0.54026	0.590	0.877
2.3 BLcerv	36	7.5419	0.80867	0.570	0.901
3.3 MDcrn	26	6.5600	0.38391	0.653	0.788
3.3 BLcrn	51	7.6022	0.61132	0.631	0.820
3.3 MDcerv	44	5.1023	0.49470	0.639	0.808
3.3 BLcerv	47	7.4487	0.68353	0.592	0.875
4.3 MDcrn	30	6.4570	0.37187	0.814	0.521
4.3 BLcrn	53	7.5306	0.64590	0.535	0.937
4.3 MDcerv	48	5.0167	0.51436	0.434	0.992
4.3 BLcerv	49	7.3380	0.69694	0.521	0.949
1.6 MDcrn	16	10.3006	0.48179	0.776	0.583
1.6 BLcrn	34	11.1756	0.65916	0.523	0.947
1.6 MDcerv	15	7.9827	0.89146	0.806	0.535
1.6 BLcerv	26	10.2438	0.67832	0.642	0.804
2.6 MDcrn	11	10.5564	0.59562	0.535	0.937
2.6 BLcrn	24	11.2550	0.74505	0.526	0.945
2.6 MDcerv	8	8.2988	0.87965	0.656	0.783
2.6 BLcerv	16	10.4613	0.66805	0.594	0.872

N number of observations. SD standard deviation of the mean.
Z Kolmogorov-Smirnov Z value. No values significant at the $P \leq 0.05$ level.

Appendix 2

The predicted values and standard residuals of each model for all individuals in the adult sample.

Individual	Model 1		Model 2		Model 3	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.00000	0.00000	0.00000	0.00000	0.06004	-0.25273
S051V0059						
S053V0290			0.00000	-0.00021		
S059V0133						
S060V0037	0.19069	-0.48541	0.00010	-0.00980		
S088V0094						
S092V0124						
S093V0126						
S100V0159						
S101V0131						
S149V0280			0.00001	-0.00283		
S151V0666	0.00000	-0.00023	0.00018	-0.01350		
S153V0435	0.99988	0.01109			0.73282	0.60382
S155V1509						
S158V0427						
S160V0613			0.00000	0.00000		
S174V0408			0.44715	-0.89934		
S192V0636						
S194V0440	0.99974	0.01607				
S195V0588						
S202V0284					0.64094	-1.33605
S236V0335	0.96475	0.19114			0.73916	0.59404
S239V0369	0.96133	0.20055				
S243V0381	0.00000	-0.00062				
S246V0396	0.77061	0.54560	0.38895	1.25341	0.86210	0.39994
S285V0452						
S302V0509						
S303V0520						
S306V0561					0.67311	0.69687
S310V550			0.43004	1.15125	0.20606	1.96290
S313V0926						
S317V0649						
S324V0671						
S325V0676						
S327V0758	0.00033	-0.01812				
S337V0714					0.36310	1.32441
S338V0721	0.00000	-0.00026	0.00000	0.00000		
S339V0728						

Individual	Model 1		Model 2		Model 3	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S340V0724	0.99702	0.05465				
S342V0737						
S344V0730	0.00000	-0.00022	0.00000	0.00000		
S345V0757	0.00213	-0.04615	0.00028	-0.01669		
S346V0733						
S347V0741					0.64686	0.73887
S350V0844	0.37760	-0.77890	0.11292	-0.35677	0.26063	-0.59372
S359V0760						
S363V0766						
S369V0886	0.77167	-1.83839	0.06390	-0.26127		
S383V0880	0.00000	-0.00004			0.06151	-0.25601
S385V0874	0.00000	-0.00156	0.00000	-0.00010	0.34772	-0.73013
S386V0848			0.00000	0.00000	0.75781	-1.76889
S387V0914						
S388V0952						
S390V0831					0.06944	-0.27317
S394V0869						
S413V0896	0.00000	-0.00032	0.00000	0.00000		
S415V9999						
S422V0962	0.00000	-0.00001	0.00000	0.00000		
S427V0938			1.00000		0.99911	0.02985
S430V0965	0.00002	-0.00441	0.54878	-1.10283		
S435V0929						
S436V0911						
S441V0932	0.00000	-0.00003	0.00000	0.00000		
S446V0944	0.64056	0.74909	1.00000	0.00000		
S454V0963	0.28726	1.57517	0.99992	0.00895	0.48267	1.03527
S461V0990	0.00002	-0.00447	0.00000	0.00000	0.19899	-0.49842
S466V1010						
S473V1003						
S476V1054	0.08922	-0.31299	0.00790	-0.08922	0.49293	-0.98595
S482V1048						
S486V1088						
S487V1096					0.40500	-0.82502
S501V1097	0.00000	-0.00030	0.00000	0.00000		
S502V1062	1.00000	0.00092	0.99988	0.01102		
S521V1150						
S544V1510	0.94716	0.23619				

Individual	Model 4		Model 5		Model 6	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.00210	-0.04586	0.01145	-0.10765	0.00000	0.00000
S051V0059						
S053V0290						
S059V0133			0.44236	1.12277		
S060V0037	0.65072	-1.36493	0.51822	-1.03712	0.19902	-0.49847
S088V0094	0.00828	-0.09138				
S092V0124						
S093V0126			0.49914	1.00171		
S100V0159						
S101V0131						
S149V0280	0.19621	-0.49407				
S151V0666	0.42034	-0.85155	0.10365	-0.34006	0.00000	-0.00026
S153V0435	0.78781	0.51898	0.93993	0.25281	0.99984	0.01276
S155V1509						
S158V0427						
S160V0613						
S174V0408			0.80952	-2.06151		
S192V0636	0.07852	-0.29191				
S194V0440	0.82194	0.46545	0.97351	0.16495	0.99948	0.02280
S195V0588						
S202V0284						
S236V0335	0.28123	1.59869	0.79909	0.50143	0.95890	0.20703
S239V0369	0.42034	1.17433	0.84018	0.43614	0.94882	0.23224
S243V0381	0.09269	-0.31963	0.07868	-0.29223	0.00000	-0.00078
S246V0396	0.75590	0.56827	0.74446	0.58588	0.83541	0.44386
S285V0452						
S302V0509	0.00991	-0.10006	0.00149	-0.03865		
S303V0520	0.05407	-0.23909	0.07789	-0.29065		
S306V0561	0.68295	0.68134	0.92441	0.28596		
S310V550	0.36839	1.30941				
S313V0926						
S317V0649						
S324V0671			0.18460	2.10172		
S325V0676	0.92728	0.28003				
S327V0758	0.31932	-0.68492	0.16139	-0.43868	0.00039	-0.01970
S337V0714						
S338V0721	0.03325	-0.18545	0.05629	-0.24423	0.00000	-0.00031
S339V0728	0.00482	-0.06961				
S340V0724	0.91117	0.31224	0.78279	0.52677	0.99753	0.04975
S342V0737						
S344V0730	0.05048	-0.23056	0.01899	-0.13913	0.00000	-0.00032
S345V0757	0.15440	-0.42731	0.37780	-0.77923	0.00177	-0.04205
S346V0733	0.13216	-0.39024				
S347V0741	0.39406	1.24003				

Individual	Model 4		Model 5		Model 6	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.51945	-1.03968	0.74837	-1.72455	0.37350	-0.77212
S359V0760	0.01637	-0.12900				
S363V0766						
S369V0886	0.33530	-0.71023	0.66441	-1.40707	0.78103	-1.88862
S383V0880	0.02508	-0.16039	0.02406	-0.15703	0.00000	-0.00005
S385V0874	0.25284	-0.58172	0.18374	-0.47445	0.00000	-0.00153
S386V0848	0.01696	-0.13137	0.16209	-0.43982		
S387V0914						
S388V0952						
S390V0831						
S394V0869	0.01028	-0.10190				
S413V0896	0.03210	-0.18211	0.02260	-0.15205	0.00000	-0.00046
S415V9999						
S422V0962	0.00667	-0.08196	0.01086	-0.10480	0.00000	-0.00002
S427V0938	0.87311	0.38122	0.94079	0.25086		
S430V0965	0.24604	-0.57126	0.17783	-0.46507	0.00002	-0.00414
S435V0929						
S436V0911			0.04374	-0.21388		
S441V0932	0.06414	-0.26179	0.04565	-0.21870	0.00000	-0.00003
S446V0944	0.82719	0.45708	0.85535	0.41123	0.58279	0.84609
S454V0963	0.17432	2.17639	0.47546	1.05034	0.30486	1.51004
S461V0990	0.34343	-0.72324	0.19203	-0.48751	0.00005	-0.00695
S466V1010						
S473V1003						
S476V1054	0.33530	-0.71023	0.56846	-1.14772	0.08306	-0.30097
S482V1048	0.51945	0.96184	0.67403	0.69542		
S486V1088						
S487V1096						
S501V1097	0.04100	-0.20678	0.05442	-0.23989	0.00000	-0.00036
S502V1062	0.91117	0.31224	0.97995	0.14302	1.00000	0.00092
S521V1150	0.01230	8.96274	0.13558	2.52502		
S544V1510	0.09269	3.12863	0.69474	0.66286	0.93353	0.26684

Individual	Model 7		Model 8		Model 9	
	Predicted Value	**Standardized Residual**	**Predicted Value**	**Standardized Residual**	**Predicted Value**	**Standardized Residual**
S047V0045	0.00262	-0.05122	0.00000	-0.00058	0.00838	-0.09192
S051V0059					0.96586	0.18801
S053V0290			0.00169	-0.04113	0.13200	-0.38996
S059V0133					0.47505	1.05121
S060V0037	0.27644	-0.61810	0.16268	-0.44079	0.40844	-0.83092
S088V0094			0.00000	-0.00072		
S092V0124						
S093V0126					0.77603	0.53723
S100V0159					0.10128	2.97881
S101V0131					0.11981	-0.36894
S149V0280			0.01267	-0.11327	0.23535	-0.55478
S151V0666	0.01485	-0.12277	0.02488	-0.15975	0.27494	-0.61578
S153V0435	0.93775	0.25765			0.90449	0.32496
S155V1509			0.00005	-0.00739		
S158V0427						
S160V0613			0.00000	-0.00174	0.02046	-0.14454
S174V0408			0.45163	-0.90751	0.66412	-1.40616
S192V0636			0.00201	-0.04486		
S194V0440	0.94980	0.22990			0.95217	0.22412
S195V0588					0.79409	-1.96379
S202V0284			0.00000	-0.00067		
S236V0335	0.88746	0.35610				
S239V0369	0.85311	0.41494			0.95828	0.20864
S243V0381	0.06108	-0.25507	0.00004	-0.00643		
S246V0396	0.79826	0.50272	0.50878	0.98259	0.56905	0.87024
S285V0452					0.92842	0.27767
S302V0509	0.00047	-0.02179			0.00314	-0.05612
S303V0520	0.07180	-0.27812			0.11309	-0.35709
S306V0561	0.86692	0.39180			0.84078	0.43517
S310V550			0.07570	3.49437	0.56905	0.87024
S313V0926					0.07833	3.43029
S317V0649					0.96922	0.17819
S324V0671						
S325V0676					0.97184	0.17023
S327V0758	0.05246	-0.23529			0.27509	-0.61602
S337V0714					0.93415	0.26550
S338V0721	0.05235	-0.23503	0.00009	-0.00964	0.04156	-0.20825
S339V0728			0.00000	-0.00142		
S340V0724	0.62389	0.77643			0.98172	0.13646
S342V0737					0.99726	0.05242
S344V0730	0.01482	-0.12263	0.00002	-0.00462	0.01295	-0.11453
S345V0757	0.25905	-0.59128	0.05963	-0.25183	0.25168	-0.57993
S346V0733						
S347V0741					0.84407	0.42981

Individual	Model 7		Model 8		Model 9	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.73695	-1.67381	0.02816	-0.17024	0.78273	-1.89807
S359V0760						
S363V0766					0.76182	0.55914
S369V0886	0.75682	-1.76412	0.22267	-0.53522	0.45109	-0.90653
S383V0880	0.01133	-0.10705				
S385V0874	0.04254	-0.21079	0.00980	-0.09947	0.09003	-0.31454
S386V0848	0.23990	-0.56179	0.00088	-0.02973	0.07994	-0.29477
S387V0914						
S388V0952						
S390V0831						
S394V0869	0.00936	-0.09719				
S413V0896	0.02554	-0.16189	0.00006	-0.00758	0.04362	-0.21357
S415V9999						
S422V0962	0.00567	-0.07551	0.00000	-0.00033	0.00160	-0.04009
S427V0938	0.98144	0.13751	0.99999	0.00323	0.99406	0.07730
S430V0965	0.03648	-0.19459	0.17959	-0.46787	0.45625	-0.91601
S435V0929						
S436V0911					0.10415	-0.34096
S441V0932	0.00682	-0.08289	0.00003	-0.00534	0.04068	-0.20593
S446V0944	0.64821	0.73668	0.99265	0.08604	0.97813	0.14955
S454V0963	0.60319	0.81109	0.85391	0.41362	0.61073	0.79836
S461V0990	0.27894	-0.62197	0.00027	-0.01639	0.11449	-0.35957
S466V1010			0.11162	-0.35446		
S473V1003					0.90356	0.32670
S476V1054	0.49100	-0.98215	0.42036	-0.85159	0.18735	-0.48015
S482V1048	0.46478	1.07310			0.97467	0.16120
S486V1088						
S487V1096						
S501V1097	0.04280	-0.21147	0.00000	-0.00043	0.01737	-0.13294
S502V1062	0.99223	0.08849	0.88012	0.36907	0.91156	0.31148
S521V1150	0.03814	5.02216			0.32401	1.44441
S544V1510	0.86473	0.39551				

Individual	Model 10		Model 11		Model 12	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.00000	0.00000	0.24333	-0.56708	0.14480	-0.41148
S051V0059						
S053V0290	0.00000	-0.00085				
S059V0133			0.25604	1.70458		
S060V0037	0.00100	-0.03169				
S088V0094			0.30283	-0.65907		
S092V0124						
S093V0126						
S100V0159						
S101V0131						
S149V0280	0.00010	-0.01012				
S151V0666	0.00317	-0.05637	0.68471	-1.47366		
S153V0435			0.41037	1.19867	0.71494	0.63144
S155V1509	0.00000	-0.00001				
S158V0427						
S160V0613	0.00000	0.00000				
S174V0408	0.30180	-0.65745				
S192V0636			0.34963	-0.73321		
S194V0440						
S195V0588						
S202V0284			0.53960	-1.08260	0.64030	-1.33419
S236V0335			0.37668	1.28639	0.80617	0.49035
S239V0369						
S243V0381			0.31807	-0.68295		
S246V0396	0.46281	1.07735	0.60140	0.81411	0.55446	0.89641
S285V0452						
S302V0509						
S303V0520						
S306V0561			0.53706	0.92844	0.15531	2.33214
S310V550	0.34068	1.39114	0.38448	1.26527	0.38659	1.25966
S313V0926						
S317V0649						
S324V0671						
S325V0676						
S327V0758						
S337V0714			0.44395	1.11915	0.48950	1.02123
S338V0721	0.00000	0.00000				
S339V0728						
S340V0724						
S342V0737					0.87915	0.37075
S344V0730	0.00000	0.00000			0.09107	-0.31653
S345V0757	0.00236	-0.04865				
S346V0733						
S347V0741			0.48992	1.02036	0.56495	0.87753

Individual	Model 10		Model 11		Model 12	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.13467	-0.39450	0.43265	-0.87326	0.48650	-0.97336
S359V0760						
S363V0766						
S369V0886	0.06779	-0.26968				
S383V0880			0.21502	-0.52337	0.14166	-0.40625
S385V0874	0.00000	-0.00127	0.40682	-0.82815	0.51484	-1.03013
S386V0848	0.00000	-0.00005	0.37962	-0.78225	0.66381	-1.40516
S387V0914						
S388V0952						
S390V0831			0.33556	-0.71065	0.09005	-0.31459
S394V0869						
S413V0896	0.00000	0.00000				
S415V9999						
S422V0962	0.00000	0.00000				
S427V0938	1.00000	0.00000	0.88056	0.36829	0.96846	0.18046
S430V0965	0.63908	-1.33068				
S435V0929						
S436V0911						
S441V0932	0.00000	0.00000	0.43768	-0.88224		
S446V0944	1.00000	0.00002				
S454V0963	0.99936	0.02534	0.30127	1.52293	0.62829	0.76916
S461V0990	0.00000	-0.00002	0.37722	-0.77826	0.15025	-0.42050
S466V1010						
S473V1003						
S476V1054	0.04861	-0.22605	0.41610	-0.84418	0.54961	-1.10467
S482V1048						
S486V1088						
S487V1096			0.40642	-0.82746	0.37930	-0.78171
S501V1097	0.00000	0.00000				
S502V1062	0.99855	0.03810				
S521V1150						
S544V1510			0.47301	1.05551		

Individual	Model 13		Model 14		Model 15	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.19247	-0.48820	0.02930	-0.17374	0.00688	-0.08326
S051V0059						
S053V0290						
S059V0133	0.18992	2.06526				
S060V0037			0.17839	-0.46597	0.71734	-1.59305
S088V0094	0.26980	-0.60786	0.08156	-0.29800	0.02249	-0.15167
S092V0124						
S093V0126						
S100V0159						
S101V0131	0.29245	-0.64291				
S149V0280					0.30229	-0.65823
S151V0666	0.76073	-1.78308	0.36802	-0.76311	0.52762	-1.05685
S153V0435	0.43202	1.14660	0.43457	1.14068	0.82216	0.46509
S155V1509						
S158V0427						
S160V0613						
S174V0408			0.52665	-1.05480		
S192V0636			0.53433	-1.07119	0.14782	-0.41648
S194V0440			0.65268	0.72949	0.84819	0.42307
S195V0588						
S202V0284	0.63457	-1.31778				
S236V0335	0.40911	1.20179	0.33295	1.41543	0.39507	1.23741
S239V0369			0.37522	1.29039	0.52762	0.94621
S243V0381	0.25133	-0.57940	0.25639	-0.58719	0.16882	-0.45068
S246V0396	0.58791	0.83721	0.65268	0.72949	0.79790	0.50328
S285V0452						
S302V0509					0.02623	-0.16413
S303V0520					0.10919	-0.35011
S306V0561	0.38820	1.25538			0.74223	0.58932
S310V550	0.45941	1.08476			0.48031	1.04019
S313V0926						
S317V0649						
S324V0671						
S325V0676					0.93114	0.27195
S327V0758			0.07279	-0.28018	0.43335	-0.87450
S337V0714	0.50364	0.99275				
S338V0721			0.19719	-0.49560	0.07304	-0.28070
S339V0728	0.14742	-0.41582	0.14892	-0.41831	0.01413	-0.11970
S340V0724	0.47553	1.05019	0.23359	1.81135	0.91796	0.29896
S342V0737						
S344V0730			0.06125	-0.25543	0.10320	-0.33923
S345V0757			0.17392	-0.45884	0.25182	-0.58015
S346V0733					0.22326	-0.53613
S347V0741	0.52167	0.95756			0.50397	0.99209

Individual	Model 13		Model 14		Model 15	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.50141	-1.00283	0.54964	-1.10473	0.61250	-1.25722
S359V0760					0.04022	-0.20471
S363V0766						
S369V0886			0.28060	-0.62453	0.44891	-0.90254
S383V0880	0.15062	-0.42110	0.10490	-0.34234	0.05768	-0.24740
S385V0874	0.44588	-0.89703	0.24481	-0.56936	0.36533	-0.75870
S386V0848	0.36966	-0.76579			0.04146	-0.20797
S387V0914						
S388V0952	0.19232	-0.48797	0.29321	-0.64409		
S390V0831	0.26742	-0.60419				
S394V0869					0.02705	-0.16674
S413V0896			0.06489	-0.26342	0.07093	-0.27630
S415V9999						
S422V0962			0.03610	-0.19354	0.01868	-0.13797
S427V0938	0.93989	0.25290			0.88772	0.35564
S430V0965			0.10784	-0.34766	0.35804	-0.74682
S435V0929						
S436V0911						
S441V0932	0.43371	-0.87515	0.16524	-0.44491	0.12552	-0.37885
S446V0944			0.62422	0.77589	0.85221	0.41644
S454V0963	0.31501	1.47462	0.15287	2.35403	0.27635	1.61822
S461V0990	0.29318	-0.64404	0.68008	-1.45801	0.45673	-0.91690
S466V1010						
S473V1003						
S476V1054	0.46934	-0.94046	0.26832	-0.60557	0.44891	-0.90254
S482V1048					0.61250	0.79540
S486V1088			0.04441	-0.21559		
S487V1096	0.41775	-0.84704				
S501V1097	0.06817	-0.27047	0.16524	-0.44491	0.08694	-0.30858
S502V1062			0.76598	0.55273	0.91796	0.29896
S521V1150					0.03153	5.54228
S544V1510	0.61943	0.78383	0.14127	2.46545	0.16882	2.21885

Individual	Model 16		Model 17		Model 18	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.03927	-0.20217	0.03473	-0.18968	0.00788	-0.08914
S051V0059			0.97109	0.17253		
S053V0290					0.00860	-0.09313
S059V0133	0.66960	0.70244	0.37474	1.29171		
S060V0037	0.23792	-0.55874	0.74269	-1.69892	0.40946	-0.83269
S088V0094					0.03705	-0.19615
S092V0124						
S093V0126	0.61358	0.79358	0.48234	1.03598		
S100V0159						
S101V0131						
S149V0280					0.05627	-0.24418
S151V0666	0.02756	-0.16836	0.44804	-0.90097	0.14577	-0.41309
S153V0435	0.84321	0.43121	0.93114	0.27194		
S155V1509					0.20915	-0.51426
S158V0427						
S160V0613			0.12500	-0.37796	0.05627	-0.24418
S174V0408	0.92383	-3.48264	0.61142	-1.25438	0.22401	-0.53729
S192V0636					0.16291	-0.44115
S194V0440	0.85867	0.40570	0.97333	0.16554		
S195V0588						
S202V0284					0.06111	-0.25512
S236V0335	0.86353	0.39753	0.68659	0.67563		
S239V0369	0.80181	0.49717	0.80510	0.49201		
S243V0381	0.15015	-0.42033	0.14768	-0.41626	0.20915	-0.51426
S246V0396	0.65137	0.73159	0.76320	0.55702	0.29071	1.56200
S285V0452						
S302V0509	0.00601	-0.07775	0.00923	-0.09651		
S303V0520	0.19653	-0.49457	0.12201	-0.37278		
S306V0561	0.76018	0.56167	0.93289	0.26822		
S310V550					0.06633	3.75171
S313V0926			0.07225	3.58346		
S317V0649						
S324V0671	0.12601	2.63366	0.40092	1.22239		
S325V0676			0.89941	0.33443		
S327V0758	0.08759	-0.30984	0.41424	-0.84094		
S337V0714			0.91342	0.30787		
S338V0721	0.17798	-0.46531	0.08850	-0.31160	0.05864	-0.24959
S339V0728			0.01389	-0.11869	0.12055	-0.37023
S340V0724	0.38394	1.26672	0.89431	0.34378		
S342V0737	0.97215	0.16924				
S344V0730	0.06009	-0.25284	0.04897	-0.22692	0.05627	-0.24418
S345V0757	0.37437	-0.77355	0.48234	-0.96527	0.07197	-0.27848
S346V0733						
S347V0741			0.64993	0.73391		

Individual	Model 16		Model 17		Model 18	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.66054	-1.39495	0.76320	-1.79526	0.45241	-0.90895
S359V0760						
S363V0766			0.85188	0.41698		
S369V0886	0.72929	-1.64134	0.61142	-1.25438	0.06633	-0.26654
S383V0880	0.06009	-0.25284	0.06519	-0.26408		
S385V0874	0.08131	-0.29750	0.47545	-0.95206	0.09919	-0.33182
S386V0848	0.53422	-1.07095	0.12804	-0.38320	0.03864	-0.20049
S387V0914	0.10526	-0.34300				
S388V0952					0.15702	-0.43159
S390V0831	0.01928	-0.14021				
S394V0869	0.06736	-0.26874			0.02041	-0.14434
S413V0896	0.10526	-0.34300	0.04181	-0.20890	0.02753	-0.16826
S415V9999						
S422V0962	0.05153	-0.23309	0.02730	-0.16754	0.00187	-0.04326
S427V0938	0.92664	0.28136	0.88898	0.35339	0.63503	0.75810
S430V0965	0.07265	-0.27989	0.48234	-0.96527	0.08805	-0.31072
S435V0929						
S436V0911	0.18400	-0.47487	0.06353	-0.26047		
S441V0932	0.02983	-0.17535	0.20313	-0.50489	0.04571	-0.21885
S446V0944	0.46318	1.07656	0.92752	0.27955	0.61450	0.79205
S454V0963	0.66054	0.71687	0.42095	1.17286	0.05627	4.09529
S461V0990	0.31919	-0.68472	0.25142	-0.57954	0.36785	-0.76283
S466V1010					0.27298	-0.61276
S473V1003			0.55109	0.90254		
S476V1054	0.50378	-1.00758	0.63728	-1.32551	0.05627	-0.24418
S482V1048	0.42306	1.16778	0.79179	0.51279		
S486V1088						
S487V1096						
S501V1097	0.14503	-0.41187	0.09782	-0.32927	0.06367	-0.26077
S502V1062	0.95876	0.20740	0.95692	0.21217	0.68416	0.67945
S521V1150	0.18400	2.10586	0.24118	1.77376		
S544V1510	0.88982	0.35188	0.48234	1.03598		

Individual	Model 19		Model 20		Model 21	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.01901	-0.13920	0.10986	-0.35131	0.00743	-0.08653
S051V0059			0.92194	0.29097	0.96201	0.19872
S053V0290	0.12696	-0.38134	0.45560	-0.91481	0.10911	-0.34996
S059V0133	0.46501	1.07262	0.62580	0.77327	0.44818	1.10961
S060V0037	0.63239	-1.31158	0.45560	-0.91481	0.42702	-0.86329
S088V0094	0.03362	-0.18652				
S092V0124	0.45693	1.09019			0.55526	0.89496
S093V0126	0.47310	1.05532	0.68983	0.67055	0.78530	0.52287
S100V0159			0.15115	2.36984	0.12699	2.62200
S101V0131	0.25295	-0.58190	0.21109	-0.51727	0.13683	-0.39814
S149V0280	0.29831	-0.65202	0.22497	-0.53877	0.28988	-0.63892
S151V0666	0.40908	-0.83204	0.03650	-0.19464	0.51249	-1.02529
S153V0435	0.69051	0.66947	0.93539	0.26281	0.86481	0.39538
S155V1509	0.15868	-0.43429			0.12230	-0.37328
S158V0427	0.85135	0.41785			0.96913	0.17847
S160V0613	0.05701	-0.24588	0.23215	-0.54985	0.01597	-0.12741
S174V0408	0.64737	-1.35492	0.86508	-2.53219	0.56585	-1.14164
S192V0636	0.28488	-0.63116				
S194V0440	0.82020	0.46821	0.79729	0.50422	0.96041	0.20304
S195V0588	0.80537	-2.03417	0.71535	-1.58526	0.79945	-1.99655
S202V0284	0.03811	-0.19905				
S236V0335	0.56215	0.88254				
S239V0369	0.77862	0.53322	0.81633	0.47434	0.96503	0.19036
S243V0381	0.15019	-0.42040				
S246V0396	0.68352	0.68044	0.45560	1.09313	0.61773	0.78666
S285V0452	0.74934	0.57837	0.91588	0.30307	0.91123	0.31212
S302V0509	0.04594	-0.21943	0.00833	-0.09163	0.00575	-0.07606
S303V0520	0.22349	-0.53648	0.33922	-0.71649	0.10500	-0.34252
S306V0561	0.72413	0.61722	0.89504	0.34245	0.79246	0.51175
S310V550	0.40908	1.20187	0.45560	1.09313	0.61773	0.78666
S313V0926	0.11654	2.75338	0.07881	3.41900	0.11775	2.73719
S317V0649	0.88471	0.36099	0.74732	0.58147	0.97883	0.14707
S324V0671	0.02612	6.10599				
S325V0676	0.92811	0.27832	0.83950	0.43724	0.97697	0.15354
S327V0758	0.48933	-0.97889	0.13141	-0.38896	0.39580	-0.80937
S337V0714	0.74934	0.57837	0.84492	0.42843	0.93540	0.26280
S338V0721	0.11992	-0.36914	0.08181	-0.29850	0.05568	-0.24283
S339V0728	0.06640	-0.26669				
S340V0724	0.95017	0.22900	0.60654	0.80542	0.99093	0.09568
S342V0737	0.97336	0.16545	0.98716	0.11404	0.99646	0.05962
S344V0730	0.08701	-0.30871	0.17325	-0.45778	0.01045	-0.10276
S345V0757	0.40125	-0.81862	0.14099	-0.40513	0.35549	-0.74268
S346V0733	0.30516	-0.66270	0.09147	-0.31730		
S347V0741	0.73058	0.60727	0.28677	1.57707	0.91795	0.29898

Individual	Model 19		Model 20		Model 21	
	Predicted Value	**Standardized Residual**	**Predicted Value**	**Standardized Residual**	**Predicted Value**	**Standardized Residual**
S350V0844	0.53802	-1.07916	0.60654	-1.24159	0.81288	-2.08427
S359V0760			0.15645	-0.43065		
S363V0766	0.63991	0.75015	0.85529	0.41133	0.70410	0.64827
S369V0886	0.48121	-0.96310	0.63529	-1.31980	0.41654	-0.84493
S383V0880						
S385V0874	0.32622	-0.69581	0.03119	-0.17942	0.18274	-0.47286
S386V0848	0.16756	-0.44864	0.32121	-0.68789	0.07091	-0.27626
S387V0914						
S388V0952			0.54697	-1.09880		
S390V0831	0.04317	-0.21242			0.01090	-0.10500
S394V0869			0.07881	-0.29248		
S413V0896	0.08701	-0.30871	0.22497	-0.53877	0.04013	-0.20446
S415V9999					0.23202	1.81933
S422V0962	0.00939	-0.09735	0.00867	-0.09351	0.00255	-0.05055
S427V0938	0.97998	0.14293	0.98005	0.14267	0.99235	0.08779
S430V0965	0.48933	-0.97889	0.11808	-0.36591	0.63782	-1.32705
S435V0929	0.96350	0.19465			0.99235	0.08779
S436V0911			0.24698	-0.57271	0.10911	-0.34996
S441V0932	0.08701	-0.30871	0.06776	-0.26960	0.05799	-0.24811
S446V0944	0.91136	0.31186	0.68983	0.67055	0.98725	0.11363
S454V0963	0.63991	0.75015	0.33015	1.42441	0.70410	0.64827
S461V0990	0.25914	-0.59143	0.08492	-0.30464	0.17640	-0.46280
S466V1010	0.56215	-1.13309	0.63529	-1.31980		
S473V1003	0.77296	0.54196	0.92482	0.28511	0.86976	0.38697
S476V1054	0.52182	-1.04464	0.06776	-0.26960	0.31713	-0.68148
S482V1048	0.86707	0.39155	0.82235	0.46478	0.98054	0.14088
S486V1088						
S487V1096						
S501V1097	0.03157	-0.18055	0.05817	-0.24852	0.02239	-0.15132
S502V1062	0.85135	0.41785	0.93029	0.27373	0.87919	0.37069
S521V1150	0.09507	3.08517	0.25464	1.71090	0.39580	1.23553
S544V1510						

Individual	Model 22		Model 23	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.38055	-0.78379	0.13471	-0.39457
S051V0059				
S053V0290				
S059V0133	0.30704	1.50231	0.13471	2.53441
S060V0037				
S088V0094	0.31895	-0.68433	0.20025	-0.50040
S092V0124				
S093V0126				
S100V0159				
S101V0131			0.32553	-0.69473
S149V0280				
S151V0666	0.50143	-1.00287	0.88645	-2.79400
S153V0435	0.47100	1.05978	0.44819	1.10959
S155V1509				
S158V0427				
S160V0613			0.22683	-0.54164
S174V0408				
S192V0636	0.47376	-0.94883	0.33052	-0.70264
S194V0440				
S195V0588				
S202V0284	0.47376	-0.94883	0.69218	-1.49956
S236V0335	0.54287	0.91764	0.40941	1.20105
S239V0369				
S243V0381	0.34223	-0.72131	0.23082	-0.54781
S246V0396	0.49174	1.01666	0.79098	0.51406
S285V0452				
S302V0509				
S303V0520				
S306V0561	0.37664	1.28649	0.65233	0.73004
S310V550	0.35353	1.35225	0.35085	1.36023
S313V0926				
S317V0649				
S324V0671				
S325V0676			0.47067	1.06049
S327V0758				
S337V0714	0.47376	1.05393	0.51589	0.96871
S338V0721				
S339V0728			0.13471	-0.39457
S340V0724			0.49326	1.01356
S342V0737				
S344V0730				
S345V0757			0.26451	-0.59970
S346V0733				
S347V0741	0.46135	1.08053	0.59941	0.81750

Individual	Model 22		Model 23	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.54287	-1.08975	0.52154	-1.04404
S359V0760				
S363V0766			0.75111	0.57565
S369V0886			0.39852	-0.81398
S383V0880	0.39105	-0.80135	0.10395	-0.34060
S385V0874	0.51666	-1.03389	0.45941	-0.92186
S386V0848	0.38973	-0.79913	0.35602	-0.74354
S387V0914				
S388V0952			0.08468	-0.30416
S390V0831	0.38316	-0.78815	0.27341	-0.61343
S394V0869				
S413V0896				
S415V9999				
S422V0962				
S427V0938	0.47238	1.05685	0.98770	0.11159
S430V0965				
S435V0929				
S436V0911				
S441V0932	0.36758	-0.76238	0.45941	-0.92186
S446V0944			0.81933	0.46958
S454V0963	0.39105	1.24789	0.21899	1.88849
S461V0990	0.41503	-0.84231	0.36123	-0.75200
S466V1010				
S473V1003				
S476V1054	0.38841	-0.79692	0.42592	-0.86136
S482V1048			0.74255	0.58882
S486V1088				
S487V1096	0.34223	-0.72131	0.38772	-0.79577
S501V1097			0.02539	-0.16141
S502V1062			0.77562	0.53786
S521V1150				
S544V1510	0.43125	1.14841	0.55527	0.89495

Individual	Model 24		Model 25	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S047V0045	0.39693	-0.81128	0.18626	-0.47843
S051V0059				
S053V0290				
S059V0133			0.18108	2.12663
S060V0037				
S088V0094			0.29188	-0.64202
S092V0124				
S093V0126				
S100V0159				
S101V0131			0.24765	-0.57374
S149V0280				
S151V0666			0.77284	-1.84448
S153V0435	0.74626	0.58311	0.45159	1.10199
S155V1509				
S158V0427				
S160V0613				
S174V0408				
S192V0636				
S194V0440				
S195V0588				
S202V0284	0.36303	-0.75494	0.69561	-1.51171
S236V0335	0.85322	0.41476	0.43027	1.15070
S239V0369			0.45588	1.09250
S243V0381			0.22882	-0.54472
S246V0396	0.49619	1.00764	0.51199	0.97630
S285V0452				
S302V0509				
S303V0520				
S306V0561	0.33045	1.42345	0.23811	1.78878
S310V550	0.24445	1.75805	0.58049	0.85011
S313V0926				
S317V0649				
S324V0671				
S325V0676				
S327V0758				
S337V0714	0.37196	1.29942	0.55502	0.89540
S338V0721			0.25416	-0.58375
S339V0728			0.10441	-0.34143
S340V0724			0.51199	0.97630
S342V0737	0.60955	0.80035	0.78766	0.51921
S344V0730	0.35640	-0.74416	0.13938	-0.40243
S345V0757				
S346V0733				
S347V0741	0.48181	1.03707	0.53357	0.93498

Individual	Model 24		Model 25	
	Predicted Value	Standardized Residual	Predicted Value	Standardized Residual
S350V0844	0.37871	-0.78073	0.54646	-1.09766
S359V0760				
S363V0766				
S369V0886				
S383V0880	0.51538	-1.03125	0.13128	-0.38874
S385V0874	0.49859	-0.99719	0.47309	-0.94756
S386V0848	0.75877	-1.77353	0.38437	-0.79016
S387V0914				
S388V0952			0.24765	-0.57374
S390V0831	0.20289	-0.50451	0.23189	-0.54945
S394V0869				
S413V0896				
S415V9999				
S422V0962				
S427V0938	0.52257	0.95584	0.94965	0.23025
S430V0965				
S435V0929				
S436V0911				
S441V0932			0.44731	-0.89963
S446V0944				
S454V0963	0.73333	0.60302	0.37621	1.28766
S461V0990	0.33471	-0.70929	0.22882	-0.54472
S466V1010				
S473V1003				
S476V1054	0.44840	-0.90161	0.54646	-1.09766
S482V1048				
S486V1088				
S487V1096	0.35640	-0.74416	0.46447	-0.93130
S501V1097			0.06281	-0.25888
S502V1062				
S521V1150				
S544V1510			0.75088	0.57600