Tracking Them All: Exploring Age-Related Variation in Sexual Dimorphism of the Human Pelvis

by

Jose Sanchez

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

DOCTORATE OF PHILOSPHY

Department of Anthropology

University of Manitoba

Winnipeg, MB, Canada

Abstract

Determining the biological characteristics of a skeleton, which include age, sex, and stature, is an integral step in skeletal analysis for bioarchaeologists and forensic anthropologists. Although sex determination in adult skeletons is highly accurate, particularly when conducted on the pelvis, this type of analysis continues to elude the field of juvenile osteology. It has been proposed that sex traits on the pelvis do not become sufficiently dimorphic until adulthood. This research examined age-related variation in the expression of sexual dimorphism in the human pelvis from an ontogenetic perspective. The main objective was to identify the age of appearance and stabilization of morphological and metric sex differences in the pelvis. This research also explored the relationship between sexually dimorphic pelvic traits and the attainment of puberty. Eighteen morphological traits and nine logistic regression equations were examined on 128 subadults (51 females and 77 males), aged 4 months to 20 years, from the Hamann-Todd and Terry Skeletal Collections. Pubertal stage assessment was also conducted based on skeletal indicators. This research showed that age-related trends exist in the appearance and stabilization of morphological pelvic traits. Three general patterns emerged from this research: traits either showed a male "default" expression, a female "default" expression, or concurrent sex expressions by birth or the time of pelvic fusion. Ten of the 18 morphological traits examined had an accuracy of 80% or above in individuals 17-20 years of age. Additionally, this research showed that the post-pubertal period is not required for the full expression of sexual dimorphism for all morphological traits. Instead, surpassing peak height velocity was shown to be more important since four traits and overall sex assessments showed substantial dimorphism occurring in the deceleration pubertal stage. Only one of the logistic regression equations tested in this research proved to be effective for sex assessment in subadults. Moreover, metric methods appeared to be best employed with age estimates as opposed to developmental (pubertal) stages. The novel strategy employed in this research to address subadult sexual dimorphism proved insightful to understand the complex nature of sexual dimorphism in the human skeleton.

Acknowledgments

First and foremost, a kraken sized thank you to my supervisor, Rob Hoppa. It's been almost a decade that I approached you (with the help of Em of course) as a nervous and awkward undergrad to chat about our shared research interests. Little did I know that a few years later you'd take a chance and take me on as a student, which is something I will be eternally grateful for. Your patience, encouragement, and support throughout these years have been invaluable, especially during my numerous nervous fits before each oral defense. You always knew what to say when I felt a little lost and unsure of myself, and always got me back on the right path. Thanks to you, I can say that I've certainly grown as a scholar and a professional. In addition to the professional development you have provided, the warm and welcoming environment you create for your students makes them (us) feel safe and at home, and for that I am extremely grateful. Should I have the opportunity to supervise students in the future, I am beyond glad that I have the perfect example of what a supportive supervisor is that I can model myself after. I feel incredibly fortunate to have had the honor and pleasure to have you as a supervisor (and mentor) through this whole process.

A big thank you also goes out to my committee members Drs. Emily Holland and Greg Smith. Your guidance, support, thought provoking questions and feedback have also been invaluable not just for this research, but also for my growth as a scholar. You have helped me consider aspects of my research from multiple perspectives, which has only served to strengthened it. I would also like to thank Dr. Daniel Wescott from Texas State University for his thought provoking feedback and questions during defense. You have certainly given me food for thought that will hopefully become mini projects in the future. Thank you to the Anthropology Department at the University of Manitoba for all their support, and a special thanks to Drs. Stacie Burke, Kathy Buddle, and Kent Fowler for their support and always checking up to see how things were going and taking the time to chat about anything and everything. Thank you to the funding sources that contributed to this research: the University of Manitoba Graduate Fellowship and the Faculty of Arts JG Fletcher Research Award.

My gratitude also goes out to Lyman Jellema and Dr. David Hunt for access to the Hamann-Todd and Terry Collections, respectively, and for their support during the data collection process.

I feel *extremely* fortunate to have also had the support of great friends during this process that I could count on to be sounding boards or to lean on. Emily Holland – I honestly don't know what to say. Where would I be without you? Words cannot describe how grateful I am to have met you back during my undergrad as that goofy student, although let's face it I'm still pretty goofy. It means the world that you reeaally took a chance on me so early on and have been there every step of the way. Your unwavering support and confidence in me over the past decade is one of the reasons I made it this far. From impromptu meetings, to being out in the field or in the lab with you, to laughing uncontrollably at the most inopportune moments, or just catching up, it has all meant so much to me and to know you have always been in my corner. I am the scholar, instructor, and professional I am today thanks to you. I am so honoured to be able to call you not just a mentor, but also a great friend. So a million times thank you and as I've said before, you're stuck with me forever ☺. Alexandra MacKay Brown – From day one, you have always been a delight! Thank you for being the best research assistant I could have asked for! I know all that shelf cranking and running back and forth was not easy, so I am beyond glad to have had that extra help. I look forward to being in your corner during your pursuit of a PhD. Tribunal – Laura Kelvin and Colleen Haukaas, y'all have always been the greatest cheerleaders and supporters a guy could ask for. Beer O'clocks are always a delight, even though all three of us haven't been in the same city for ages. I know both of you will always be there and I can't wait until we can have our next Tribunal reunion in person! The House of Hoppa: Amy Scott – You have no idea how grateful I am, and will always be, that you were pretty much my life support during my first year in Winnipeg. Thank you for always being there, for always being a huge support, and of course, for all the uncontrollable laughter! I am so grateful to know such a brilliant mind with a huge heart. Sarah Duignan – my fellow PhD-er in the trenches writing up a dissertation. Thank the cosmos our paths crossed and that we hit it off instantly! Thank you for always being there to commiserate with during this process and lending an ear. While COVID definitely ruined our plans of writing dates, knowing that

you're only a text away has been a huge comfort. Also, thanks for being part of the best CAPA family ever! Liz Stevens – I am sooo glad our time at UofM overlapped so much! Thank you for always being there to chat about research, bounce ideas off with, and especially for being there for me when things were going up in flames, being a calming presence, and helping come up with Plan B for data collection. Also, thanks for all the much-needed distractions; especially the road trips where we searched for those darn limestone caves that we never managed to find! Kaela Parker – thanks for being a great sounding board and for all the much-needed distractions, and of course the Mario party nights!

Naturally, I have saved the best for last. Tati, Don Oscar, and Marcelo – thank you for everything you have done. From always being pendiente to doing a day trip to Winnipeg to help with moving, I am beyond grateful for your constant support and servicios. Madrina – thank you for always checking up on me, for asking how you could help, and for just being a huge support system. To my family, lo hicimos!! I don't know what I would do without you all and your constant support, pushing me every step of the way, cheering me on, and listening to me babble on and on about stuff that sounded like gibberish. Ma y Pa, thank you for always being the biggest supports, encouraging me to follow my passion, and pushing me to be the best I can be. The sacrifice you both made 33 years ago made all of this possible! Y por todo lo que hacen, han hecho y todo sus sacrificios, un billión de gracias. Your unconditional love and support gave me strength whenever I needed it. This was done with you, for you, and because of you. Merc and Lu – I honestly don't know where I'd be without you two goof balls. Thank you for always being there, always lending an ear, and always being such huge supports that I could, and can always, lean on. From Cleveland to DC, y'all have been there with me, literally, and I wouldn't have it any other way. Silv – thanks for checking in and always making sure I could video chat with Ale and Mati during my time in Winnipeg. Seeing and hearing the boys always gave me that extra push to keep going. To anyone I may have unintentionally missed, thank you!

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1 - Introduction

The human skeleton contains a wealth of information about a person and their life. That information can be used for possible identification purposes or to understand important details about a person's interaction with the environment in which they lived. The most significant categories: sex, age, ancestry, and stature, are important characteristics that may be derived from skeletal remains since they constitute the physical attributes of a person during life. This is referred to as the biological profile (Stanojevich, 2012). Sex is one of the most important biological features to assess from skeletal remains as it determines the appropriate standards for age assessment and stature estimation in adult skeletons (Ousley, 1995; Horbaly et al. 2019). Sex is also one biological characteristic that can be determined with great accuracy from adult skeletal remains. The same, however, cannot be said for subadult skeletons, individuals who have yet to complete skeletal growth (Lewis, 2007). While sex differences in the subadult pelvis have been identified since the late 1800s, finding a reliable method for subadult sex determination remains, as Lewis (2017) states, the "holy grail" of juvenile osteology.

A number of morphological traits for subadult sex assessment have been proposed that have notably high accuracies. For example, Schutkowski's (1993) greater sciatic notch angle claims 95% accuracy, but when other researchers test these traits, comparably high accuracy rates are seldom replicated and can be as low as 54% (Vlak et al. 2008). While a degree of variability in accuracy rates for analytical methods (such as sex determination) is not uncommon in skeletal biology even for adult skeletons (Vance et al. 2011; Lesciotto and Doershuk, 2017; Horbaly et al. 2019), the variability in accuracy rates from subadult sex assessment methods tend to be quite disparate, and thus various methods have been deemed unreliable. Due to the lack of consistently high accuracy with subadult sex assessment methods, most bioarchaeologists and forensic anthropologists refrain from assessing this biological characteristic from the immature skeleton.

A number of factors possibly accounting for accuracy inconsistency for subadult sex assessment have been proposed. These include, but are not limited to: observer error; issues with adequate sampling; environmental stressors (including disease and

poor nutrition); and population differences (Scheuer, 2002; Cardoso and Saunders, 2008; Vlak et al. 2008; Wilson and Humphrey, 2017). Observer error when scoring subadult sex assessment traits is one factor that has been investigated (Scheuer, 2002; Cardoso and Saunders, 2008), but the effects of population differences and environmental stressors have yet to be explored in-depth.

Another issue that has often been emphasized, but as of yet has not been thoroughly addressed, is the fact that the methods developed for subadult sex assessment focus on regions of the skeleton that are most effective for adult sex assessment, most notably the pelvis and skull (Rissech and Malgosa, 2005; Rogers, 2009; Klales and Burnes, 2017). Since observable sex differences in these regions of the skeleton develop during puberty, specifically at the time of the adolescent growth spurt, it has been suggested that dimorphism in the pelvis and skull are minimal until after that time. Minimal research, however, has been conducted that assesses the applicability of adult sex traits to subadult individuals (under 20 years of age). Humphrey (1998) and Wilson and Humphrey (2017) have proposed that it is possible that sex differences appear before puberty but at different ages between the sexes, where one trait may be dimorphic after five years of age, for example, while other traits may be dimorphic after seven years of age. Thus, dimorphic traits of the pelvis may be problematic for subadult sex assessment because of the possibility of age-related variation in trait expression (Mittler and Sheridan, 1992; Cardoso and Saunders, 2008; Vlak et al. 2008; Wilson et al. 2015; Wilson and Humphrey, 2017). Potential age-related variation in pelvic trait expression that may exist is at present poorly understood because, with the exception of Sutherland and Suchey (1991) and more recently Klales and Burns (2017), there is a lack of research attempting to identify when dimorphic features on the pelvis begin to appear. It is important then, to understand the age at which morphological dimorphic traits in the pelvis appear and stabilize that may subsequently allow for accurate sex assessment in subadults.

Furthermore, the morphological traits used for adult sex assessment result from changes in size and shape of the pelvis. Due to the changes in size of the pelvis, different measurements from the pelvic bone have also been noted to be useful for adult sex assessment (Camacho et al. 1993; Arsuaga and Carretero, 1994; Albanese, 2003).

Similar to morphological changes to the pelvis, little is known as to when these metric differences emerge. Testing the applicability of metric methods of adult sex determination on subadults also has the potential to identify the minimum age at which measureable differences between the sexes provide a reliable sex determination. Through this sort of testing it will be possible to understand if, as morphological dimorphic traits appear and develop, corresponding measurable differences also appear.

1.1 Purpose and Significance

The purpose of this research is to examine age-related variation in the expression of sexual dimorphism in the human pelvis from an ontogenetic perspective. In order to be as comprehensive as possible, this study examines 18 morphological traits of the pelvis, which were derived from commonly used traits for adult sex assessment and traits proposed for subadult sex assessment. In addition to the morphological features, seven measurements required for Albanese's (2003) metric method of adult sex assessment were examined. The seven measurements were then applied to nine logistic regression equations generated in Albanese's (2003) study to achieve a sex determination. The morphological and metric sex assessments were conducted on individuals from the Hamann-Todd and Terry Collections, two late nineteenth to early twentieth century skeletal collections in the United States with documented sex and age-at-death, between the ages of 4 months and 20 years. By using an ontogenetic approach, this research traces when sexually dimorphic changes to the pelvis, both visual and measurable, begin to appear and subsequently stabilize. Tracking the development of sexual dimorphism in the pelvis is important for subadult sex assessment research, as it will highlight whether or not transitions in dimorphic features exist and provide a better understanding of age thresholds that restrict the applicability of sex assessment methods currently available.

Since this project applies a developmental perspective to understanding the appearance of sexual dimorphism in the pelvis, this research will also examine what relationship exists between the pubertal stage of an individual, determined using osteological features, and the appearance of sexual dimorphism in the pelvis. This

relationship may be key to further understand the limits the attainment of puberty has on subadult sex assessment, while providing tangible evidence of the importance puberty attainment has on reliable sex assessment. Furthermore, understanding a more refined pubertal stage and sexual dimorphism relationship may also provide a more nuanced understanding of what variables are required to determine at what age sex assessment can be confidently conducted. It is possible that the stage of development, such as pubertal stage, is more important for determining whether subadult sex assessment can be conducted, as opposed to chronological age.

This research is significant because it addresses the issue of subadult sex assessment from a developmental perspective. By examining age-related variation in the expression of sexual dimorphism in the pelvis, this research will contribute a more thorough understanding of the ontogeny of sexual dimorphism in the pelvis. A study tracking the appearance of a comprehensive list of sexually dimorphic pelvic traits has not been undertaken in forensic anthropology or bioarchaeology. Understanding when different dimorphic traits develop and how they are expressed at different ages may help identify avenues to develop improved techniques for subadult sex assessment. By examining a comprehensive list of traits across a large age range, this research has identified a sexually dimorphic trait that merits future investigation. This research is also the first to combine the applicability of determining the pubertal stage of an individual with the appearance of sexual dimorphism. Given that the attainment of puberty is an oft-cited issue for subadult sex determination, linking these two variables is a crucial step for understanding the limits associated with subadult sex assessment. Since the attainment of puberty is variable between and within sexes, understanding the relationship between the attainment of puberty and the development of sexual dimorphism may effectively encompass those individuals who fall within younger chronological age ranges and therefore provide a better threshold for when reliable sex assessment can be conducted.

In contemporary forensic contexts, the ability to accurately assess the sex of unknown skeletal remains is crucial as this biological characteristic has the potential to reduce the number of individuals on a missing persons list, since eliminating all persons of the opposite sex could reduce the missing persons list by half. In doing so, the

possibility of an investigator matching unknown skeletal remains to someone in a missing person's list may ultimately lead to a positive identification. Archaeologically, being able to confidently determine the sex of subadult remains will allow for better and more detailed reconstruction of how childhood may have been experienced in past populations. Knowing the sex of subadults in archaeological populations will provide the opportunity to examine potential gendered differences in childcare, children's status, disease susceptibility, individuality, and treatment in death and burial. Moreover, adolescent individuals have seldom been included or investigated in studies of sex assessment, therefore little is known as to which methods are appropriate and reliable for adolescent sex assessment, with few exceptions (ex. Rogers 2009). By incorporating adolescents among the subadults examined in this study, this research hopes to identify the best possible pelvic traits for adolescent sex assessment as well. This information can subsequently be used to understand nuanced sex differences in the processes of adolescent socialization in past societies.

1.2 Research Objectives and Hypotheses

Guided by the main research question, "does age-related variation exist in the expression of sexual dimorphism in the pelvis?", two main objectives emerge. **The first, and over-arching objective is to identify the age of appearance and stabilization of morphological and metric sex differences in the pelvis.** By identifying when sexual dimorphism (both visual and measurable) appears in the growing pelvis, it may be possible to determine a minimum age at which different morphological traits are most reliable for skeletal sex assessment of subadult remains. Moreover, it may be possible to determine if metric methods applied to adult skeletons can also be reliably applied to subadult skeletons. Through this objective, this research will identify when certain traits appear and when the traits become reliable indicators of sex. With this understanding, it may be possible to subsequently examine if any transitions exist between trait expression that may further aid in identifying the most effective traits for sex determination at particular ages or age ranges. By investigating when measurable differences between the sexes in the dimensions of the pelvis appear in relation to morphological changes in shape, this study seeks to provide an understanding of the

possible sequence of changes that occur in the pelvis as a result of sexual dimorphism throughout skeletal growth.

The second objective of this research is to explore the relationship that exists between sexually dimorphic traits in the pelvis with the attainment of puberty. By understanding this relationship, this study aims to determine what method of age categorization may act as a better threshold for identifying when subadult sex assessment is most applicable: chronological age or developmental stage (based on puberty).

In order to address the first objective, gross examination of 18 morphological pelvic traits was conducted on subadult individuals (identified as being 20 years of age and younger) from the Hamann-Todd and Terry collections. The 18 pelvic traits used were derived from 15 traits commonly used for adult sex assessment summarized by Rogers and Saunders (1994) and three traits proposed for subadult sex assessment by Schutkowski (1993). The nine logistic regression models generated by Albanese (2003) utilizing metrics from the pelvis and femur were also applied to subadult skeletons between the ages of 12-20 years. Since possible age-related variation in the expression of sexual dimorphism in the skeleton has been proposed by Humphrey (1998) and Wilson and Humphrey (2017), the first null hypothesis in this research states,

$H0^1$ = Age-related variation does not exist in the expression of sexual dimorphism

HA^1 = Age-related variation exists in the expression of sexual dimorphism

This first null hypothesis ($H0^1$) will be rejected if statistical tests of association demonstrate that a significant relationship ($p< 0.05$) exists between the appearance of different dimorphic pelvic traits at a particular age or age range. The first null hypothesis will be further rejected if tests of association indicate a significant relationship ($p<0.05$) between high accuracy of correct sex classification of the regression models at a particular age or age category. Falsification of this null hypothesis would then provide evidence that sexual dimorphism in the pelvis is not a uniform process across all traits, including metrics, and, therefore, sexual dimorphism in the growing skeleton should be examined cautiously.

To address the second objective, pubertal stage was determined for each individual between 8-20 years of age following the standards of Shapland and Lewis (2013). The pubertal stages were then examined in relation to correct sex allocation for the 18 morphological pelvic traits and nine logistic regression equations assessed. It is believed that most of the sexually dimorphic traits used for adult sex assessment can only be scored in adulthood, or after puberty, when they are fully expressed (Lewis, 2007; Rogers, 2009; Klales and Burns, 2017). The second null hypotheses tested, then, states,

$H0^2$ = Sexual dimorphic pelvic traits cannot be scored until post-puberty

HA^2 = Sexual dimorphic pelvic traits can be scored prior to post-puberty

Rejection of the second null hypothesis ($H0^2$) will occur if statistical tests of association demonstrate a significant relationship ($p<0.05$) between high correct sex classifications of pelvic traits and pubertal stages that precede the post-puberty stage. The second null hypothesis will be further rejected if tests of association also demonstrate a significant relationship ($p<0.05$) between high correct sex classification of the nine logistic regression equations and pubertal stages occurring before the post-puberty period.

1.3 Ethics

While the Hamann-Todd and Terry collections were both amassed specifically for research purposes, this research using human skeletal remains was undertaken with respect and consideration. The ethical guidelines provided by the American Anthropological Association for the treatment of human remains were followed, where permission to work with the two collections was granted by the respective collection's curator. To avoid any risk of causing damage to the skeletal remains, they were placed on pieces of foam at all times. Despite the availability of each individual's name, they were not recorded, as this information was not required for analysis. This research received formal approval [HS21920 (J2018:039)] from the Joint Faculty Research Ethics Board in the Office of Research Ethics and Compliance on May 24, 2018.

1.4 Dissertation Structure

Chapter 2 begins with a review of the relevant literature, which examines sexual dimorphism in humans and explores how the human pelvis grows and remodels in relation to sexual dimorphism. Chapter 2 also provides information on the characteristics associated with puberty and osteological indicators that have been identified as being useful for the assessment of pubertal stage from the skeleton. Chapter 3 begins with an historical contextualization of the United States in the late nineteenth and early twentieth centuries. This chapter then presents the materials and methods used in this research, including information on the skeletal collections used, how the sample was selected, the pelvic traits used, and the pelvic measurements taken. Chapter 4 presents the results and statistical analyses used to test the null hypotheses put forth by this research. Chapter 5 provides a discussion of the results outlining which pelvic traits are effective for subadult individuals, when dimorphism in the pelvis becomes apparent, proposing a minimum age at which sex can be assessed, and the relationship between pubertal stage and the appearance of sexual dimorphism in the pelvis. Finally, Chapter 6 presents the conclusion of this research and outlines future directions for subadult sex assessment. These future directions are put forth in order to provide avenues to refine our understanding of the development of sexual dimorphism to aid in the ultimate goal of finding a reliable method of subadult sex assessment.

Chapter 2 – Literature Review

2.1 Introduction

Sex determination of skeletal remains is an integral step for archaeological interpretation of past societies and positive identification of unknown human remains in contemporary criminal investigations. Determining the skeletal sex of subadult individuals, however, has remained the most problematic area of juvenile osteology (Lewis 2007). Children and childhood as a field of study in anthropology and archaeology started to gain substantial attention after the 1970s, with the introduction of gender or feminist theory to the discipline of anthropology (Lillehammer 1989; Lewis 2007; Baxter 2008; Halcrow and Tayles 2008; Thompson et al. 2014). It has been argued that prior to the introduction of feminist theory to anthropology, studies on children and childhood were lacking because children were seen as incidental to social life and lacked the agency needed to influence and create changes in society (Lewis 2007; Baxter 2005; Baxter 2008; Halcrow and Tayles 2008; Thompson et al. 2014). In physical anthropology, studies focusing on children emerged earlier than those in archaeology and social and cultural anthropology. Most early studies on children were descriptive in nature and attempted to develop standards that helped with establishing a biological profile from subadult remains, specifically developing age estimation and sex assessment techniques (Boucher 1955, 1957; Fazekas and Kosa 1978; Weaver 1980). There was a focus on studies exploring morphological methods of sex determination from the subadult skeleton in the 1990s (Hunt 1990; Schutkowski 1993; Mittler and Sheridan 1992; Holcomb and Koingsberg 1995; Molleson et al. 1998), but subsequent testing of those proposed methods in the following decade seldom achieved comparable high accuracies, rendering those methods unreliable. While morphological methods of sex assessment have proven to be problematic for subadult sex assessment, subadult sex assessment has had some promising advances in the past decade, with biochemical methods (such as ancient DNA and amelogenin analysis) providing promising results (Żądzińska et al. 2008; Stewart et al. 2017; Lewis 2019; Parker et al. 2019). These methods of sex assessment, however, are not without their own issues and include: expense of testing, consistency in results, false negatives and

positives, and further testing for refinement (Stewart et al. 2017; Lewis 2018). While these biochemical analyses show the most promise, morphological methods remain vital for quick, non-destructive, accessible but accurate determinations of sex. It has not been until the past four years that subadult sex determination has re-emerged in juvenile osteology in an attempt to develop reliable morphological methods of sex determination for subadults (Olivares and Aguilera 2016; Klales and Burns 2017; Luna et al. 2017; Stull et al. 2017; Wilson and Humphrey, 2017). One avenue of investigation that is required in order to develop improved morphological techniques for subadult sex assessment, is to explore age-related trends in the appearance of sexually dimorphic features to understand when they are first expressed and the age at which they display sufficient dimorphism for skeletal sex assessment, as is analyzed in this research.

2.2 Sex Assessment from the Skeleton

Sex assessment in the adult skeleton can be conducted with a number of skeletal elements using morphological and/or metric techniques. Morphological methods of skeletal sex assessment are based on gross observations of size and shape differences of particular traits whereas metric methods are based on measurements from bones that relate to size differences between the sexes (Rogers and Saunders, 1994; Albanese 2003). In terms of morphological methods of sex assessment, the most commonly used skeletal elements (in order from highest accuracy) include the pelvis, skull, and humerus. Sex assessment from the pelvis can provide results with accuracy of 96% and over (Rogers and Saunders 1994), an accuracy of approximately 90-96% can be achieved from the skull (Rogers 2005; Williams and Rogers 2006), and accuracy between 80-92% can be obtained using the humerus (Rogers 1999; Falys et al. 2005).

The pelvis is the most accurate skeletal element for adult sex assessment because of the reproductive and hormonal differences that exists between the sexes (Dunsworth 2020) and has been an area of the skeleton that has garnered substantial investigation for subadult sex assessment. Many of the early studies on the adult human pelvis examined the structure and size of the pelvic inlet since it directly relates to reproductive differences between males and female (Blake 2011). Features from the entire pelvis, however, were later incorporated to examine if sex differences can be observed on other features of the pelvis and not limit the examination of sex to the

pelvic inlet. Some pelvic features noted to express dimorphism include: subpubic angle, ventral arc, pubic bone length, medial aspect of the ischiopubic ramus, greater sciatic notch breadth, subpubic concavity, preauricular sulcus, auricular surface height, ilium shape and the ischio-pubic index (the ratio between pubic length and ischial length) (Washburn 1948; Genovés 1959; Krogman 1962; Phenice 1969; Anderson 1990; Budinoff and Tague 1990).

Of the traits outlined above, three-gained popularity due to the research conducted by Phenice (1969) who examined the presence of sexual dimorphism in the pubic bone. In his study, Phenice (1969) identified three traits (subpubic concavity, ventral arc, and medial aspect of the ischio-pubic ramus) that, when used in combination, could provide a sex assessment with 96% accuracy. Tests of the Phenice traits have proven that they are, indeed, highly accurate though not to the level that was originally achieved by Phenice (1969). Researchers have proposed that the differing levels of accuracy may be impacted by observer error (both intra- and interobserver) and by the population that is being used (Lovell 1989; Sutherland and Suchey 1991; Ubelaker and Volk 2002; Garvin 2012). Despite the impact observer error and inter-population variation may introduce, Klales (2020) has shown that the Phenice (1969) traits remain the most preferred method of sex assessment used by forensic anthropologist and bioarchaeologists. Rogers and Saunders (1994) conducted a comprehensive study on 17 pelvic traits for adult sex assessment and tested both the accuracy and precision of each individual trait and combinations of traits. The authors demonstrated that six traits independently produced accuracies above 80% and had intra-observer error below 10%, which included: the ventral arc, obturator foramen size and shape, true pelvis shape, sacrum shape, subpubic concavity, and pubis shape (Rogers and Saunders 1994). The other 11 traits they examined scored either high on precision but poor on accuracy, or vice versa, and therefore suggested caution in their reliability for sex determination.

Metric methods of adult sex determination from the pelvis have also garnered substantial attention. Washburn's (1948) pioneering study on sex differences in the pubic and ischial bones found that sex could be determined using the ischio-pubic index with an accuracy of over 90%. His study also showed that, when using the sciatic notch

and the ischio-pubic index in tandem, even higher accuracy could be achieved. Washburn's ischio-pubic index was found to be difficult to replicate between studies, which suggested that identifying the proper landmarks from which to take the measurements was difficult (Albanese, 2003; İşcan and Steyn, 2013). Albanese (2003) proposed a new measurement (the superior pubis ramus length) to circumvent the difficulty of identifying the necessary landmarks used for Washburn's (1948) ischio-pubic index. Additionally, Albanese (2003) developed 17 logistic regression equations based on six measurements from the pelvis and femur (superior pubis ramus length, acetabular ischium length, hipbone height, iliac breadth, maximum diameter of femur head, epicondylar breadth of femur). Intra-observer error for the superior pubis ramus length was 0.57%, compared to 2.7% error found with traditional pubis length measurements (Albanese 2003). Moreover, when using the superior pubis ramus length with other measurements of the pelvis and femur, the accuracy of the 17 logistic regression models ranged between 90 and 98.5%. In addition to the high accuracy and reproducibility of the superior pubis ramus length measurement, the logistic regression equations produced by Albanese (2003) are not population specific and are therefore widely applicable. Sex differences in the measurements of the pelvis and femur examined by Albanese (2003) exist because of the relationship between measurements and biomechanics, where metrics take into account reproduction and bipedal adaptations in females but only bipedalism in males (Silva Braz, 2009). For example, the superior pubis ramus length is expected to be larger in females than males because reproductive requirements for a broader pelvic inlet result in a wider pubic bone in females (Anderson 1990). Research is lacking that attempts to understand when these measureable differences in the pelvis between the sexes occur and, therefore, the earliest age at which metric methods of sex determination can be performed is unknown. It is possible that the measurements used in Albanese's (2003) method have the potential to be applied to older subadults, such as adolescents, once fusion of the pelvic elements has occurred (usually around 11-15 years old in females and 14-17 years in males).

While traits for adult sex assessment have been tested at length, very little attention has been given to examining when during growth and development these

dimorphic traits in the pelvis begin to appear. Research of this nature has recently been conducted by Stock (2018) who examined the age of full expression of sexually dimorphic cranial traits. Her study concluded that three traits displayed full expression before full dental maturity, however, this research was conducted by applying a metric value to quantify the size of morphological cranial traits from radiographs, which may not compare well with dry bone observations (DiGangi and Moore 2013, p.107). In a similar vein of applying adult sex determination techniques to subadults, Rogers (2009) tested the applicability of four morphological traits on the distal humerus that are accurate (80-92% accuracy) for adult sex determination on adolescent skeletons (aged 11-20 years). Her study showed that the distal humerus traits used for adult sex assessment can be applied and achieve an accuracy of 81% once the trochlea of the humerus has fused, which occurs at approximately 11 years of age (Rogers 2009), suggesting age-related sex differences exist in some areas of the skeleton prior to adulthood. In terms of examining the appearance of sexually dimorphic pelvic traits, Sutherland and Suchey (1991) observed that the ventral arc is not recognizable until the mid-20s, but that a precursor form of the ventral arc (precursor arc) can be seen as early as 14 years of age and is the most frequent expression at 20 years of age. The precursor arc is defined as a faint line found on the ventral surface of the pubis that occurs as the lower extremity fills in with fine dense bone (Sutherland and Suchey 1991). More recently, Klales and Burns (2017) tested the applicability of using a modified method of scoring Phenice traits to subadult individuals. Their study showed that the modified method could be applied successfully to early adolescents (12.6-15.5 years old) and late adolescents (15.6-20.5 years of age) with an accuracy of 85 and 97%, respectively (Klales and Burns 2017). While demonstrating promising results, Klales and Burns' (2017) observations utilized radiographs, which can be problematic since radiographic observations do not necessarily compare well with dry bone (DiGangi and Moore 2013, p.107). Apart from the studies on Phenice traits by Sutherland and Suchey (1991) and Klales and Burns (2017), there is yet to be a comprehensive study examining the age of appearance of multiple sexually dimorphic traits from the entire pelvis. Understanding how sexual dimorphism emerges in humans

and the growth of the human pelvis are crucial to explore age-related variation in sexual dimorphism.

2.3 Sexual Dimorphism and Pelvic Growth

Sex can be assessed from the human skeleton because sexual dimorphism exists in humans. Sexual dimorphism, or any physical differences between sexes, is associated with sexual selection that emerged via male competition and mate choice in non-human primates (Plavcan 2012). Vestiges of sexual dimorphism can be found in modern humans with differences in body weight and stature, which result from differing rates and duration of growth between males and females that diverge during adolescence (Humphrey 1998). For example, body size is a sexually dimorphic characteristic that can be seen overtly as adult males are, on average, approximately 9% taller and 15% heavier than adult females (Stulp and Barrett 2016; Horbaly et al. 2019). Males, then, tend to be built more robust than females, resulting in thicker bones with larger muscle markings. Morphology in the skeleton also has the potential to reflect adaptations associated with differing reproductive functions, where child bearing in females requires broader hips and a wider pelvic inlet (Nettle 2002; Plavcan 2012). While sex differences in the skeleton exist, differences in the rate, onset, and duration of skeletal growth between the sexes affects the degree of variation in sexual dimorphism that is observed in adults (Humphrey 1998; Horbaly et al. 2019).

Biologists generally differentiate between two types or periods of sexual dimorphism; primary sexual dimorphism and secondary sexual dimorphism. Primary sexual dimorphism is directly associated with successful reproduction and includes: number, size and mobility of gametes, structure of reproductive organs, and obstetrically related differences in the pelvis (Kirchengast 2014). For example, female gametes, the ova, are much less numerous but larger than the male sperm since the ova contain nutrients that are essential for the development of an embryo. Primary sex differentiation occurs during the prenatal phase of life and is a process that is genetically programmed (Kirchengast 2014). Sex differentiation during this intrauterine time is also dependent on sex hormones, where male differentiation depends on the secretion of anti-müllerian hormone, testosterone, and dihydrotestosterone (Carlson 2013; Kirchengast 2014). In contrast, the full development of the reproductive organs in

females depends on estrogen secreted by the fetal ovaries. The release of hormones responsible for sex differentiation is controlled via the production of gonadotropin-releasing-hormone (GnRH) by the hypothalamus during fetal life, which continues into early infancy. Gonadotropin-releasing-hormone, in turn, initiates the release of luteinizing hormone (LH) and follicle stimulating hormone (FSH) from the pituitary gland (Bogin 1999). LH and FSH travel to the ovaries and testes via the blood stream and stimulate the production of estrogen and androgens (Bogin 1999). The hypothalamus is inhibited by late infancy, which results in the cessation of GnRH secretion and a reduction in sex hormones leading to the suspension of reproductive maturity. The hypothalamus is not reactivated until adolescence where GnRH production starts again.

Secondary sexual dimorphism relates to features indirectly related to reproductive success such as: voice change, body shape, body composition, stature, and body and facial hair (Sinclair 1989; Bogin 1999; Shapland and Lewis 2013). These features are not necessarily essential to successful reproduction in a direct manner, but may enhance reproductive success. Secondary sexual dimorphism develops during the post-natal phase, particularly during puberty and adolescence, and is driven by increased secretion of sex hormones associated with this life-stage (Bogin 1999; Kirchengast 2014). Puberty, also referred to as gonadarche, is the re-initiation of activity in the hypothalamic-pituitary-gonadal system of hormone production (Bogin 1999). After being inactive from sexual development since about two years of age, the hypothalamus is stimulated again to produce GnRH. The transition into adolescence is marked by this renewed production of GnRH, and the secretion of this hormone must occur in pulses (Bogin 1999). Gonadarche, then, is triggered when the pulsatile secretion reaches both the necessary amplitude and frequency (Bogin 1999 p. 83). Similar to the hormonal surge that occurs during fetal life and early infancy, the ovaries in females are stimulated to produce high levels of estrogen, while the testes in males are signalled to produce high levels of testosterone. The effect of gonadarche is easily seen because of the visible signs of sexual maturity, which include: an increase in density of pubic hair, development of breasts in girls, maturation of external genitalia, deepening of voice in boys, and sex differences in body size and composition (Sinclair 1989; Bogin 1999; Shapland and Lewis 2013; Kirchengast 2014; Henderson and Padez

2017). Sexual dimorphism in the pelvis, then, is associated with secondary sexual dimorphism.

The pelvis is the skeletal element that shows the greatest level of sexual dimorphism and is considered the best indicator for sex assessment in adults (Scheuer and Black 2000; Walker 2005; Garvin et al. 2014; Klales and Burns 2017). The high level of sexual dimorphism in the pelvis has largely been attributed to the difference in parturition between females and males, but high levels of estrogen in females may also contribute to the sexual dimorphism observed in humans (Dunsworth 2020). The human pelvic bones are large, irregularly shaped bones that articulate with their counterpart anteriorly through a secondary cartilaginous joint at the pubic symphysis and with the sacrum posteriorly. The adult pelvic bone is formed from the fusion of three main separate bones, the pubis (inferio-anterior), ischium (inferio-posterior), and ilium (superior) (Scheuer and Black 2000). The three pelvic bones develop from a cartilaginous precursor and all three bones are well developed and identifiable at birth. While there is little change in the morphology of the bones in the first few years after birth, there is rapid growth of each bone in the first three months which then slows until 2-3 years of age, with a further reduction until puberty where growth increases again (Miles and Bulman 1995; Scheuer and Black 2000). Fusion of the ischium to the pubis can occur as early at 3 years of age, but generally occurs between 5-8 years (Caffey 1993; Scheuer and Black 2000). The ilium then proceeds to fuse to the already fused ischium and pubis at the acetabulum at about 15 years of age. The fusion patterns of the pelvis and its elements are important to consider when examining age-related variation in sexual dimorphism since they will impact the earliest age at which certain sexually dimorphic features, such as the acetabulum size and orientation, subpubic concavity, ischiopubic ramus ridge, and obturator foramen shape, can be confidently assessed from dry bone.

The ilium is initially flat but becomes curved due to hormonal and functional influences (Blake 2011). By approximately nine years of age, the ilium's width has reached 70% of its adult size (Humphrey 1998) and could, therefore, show substantial sexual dimorphism in size and shape around this age. Sexual dimorphism, however, does not occur equally for pelvic bones since growth occurs at different rates in multiple

areas for males and females (Coleman 1969; Tague 1994; Bogin 1999). For example, the pubis shows different growth between males and females, where females exhibit prolonged growth along the pubic symphysis. The medial and ventral portions of the pubic symphysis grow into early adulthood in females, which results in a longer (medially projecting) pubic length (Anderson 1990; Budinoff and Tague 1990; Sutherland and Suchey 1991; Tague 1994). As a result, adult females have a broader, rectangular shaped pubic bone. Growth of the pubic bone in males, on the other hand, is complete at approximately 18 years of age (Tague 1989; Budinoff and Taugue 1990) and results in a narrower shaped pubis.

During the hormonal surge that occurs during puberty, defining characteristics in the pelvis are distinguishable between males and females. Moreover, the pubertal growth spurt is a time where pelvic remodeling occurs in anticipation for reproduction as growth increases rapidly to prepare the body for its adult form (Bogin 1999). Dunsworth (2020), however, has suggested that pelvic remodelling, particularly of the pelvic inlet, might also occur in response to an increase in volume during adolescence that female internal organs (reproductive and otherwise) occupy, which greatly contrasts that of males. It has also been proposed that differentiation in female pelvic growth during puberty is restricted to the regions directly involved in the development of the pelvic inlet. These alterations to the pelvis have been shown to occur rapidly with hormone induction over a period as short as 18 months in clinical studies (Scheuer and Black, 2000). Several studies have, in fact, shown that females tend to have larger pelvic cavity and birth canal dimensions than males among 18 year olds and that the size of the true pelvis in females expands more than that of males between 8 and 18 years of age (Wood and Chamberlain 1986; Arsuaga and Carretero 1994; LaVelle 1995). Scheuer and Black (2000), however, have suggested that secondary sex differences of the pelvis should be sufficiently advanced from around mid-puberty. Further expansion of the true pelvis in females does continue until about 25-30 years of age (Huseynov et al. 2016), likely resulting in stronger dimorphism of this pelvic feature in adults.

The ventral arc, located on the anterior surface of the pubic bones, is one of the most common features used to differentiate between adult males and females. Differences in the expression of this pelvic trait are believed to be the result of variation

in muscle attachment sites for the adductor mangus and adductor brevis muscles, and the gracilis muscle (Budinoff and Tague 1990; Anderson 1990). In females, these muscles are more laterally and cranially positioned compared to males. This relationship between muscle position on the pubic bone begins around puberty, when hormonal activity results in the lengthening of the female pubis (Anderson 1990; Klales et al. 2012). By contrast, in males the attachment of these muscles is more medial and parallels the pubic symphysis and inferior ramus, potentially resulting in a dense symphyseal margin. Phenice (1969) suggested that the ventral arc does not appear in females until 20-23 years of age, but Sutherland and Suchey (1991) have noted a variation to this trait (a precursor arc).

Developmental differences occurring between males and females during puberty have also been proposed to influence the dimorphic expression of another one of Phenice's (1969) traits: the subpubic angle. These differences include different directional growth occurring at the middle of the ischiopubic ramus between males and females and increased growth in regions of the pubis and ischium in females that results in an obtuse subpubic angle in females and an acute angle in males (Coleman 1969; LaVelle 1995; Klales et al. 2012). While Phenice (1969) originally proposed that this trait was not well developed in females until 20 years of age, Klales and Burns (2017) have provided promising results that indicate this trait might be applicable to individuals as young as 12.6 years of age. Little research has been conducted to understand the dimorphism displayed by Phenice's third trait; the ischio-pubic ramus ridge, but it has been suggested that females tend to have narrower ischio-pubic rami compared to males due to medial-lateral lengthening of the pubic bone that occurs during adolescent growth in females (Coleman 1969; Klales and Burns 2017). While the majority of adult pelvic traits have been suggested to only be useful after 20 years of age, there has yet to be a comprehensive study examining whether substantial dimorphism in these traits can be observed in individuals younger than 20.

Fehling (1876) was one of the first to suggest the potential presence of sexually dimorphic traits in the subadult pelvis (Schutkowski 1993; Lewis 2007). Since his early study the search for reliable sex determination methods of subadult skeletal remains has resulted in the development of a number of metric, morphological, and

morphometric methods throughout the skeleton (Boucher 1955, 1957; Hunt and Gleiser 1955; Sundick 1977; Fazekas and Kosa 1978; Weaver 1980; De Vito and Saunders 1990; Schutkowski 1993; Holcomb and Konigsberg 1995; Molleson et al. 1998; Wilson et al. 2008, 2011; Rogers 2009; Veroni et al. 2010; Klales and Burns 2017; Luna et al. 2017; Stull et al. 2017). While the studies that first developed and proposed sex determination methods reported successful results, subsequent testing of those methods proved otherwise (Hunt 1990; Mittler and Sheridan 1992; Loth and Henneberg 2001; Scheuer 2002; Sutter 2003; Cardoso and Saunders 2008; Vlak et al. 2008; Olivares and Aguilera 2016; Lamer et al. 2021). One proposed issue with subadult sex assessment is that the methods currently available primarily focus on skeletal elements that demonstrate a high degree of dimorphism in adult skeletons, namely the pelvis and skull. However, since the observable sex differences in these areas of the skeleton develop at the time of the adolescent growth spurt (Rissech and Malgosa 2005; Rogers 2009; Wilson et al. 2015; Lewis 2018) sexual dimorphism in the skeleton is believed to be minimal until after the growth spurt. Research has conversely suggested that sex differences in hormonal levels do exist in-utero and after birth that are sufficient enough to result in some dimorphic differences in body composition and morphology, which extends to the skeletal system (Boucher 1957; Knickmeyer and Baron-Cohen 2006; Kirchengast 2014; Stull et al. 2017; Dunsworth 2020). It has also been suggested that age-related variation in trait expression may exist in the growing skeleton (Humphrey 1998; Wilson and Humphrey, 2017). Thus, it is believed that differences in the skeleton between sexes may appear at different ages (Humphrey 1998), but age related-variation in trait expression is not well understood since the ontogeny of sexual dimorphism has not been examined extensively (Wilson and Humphrey 2017). Wilson and colleagues (2015) attempted to examine age-related variation and hypothesized that differences in female and male pelves might arise during growth in three ways. The first is through ontogenetic scaling, which supposes that males and females have the same shape at the onset of development and follow the same growth trajectory, where the relationship between size and shape are maintained, but would differ in the length of growth. Second, parallel trajectory suggests that males and females have different shapes at the onset and subsequently follow the same ontogeny of shape but never

resemble one another. Third, divergent trajectory supposes that shape changes would be confined to later stages of ontogeny with no differences in shape seen between the sexes at the youngest stage (Wilson et al. 2015). When examining growth trajectories of three pelvic traits (greater sciatic notch, auricular surface, and iliac crest), Wilson and colleagues (2015) discovered evidence of divergent trajectories contributing to the dimorphism in those features. More extensive research has not been conducted on age-related variation in trait expression but is required to examine if most, or all, pelvic traits follow divergent ontogenetic trajectories.

Due to the high level of accuracy for sex assessment in the adult pelvis, the ilium of subadults has garnered much attention. Early methods of subadult sex determination using the pelvis relied heavily on measurements of the pelvis (Reynolds 1945, 1947; Boucher, 1957; Coleman, 1969). The accuracy rates for these metric methods varied considerably, an observation that was attributed to a difficulty in correctly identifying the landmarks required for measurements. Weaver (1980), however, was the first to propose a qualitative trait of the ilium, elevation of the auricular surface, for subadult sex assessment. Based on a study of 153 subadults, Weaver (1980) achieved an accuracy ranging from 73.1% and 91.7% for males and 75% and under for females (Weaver 1980). Males in the fetal category and those six-months of age had the highest accuracy rates, whereas accuracy for females consistently decreased from fetal to 6 months of age. Mittler and Sheridan (1992) determined that this trait fared well for males (85.3%) but not females (58.3%). Schutkowski (1993) subsequently proposed a list of pelvic features for subadult sex assessment. These traits included: greater sciatic notch angle, greater sciatic notch depth, arch criterion, and curvature of the iliac crest. His study, using the Christ Church Spitalfield Collection, achieved a 95% accuracy rate for males using the greater sciatic notch angle, 81.2% accuracy for the greater sciatic notch depth in males, 73.3% accuracy for arch criterion in males, and 81.2% accuracy for iliac crest shape in males (Schutkowski 1993). These traits, however, did not prove to be as successful in females (71.4%, 76.5%, 70.6%, and 62.1% respectively). Subsequent testing of these traits by Sutter (2003), Cardoso and Saunders (2008), and Vlak and colleagues (2008) have provided varying accuracy rates seldom achieving such high accuracy. These traits, however, still merit examination in the context of skeletal growth

to provide a thorough understanding of whether or not there is a specific age or age range that these traits are most effective. When exploring sexual dimorphism in the human pelvis, it is clear that adolescence, in particular puberty, is believed to be integral for the development of sex differences. Since puberty is a process, and not a singular event, it remains to be understood when, during pubertal growth, sexual dimorphism becomes sufficiently pronounced for skeletal sex assessment. This period of post-natal life, therefore, requires consideration and attention.

2.4 Adolescence and Puberty

Adolescence is a life stage where social and sexual maturation occurs. This life stage begins with puberty, a biological process that marks a person's sexual maturity and was historically a marker of when a child would gain adult status, particularly when date-of-birth (for chronological age) was not rigorously recorded (Rogers 2009; Shapland and Lewis 2013; Arthur et al. 2013). While adolescence is not a life-stage unique to humans, the length of time between puberty and adulthood (5-8 years) and the rapid acceleration in growth velocity of essentially all skeletal tissue, known as the adolescent or pubertal growth spurt, are unique (Bogin 1999). The physical changes that occur during puberty follow a specific sequence, where the pubertal growth spurt begins with the acceleration of growth velocity, which continues to accelerate until peak height velocity (PHV) is reached. Peak height velocity is followed by a deceleration period, where growth returns to a slower rate, and ending in full maturation that is marked by epiphyseal fusion of long bones (Hägg and Taranger 1982; Rogol et al. 2000; Shapland and Lewis 2013; Arthur et al. 2016). Girls tend to enter and complete each stage of puberty earlier than boys, and reach PHV one to two years before boys (Bogin 1999; Rogol et al. 2000). Identifying PHV is of particular interest in bioarchaeology since it coincides with the obvious physical changes associated with puberty, such as breast development, and precedes menarche in girls and voice breaking in boys by about one year, that would mark the transition into adulthood (Shapland and Lewis 2013; Lewis, 2018). In females, timing of menarche is particularly significant because of its association with the timing of marriage in past societies and biological processes such as adult risk of osteoporosis (Bogin 1999; Karapanou and Papadimitriou 2010; Arthur et al. 2016). The timing and tempo of puberty, however, vary

widely between individuals and populations (even among healthy children), which is a consequence of genetic and environmental factors (Bogin 1999; Rogol et al. 2000; Arthur et al. 2016). Modern standards have put the onset of the pubertal growth spurt at about 10 years of age for females and approximately 12 years of age for males, with PHV attained at 12 and 14 years of age for females and males, respectively (Aksglaede et al. 2008). The different stages of pubertal growth may provide some indication as to when the onset of morphological differentiation of the pelvis between the sexes occurs and if it occurs, as Scheuer and Black (2000) suggest, by mid-puberty.

Given the social importance of the adolescent period, bioarchaeology has developed a new focus to identify the different stages of the growth spurt during puberty in order to link the biological process of puberty to social age and transitions in past societies (Shapland and Lewis 2013, 2014; Arthur et al. 2016; Lewis et al. 2016 a,b; Doe et al. 2017; Henderson and Padez 2017). The signs that are characteristic of puberty, however, are primarily related to changes in the soft tissue and, therefore, are not visible in skeletal remains. Fortunately, clinical studies have shown that skeletal maturation of certain elements are closely related to sexual maturation since fusion of the epiphyses is triggered by the release of estrogen in females, which gets converted to androgen in males, during later adolescence (Hassell and Farman 1995; Cutler 1997; Grumbach 2000; Dunsworth 2020). Due to the close relationship between skeletal maturation and sexual maturation, research has been conducted exploring the use of certain skeletal elements for pubertal assessment in past societies, indicating it is possible to identify puberty from the skeleton (Shapland and Lewis 2013, 2014). When indicators of pubertal stage are combined with skeletal indicators of sex, it may be possible to explore the relationship between skeletal expression of sexual dimorphism and pubertal development, thereby identifying a critical period during the pubertal growth spurt where dimorphism becomes sufficiently pronounced for skeletal sex assessment.

2.4.1 Skeletal Indicators of Pubertal Stage

Studying the attainment of puberty in the past has become a significant area of inquiry in bioarchaeology because of the contribution it may have to inform our current understandings of the decreasing secular trend in the timing of puberty, which in and of

itself, has health concerns for modern populations (Arthur et al. 2016). While overt signs of puberty are not visible in skeletal remains, clinical studies have identified a number of developmental markers in the skeleton that are strongly correlated with the different stages of the pubertal growth spurt (Hewitt and Acheson 1961; Grave and Brown 1976; Chertkow 1980; Hägg and Taranger 1982; Flores-Mir et al. 2006; Soegiharto et al. 2008). These skeletal markers include: ossification of the hamate hook, fusion of the phalangeal epiphyses, iliac crest ossification, fusion of the distal radius epiphysis, mineralization of the mandibular canine, and cervical vertebrae maturation (Shapland and Lewis, 2013; Shapland and Lewis, 2014). Being able to link stages of puberty (through these osteological indicators) with substantial dimorphism in the pelvis can provide a means of determining whether or not full expression of dimorphic traits in the pelvis only occurs once puberty has ended (Rissech and Malgosa 2005; Silva Braz 2009; Rogers 2009; Wilson et al. 2015; Klales and Burns 2017; Luna et al. 2017; Lewis 2018).

Ossification of the hand and wrist bones is a method of skeletal age assessment but is one that is more commonly used in medical or clinical practice than in biological anthropology (Shapland and Lewis 2013). Two common methods of age assessment from the wrist include the atlas method (ex. Greulich and Pyle 1959), which uses radiographic standards that are representative of healthy children at different chronological ages for visual comparison, or by composite scoring (ex. Tanner et al. 2001), which is a process where each individual bone, observed radiographically, is rated using a biologically weighted scoring system (Chertkow 1980). The application of these age estimation methods are not easily applicable to archaeological contexts since it is difficult to differentiate between the lack of ossification from post-mortem loss, particularly when full recovery of the epiphyses of the wrist and hands has been overlooked. Research has shown, however, that ossification of the hamate hook and the epiphyseal fusion of the distal phalanges have been linked to different stages of the pubertal growth spurt and so are useful indicators of pubertal stage in bioarchaeological contexts.

From a clinical system devised by Tanner and colleagues (2001), stage G for hamate hook development (no hook developing) is indicative of an individual being pre-

pubescent. Ossification of the hamate hook (Stages H and H.5) occurs during the acceleration phase of the growth spurt in both girls and boys, and is completely developed (Stage I) at about 6 months before PHV (Grave and Brown 1976; Chertkow 1980; Lewis et al. 2016). While additional clinical studies have not produced research linking the biological mechanisms of hamate hook development to pubertal stage, this indicator has still shown promise in identifying pubertal stage in modern populations (Shapland and Lewis 2013). In modern pediatrics, the appearance and fusion of the phalangeal epiphyses tends to be used to examine pubertal stage (Houston 1980; Özer et al. 2006). Grave and Brown (1976) suggest that when the width of the phalangeal epiphysis is equal to the diaphysis, the individual is in the acceleration phase of the growth spurt. Moreover, the authors further showed that epiphyseal capping (when both length and width of the epiphysis is equal to the diaphysis) occurs at about the time of PHV (Grave and Brown 1976). In archaeological contexts, however, these two stages of epiphyseal development may be difficult to ascertain, if at all, especially if the epiphyses of the phalanges are not recovered. Fusion of the phalangeal epiphyses, however, can be assessed in archaeological contexts and has been shown to occur in the deceleration stage, shortly after PHV has been achieved and menarche in girls (Grave and Brown 1976; Hägg and Taranger 1982; Houston 1980). Epiphyseal fusion of the phalanges has been shown to progress from the distal phalanges to the proximal (Grave and Brown 1976), thus fusion of the distal phalanges is more closely associated with PHV. Furthermore, menarche has been shown to be closely associated with epiphyseal fusion of the second distal hand phalanx (Lewis et al. 2016). Consequently, the hamate hook development is used in this study as it helps delineate the beginning and end of the acceleration phase, while fusion of the distal phalangeal epiphyses is used in to identify the period between PHV and the deceleration stage.

The distal radius is commonly included in the assessment of skeletal age based on hand and wrist development. As with hamate hook development and the fusion of the phalangeal epiphyses, a study by Hägg and Taranger (1982) on 212 Swedish children was the first to demonstrate that the fusion of the radius' distal epiphysis is correlated with the pubertal growth spurt. The beginning of fusion of the radial distal epiphysis is associated with the deceleration stage of the growth spurt and has shown

to occur approximately one year before the end of the growth spurt (Hägg and Taranger 1982). Hägg and Taranger (1982) also showed that full fusion of the radial distal epiphysis occurred only at the end of the pubertal growth spurt. Fusion of the distal epiphysis of the radius begins at about 14.8 years in girls and 16.5 years in males and is completely fused about two years later in both males and females (Hägg and Taranger 1982). A more recently study by Zhang and colleagues (2008) on 14, 757 Han urban children has confirmed the observations made by Hägg and Taranger (1982). Fusion of the distal radius is, therefore, used in this study to provide an end limit to the growth spurt.

The ossification of the iliac crest, known as the Risser sign in clinical studies, is a commonly used indicator of puberty in orthopedics, since the relationship between the two has been noted since the 1940s (Shapland and Lewis 2013). Ossification of the iliac crest has also been used as a strong signal of menarche since clinical studies have shown that, on average, there is a 6-month gap between the beginning of iliac crest ossification and the start of menarche, and these two events rarely occur more than a year apart from one another (Shapland and Lewis 2013). As with the epiphyses of the hand and wrist, recovery of an ossified but unfused iliac crest epiphyses may rarely occur, which complicates scoring when absent (not ossified versus lost post-mortem). If, however, an ossified iliac crest epiphysis is recovered or has begun fusing to the ilium, this would be indicative that PHV has been achieved and, in females, menarche has been achieved (Shapland and Lewis 2013; Lewis et al. 2016). This indicator is therefore used in the present study to identify the outside limits of pubertal growth (deceleration and post-puberty).

Dental development is a process more commonly associated with chronological age, although the variability in the mineralization and eruption of the canine has proven to be least accurate for age estimation, next to the third molars (Liversidge et al. 2010), as opposed to hormonal changes that result from puberty (Flores-Mir et al. 2006). Research has shown that hormones may influence the variability in canine growth. In fact, in 1959 Meredith observed that the eruption of the mandibular canine had a closer association to the adolescent growth spurt compared to the other teeth. While this trait is not used in the present study, it may be used to provide closer identification of the

onset of pubertal growth and the end of the acceleration phase. The final osteological indicator that has been used in bioarchaeological studies for assessing pubertal stage is the maturation state of the cervical vertebrae. Using cervical vertebrae maturation (CVM) as an alternative method of age assessment to hand-wrist maturation was first suggested by Lamparski (1972), who devised a six stage standard to describe changes in morphology of the vertebrae throughout adolescence (Shapland and Lewis 2014). Hassel and Farmer (1995) subsequently adopted the six-stage system proposed by Lamparski and calculated the average position each CVM stage was along the pubertal growth spurt. While CVM is not applied to the present study, Hassel and Farmer (1995) showed that the different CVM stages has the potential to identify all the stages of pubertal growth from the acceleration of growth to the end of puberty.

Chapter 3 – Materials and Methods

3.1 Late 19th Century & Early 20th Century United States

This research was conducted on subadult individuals from two documented skeletal collections in the United States: the Hamann-Todd and Terry Collections. These two collections are composed of individuals that lived in the U.S. Mid-west during the late nineteenth to early twentieth century. In order to understand the environment at the time these individuals were alive, it is important to understand the historical (social and economic) context of the Mid-West during the late 1800s and early 1900s. Two major events in U.S. history occurred during this time period that altered the social and/or economic aspects of society; the Great Northern Migration and the Great Depression. These events likely had an impact on the well-being of the children and adolescents in the collections and contributed in some manner to the overall demographic composition of the skeletal collections.

The abolition of slavery, which occurred in 1865, provided newly freed African Americans in the Southern U.S. with many new opportunities, such as the freedom to relocate to areas with improved living circumstances, resulting in the Great Northern Migration (Tolnay and Beck 1990, 1992). The Great Northern Migration occurred between 1910 and 1970 and involved the internal migration of over six million African Americans from the rural Southern States to the United States' urban North, Midwest, and West (Tolnay and Beck 1992; Phillips 1999; Tolnay 2003; Massey et al. 2009). This migration of African Americans was driven by the lack of political and economic opportunities available in the South, but more importantly, the need to escape Jim Crow laws (laws of racial segregation) and racial violence (Marks 1989; Tolnay and Beck 1990; Tolnay 2003; Muller et al. 2017). On an economic level, African Americans in the rural South remained a part of the plantation economy post-Emancipation, and continued to have little prospect of benefitting economically. This was exacerbated by the fact that African Americans were caught between competing class interests of the white community, where white employers benefitted from cheap labour by Black labourers, but poor whites competed with Black labour for jobs. Such a scenario often generated violent conflict between poor whites and African Americans (Tolnay and Beck

1990). While there was considerable migration of African Americans between Southern states and from Southern rural to urban centres, the Black community still faced a dismal economic situation. Opportunities for African Americans improved as major urban manufacturing centers emerged in cities like New York, Cleveland, and St. Louis, that opened a market for unskilled labourers during the industrial boom following the First World War (Tolnay and Beck 1990; Phillips 1999; Tolnay 2003; Griffin 2005; Massey et al. 2009; Wilkerson 2010). In turn, there were increased job opportunities and wages for African Americans that were not previously available. These improved economic options initiated two large migration waves, the first between 1910 and1940, and the second between 1940 and 1960.

The freedom to relocate afforded to the African American community post-Emancipation became an asset, particularly from 1877 onward, when state and local laws were introduced that enforced racial segregation in the Southern U.S. known as the Jim Crow laws (Tolnay and Beck 1990). The Jim Crow laws were put in place to enforce the perceived inferior caste position of African Americans. These laws also allowed for restrictive voting statues, which curtailed the Black vote, and the allocation of vastly unequal financial support for Black and white schools by state legislatures (Kousser 1980). In addition to laws that were put in place to ensure African Americans stayed an inferior caste, racial violence in the form of lynching, mob violence, and legal executions was all too common (Tolnay and Beck 1990; Phillips 1999). Thus, social abuses and violence, in addition to economic hopelessness, all provided the impetus for out-migration from the rural south to the urban North and Mid-West. However, the large influx of labourers from the Southern U.S. to the North resulted in overcrowded and unsanitary conditions in urban neighbourhoods (Kusmer 1978; Phillips 1999; Muller et al. 2017). Moreover, the majority of African Americans moving to the North and Mid-West lacked sufficient social and economic support systems, and in instances of illness or poor nutrition, were forced to find public assistance or hospital care, which were primary sources of unclaimed bodies between the 1830s and 1968 (Muller et al. 2017). During this time, unclaimed bodies, a high proportion of which were African American, were subsequently sent to medical schools instead of being buried at the expense of taxpayers, which reinforced the structural violence these individuals experienced during

life (Muller et al. 2017; de la Cova 2019). Subsequently, there are a high number of African American individuals in the Hamann-Todd and Terry collections, many whom were of low socio-economic status.

The second major event that occurred in the early twentieth century that changed the economic landscape in the U.S was the Great Depression. While the 1920s is generally considered to be a decade of economic prosperity and industrial growth, the end of the decade was anything but prosperous. The Great Depression started with the U.S. stock market crash on “Black Tuesday” (September 4, 1929) and lasted until 1939 (Stuckler et al. 2011). The early years of the Great Depression, from 1929 to 1933, were when the Great Depression was at its worst; unemployment increased sevenfold while industrial output (the quantity of goods produced in a given period) fell approximately 30 percent and more than 3,000 banks collapsed (Cole and Ohanian 2007; Stuckler et al. 2011). Despite a relatively quick economic recovery that began in 1934, both employment and industrial output remained well below the pre-1929 levels (Cole and Ohanian 2007). Thus, people living through the Great Depression lived through income reductions and deprivations, which resulted in social unrest (Granados and Roux 2009). Generally, little attention has been given to the effects the Great Depression had on population health, with the exception of studies by Granados and Roux (2009) and Stuckler and colleagues (2011). While one would think that an economic crisis would result in a mortality crisis, the opposite has been shown (albeit with some exceptions). A drop in overall mortality was shown during the first four years of the Great Depression, where mortality was 1,273.4/100,000 in 1929 and went down to 1,148/100,000 in 1933 (Stuckler et al. 2011). Infectious diseases such as pneumonia, influenza, and respiratory tuberculosis declined between 1929 and 1932, but chronic diseases (i.e. heart disease, cancer, diabetes) rose (Granados and Roux 2009; Stuckler et al. 2011). Stuckler and colleagues (2011) suggest that, while the exact reasons for the decrease in mortality and infectious disease cannot be ascertained, possible reasons for the decrease include: nation-wide expansion of financial relief for the unemployed; prohibition on alcohol that was enacted in 1919; sanitation and hygiene improvements; and advances in medical therapy. While health may not have been greatly impacted in a negative way during the Great Depression, the financial difficulties it created likely made

it difficult for families to afford cemetery burials for deceased family members. The financial distress experienced during the Great Depression, then, likely provided an increased opportunity to procure cadavers whose skeletons would ultimately be incorporated into the Hamann-Todd and Terry collections. Thus, the combination of large events, such as the Great Northern Migration and Great Depression, likely influenced the demographic profile of those individuals whose bodies ended up being unclaimed at morgues and hospitals in Mid-Western urban centres, such as St. Louis and Cleveland, that were subsequently incorporated in skeletal collections.

3.2 Materials

3.2.1 The Osteological Collections

The Terry Collection was created by then Demonstrator of Anatomy, Robert J. Terry at the Missouri Medical College (later known as Washington University Medical School). Influenced by mentors George S. Huntington and Sir William Turner, Terry was aware that there was an absence of documented osteological and anatomical collections that could be used for research (Hunt and Albanese 2005; Muller et al. 2017). Terry's initial collection of human skeletal remains began in 1898 and expanded in 1900, followed by a second collection that was amassed shortly after the first collection was destroyed in a fire. The current Terry Collection represents a third attempt at establishing a skeletal collection (Hunt and Albanese 2005). While the first collection was destroyed in a fire, the second collection was rendered useless due to a disruption, mainly the result of commingling of skeletons during Terry's year of absence at Harvard University. In the 1920s, Terry devoted more time and resources toward amassing human skeletons from cadavers initially used in anatomy classes than his previous attempts. The cadavers were primarily obtained from local hospitals in St. Louis, as well as coroner facilities, although a small portion of cadavers was obtained from institutions, such as poor houses and clinics, throughout Missouri (Muller et al. 2017). The cadavers obtained consisted mainly of individuals who were unclaimed by their relatives and subsequently became property of the state to be buried at the taxpayers' expense (Hunt and Albanese 2005). In order to avoid the costs of burial, it was common practice prior to World War II for unclaimed cadavers to be used in medical school anatomy classes for most jurisdictions in the United States and Canada

(Hunt and Albanese 2005; Muller et al. 2017). Terry accepted all unclaimed bodies that were available and insisted on representing the complete range of human skeletal variation. He was diligent with including all forms of “normal” individuals and did not focus only on individuals with pathological conditions, as was often the case when procuring cadavers for medical school anatomy collections, and this is reflected in the current collection.

Upon Terry’s retirement in 1941, Mildred Trotter took over and continued collecting skeletons until her retirement in 1967. Trotter contributed a concerted effort to balance the collection’s demographic composition and focused her collecting efforts on white females and younger individuals (Trotter 1981; Hunt and Albanese, 2005; Muller et al. 2017). Prior to Trotter’s focus on females, social and economic factors affected the number of females available in the early twentieth century, which resulted in a low number of female cadavers during Terry’s collection process (Hunt and Albanese 2005). Under Trotter’s direction, the collection’s name changed from the Washington University Collection to the Terry Collection. Decreasing interest in continuing support for the Terry Collection by Washington University, in addition to Trotter’s anticipated retirement, led Trotter to contact T. Dale Stewart of the Smithsonian Institution with regards to transferring the Terry Collection for permanent curation at the Smithsonian Institution, which occurred in 1967 (Hunt and Albanese 2005; Muller et al. 2017).

The current iteration of the Terry collection is at the Smithsonian Institution’s National Museum of Natural History (NMNH) in Washington, D.C. The collection includes individuals that were born between 1828 and 1943 and who died between 1917 and 1965 (Hunt and Albanese, 2005; Muller et al. 2017). The collection is comprised of approximately 1,728 individuals who were collected between 1910 and 1967 in St. Louis, Missouri. The age-at-death ranges from 14 to 102 years and has a male to female ratio of 1.4:1 (Hunt and Albanese, 2005). Table 3.1 outlines the number of individuals in the Terry Collection with known age and sex as per data collected by Muller and colleagues (2017). Sixteen individuals are under 20 years of age and an additional eight individuals are 20 years of age, which marks the age limit set out for this research, for a total of 24 subadults.

Table 3.1: Number of individuals in the Terry Collection by age and sex

Age Cohort	Black Males	White Males	Black females	White Females	Total
<20	6	2	7	1	16
20-29	69	8	49	3	129
30-39	99	27	57	12	195
40-49	109	68	59	22	258
50-59	96	94	62	54	306
60-69	79	131	49	82	341
70-79	37	89	39	77	242
80-89	11	17	25	46	99
90-99	2	0	6	6	14
100+	0	0	2	0	2
Total	508	436	355	303	1602

The Hamann-Todd skeletal collection was initiated by Anatomy professor Carl A. Hamann at Western Reserve University (WRU), now Case Western Reserve University, in 1893. Hamann's training at the University of Pennsylvania Medical School instilled the importance of anatomical comparative collections and he sought to build one upon being hired at WRU (Muller et al. 2017). After 19 years at WRU, Hamann had collected, processed and catalogued approximately 100 human skeletons that were mainly unclaimed bodies (Quigley 2001).

In 1912, Hamann stopped processing cadaver skeletons because of his appointment to Professor of applied anatomy and clinical surgery (Muller et al. 2017). T. Wingate Todd replaced Hamann as the chair of Anatomy and expanded Hamann's collection by acquiring over 3,000 human skeletons during his 26 years at WRU. Expansion of the collection was made easier for Todd than Hamann, due to revisions to the Anatomical Laws of the State of Ohio requiring superintendents of regional mortuaries, charity institutions, mental institutions, and city hospitals to notify Todd of unclaimed bodies in their possession (Quigley 2001; Muller et al. 2017). Many of the individuals acquired by Todd had documents with personal information and files containing death certificate data such as age, name, sex, race, birthplace, occupation, and cause of death. Todd acknowledged the important role the environment and culture played in growth, development, and overall health (Muller et al. 2017), a stark contrast from his contemporaries, which resulted in the presence of a substantial number of

infants, children, and adolescents in the collection. The collection process ended in 1938 upon Todd's death.

The Hamann-Todd collection is curated at the Cleveland Museum of Natural History (CMNH) in Cleveland, Ohio. The collection consists of approximately 3,100 African American and Caucasian individuals who were collected between 1893 and 1938 in Cleveland, Ohio. The collection includes individuals who were born between 1825 and 1910 and died between 1912 and 1938 (Lovejoy et al. 1985; Muller et al. 2017). All individuals in this collection are unclaimed bodies, many transient individuals, from the Cleveland and Cuyahoga County morgues as well as city and charity hospitals that were sent to the WRU Medical School in Cleveland for use in teaching (Lovejoy et al. 1985; Quigley 2001). The age at death ranges from birth to 105 years and the collection has a substantially higher proportion of males (N=2,122) compared to females (n= 377). Lovejoy and colleagues (1985) have noted that many of the ascribed ages from morgue records are grossly incompatible with morphological assessment and they noted spikes in 5-year-increment ages, particularly for older age categories, suggesting that some family members of the deceased, or the deceased person themselves, reported erroneous ages upon admission to hospitals. The collection contains a total of 106 individuals between 0 and 20 years of age with both the cranium and post-cranial skeletal elements available, providing a large sample of subadults for the present research.

3.2.2 The Study Sample

The term "subadult" is generally used as a broad categorization for individuals who are considered non-adults, or have yet to complete skeletal growth (Lewis 2007). The umbrella term of subadults, then, encompasses general age categories commonly used in bioarchaeology such as infants (0-3 years of age), children (3-12 years of age), and adolescents (12-20 years of age) (Buikstra and Ubelaker, 1994). For the purpose of this research, the term "subadult" is used in its broadest sense to refer to all individuals 20 years of age and younger. A total of 134 subadult individuals were available for study, 106 from the Hamann-Todd collection (HTH) and 28 from the Terry Collection. Access was granted to an additional four subadult individuals who are part of the Terry collection, which accounts for the larger number of subadults available for this study

compared to the published data available. The study sample used in this research consists of a combined sample of 128 subadult individuals, 102 from the Hamann-Todd Collection and 26 from the Terry Collection. The presence of skeletal elements required for sex assessment was prioritized and formed the basis of sample selection. The individuals in the study sample are between the ages of 4 months and 20 years, of which 51 are female and 77 are male (Figures 3.1 and 3.2) and have a year of death ranging between 1911 and 1948 (Figure 3.3). Only three individuals have a year of death after 1948. Four-peak year of deaths are seen from the entire study sample at 1918, 1924, 1929, and 1931.

Using the digital databases for the Hamann-Todd and Terry collections, all individuals between 0 and 20 years of age were selected and the column with information regarding sex was removed to ensure there was no knowledge of this biological characteristic prior to all skeletal analysis. The catalogue number of all individuals falling within the stipulated age range was collected and was the only piece of information known at the time of analysis.

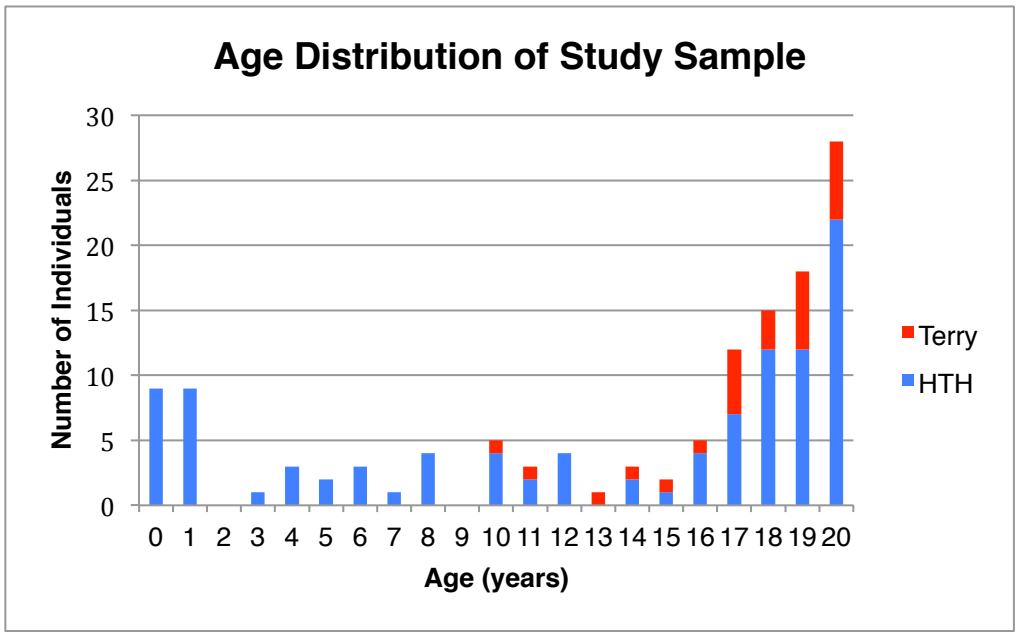

Figure 3.1: Age distribution of the study same separated by collection

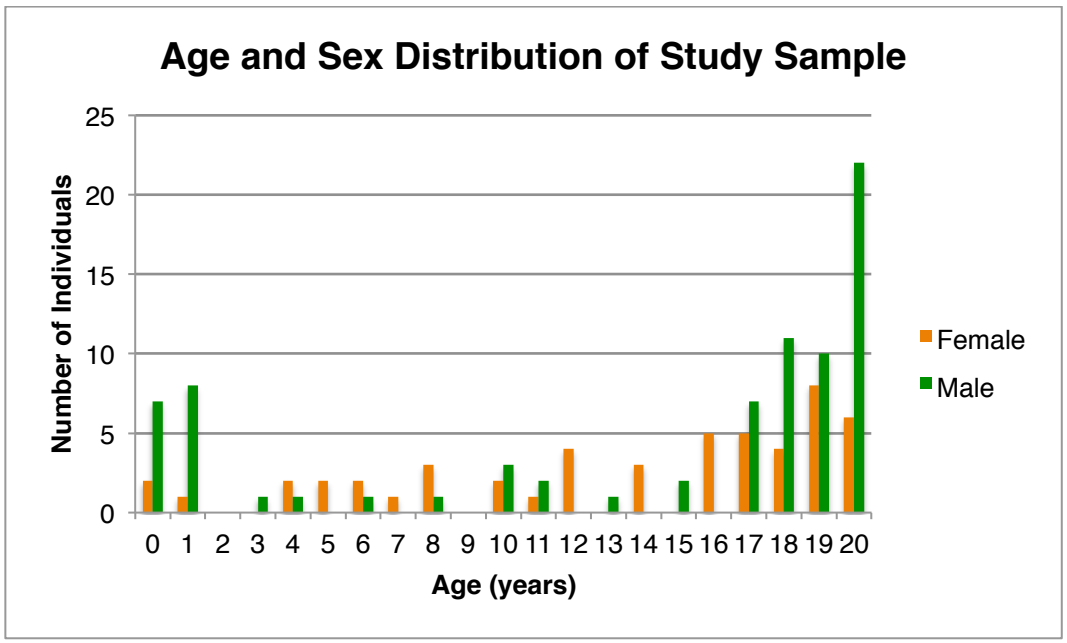

Figure 3.2: Distribution of the study sample based on chronological age and documented sex

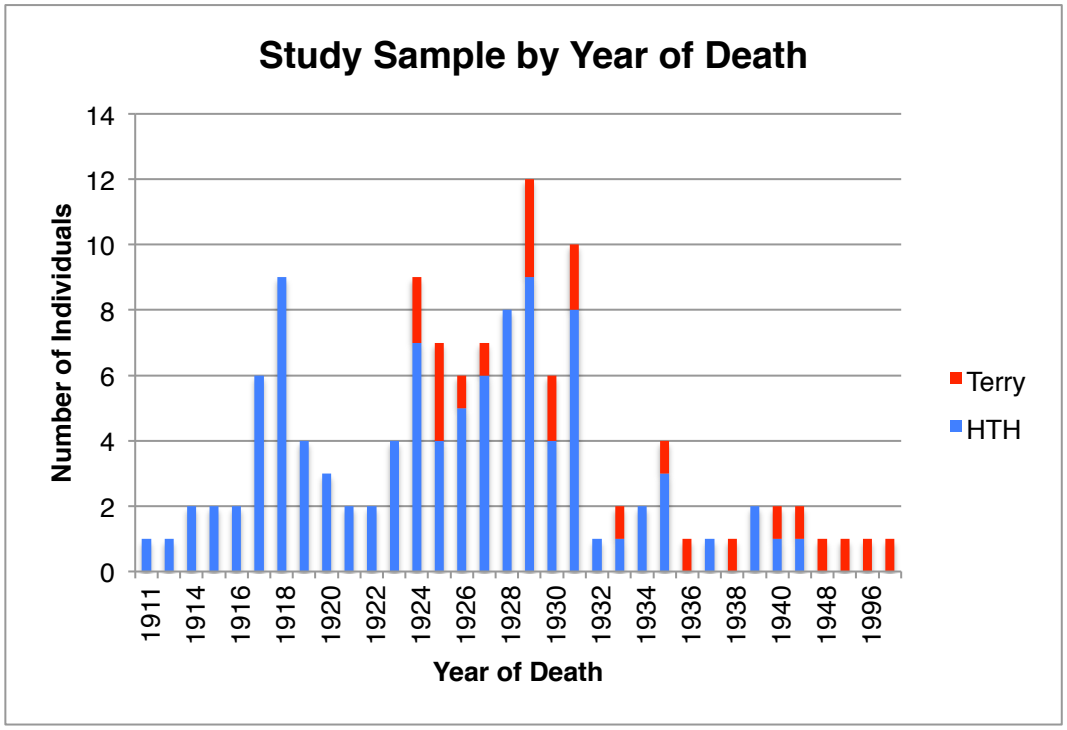

Figure 3.3: Year of death distribution of subadults in the study sample

The study sample was selected based on the two aspects examined in this research: 1) pelvic sex assessment and 2) pubertal stage assessment. Since the

examination of sexual dimorphism was the primary variable assessed in this research, an individual had to have, at minimum, a complete pelvis, or its unfused elements, to be included in this study and contributed to the total sample size (n=128). Ideally, a single individual was represented in each part of analysis in order to provide a robust sample to compare how sex assessment relates to pubertal stage.

In order to be included in the first part of this research regarding sex assessment, the subadult individuals' os coxa, or its unfused elements, was required to have little to no damage. Given the extensive list of morphological traits that were assessed, there was some flexibility to the degree and location of damage when present. If, however, the damage was to an extent that the majority of the morphological traits were unable to be scored, the individual was not selected for any part of this study. The flexibility in the level of preservation seen on the os coxa was also used when selecting individuals for metric sex assessment. The metric method applied in this study, Albanese (2003), was created in such a way that preservation issues often encountered in bioarchaeology and forensic anthropology was accounted for. Thus, in order to be included for metric sex assessment, the minimal requirement was the presence of an os coxa and a femur such that the preservation of each element allowed for a combination of measurements that would fit any one of the nine logistic regression models generated by Albanese (2003) tested in this study. Based on the criteria outlined above, 82 individuals met the requirements for metric sex assessment and were included in this part of the study.

While adolescence is a developmental period often defined as between the ages of 12 and 20 years (Buikstra and Ubelaker 1994; Bogin 1999; Lewis 2007), this study included individuals as young as eight years of age for pubertal stage analysis to encompass the possibility of inter-individual and population variation in the timing of puberty (Bogin 1999). Moreover, since it has been noted that puberty can be influenced by a number of factors including genetics, social status, nutrition, and physiological stress (Bogin 1999: Louis et al. 2008; Arthur et al. 2016; Lewis et al. 2016), a lower than usual age minimum was included. Shapland and Lewis (2013; 2014) outlined six osteological markers that have been shown clinically to be closely associated with specific stages of puberty and include: mandibular canine development, hamate hook development, epiphyseal fusion of the phalanges, iliac crest ossification, fusion of the

distal radius, and cervical vertebrae maturation. Dental radiographs for the adolescents in the Hamann-Todd and Terry collections were not available and, therefore, canine development was not assessed. Additionally, due to time constraints cervical vertebrae maturation was also not assessed. Consequently, four of the six osteological markers of puberty were assessed for each adolescent, aged 8-20 years for this part of the research, in the sample and include: hamate hook development, epiphyseal fusion of the phalanges, iliac crest ossification, and fusion of the distal radius. In order to be included in the analysis of pubertal stage, individuals were required to have a minimum of three of these osteological markers present with little to no damage, as per Lewis and colleagues (2016). While it was ideal if both the left and right elements were present for analysis, individuals were still included in this part of the analysis if only unilateral assessment was possible. Based on the criteria outlined above, 96 individuals were included in this part of analysis (38 females and 58 males). Given that each part of this research includes a distinct number of individuals, Table 3.2 outlines the total number of individuals and the number of females and males included in each aspect of analysis.

Table 3.2: Breakdown of the number of individuals in each part of analysis in this research

Part of Analysis	Total Number of Individuals	Number of Females	Number of Males
Morphological sex assessment	128	51	77
Metric Sex Assessment	82	34	48
Puberty Assessment	96	38	58

3.3 Methodology

3.3.1 Sex Assessment Methods

In order to ensure this research was conducted blind, meaning that the main variable in this study (sex) was unknown during data collection; two procedures for retrieving the remains were followed. The individuals at the Hamann-Todd collection had basic information (catalogue number, sex, age, and “race”) recorded on the side of the tray containing the remains. Opaque paper was placed over this information during the retrieval process. For a large sub-set of individuals at the Hamann-Todd collection,

a research assistant was available and ensured any information on the plastic trays regarding the sex of an individual was covered using opaque paper. Some younger individuals had signs indicating male or female drawn on the bone, which was covered with opaque paper by the research assistant. When retrieving individuals myself or by the research assistant, three individuals were selected at once for analysis. Selecting three individuals at once was done to randomize which individual was being assessed in the event that the recorded sex of an individual was seen. Each individual in the Terry collection contained basic information (age, sex, catalogue number) on a separate piece of paper and was flipped over upon retrieval of the skeleton. Curation of the individuals at the Terry collection was distinct from Hamann-Todd collection in that the individuals were not in plastic trays but instead were on metal shelves. Only the skeletal elements needed for analysis were retrieved. As with the Hamann-Todd collection, three individuals were chosen at the same time during analysis of the Terry collection to further ensure sex assessment was conducted blind.

Sex assessment was conducted for the total sample of 128 individuals (n= 102 HTH; n= 26 Terry) using the 18 pelvic sex assessment traits listed in Table 3.3. Fifteen of the traits correspond to pelvic traits used in adult sex assessment that have been tested by Rogers and Saunders (1994) for accuracy and precision. The remaining three traits have been proposed for subadult sex assessment by Schutkowski (1993) that are not traits used in adult sex assessment (ex. greater sciatic notch depth and auricular surface height/elevation). Ideally all 18 traits were scored for each individual, but when damage made clear visualization of a trait difficult, the trait was not assessed. Every effort was made to assess each trait independent of one another so as to not influence the assessment of subsequent traits. Though, due to the proximity of some traits to each other, it was not possible to completely separate each trait. The scoring of one trait in particular, the ventral arc, was modified for the purpose of this study in order to encompass a wider form of expression. While Phenice (1969) defines the presence of this trait as “the presence of a slightly elevated ridge of bone extending from the pubic crest and arcs inferiorly across the ventral surface to the lateral most extension of the subpubic concavity” (pg. 298), Sutherland and Suchey (1991) noted the presence of a precursor arc in predominantly adolescent females. The precursor arc is defined as a

faint line found on the ventral bone taking the same course as the ventral arc defined by Phenice and occurs as “the lower extremity fills in with fine dense bone before the symphyseal rims becomes fully defined” (Sutherland and Suchey 1991, p. 504). As a result, the ventral arc trait was scored as “present” if either a full ventral arc as defined by Phenice (1969) or the precursor arc defined by Sutherland and Suchey (1991) were expressed and it was noted which of the two forms of expression was observed. A second round of sex assessment on 96 individuals was conducted six days after the end of the first round of sex assessment in order to examine intra-observer error. A final sex assessment was made for each individual based on “majority rule” of trait expression, where sex was determined from the category (male or female) in which most features were assigned. For younger individuals, 16 years old and younger, a second final assessment was made where more weight was placed on traits proposed solely for subadult sex assessment.

Table 3.3: List of 18 morphological traits used for pelvic sex assessment

Trait	Male Expression	Female Expression
Subpubic concavity angle	V-shaped	U-shaped
Ischiopubic ramus ridge	Ridge absent	Ridge present
Ventral arc presence	Arc absent	Arc present
Shape of pubic bone	Narrow	Broad & rectangular
Doral pubic pitting	Absent	Present
Sciatic notch shape and size	Small, close, deep	Wide, shallow
Auricular surface height	Not raised	Raised
Preauricular sulcus presence and shape	Absent or thin grooves	Large, circular depression
Ilium shape	High and vertical	Laterally divergent
Pelvic inlet shape	Heart shaped	Elliptical
True pelvis size and shape	Small, narrow	Shallow and spacious
Obturator foramen shape	Large and ovoid	Small and triangular
Acetabulum size and orientation	Large, directed laterally	Small, directed antero-laterally
Development of muscle markings	Marked and rugged	Gracile and smooth
Sacrum Shape	Long, Narrow	Short, Broad
Greater Sciatic Notch Angle	The angle of the sciatic notch is approximately 90°	The angle of the sciatic notch is greater than 90°
Arch Criterion	When drawing a cranial extension from the vertical side of the greater sciatic notch, the arch leads into the lateral rim of the auricular surface	When drawing a cranial extension from the vertical side of the greater sciatic notch, the arch formed crosses the auricular surface
Iliac Crest Curvature	When viewed, the iliac crest exhibits a marked S-shape	When viewed, the iliac crest exhibits a faint S-shape

Once the morphological sex assessment was completed for each individual, measurements were taken of each os coxa and femur following the six measurements outlined by Albanese (2003) (see Table 3.4 for descriptions). Digital calipers were used to measure the superior pubis ramus length, acetabular ischium length, maximum diameter of the femoral head, and epicondylar breadth of the femur to the nearest millimeter. An osteometric board was used to measure hipbone height and iliac breadth

to the nearest millimeter. The pelvic measurements were taken after morphological sex assessment was conducted in order to avoid any potential bias or observable patterns in metrics that may have influenced morphological assessments of sex. The six measurements were then applied to the logistic regression model:

$$P = \frac{1}{(1+e^{-Z})}$$

where P is the probability of an event, for this equation the event is “male or not male”, and Z is a linear combination of independent variables (Albanese 2003) or:

$$Z = \beta_o + \beta_1 X_1 + \beta_2 X_2 \ldots$$

β_o represents the constant, β_x the measurement coefficient, and X is the value of each measurement. The probabilities generated by the logistic regression models are between 0 and 1, where values greater than 0.5 are considered male while values less than 0.5 are female (Albanese 2003). Table 3.5 presents the coefficients of each model.

Table 3.4: List and description of measurements taken for pelvic sex assessment per Albanese (2003)

Measurement	Description
Superior Pubis Ramus Length (SPRL)	Superior margin of the pubic symphysis to the superior- anterior apex of the lunate surface in the acetabulum. Using sliding calipers, nearest mm.
Acetabular Ischium Length (AIL)	Most inferior point on the ischium to the superior-anterior apex of the lunate surface in the acetabulum. Ensuring this measurement is not perpendicular to SPRL. Measured to the nearest mm using sliding calipers.
Hipbone Height	Distance from the most inferior point of the ischial tuberosity to the most superior point on the iliac crest (osteometric board, nearest mm).
Iliac Breadth	Distance from the anterior superior iliac spine to the posterior superior iliac spine (osteometric board, nearest mm).
Maximum diameter of femur head	Maximum diameter of the head of the femur at the border of the articular surface (sliding caliper, nearest mm).
Epicondylar breadth of femur	Distance between the most projecting points on the medial and lateral condyles (sliding calipers, nearest mm).

Table 3.5: Coefficients for Albanese (2003) logistic regression models

Model	Hip bone height	Iliac Breadth	SPRL	AIL	Max. Di of Femur Head	Epicondylar Breadth	Constant
1	0.5950	-0.5192	-1.1104		1.1696	0.5893	-61.5345
2		-0.1600	-0.5951	0.2920	1.0365	0.3901	-30.5291
3	0.2572		-0.9852		0.7303	0.3177	-40.5313
4	0.4323	-.02217	-0.7404	0.3412			-30.3590
5	0.3084		-.8092	0.2657			-28.3111
8	0.4868	-0.4903	-1.0597	0.2901	1.6241		-45.2528
10	0.5267	-0.3785	-0.8156			0.7758	-52.8262
11	0.2896		-0.8794			0.5783	-42.6362
20	0.2007	-0.4445		0.1734	0.5697	0.3915	-41.9071

3.3.2 Pubertal Stage Assessment

In order to examine the relationship between the appearance of sexually dimorphic traits in the pelvis and the attainment of puberty, individuals in the study sample between 8-20 years of age were assessed for their pubertal stage using four osteological markers; hamate hook development, epiphyseal fusion of the distal phalanges, iliac crest ossification, fusion of the distal radius (Shapland and Lewis 2013), after sex assessment was completed. In order to capture the most accurate representation of pubertal stage, at least three out of the four osteological markers were required to be present and sufficiently preserved for scoring. The fusion of the distal phalangeal epiphyses, iliac crest, and distal radius epiphysis was scored using a three-stage system of “unfused”, “partially fused”, and “fused,” while the development of the hamate hook was scored using a four stage scoring system of G, H, H.5, and I all outlined by Shapland and Lewis (2013). The left and right elements were assessed to account for possible bilateral asymmetry, but if individuals only had a left or right element present, fusion and developmental states were still scored. In the event that asymmetry in scores was observed, the more advanced stage was used as per Shapland and Lewis (2013). A final pubertal stage assessment was made for each individual based on the scores of all osteological indicators using the criteria outlined in Table 3.6.

Table 3.6: Skeletal features associated with pubertal stage. Modified from Lewis et al. (2016) and Arthur et al. (2016)

Stage	Hamate Hook	Distal Phalangeal epi fusion	Distal radius epiphysis	Iliac crest ossification
Pre-puberty	Stage G: hook absent	Distal hand phalanges unfused	unfused	Epiphysis not present
Acceleration	Stages H or H.5: hook appearing or developing	Distal hand phalanges unfused	unfused	Epiphysis 50% complete, unfused
PHV or transition	Stage I: Hook complete	Distal hand phalanges unfused	unfused	Epiphysis 50-75% complete, unfused
Deceleration	Stage I: Hook complete	Distal hand phalanges fusing	unfused	Epiphysis 75-100% complete, non to partial fusion
Maturation	Stage I: Hook complete	Distal hand phalanges fusing/fused	partially fused	Epiphysis 100% complete, partial fusion
Post-puberty	Stage I: Hook complete	Distal hand phalanges fused	fused	Epiphysis fused

3.4 Statistical Analysis

Intra-observer Error

Intra-observer error was conducted on a total of 96 individuals from the Hamann-Todd study sample. All 18 morphological pelvic traits were scored for each of the 96 individuals six days after the initial assessment. Intra-observer error is a way to assess the precision of a method or the precision by which a trait was scored, which in turn measures an observer's ability to reproduce their results. The ability to reproduce results reflects both the researcher's capabilities but also reflects the nature of the features examined (Rogers 1999). Morphological traits can take on a wide range of expression; thus, intra-observer error stresses the importance of getting familiarized with the range of variation (Wood 2012). The usefulness of the method to score a trait, then, can be observed based on the consistency at which it can be scored by a researcher. Intra-observer error was calculated by examining the percentage of disagreement in score for all 18 morphological traits of the pelvis between the two

rounds of assessment. It is generally accepted that 10% is the highest acceptable level of error in order to be considered reliable (Nichol and Turner 1986; Rogers and Saunders 1994).

Sex Classification Accuracy

The accuracy of each pelvic trait was determined by comparing the blind assessments of the trait expression (as either male or female) to the documented sex. Accuracy of each trait was initially calculated for subadults, as a broad category, to first test the utility and applicability of each of the 18 pelvic traits for sex determination of individuals who are 20 years and younger. A final assessment of sex using all 18 traits was also determined two ways. First, each trait was weighed equally and a 'majority rule' approach was applied, where sex was determined based on the category (male or female) in which most features were assigned. Second, a final sex determination was also conducted by applying more weight to traits specifically identified for subadult sex assessment in individuals 16 years and younger. The accuracy of each final decision regarding sex was determined by comparing the final sex estimate of each individual to their documented sex.

Each pelvic trait was then assessed based on the chronological age of each individual in an attempt to identify the earliest chronological age at which male and female expressions of each pelvic trait appears in relationship to known sex. Examining trait expression by chronological age was also done in order to determine the age at which each trait can be applied with high accuracy for males and females. In order to obtain a more robust sample sizes for age analysis, the subadults were grouped into four "traditional" age categories used in bioarchaeology, which include: infant (0-2.9 years), child (3-11.9 years), adolescents (12-16.9 years), and late adolescents (17-20 years). The infant and child categories follow Bogin's (1999) post-natal life stages, while the adolescent and late adolescent categories are modified versions of Baker and colleagues' (2005) proposed age categories that take into account developmental changes during adolescence, such as brain development and hormonal changes. The accuracies of each trait and final sex estimates were calculated for each age cohort to determine if particular age ranges show a high degree of correct sex allocation. Chi-

square test of independence was conducted for the age cohorts with high sex allocation accuracy using SPSS 20.0.

In order to determine the relationship that exists between pelvic trait expression and the attainment of puberty, individuals aged 3-20 years were separated into six pubertal stages: pre-puberty, acceleration, peak height velocity (PHV), deceleration, maturation, and post-puberty. To increase the sample size of the pre-puberty stage, individuals 3-7 years of age were included, and automatically placed in this stage, as they represent the age range at which there is no hormonal release from the hypothalamus (Bogin 1999). Classification of pubertal stage for individuals 8-20 years of age was based on development scores of skeletal pubertal markers outlined by Shapland and Lewis (2013) and Table 3.6. Accuracies of each pelvic trait and final sex assessments were then calculated for each pubertal stage category.

Pelvic measurements from 82 individuals were entered into nine out of 17 logistic regression models generated by Albanese (2003) (Table 3.5). Models 1, 2, 3, 4, 5, 8, 10, 11, and 20 were chosen for metric sex assessment because they have the highest correct sex allocation accuracies (95.5-98.5%). Similar to the morphological traits, accuracy of sex allocation for each linear regression model was calculated for subadults, the same four age categories used for morphological traits, and for pubertal stages.

Chi-square Test of Independence

In order to ensure that the accuracies seen in this study are not the result of chance, statistical testing (chi-square test) was applied to examine the significance of correct sex classification accuracies. Chi-square test of independence uses categorical data and frequencies to test whether a relationship exists between two or more variables (Shennan 1997). Each individual in a sample is divided into mutually exclusive categories, in this research male or female trait expression, and proportions for each category are calculated (Drennan 2010). Expected frequencies are calculated based on an ideal hypothetical sample distribution obtained if sample proportions were in agreement with those specified in the null-hypothesis (Gravetter and Wallnau 2004). Chi-square test, then, is a measure of how much deviation exists between observed values from expected values (Moore et al. 2009). A large discrepancy between the

observed data and the expected values would therefore lead to a rejection of the null hypothesis. One principal concern with the chi-square test, however, is that a robust sample is required for it to reliably approximate the real probabilities (Drennan 2010). A commonly applied middle-ground standard is that no expected values can be less than 1 and no more than 20% of expected values can be less than 5 (Drennan 2010, p.192). If two-by-two tables have low expected values rendering the results of a chi-square unreliable, Fisher's exact test can be calculated and applied regardless of how low expected values are (Drennan 2010). Fisher's exact test directly calculates the significance probability and does not require a minimum size for expected values, as they are not required for calculation (Drennan 2010).

Chapter 4 - Results

Overview

This chapter presents the results of this research beginning with the overall accuracy achieved for each of the 18 morphological traits and nine logistic regression equations for all subadults (≤ 20 years). Intra-observer error is then presented for the 18 morphological traits and each morphological trait is subsequently ranked based on accuracy and precision. The ages of appearance and stabilization of all 18 morphological traits are presented for both males and females. Then, the efficacy of morphological traits and logistic regression models are presented for pre-determined age categories. Finally, the relationships between the efficacy of morphological traits and logistic regression models with pubertal stages are presented.

4.1 Accuracy of Morphological Traits and Logistic Regression Models

Overall, six morphological pelvic traits meet or exceed 80% accuracy for sex assessment in subadults, nine pelvic traits have an overall accuracy between 70 and 79%, and only three traits have an accuracy between 60 and 67% (Table 4.1). Three of the six pelvic traits that exceed 80% accuracy perform well for both males and females and include: pelvic inlet shape, true pelvis size and shape, and acetabulum size and orientation. Two of the six pelvic traits (pubic bone shape and subpubic concavity) performed better on males compared to females, while one of the six traits (ilium shape) performed better for females than males (Table 4.1). In terms of the morphological traits that achieved between 70 and 79% accuracy, all but three traits (greater sciatic notch angle, greater sciatic notch size & shape, and development of muscle markings) performed better for males than females. Conversely, the greater sciatic notch angle, greater sciatic notch size & shape, and development of muscle markings performed better on females than males. Of the three least accurate traits, iliac crest curvature and arc criterion performed better for females than males, while dorsal pubic pitting performed the least accurate for females and markedly better for males.

Table 4.1: Summary of overall correct sex classification and classification accuracies for females and males for 18 morphological pelvic traits

Trait	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pelvic Inlet Shape**	81	75	92.6	30	28	93.3	51	47	92.2
True Pelvis Size & Shape**	82	75	91.5	31	27	87.1	51	48	94.1
Acetabulum Size & Orientation**	93	84	90.3	36	33	88.9	57	52	91.2
Pubic Bone Shape*	122	101	82.8	49	30	61.2	73	71	97.3
Ilium Shape*	125	101	80.8	50	42	84	75	59	78.7
Subpubic Concavity*	119	96	80.7	48	26	54.2	71	70	98.6
Preauricular Sulcus	122	96	78.7	49	25	51	73	71	97.3
Sacrum Shape	84	66	78.6	31	19	61.3	53	47	88.7
Auricular Surface Height	119	93	78.2	46	23	50	73	70	95.9
Greater Sciatic Notch Angle	126	97	77	51	48	94.1	75	49	65.3
Obturator Foramen Shape	112	88	76.5	48	26	54.2	69	62	92.5
Greater Sciatic Notch Size & Shape	127	97	76.4	51	46	90.2	76	51	67.1
Ventral Arc	110	83	75.5	41	14	34.1	69	69	100
Development of Muscle Markings	98	72	73.5	40	38	95	58	34	58.6
Ischiopubic ramus	115	81	70.4	28	20	71.4	67	61	91
Iliac Crest Curvature	126	84	66.7	51	42	82.4	75	42	56
Arc Criterion	126	83	65.9	51	46	90.2	75	37	49.3
Dorsal Pubic Pitting	120	73	60.8	47	2	4.3	73	71	97.3

N = total number of individuals, N_f and N_m = total number of females and males, respectively, and n_f and n_m = number of females and males, respectively, with correct sex classification
*traits that have overall accuracy 80% and above
**traits that have an overall accuracy of 80% and above and perform well for females and males

A final sex assessment was conducted taking into account all 18 pelvic traits and using a majority rules approach. Overall sex assessment accuracy for subadults was 79.7%. Males were more accurately assessed compared to females (~90% and ~65%, respectively), while 8.6% (11/128) of individuals had an indeterminate final sex assessment. A second final sex assessment was conducted that placed more weight on traits proposed in previous research for subadult sex assessment. Overall accuracy increased to 82% when more weight was placed on pelvic traits proposed specifically for subadult sex assessment. Using this approach, female accuracy was higher than

that of males (98% and 71.4%, respectively), while 3.9% (5/128) of individuals had an indeterminate final sex assessment.

All nine logistic regression models produced by Albanese (2003) achieved or exceeded a minimum overall accuracy of 80% in this research (Table 4.2). The nine regression models tested in this study, however, show a decrease in accuracy compared to those achieved by Albanese (2003), where eight of the nine regression models show a marked decrease in accuracy of about 10% or more. While overall accuracy met or surpassed the 80% accuracy minimum, all regression models performed better on males compared to females. Models 5 and 20, however, showed the most comparable levels of correct sex classification for both males and females (a difference of 1.7 and 5.3% respectively).

Table 4.2: Summary of overall correct sex classification and classification accuracies for females and males for nine logistic regression models (see Table 3.5 in Chapter 3)

Regression Model	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Model 20	76	71	93.4	31	28	90.3	45	43	95.6
Model 8	74	64	86.5	28	20	71.4	46	44	95.7
Model 3	73	63	86.3	28	20	71.4	45	43	95.6
Model 4	76	65	85.5	29	23	79.3	47	42	89.4
Model 2	75	63	84	29	19	65.5	46	44	95.7
Model 1	74	62	83.8	29	20	69	45	42	93
Model 10	75	62	82.7	29	21	71.4	45	43	95.6
Model 11	75	62	82.7	29	20	69	46	42	91.3
Model 5	76	61	80.3	29	23	79.3	47	38	81

4.1.2 Intra-observer Error and Ranking of Morphological Traits

Five morphological traits produced a percentage of error that fell below the acceptable error level of 10% (Nichol and Turner 1986; Rogers and Saunders 1994) (Table 4.3). True pelvis size & shape and preauricular sulcus had a percentage of error that marginally exceeds the 10% threshold (10.7% and 10.8%, respectively), while obturator foramen shape and subpubic concavity exceeded the percentage of error threshold by 1% (11.1 and 11.2 percent, respectively). The remaining nine morphological traits failed to meet the acceptable level of error, where intraobserver

error is between 12-26%. Cohen's kappa was also calculated to examine inter-rater reliability, including same rater agreement. The kappa statistic measures the proportion of agreement between two raters, or rounds of rating, while taking into account the possibility that a level of agreement between ratings may occur by chance (Ranganathan et al. 2017). Cohen's kappa values can range from 0 – 1.0, where 0 representing the level of agreement that occurs by chance and 1 representing perfect agreement between raters or scoring rounds (Ranganathan et al. 2017). While the hierarchy of values for interpreting the kappa statistic is arbitrary, general interpretation is as follows: 0 = poor/equivalent to chance; 0.10-0.20 = slight agreement; 0.21-0.40 = fair agreement; 0.41-0.60 = moderate agreement; 0.61-0.80 = substantial agreement; 0.81-0.99 = near perfect; and 1 = perfect agreement (Landis and Koch 1977; Ranganathan et al. 2017). Overall, the results of the kappa statistic ranged between 0.449 and 0.931 for the morphological pelvic traits, with the majority of kappa values falling between 0.632 and 0.931 (Table 4.3). Three pelvic traits had a kappa value above 0.81 and include auricular surface height (κ_c = 0.931), pubic bone shape (κ_c = 0.828), and greater sciatic notch size & shape (κ_c = 0.812). The bulk of traits (n=10) had substantial agreement between trial 1 and trial 2 assessments, while five traits only had moderate agreement between the two trials.

Table 4.3: Percentage of intraobserver agreement, error, and Cohen's Kappa values for the 18 morphological pelvic sex traits

Trait	n	Error #	% Agreement	% Error	κ_c
Auricular Surface Height	90	2	97.8	2.2	0.931
Dorsal Pubic Pitting	91	2	97.8	2.2	0.655
Ventral Arc	78	5	93.6	6.4	0.778
Pubic Bone Shape	93	5	93.5	6.5	0.828
Greater Sciatic Notch Size & Shape	96	9	90.6	9.4	0.812
True Pelvis Size & Shape	56	6	89.3	10.7	0.768
Preauricular Sulcus	93	10	89.2	10.8	0.704
Obturator Foramen Shape	81	9	88.9	11.1	0.749
Subpubic Concavity	89	10	88.8	11.2	0.689
Ischiopubic ramus	81	9	87.7	12.3	0.632
Pelvic Inlet Shape	56	7	87.5	12.5	0.741
Development of Muscle Markings	77	10	87	13	0.719
Acetabulum Size & Orientation	65	9	86.2	13.8	0.699
Sacrum Shape	58	9	84.5	15.5	0.570
Greater Sciatic Notch Angle	96	19	80.2	19.8	0.602
Ilium Shape	93	23	75.3	24.7	0.505
Arc Criterion	96	24	74	26	0.502
Iliac Crest Curvature	95	25	73.7	26.3	0.449

The overall rankings of each pelvic trait that take into account both the degree of intra-observer error and level of accuracy are provided in Table 4.4. The top two traits that appear to be most effective when applied to subadults are pubic bone shape and true pelvis size & shape. The least effective traits when applied to subadults include iliac crest curvature and arc criterion.

Table 4.4: Ranking of effectiveness of pelvis traits (accuracy and precision combined)

Trait	Accuracy Score	Precision Score	Overall Score	Rank
Pubic Bone Shape	4	4	8	1
True Pelvis Size & Shape	2	6	8	1
Auricular Surface Height	9	1	10	3
Pelvic Inlet Shape	1	11	12	4
Preauricular Sulcus	7	7	14	5
Subpubic Concavity	6	9	15	6
Acetabulum Size & Orientation	3	13	16	7
Ventral Arc	13	3	16	7
Greater Sciatic Notch Size & Shape	12	5	17	9
Obturator Foramen Shape	11	8	19	10
Dorsal Pubic Pitting	18	1	19	10
Ilium Shape	5	16	21	12
Sacrum Shape	8	14	22	13
Greater Sciatic Notch Angle	10	15	25	14
Ischiopubic ramus	15	10	25	14
Development of Muscle Markings	14	12	26	16
Iliac Crest Curvature	16	18	34	17
Arc Criterion	17	17	34	17

4.2 Age-Related Trends: Trait Appearance and Stabilization

The results in the section above provide clear evidence that there are some sexually dimorphic traits in the pelvis that, while normally used for adult sex assessment, can be applied to individuals 20 years old and younger. Table 4.5 outlines the age at which each pelvic trait appears and stabilizes. Age of appearance refers to the age at which the male or female form of a trait is first expressed for the correct sex. The mean age of stabilization refers to the average age at which a trait has high accuracy (80% and above) of sex classification with a sample size above three. Minimum age of stabilization, then, refers to the earliest age at which a trait demonstrates high correct sex classification given a minimum sample size of three.

Four traits (greater sciatic notch size & shape, ilium shape, greater sciatic notch angle, and arc criterion) have both male and female expressions first appearing at birth. Two traits, development of muscle markings and iliac crest curvature, have the female

expression present at birth and the male expression appearing at a later age (17 and 6 years old, respectively) than females. The preauricular sulcus in females was first expressed at the age of 6, while auricular surface height in females was first expressed at 8 years of age. Both traits, however, had a male expression appearing at birth. Six pelvic traits had a female expression first appearing at 12 years of age, of which five had a male expression first appearing one year later at 13 years old, and one (pubic bone shape) was first expressed at birth for males (Table 4.5). Three traits (subpubic concavity, ischiopubic ramus ridge, and ventral arc) had the female expression first appearing at 14 years of age but all three traits were first expressed at birth for males. Dorsal pubic pitting was the only trait with a late first appearance for the female expression at 18 years old but the first male expression appeared at birth.

Seven pelvic traits had an earlier average age of stabilization in males compared to females (Table 4.5). Two traits, subpubic concavity and ischiopubic ramus ridge, had an average age of stabilization one year earlier than females. Preauricular sulcus and pubic bone shape stabilized, on average, two years earlier in males than females, while ventral arc and auricular surface height had an average age of stabilization that was three years earlier in males than females. Dorsal pubic pitting was the only trait with a vastly different average age of stabilization between males and females, a difference of over seven years.

Ten pelvic traits had an earlier average age of stabilization in females compared to males. Four traits (pubic inlet shape, true pelvis size & shape, acetabulum size and orientation, and sacrum shape) had an average age of stabilization that occurs one year earlier in females compared to males. Ilium shape had a mean age of stabilization that was two years earlier in females, while three traits (greater sciatic notch angle, development of muscle markings, and iliac crest curvature) stabilized three years earlier in females compared to males. The arch criterion had an average age of stabilization that was approximately four years earlier in females than males, and the greater sciatic notch size & shape stabilizes in females five years earlier than males. Only one trait, obturator foramen shape, had an average age of stabilization that was comparable between females and males (18.1 and 18.8 years, respectively).

Three traits (ventral arc, dorsal pubic pitting, and sacrum shape) had comparable minimum and mean ages of stabilization in females, while development of muscle markings, iliac crest curvature, and arc criterion had comparable minimum and mean age of stabilization in males. For all six traits, the similar minimum and mean ages of stabilization occurred at 18 years and above.

Table 4.5: Summary of age of appearance and stabilization of dimorphic traits of the pelvis

Trait	Age of Appear. (F)	Avg Age of Stab. (F)	Min. Age of Stab. (F)	Age of Appear. (M)	Avg Age of Stab. (M)	Min. Age of Stab. (M)
Subpubic Concavity	14 yrs	17.8 yrs	14 yrs	birth	16.1 yrs	1 yr
Ischiopubic ramus	14 yrs	19.1 yrs	18 yrs	birth	18.3 yrs	10 yrs
Ventral Arc *	14 yrs	19.6 yrs	19 yrs	birth	15.9 yrs	1 yr
Pubic Bone Shape	12 yrs	18.3 yrs	16 yrs	birth	15.8 yrs	1 yr
Dorsal Pubic Pitting	18 yrs	20 yrs	20 yrs	birth	12.7 yrs	birth
Greater Sciatic Notch Size & Shape	birth	14.2 yrs	4 yrs	birth	18.9 yrs	15 yrs
Auricular Surface Height	8 yrs	17.9 yrs	16 yrs	birth	14.8 yrs	birth
Preauricular Sulcus	6 yrs	17.5 yrs	16 yrs	birth	15.2 yrs	birth
Ilium Shape**	birth	16.5 yrs	12 yrs	birth	19 yrs	17 yrs
Pelvic Inlet Shape	12 yrs	17.8 yrs	14-16 yrs	13 yrs	19 yrs	17 yrs
True Pelvis Size & Shape	12 yrs	18.3 yrs	17 yrs	13 yrs	19 yrs	17 yrs
Obturator Foramen Shape	12 yrs	18.1 yrs	16 yrs	13 yrs	18.8 yrs	15 yrs
Acetabulum Size & Orientation°	12 yrs	17.9 yrs	16 yrs	13 yrs	19 yrs	17 yrs
Development of Muscle Markings	birth	17.1 yrs	12 yrs	17 yrs	19.7 yrs	19 yrs
Sacrum Shape	12 yrs	18.1 yrs	18 yrs	13 yrs	19 yrs	17 yrs
Greater Sciatic Notch Angle^	birth	16.3 yrs	8 yrs	birth	19 yrs	17 yrs
Arc Criterion	birth	15.9 yrs	8 yrs	birth	19.7 yrs	19 yrs
Iliac Crest Curvature	birth	15.7 yrs	8 yrs	6 yrs	19.4 yrs	18 yrs

*a form of the precursor arc may appear earlier, see Chapter 5 discussion on symphyseal face "twisting"

**correct assessment of ilium shape exceeds incorrect assessment for females at 8 yrs old, but sample size is low (n=3). Male stabilization of this trait might be earlier but no males 14-16 years old available in sample

°correct assessment exceeds incorrect at 14, but sample size is low (n=3)

^female scoring is almost consistently reliable, but small sample size (n=2) until 12 years old; Male stabilization may be earlier, but small sample sizes (n=1-3) makes it difficult to assess

4.2.1 Pre-determined Age Category Accuracy

Overall, ten morphological traits (ilium shape, pelvic inlet shape, greater sciatic notch size & shape, obturator foramen, greater sciatic notch angle, sacrum shape, arc criterion, development of muscle markings, iliac crest curvature, and true pelvis size & shape) showed a consistent increase in accuracy from younger age categories to older age categories. Six traits (subpubic concavity, pubic bone shape, auricular surface, preauricular sulcus, ventral arc, and ischiopubic ramus) showed decreasing accuracy between the infant and child age categories, with a subsequent gradual increase in accuracy from the child category to the late adolescent category. Dorsal pubic pitting was the only trait that did not conform to either aforementioned pattern. Instead, a gradual decrease in accuracy was observed from the infant category to the adolescent category, with a subsequent increase in accuracy from adolescent to late adolescent categories.

Two traits (ilium shape and pelvic inlet shape) showed high accuracies in the adolescent (90-93%) and late adolescent (92-94%) categories. Both traits performed well on males and females in both the adolescent and late adolescent categories with accuracy at or exceeding 88% (Table 4.6-4.7). A Fisher's exact test demonstrated that the high accuracy of correct sex classification of the ilium shape in the adolescent category (χ^2= 10.313, df = 1, p= 0.009) and late adolescent category (χ^2= 48.548, df=1, p<0.001) were statistically significant. In terms of the pelvic inlet shape, despite the adolescent category having a high accuracy (90%), the results of a Fisher's exact test did not reveal a statistical significance for this age category (χ^2= 4.444, df = 1, p= 0.200), but did indicate a statistically significant relationship for the late adolescent category (χ^2= 52.058, df = 1, p<0.001).

Table 4.6: Accuracy rates for ilium shape by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	16	7	43.8	3	1	33.3	13	6	46.2
	Child	21	13	61.9	12	9	75	9	4	44.4
	Adolescent	15	14	93.4	12	11	91.7	3	3	100
	Late Adolescent	73	67	91.8	23	21	91.3	50	46	92

Table 4.7: Accuracy rates for pelvic inlet shape by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	1	0	0	-	-	-	1	0	0
	Child	1	1	100	-	-	-	1	1	100
	Adolescent	10	9	90	9	8	88.9	1	1	100
	Late Adolescent	69	65	94.2	21	20	95.2	48	45	93.8

Seven traits produced high accuracy ranging from 84-97% in the late adolescent categories, which include: subpubic concavity, pubic bone shape, greater sciatic notch size and shape, true pelvis size & shape, obturator foramen shape, acetabulum size & orientation, and greater sciatic notch angle (Tables 4.8-4.14). All seven traits performed well for both males and females, where sex classification for both sexes met or exceeded a minimum of 80%. The results of Fisher's exact tests revealed that the high accuracy obtained for the late adolescent category was statistically significant ($p<0.001$) for five traits (subpubic concavity, pubic bone shape, greater sciatic notch size & shape, obturator foramen shape, and greater sciatic notch angle) (Tables 4.8- 4.10, 4.12, 4.14). Additionally, a Pearson's chi square test indicated that the correct sex classification observed in late adolescents category for true pelvis size (Table 4.11) and acetabulum size & orientation (Table 4.13) was statistically significant ($p<0.001$).

Table 4.8: Accuracy rates for subpubic concavity by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	10	9	90	1	0	0	9	9	100
	Child	21	9	42.9	12	0	0	9	9	100
	Adolescent	15	10	66.7	12	7	58.3	3	3	100
	Late Adolescent	73	68	93.2[a]	23	19	82.6	50	49	98

[a] $\chi^2 = 51.461$, df = 1, $p<0.001$

Table 4.9: Accuracy rates for pubic bone shape by age category for combined and separate sexes. NF = total females, nF = correctly identified females, NM = total males, nM = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	14	13	92.9	2	1	50	12	12	100
	Child	21	9	42.9	12	0	0	9	9	100
	Adolescent	15	9	60	12	6	50	3	3	100
	Late Adolescent	72	70	97.2[a]	23	23	100	49	47	95.9

[a] χ^2= 63.536, df = 1, p<0.001

Table 4.10: Accuracy rates for greater sciatic notch & shape by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	18	7	38.9	3	3	100	15	4	26.7
	Child	21	16	76.2	13	13	100	8	3	37.5
	Adolescent	15	12	80	12	10	83.3	3	2	66.7
	Late Adolescent	73	62	84.9[a]	23	20	87	50	42	84

[a] χ^2= 40.651, df = 1, p<0.001

Table 4.11: Accuracy rates for true pelvis size & shape by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	1	1	100	-	-	-	1	1	100
	Child	1	1	100	-	-	-	1	1	100
	Adolescent	10	6	60	9	5	55.6	1	1	100
	Late Adolescent	70	67	95.7[a]	22	22	100	48	45	93.8

[a] χ^2= 57.750, df = 1, p<0.001

Table 4.12: Accuracy rates for obturator foramen shape by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Child	21	9	42.9	12	0	0	9	9	100
	Adolescent	15	10	66.7	12	7	58.3	3	3	100
	Late Adolescent	73	64	87.7[a]	23	19	82.6	50	45	90

[a] χ^2= 37.635, df = 1, p<0.001

Table 4.13: Accuracy rates for acetabulum size & orientation by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Child	4	3	75	1	1	100	3	2	66.7
	Adolescent	15	11	73.3	12	9	75	3	2	66.7
	Late Adolescent	73	69	94.5 [a]	23	22	95.7	50	47	94

[a] χ^2= 56.229, df = 1, p<0.001

Table 4.14: Accuracy rates for greater sciatic notch angle by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	17	6	35.3	3	3	100	14	3	21.4
	Child	21	16	76.2	13	12	92.3	8	4	50
	Adolescent	15	14	93.3	12	12	100	3	2	66.7
	Late Adolescent	73	61	83.6 [a]	23	21	91.3	50	40	80

[a] χ^2= 32.782, df = 1, p<0.001

Two traits, auricular surface height and preauricular sulcus, demonstrated overall high accuracies for the late adolescent categories (89 and 92%, respectively). Both traits, however, performed markedly better on males compared to females (17 and 20% difference, respectively) where female accuracy was just shy of a minimum accuracy of 80% (Table 4.15 and 4.16). The results of a Fisher's exact test indicated that correct sex classification in the late adolescent category for auricular surface height (χ^2= 38.685, df = 1, p<0.001) and preauricular sulcus (χ^2= 47.586, df = 1, p<0.001) was statistically significant.

Table 4.15: Accuracy rates for auricular surface height by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	14	12	85.7	2	0	0	12	12	100
	Child	21	9	42.9	13	1	7.7	8	8	100
	Adolescent	12	8	66.7	9	5	55.6	3	3	100
	Late Adolescent	72	64	**88.9**	22	17	77.3	50	47	94

Table 4.16: Accuracy rates for preauricular sulcus by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	13	11	84.6	1	0	0	12	11	91.7
	Child	21	10	47.6	13	2	15.4	8	8	100
	Adolescent	15	8	53.3	12	5	41.7	3	3	100
	Late Adolescent	73	67	**91.8**	23	18	78.3	50	49	98

Four traits had overall accuracies ranging from 81-90% for the late adolescent age category; thereby exceeding the 80% threshold. Upon closer examination, however, three of the traits (ventral arc, ischiopubic ramus ridge, sacrum shape) performed markedly better on males compared to females, where male accuracy exceeded that of the females between 19 and 37%, depending on the trait (Tables 4.17–4.20). Moreover, accuracy for female sex assessment using any one of these three traits was moderate at best, ranging from 63-73%. One trait, the arch criterion, demonstrated the opposite pattern in which the trait performed better on females compared to males (87% and 62%, respectively). Pearson's chi-square tests showed that the accuracies achieved for the late adolescent category for ischiopubic ramus and arch criterion were statistically significant (χ^2= 33.124, df=2, p<0.001; χ^2= 16.425, df = 2, p<0.001, respectively), despite the variable efficacy of those traits between males and females. Moreover, a Fisher's exact test showed that the accuracy in correct sex classification observed for late adolescents using sacrum shape and ventral arc were also statistically significant (χ^2= 21.333, df=1, p<0.001; χ^2= 35.579, df = 1, p<0.001, respectively). While high overall accuracy is seen for the infant category (83%) when using the ventral arc, only male sex is correctly classified and no females are correctly classified (Table 4.17).

Table 4.17: Accuracy rates for ventral arc by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	12	10	83.3	2	0	0	10	10	100
	Child	17	7	41.2	10	0	0	7	7	100
	Adolescent	13	5	45.5	10	2	20	3	3	100
	Late Adolescent	68	61	**89.7**	19	12	63.2	49	49	100

Table 4.18: Accuracy rates for ischiopubic ramus ridge by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	11	7	63.6	2	0	0	9	7	77.8
	Child	20	8	40	12	0	0	8	8	100
	Adolescent	15	7	46.6	12	4	33.3	3	3	100
	Late Adolescent	69	59	**85.5**	22	16	72.7	47	43	91.5

Table 4.19: Accuracy rates for sacrum shape by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	1	1	100	-	-	-	1	1	100
	Child	1	1	100	-	-	-	1	1	100
	Adolescent	12	7	58.3	10	5	50	2	2	100
	Late Adolescent	70	57	**81.4**	21	14	66.7	49	43	87.8

Table 4.20: Accuracy rates for arch criterion by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	17	6	35.3	3	3	100	14	3	21.4
	Child	21	15	71.4	13	13	100	8	2	25
	Adolescent	15	11	73.3	12	10	83.3	3	1	33.3
	Late Adolescent	73	51	**83.6**	23	20	87	50	31	62

The remaining three traits (dorsal pubic pitting, development of muscle markings, and iliac crest curvature) failed to achieve an accuracy that met or surpassed the 80% threshold in any of the age categories (Tables 4.21–4.23). While dorsal pubic pitting did

achieve an accuracy of 91% for the infant category, there was only one female in the age category (Table 4.21). Moreover, dorsal pubic pitting consistently performed extremely well for males but extremely poor for females. This contrasts to the development of muscle markings and iliac crest curvature where females consistently had higher sex classification accuracy compared to males, with the exception of adolescent category for iliac crest curvature (Tables 4.22–4.23). Despite not achieving a minimum accuracy of 80% in any age category, a Fisher's exact test showed that the sex classification accuracy achieved using development of muscle markings and iliac crest curvature was statistically significant for the late adolescent category (χ^2= 29.275, df=1, p<0.001; χ^2= 16.164, df = 1, p<0.001, respectively).

Table 4.21: Accuracy rates for dorsal pubic pitting by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	12	11	91.7	1	0	0	11	11	100
	Child	21	9	42.9	12	0	0	9	9	100
	Adolescent	15	3	20	12	0	0	3	3	100
	Late Adolescent	72	50	69.4	22	2	9.1	50	48	96

Table 4.22: Accuracy rates for development of muscle markings by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Child	10	4	40	5	4	80	5	0	0
	Adolescent	15	11	73.4	12	11	91.7	3	0	0
	Late Adolescent	73	57	**78.1**	23	23	100	50	34	68

Table 4.23: Accuracy rates for iliac crest curvature by age category for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded values indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	17	6	35.3	3	3	100	14	0	0
	Child	21	15	71.4	13	13	100	8	3	37.5
	Adolescent	15	11	73.3	12	8	66.7	3	3	100
	Late Adolescent	73	54	**74**	23	18	78.3	50	36	72

When examining the results of the final sex assessments when all 18 morphological traits are weighted equally, only the late adolescent category showed an accuracy that exceeded a minimum of 80% (Table 4.24). A Pearson chi-square test showed that the high correct sex classification in the late adolescent category was statistically significant (χ^2= 66.251, df = 2, p<0.001). While the adolescent category did not meet the minimum accuracy level required in this study, a Pearson chi-square test showed that the sex classification in this age category was statistically significant (χ^2= 10.313, df = 2, p = 0.006). When a final sex assessment that placed more weight on juvenile traits was applied, the adolescent and late adolescent categories exceeded an 80% accuracy minimum where accuracy was 100% and 94.5%, respectively (Table 4.25). The accuracy for adolescent sex assessment showed a marked increase when more weight was placed on pelvic traits proposed for subadults such as greater sciatic notch angle, greater sciatic notch depth, auricular surface elevation, arch criterion, and iliac crest curvature (73.3% vs. 100%). Moreover, the child age category also demonstrated an increase in accuracy when juvenile traits were weighed more than when all 18 traits are weighed equally (40.9% to 72.7%). The infant category, however, showed a marked decrease in accuracy when juvenile traits were weighed more (67% to 28%), which is attributed to a stark decrease in correct male sex classification (Table 4.25). A Pearson chi-square showed that correct sex classification in the adolescent category (χ^2= 15.000, df = 1, p = 0.002) and late adolescent category (χ^2= 62.188, df = 2, p<0.001) are statistically significant. While the child category failed to meet the 80% accuracy threshold, a Pearson chi-square test showed that the correct sex classification in this age category was statistically significant (χ= 9.346, df = 2, p = 0.009).

Table 4.24: Accuracy for final sex assessment (all traits weighted equally). Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	18	12	66.7	3	2	66.7	15	10	66.7
	Child	22	9	40.9	13	1	7.7	9	8	88.9
	Adolescent	15	11	**73.3**	12	8	66.7	3	3	100
	Late Adolescent	73	70	**95.9**	23	22	95.7	50	48	96

Table 4.25: Accuracy for final sex assessment (juvenile traits weighted more). Bolded values indicate statistical significance

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Age Category	Infant	18	5	27.8	3	3	100	15	2	13.3
	Child	22	22	**72.7**	13	13	100	9	3	33.3
	Adolescent	15	15	**100**	12	12	100	3	3	100
	Late Adolescent	73	69	**94.5**	23	22	95.7	50	47	94

When examining the accuracy of the nine logistic regression models proposed by Albanese (2003) by age category, there was an overall increase in accuracy between the adolescent and late adolescent categories with the exception of Model 8 (Tables 4.26 and 4.27). Six models had an accuracy that met or exceeded a minimum of 80% for both the adolescent and late adolescent categories (Table 4.26). Models 4, 10, and 5 performed better on females than males in the adolescent category, while Models 8 and 11 performed better on males compared to females in the adolescent category. Model 20 only had females included for the adolescent category, thus the overall high accuracy of this regression model for the adolescent category should be taken with caution. In terms of the late adolescent age category, while all six models had an overall accuracy that ranged from 80-94%, all models performed better on males compared to females. Model 20 was the only logistic regression equation that performed well for both males and females, surpassing 80% accuracy for both sexes. Despite the high accuracy for the adolescent age category for all six models, none were found to be statistically significant. All models, however, were statistically significant for the late adolescent age category (Model 20: χ^2= 50.336, df =1, p<0.001; Model 8: χ^2= 28.602, df = 1, p<0.001; Model 4: χ^2= 30.707, df = 1, p<0.001; Model 10: χ^2= 23.534, df =1, p<0.001; Model 11: χ^2= 23.534, df =1, p<0.001; Model 5: χ^2= 21.112, df = 1, p<0.001), when examined using a Fisher's exact test (Table 4.26).

Three logistic regression equations, Models 1-3, had an overall accuracy that exceeded the 80% minimum for the late adolescent age category only (Table 4.27). All three models performed extremely well on males in the late adolescent category but performed moderately, at best, for late adolescent females. Despite the male bias in correct sex classification, a Fisher's exact test indicated that sex classification was statistically significant (p<0.001) for the late adolescent age category for all three models (Table 4.27).

Table 4.26: Accuracy of Logistic Regression Models with both age categories >80%. Bolded values indicate statistical significance

Logit. Model	Age Category	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Model 20	Adolescent	9	8	88.9	9	8	88.9	-	-	-
	Late Adolescent	67	62	**93.9**	23	20	87	44	43	97.7
Model 8	Adolescent	9	8	88.9	8	7	87.5	1	1	100
	Late Adolescent	65	56	**86.2**	20	13	65	45	43	95.6
Model 4	Adolescent	10	8	80	8	7	87.5	2	1	50
	Late Adolescent	66	57	**86.4**	21	16	76.2	45	41	91.1
Model 10	Adolescent	10	8	80	8	7	87.5	2	1	50
	Late Adolescent	65	54	**83.1**	21	14	66.7	44	40	90.9
Model 11	Adolescent	10	8	80	8	6	75	2	2	100
	Late Adolescent	65	54	**83.1**	21	14	66.7	44	40	90.9
Model 5	Adolescent	10	8	80	8	7	87.5	2	1	50
	Late Adolescent	66	53	**80.3**	21	16	76.2	45	37	82.2

Table 4.27: Accuracy of Logistic Regression Models with only one age category >80%

Logit. Model	Age Category	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Model 3	Adolescent	9	7	77.8	8	6	75	1	1	100
	Late Adolescent	64	56[a]	87.5	20	14	70	44	42	95.5
Model 2	Adolescent	11	8	72.7	9	6	66.7	2	2	100
	Late Adolescent	63	55[b]	87.3	19	13	68.4	44	42	95.5
Model 1	Adolescent	9	7	77.8	8	6	75	1	1	100
	Late Adolescent	65	55[c]	84.6	21	14	66.7	44	41	93.2

[a] χ^2= 31.418, df = 1, p = ; [b] χ^2= 29.846, df = 1, p<0.001 ; [c] χ^2= 26.363, df = 1, p<0.001

4.3 Pubertal Stage Assessment

In order to examine the relationship that may exist between sexual dimorphism with the stages of pubertal growth, pubertal stage was first examined. A total of 96 individuals between 8 and 20 years of age met the criteria required for osteological pubertal stage assessment, 38 of which were females and 58 males. Hamate hook development showed some developmental variability in females, with one female having a hamate hook score of “H” (hook appearing) as late as 18 years of age and two females with a hamate hook completely developed as early at 12 years (Figure 4.1). Moreover, one female had an undeveloped hamate hook as late as 13 years old. Males showed less variability in the development of the hamate hook, with two individuals

having had a hamate hook developing (score H.5) at 17 and 19 years of age (Figure 4.2). Figure 4.3 showed that the hamate hook was complete (stage I) between 12-20 years of age in females and between 15 and 20 years, and became the dominant developmental state by 14 and 15, respectively (Figure 4.1 and 4.2). One male outlier was observed with a completely developed hamate hook at 13 years of age (Figure 4.3).

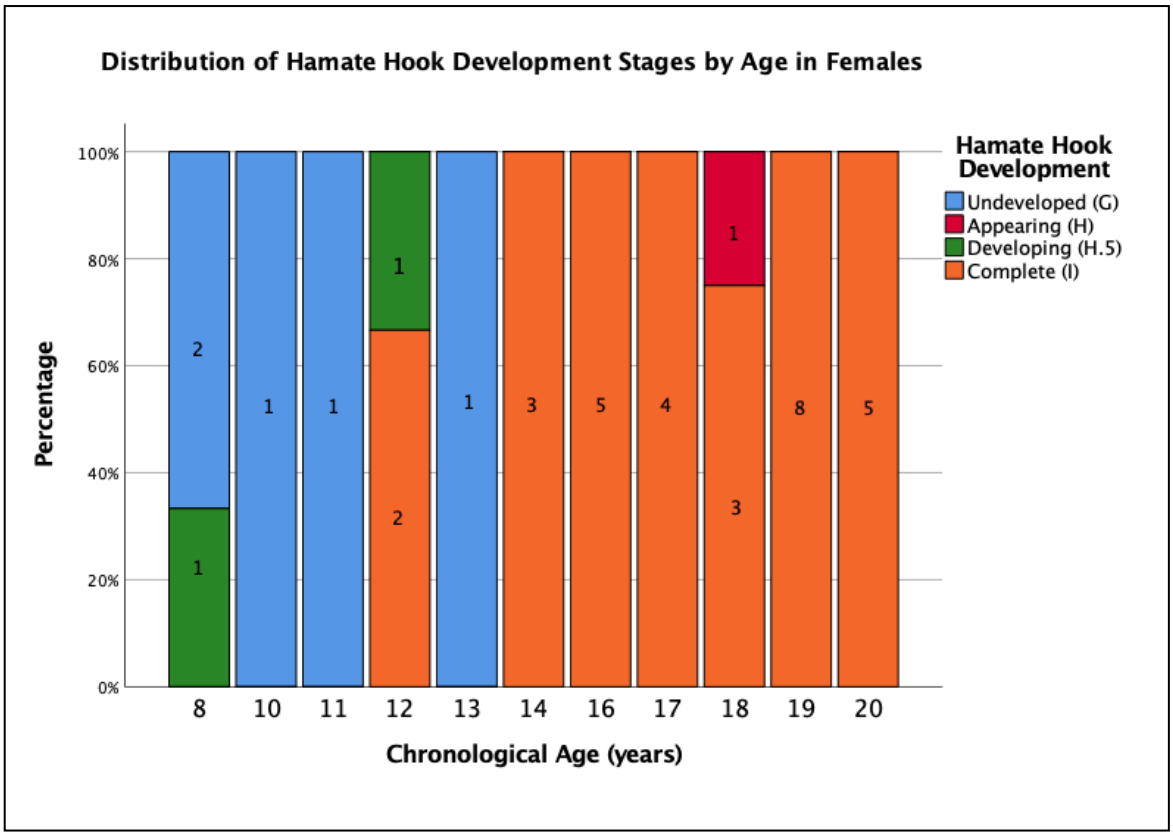

Figure 4.1: Distribution of hamate hook development scores in females between the ages of 8 and 20 years. The numbers on the bars represent the number of individuals at the stage of development.

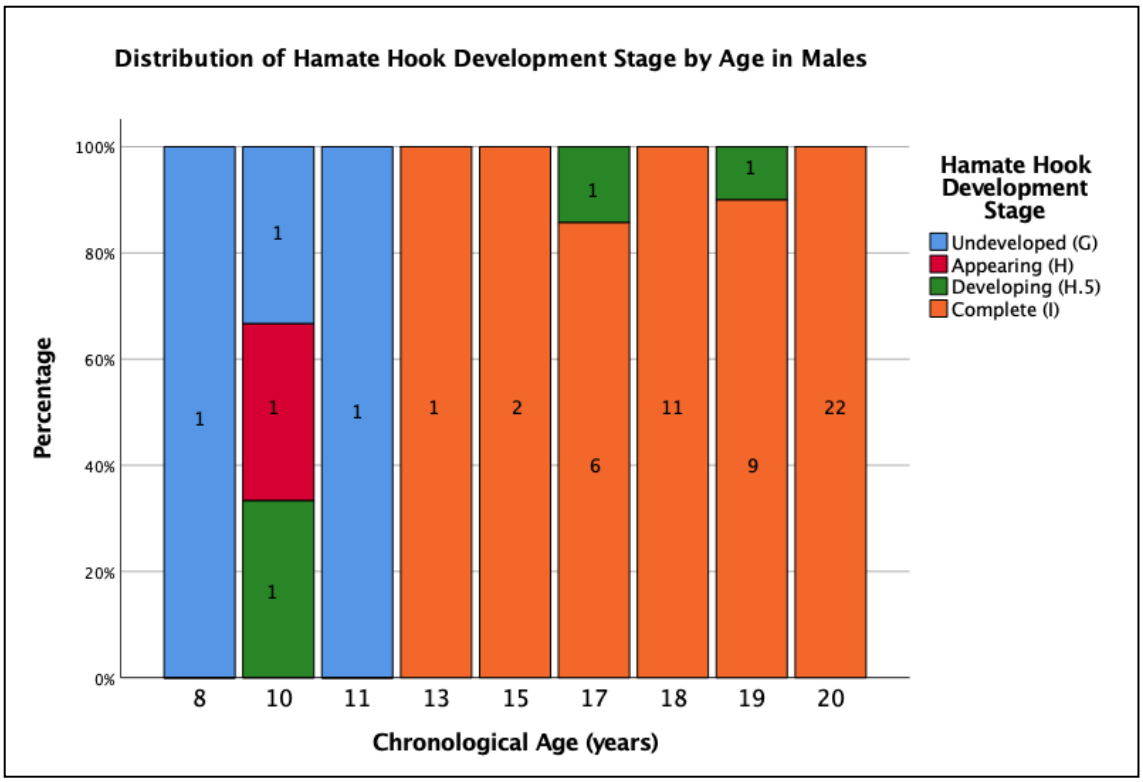

Figure 4.2: Distribution of hamate hook development scores in males between the ages of 8 and 20 years. The numbers on the bars represent the number of individuals at the stage of development

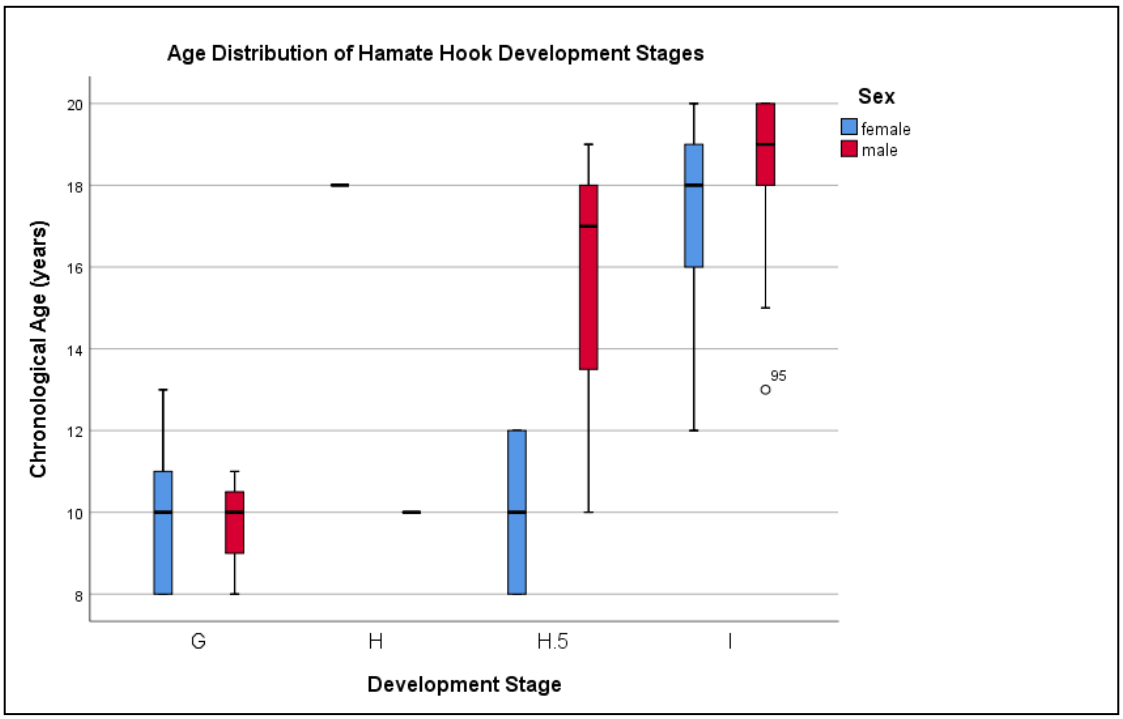

Figure 4.3: Age distribution comparison of all four hamate development stages between males and females

The fusion state of distal hand phalangeal epiphyses in females showed overall less developmental variability compared to hamate hook development. Fused phalangeal epiphyses became the dominant developmental state at 14 years of age, while unfused epiphyses was the dominant fusion state between 8 and 13 years of age (Figure 4.4). As shown in Figure 4.6, the lone eight-year-old female with fused distal phalanges epiphyses seen in Figure 4.4 was considered an outlier. In males, fused distal phalangeal epiphyses was the dominant developmental state at 17 years of age, although this particular age still contained one individual with unfused epiphyses and one with partial epiphyseal fusion (Figure 4.5). Unfused phalangeal epiphyses were the common fusion state in males between 10 and 15 years of age. While complete fusion of epiphyses of the distal phalanges in females occurred between 14 and 20 years of age, where the majority of females with fused epiphyses were between 16 and 19 years (Figure 4.6). In contrast, the majority of males had fused epiphyses on the distal phalanges between 18 and 20 years of age (Figure 4.6).

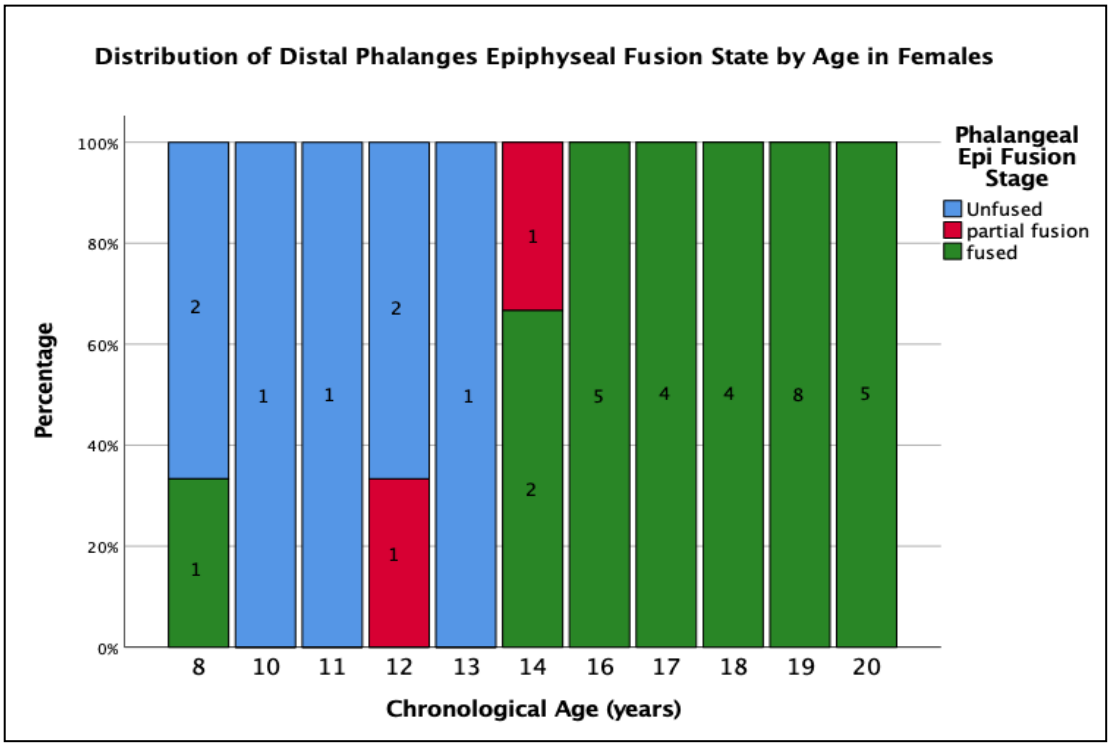

Figure 4.4: Distribution of fusion states of distal phalangeal epiphyses in females between the ages of 8 and 20 years.

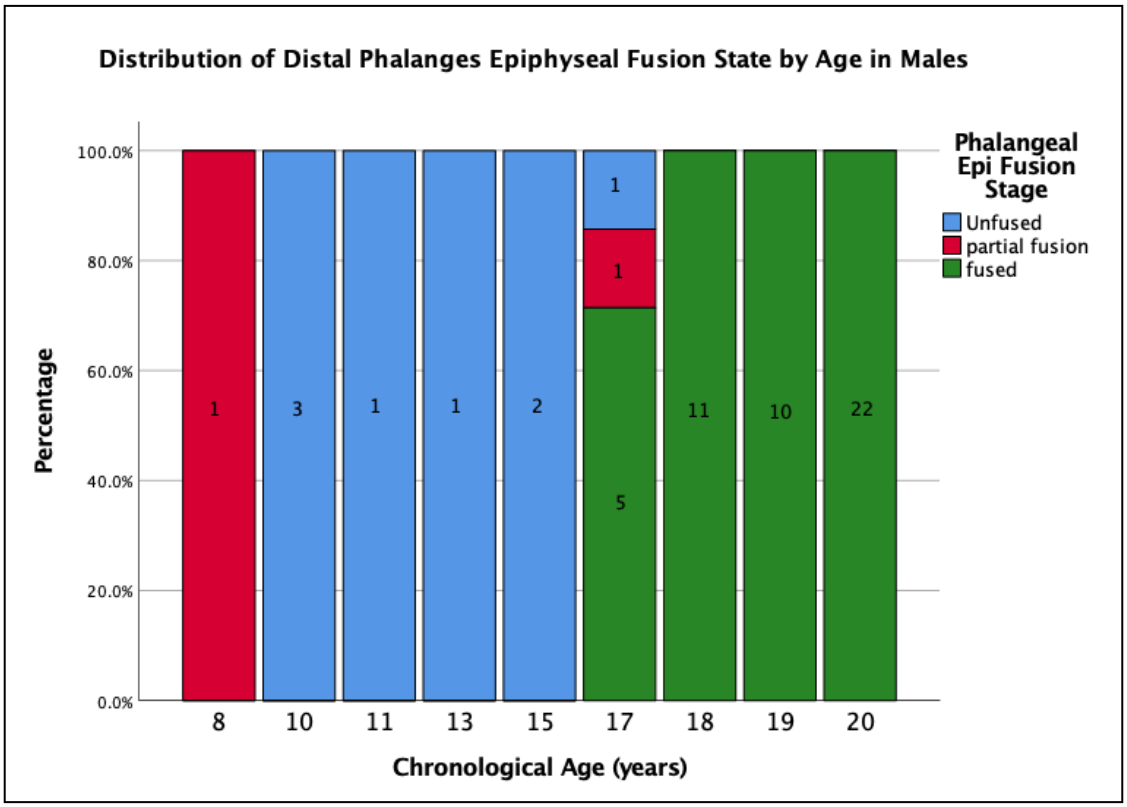

Figure 4.5: Distribution of fusion states of distal phalangeal epiphyses in males between the ages of 8 and 20 years.

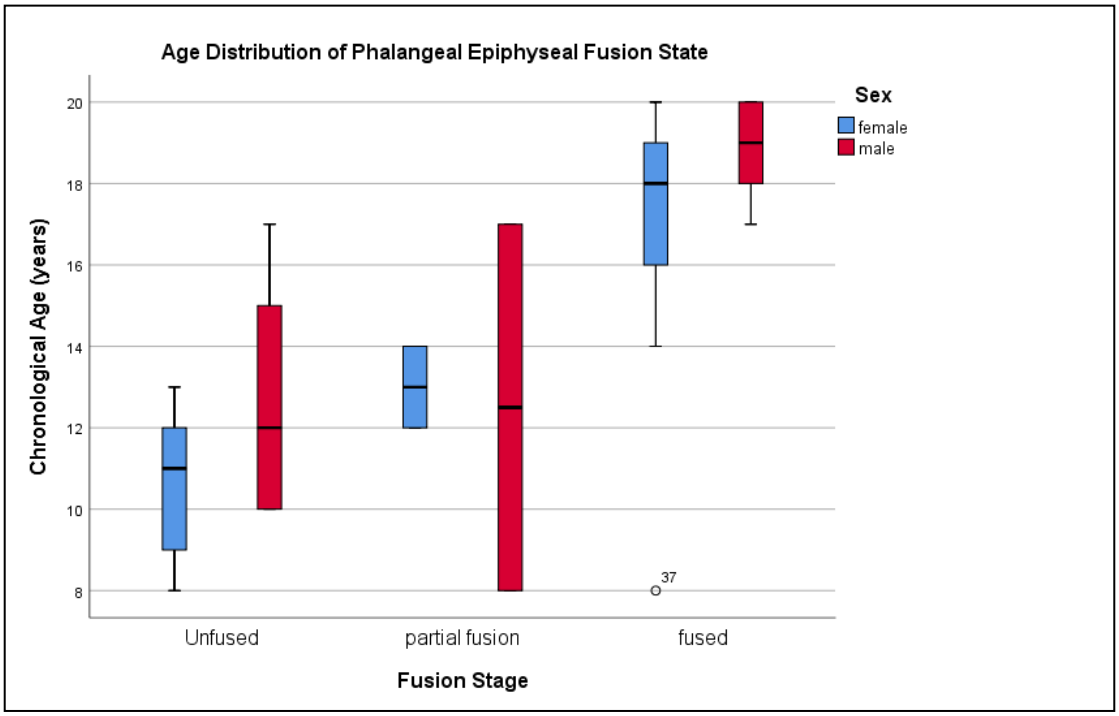

Figure 4.6: Age distribution of epiphyseal fusion stages of the distal phalanges between males and females

Epiphyseal fusion of the distal radius showed a gradual transition in females. Unfused distal radius epiphyses were the most common state between 8 and 14 years of age in females (Figure 4.7). The age of 16 appeared to have been a transition age of sorts, where distal radius epiphyses at all three-fusion states were seen. Partial and complete fusion of the distal radius epiphysis was seen at comparable quantities between 17 and 18 years of age, while a completely fused distal radius epiphysis did not become the dominant development stage until 19 years of age. In males, an unfused distal radius epiphysis was the dominant developmental state from 8-15 years of age (Figure 4.8), although males with an unfused distal radius epiphysis were seen in individuals as late as 20 years of age. Fused distal radius epiphyses became the dominant fusion state at around 19 years, but more notably by 20 years old. As seen in Figure 4.9, while fused distal radius epiphyses in females occurred from 14-20 years of age, the majority of individuals seen are between 17-20 years. In males, a fused distal radius epiphysis occurred between 18 and 20 years of age, with the majority of individuals with a fused epiphysis seen between 19 and 20 years (Figure 4.9). The 17-year-old male with a fused distal radius epiphysis seen in Figure 4.8 appears to be an outlier, as depicted in Figure 4.9.

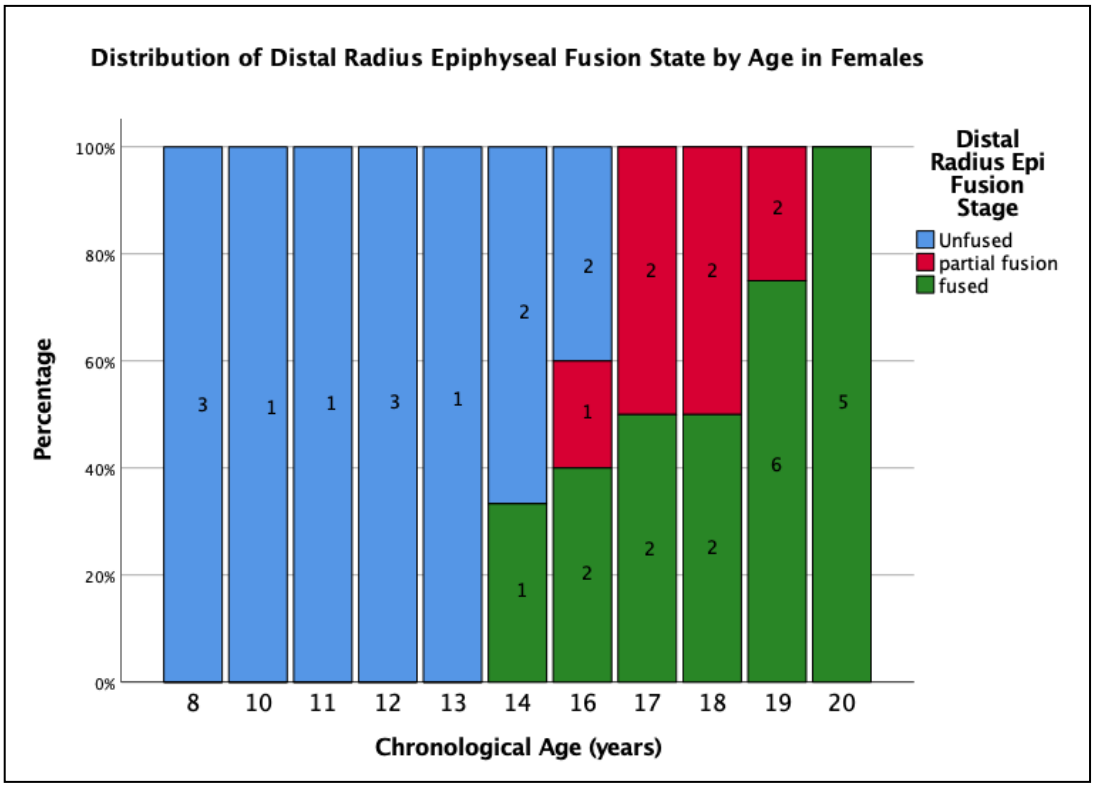

Figure 4.7: Distribution of epiphyseal fusion stages of the distal radius in females between the ages of 8 and 20 years.

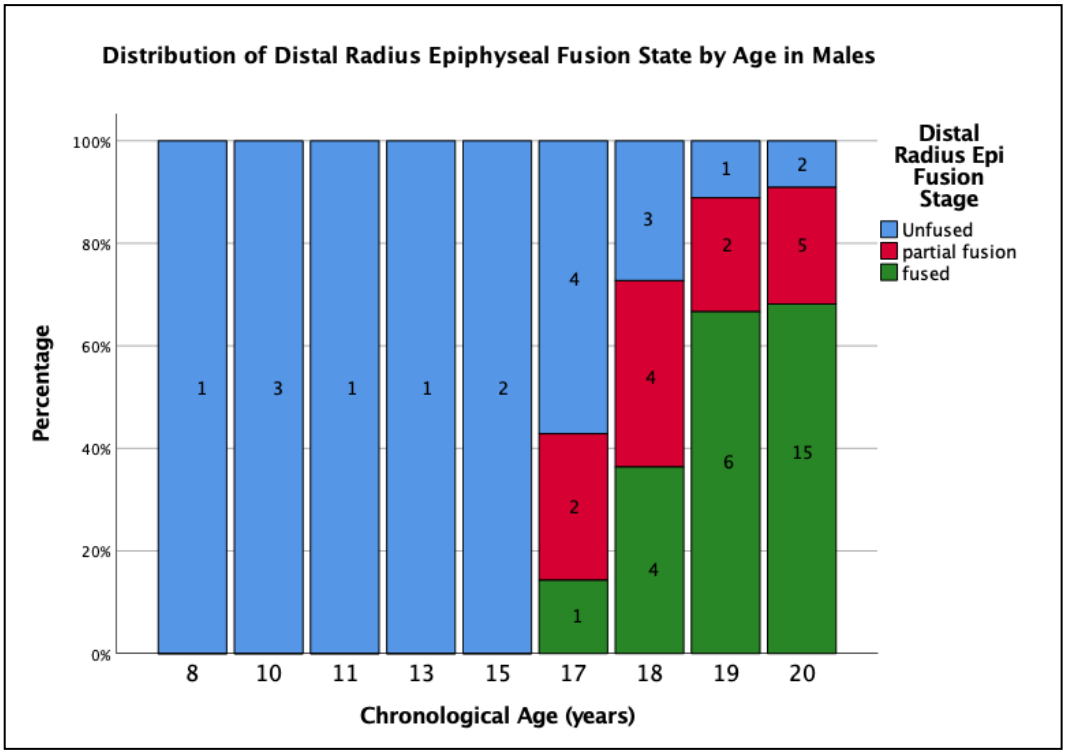

Figure 4.8: Distribution of epiphyseal fusion stages of the distal radius in males between the ages of 8 and 20 years.

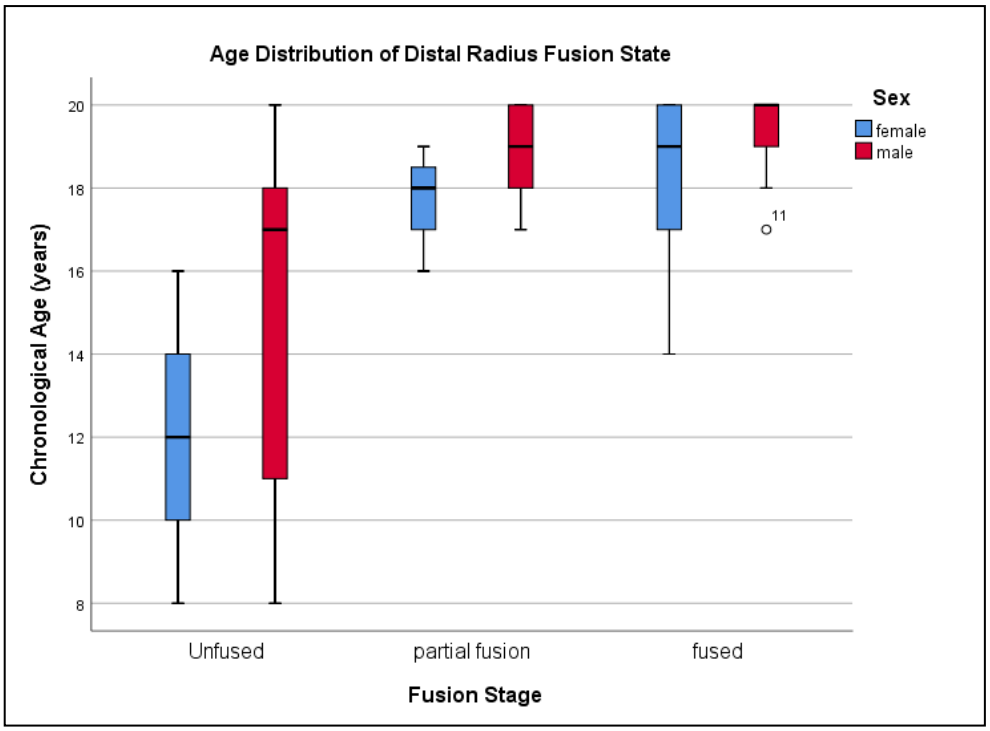

Figure 4.9: Age distribution of epiphyseal fusion stages of the distal radius between males and females

Developmental stages of the iliac crest ossification and fusion in females showed a gradual transition between fusion states. Iliac crests that were not ossified or not present were the dominant developmental state from 8 to 14 years of age, while partial fusion became more common between 16 and 18 years (Figure 4.10). Only one instance of an unfused but ossified iliac crest was seen in a 16 year old. A completely fused iliac crest epiphysis became the common fusion state by 19 years of age. Figure 4.12 demonstrates that the 14-year-old female with a partially fused iliac crest is an outlier, thereby demonstrating early epiphyseal fusion compared to the other females in this sample. Non-ossified or absent iliac crests are the dominant fusion state in males between 8 and 15 years of age, although males with unfused, un-ossified crests can be seen as late at 19 years old (Figure 4.11). Partial fusion of the iliac crest became more common at approximately 18 years of age, while partial and complete epiphyseal fusion were relatively equal between 19 and 20 years of age. While partial fusion of the iliac crest was seen between 16 and 19 years of age, the majority of females in the sample with partial fusion are between 17 and 18 years old (Figure 4.12). In contrast, the

majority of males with partial epiphyseal fusion were between 18 and 20 years, despite the range being between 17 and 20 years. The age range of full fusion of the iliac crest is the most comparable between males and females of all four pubertal traits, with the majority of both males and females being 19-20 years old (Figure 4.12).

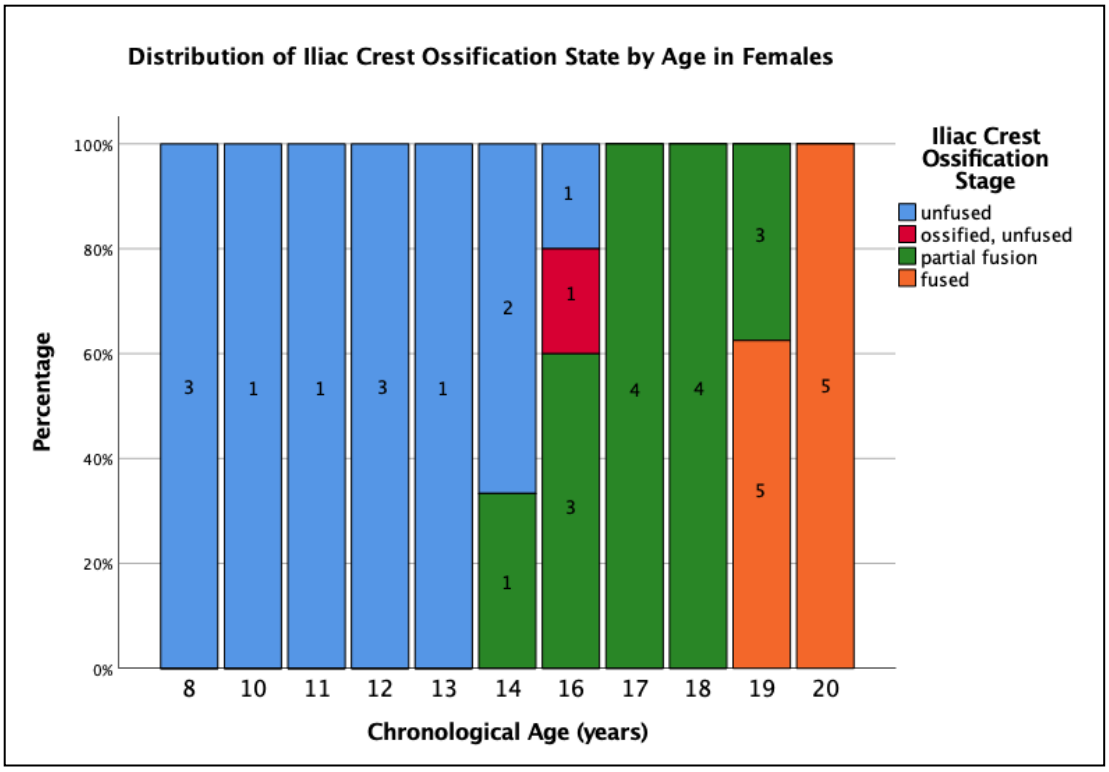

Figure 4.10: Distribution of iliac crest ossification and fusion state in females between the ages of 8 and 20 years.

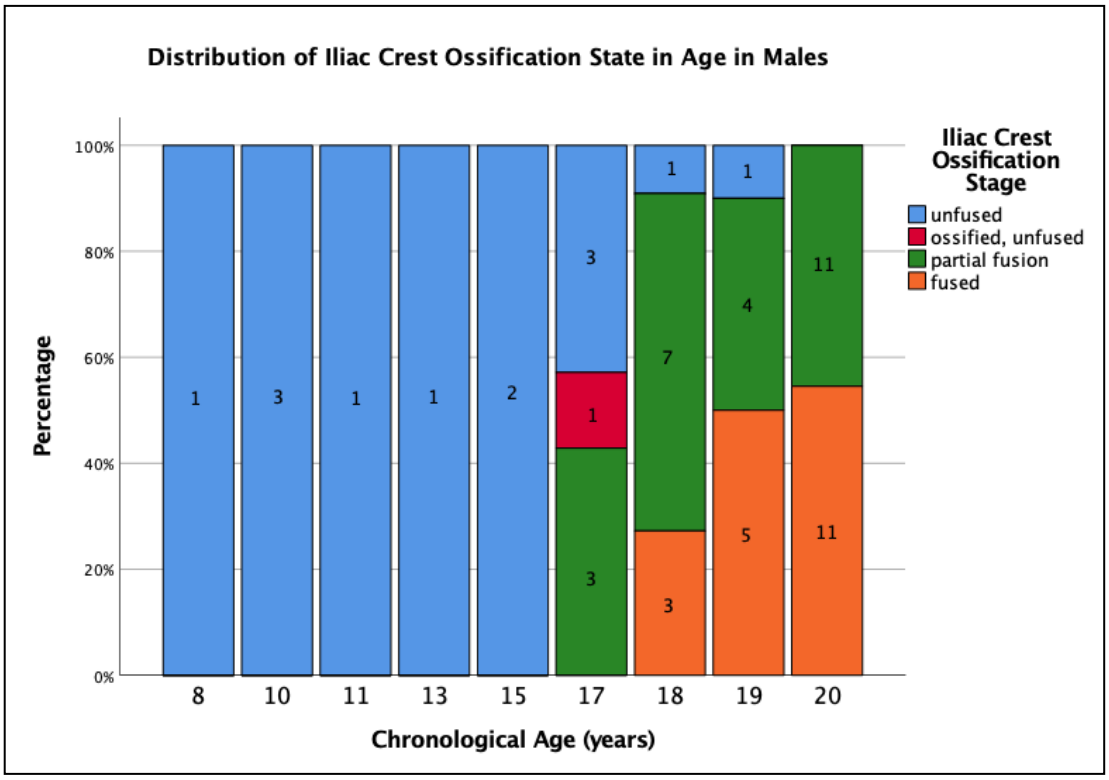

Figure 4.11: Distribution of iliac crest ossification and fusion state in males between the ages of 8 and 20 years.

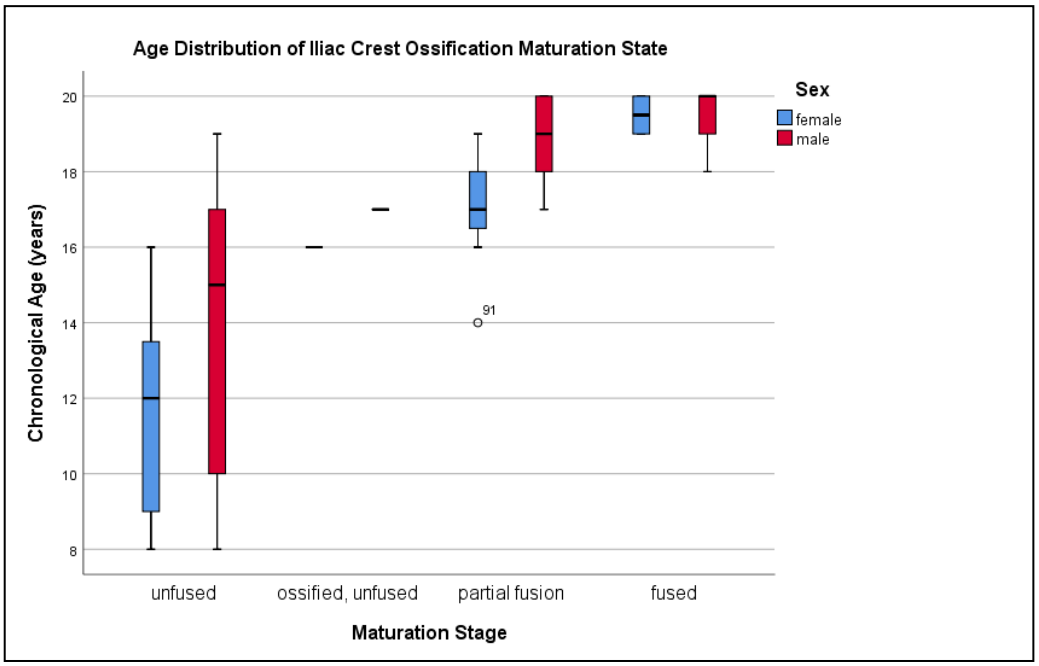

Figure 4.12: Age distribution of iliac crest epiphysis ossification and fusion between males and females

Based on the developmental stages of the four pubertal traits above, a final assessment of puberty stage for each individual was conducted. The majority of adolescents examined were in the deceleration to post-puberty stages (n=78). Ten adolescents were in the acceleration and around peak height velocity (PHV) stages (n = 5 per stages) and eight were in the pre-puberty stage (Figure 4.13). A summary of the age ranges, and mean ages, of pubertal stage attainment for females and males is provided in Table 4.28. On average, females reached each pubertal stage one to two years earlier than males with the exception of the pre-puberty stage, where males and females were at this pubertal stage at comparable ages (Figure 4.14). As illustrated in Figure 4.15, there is substantial overlap in age distribution for each pubertal stage between males and females, with the exception of the around PHV and Deceleration stages. The results of an independent samples t-test showed there was a significant difference between females and males in the mean age of attainment of the deceleration stage only (Table 4.29). While a significant difference between females and males in mean age of attainment of the post-puberty stage is also seen (Table 4.29), equal variance is not assumed for this pubertal stage.

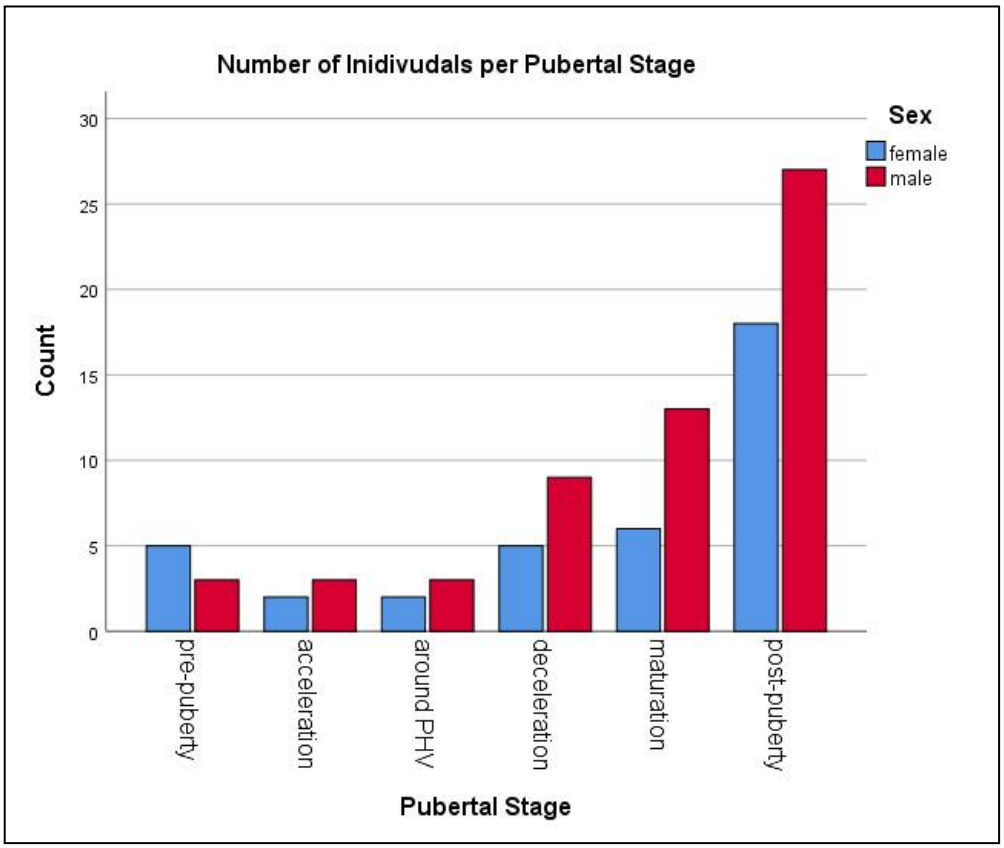

Figure 4.13: Distribution of individuals in each pubertal stage

Table 4.28: Summary of age range and mean age of pubertal stage attainment for females and males

Pubertal Stage	Female Range (yrs)	Female Mean (yrs)	Male Range (yrs)	Male Mean (yrs)
Pre-puberty	8-13	10.0	8-11	9.7
Acceleration	8-12	10.0	10-17	12.3
Around PHV	12	12.0	13-15	14.3
Deceleration	14-16	15.2	17-19	17.9
Maturation	17-19	18.0	17-20	19.0
Post-puberty	14-20	18.3	18-20	19.3

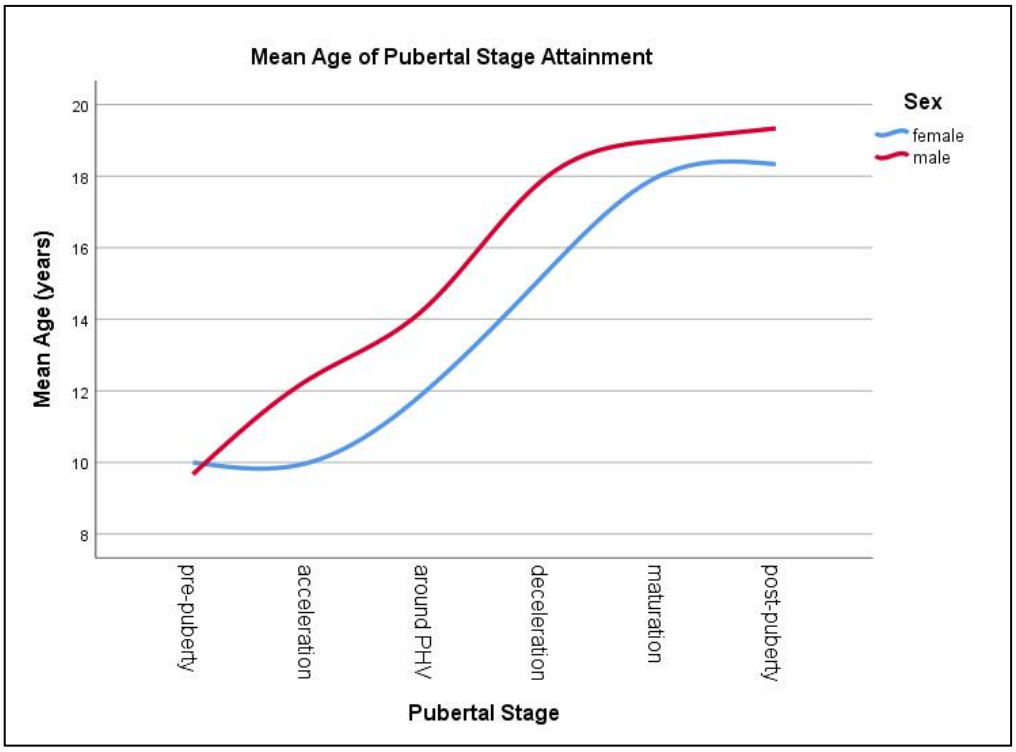

Figure 4.14: Mean age of attainment of pubertal stage for males and females

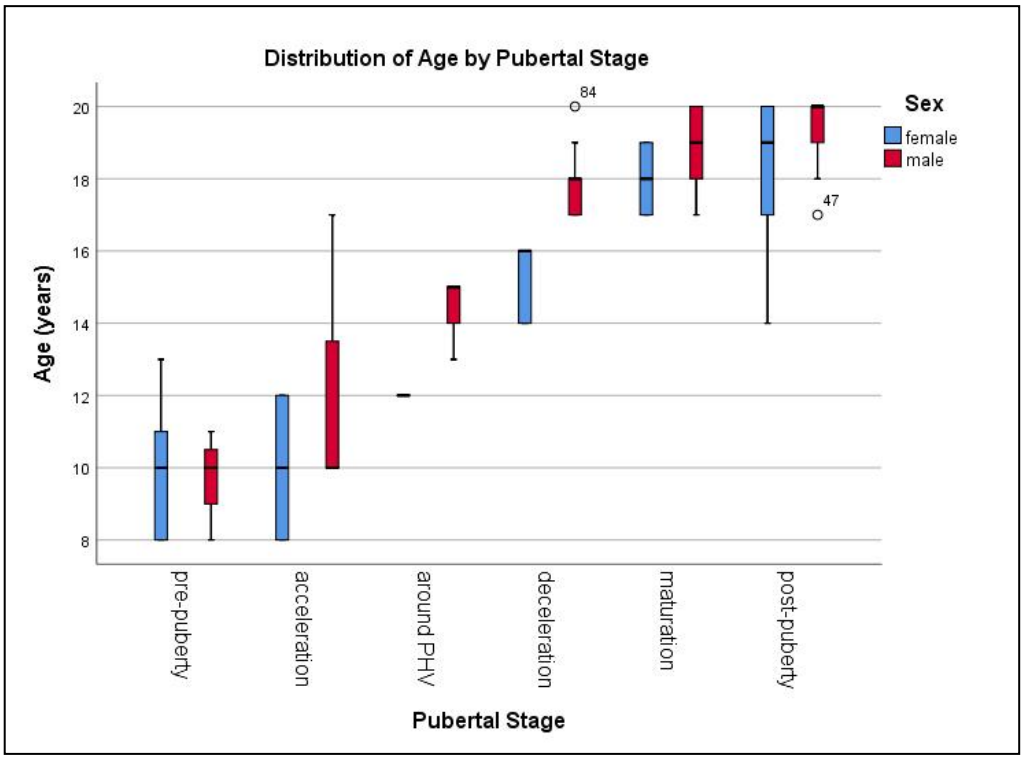

Figure 4.15: Ages distribution of each pubertal stage

Table 4.29: Independent t-test, mean age of pubertal stage attainment between females and males

Pubertal Stage	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Diff	Std. Error Diff	95% Confidence Interval of the Difference	
Pre-Puberty	0.436	0.533	0.235	6	0.822	0.333	1.419	-3.140	3.807
Acceleration	1.224	0.349	-0.694	3	0.537	-2.333	3.361	-13.030	8.363
Around PHV	9.600	0.053	-2.711	3	0.073	-2.333	0.861	-5.072	0.406
Deceleration	0.321	0.581	-4.514	12	**0.001**	-2.689	0.596	-3.987	-1.391
Maturation	1.078	0.314	-1.969	17	0.065	-1.000	0.508	-2.071	0.071
Post-Puberty	8.610	0.005	-2.283	23	**0.032**	-1.000	0.438	-1.906	-0.094

4.3.1 Puberty and Sexual Dimorphism

The accuracy of all 18 morphological pelvic traits and the nine logistic regression models were examined in relation to pubertal stage. In order to account for ontogenic processes associated with sexual dimorphism in the human body, the pre-puberty stage was extended to include individuals as young as 3 years of age. This was done to encompass the entire age range in which the hypothalamus is inactive prior to the initiation of the pubertal growth spurt (Bogin 1999). An infant category, which includes individuals from birth to 2.9 years of age, was added to incorporate the ages at which the hypothalamus is active, although to a lesser extent, during primary sexual dimorphism.

Ilium shape is the only pelvic trait that returned an overall sex classification accuracy at or above 80% for each of the five pubertal stages (acceleration, around PHV, deceleration, maturation, and post-puberty) (Table 4.30). Ilium shape performed better for females than males in the acceleration and deceleration stages, where male accuracy did not reach 80%. While ilium shape performed better on males in the maturation and post-puberty stages, female accuracy still exceeded 80% accuracy. Despite the high overall accuracy for the acceleration and around PHV stages, 80% and 100% respectively, they were not at statistically significant levels. However, a Fisher's exact test demonstrated that the high accuracies observed for correct sex classification in the deceleration (~86%), maturation (~95%), and post-puberty (93%) stages are

statistically significant (χ^2= 7.778, df = 1, p = 0.021; χ^2= 14.702, df = 1, p = 0.001; χ^2= 14.702, df = 1, p < 0.001, respectively).

Table 4.30: Accuracy rates for ilium shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; bolded values indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	16	7	43.8	3	1	33.3	13	6	46.2
	Pre-puberty	17	10	58.8	11	8	72.7	6	2	33.3
	Acceleration	5	4	80	2	2	100	3	2	66.7
	Around PHV	5	5	100	2	2	100	3	3	100
	Deceleration	14	12	**85.7**	5	5	100	9	7	77.8
	Maturation	19	18	**94.7**	6	5	83.3	13	13	100
	Post-puberty	45	42	**93.3**	18	16	88.9	27	26	96.3

Three pelvic traits had overall accuracies meeting or exceeding the 80% threshold in four pubertal stages. Pelvic inlet shape demonstrated high overall accuracies for around PHV (100%), deceleration (~85%), maturation (~89%), and post-puberty (~98%) stages. This trait performed well for both males and females for all pubertal stages, although higher male accuracy was observed in the deceleration and maturation stages (Table 4.31). A Fisher's exact test showed that the accuracy for deceleration, maturation, and post-puberty stages were statistically significant, however the around PHV stage was not (Table 4.31).

Pubic bone shape demonstrated high accuracy for the infant (~93%), around PHV (80%), maturation (100%), and post-puberty (~96%) stages. Despite the overall accuracies meeting or exceeding the necessary threshold, pubic bone shape performed exceptionally well for males but not for females in the infant and around PHV stages (Table 4.32) and correct sex classification for these two stages were not statistically significant. Pubic bone shape performed comparably well for males and females in the maturation and post-puberty stages, and Fisher's exact tests showed that overall sex classification accuracy for these pubertal stages was statistically significant (Table 4.32).

Greater sciatic notch angle had high overall accuracy for the acceleration (80%), around PHV (80%), maturation (84%), and post-puberty (91%) stages. The accuracy of this trait was much higher in females compared to males in the acceleration and around

PHV stages, where male accuracy failed to meet the necessary 80% (Table 4.33). Moreover, despite the high overall accuracy in these two pubertal stages, they were not statistically significant. The greater sciatic notch angle did perform comparably well for males and females in the maturation stage, and performed better in females than males in the post-puberty stage. In both instances, the accuracy for males and females exceeded 80%. A Fisher's exact test showed that the accuracy achieved in the maturation and post-puberty stages was statistically significant (Table 4.33).

Table 4.31: Accuracy rates for pelvic inlet shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	-	-	-	-	-	-	-	-	-
	Pre-puberty	-	-	-	-	-	-	-	-	-
	Acceleration	-	-	-	-	-	-	-	-	-
	Around PHV	2	2	100	1	1	100	1	1	100
	Deceleration	13	11	84.6[a]	5	4	80	8	7	87.5
	Maturation	18	16	88.9[b]	5	4	80	13	12	92.3
	Post-puberty	44	43	97.7[c]	17	17	100	27	26	96.3

[a] χ^2= 5.923, df=1, p = 0.032; [b] χ^2= 9.411, df=1, p = 0.008; [c] χ^2= 40.016, df=1, p < 0.001

Table 4.32: Accuracy rates for pubic bone shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	**Infant**	14	13	92.9	2	1	50	12	12	100
	Pre-puberty	18	6	33.3	12	0	0	6	6	100
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	4	80	2	1	50	3	3	100
	Deceleration	14	11	78.6	5	3	60	9	8	88.9
	Maturation	19	19	100[a]	6	6	100	13	13	100
	Post-puberty	44	42	95.5[b]	18	17	94.4	26	25	96.2

[a] χ^2= 19.000, df = 1, p<0.001; [b] χ^2= 36.115, df = 1, p<0.001

Table 4.33: Accuracy rates for greater sciatic notch angle by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	17	6	35.3	3	3	100	14	3	21.4
	Pre-puberty	17	12	70.6	12	11	91.7	5	1	20
	Acceleration	5	4	80	2	2	100	3	2	66.7
	Around PHV	5	4	80	2	2	100	3	2	66.7
	Deceleration	14	10	71.4	5	5	100	9	5	55.6
	Maturation	19	16	84.2[a]	6	5	83.3	13	11	84.6
	Post-puberty	45	41	91.1[b]	18	17	94.4	27	24	88.9

[a] $\chi^2 = 8.146$, df = 1, p = 0.010; [b] $\chi^2 = 30.375$, df = 1, p<0.001

Seven pelvic traits had accuracies that met or exceeded 80% accuracy in three pubertal stages. Three of those traits (true pelvis size & shape, obturator foramen, and preauricular sulcus) had high accuracy rates for the around PHV, maturation, and post-puberty stages. True pelvis size & shape performed exceptionally well for both males and females in the three pubertal stages. A Fisher's exact test showed that despite the high accuracy in the around PHV stage, it was not statistically significant, but the accuracies for maturation and post-puberty stages were (Table 4.34). Obturator foramen shape performed comparably well for males and females in the around PHV and maturation stages, but this trait performed much better in males than females in the post-puberty stage. Despite the variable accuracy between the sexes in the post-puberty stage, a Fisher's exact test showed that the overall accuracies in the maturation and post-puberty stages were statistically significant (Table 4.35) when using the obturator foramen shape. When examining the results for preauricular sulcus closer, this trait performed exceptionally well for males in the around PHV and maturation categories, but not for females. The preauricular sulcus still performed better for males than females in the post-puberty stage, but accuracy for females surpassed 80%. Moreover, a Fisher's exact test showed that the results attained for the maturation and post-puberty stages were statistically significant (Table 4.36).

Table 4.34: Accuracy rates for true pelvis size & shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	-	-	-	-	-	-	-	-	-
	Pre-puberty	-	-	-	-	-	-	-	-	-
	Acceleration	-	-	-	-	-	-	-	-	-
	Around PHV	2	2	100	1	1	100	1	1	100
	Deceleration	13	8	61.5	5	2	40	8	6	75
	Maturation	18	18	100 [a]	5	5	100	13	13	100
	Post-puberty	45	43	95.6 [b]	18	17	94.4	27	26	96.3

[a] χ^2= 18.00, df = 1, p<0.001; [b] χ^2= 37.052, df = 1, p<0.001

Table 4.35: Accuracy rates for obturator foramen shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	-	-	-	-	-	-	-	-	-
	Pre-puberty	18	6	33.3	12	0	0	6	6	100
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	5	100	2	2	100	3	3	100
	Deceleration	14	10	71.4	5	3	60	9	7	77.8
	Maturation	19	17	89.5[a]	6	5	83.3	13	12	92.3
	Post-puberty	45	39	86.7[b]	18	14	77.8	27	25	92.6

[a] χ^2= 10.871, df = 1, p = 0.003; [b] χ^2= 26.548, df = 1, p<0.001

Table 4.36: Accuracy rates for preauricular sulcus by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	13	11	84.6	1	0	0	12	11	91.7
	Pre-puberty	17	7	41.2	12	2	16.7	5	5	100
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	4	80	2	1	50	3	3	100
	Deceleration	14	10	71.4	5	1	20	9	9	100
	Maturation	19	17	89.5[a]	6	4	66.7	13	13	100
	Post-puberty	45	41	91.1[b]	18	15	83.3	27	26	96.3

[a] χ^2= 10.978, df = 1, p =0.004; [b] χ^2= 29.887, df = 1, p<0.001

Acetabulum size & orientation, auricular surface height, subpubic concavity, and ventral arc had 80% accuracy or above in the deceleration, maturation, and post-puberty stages. Acetabulum size & orientation performed well for both males and females in the three pubertal stages, where accuracy meets and exceeds 80% (Table 4.36). Auricular surface height performed well on males in the deceleration and maturation categories, but female accuracy did not meet the required 80% minimum for either pubertal stage. Auricular surface height did, however, perform comparably well for males and females in the post-puberty stage (Table 4.37). Conversely, subpubic concavity was comparably effective for both males and females in the deceleration and maturation stages. Female accuracy in the post-puberty stage, however, failed to meet 80% whereas males had 100% accuracy (Table 4.38). Despite having overall accuracies above 80% in the deceleration, maturation, and post-puberty, the ventral arc performed exceptionally poor for females in all three pubertal stages, but exceptionally well for males (Table 4.39). Thus, the overall high accuracy in the deceleration, maturation, and post-puberty stages are the result of high accuracy in male classification. Fisher's exact tests showed that the results attained in the deceleration, maturation, and post-puberty stages for acetabulum size & orientation, auricular surface height, and subpubic concavity were statistically significant (see Tables 4.37-4.39). Despite the poor performance of the ventral arc on females in the maturation and post-puberty stages, a Fisher's exact test showed that the accuracies observed in the two pubertal stages were statistically significant (Table 4.40).

Table 4.37: Accuracy rates for acetabulum size & orientation by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	-	-	-	-	-	-	-	-	-
	Pre-puberty	3	2	66.7	1	1	100	2	1	50
	Acceleration	4	2	50	2	1	50	2	1	50
	Around PHV	5	3	60	2	1	50	3	2	66.7
	Deceleration	14	12	85.7[a]	5	4	80	9	8	88.9
	Maturation	19	18	94.7[b]	6	5	83.3	13	13	100
	Post-puberty	45	44	97.8[c]	18	18	100	27	26	96.3

[a] χ^2= 6.644, df = 1, p = 0.023; [b] χ^2= 14.702, df = 1, p = 0.001; [c] χ^2= 41.053, df = 1, p<0.001

Table 4.38: Accuracy rates for auricular surface height by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	14	12	85.7	2	0	0	12	12	100
	Pre-puberty	17	5	29.4	12	0	0	5	5	100
	Acceleration	5	4	80	2	1	50	3	3	100
	Around PHV	4	3	75	1	0	0	3	3	100
	Deceleration	13	12	92.3[a]	4	3	75	9	9	100
	Maturation	19	15	79[b]	6	4	66.7	13	11	84.6
	Post-puberty	44	40	90.9[c]	17	14	82.4	27	26	96.3

[a] χ^2= 8.775, df = 1, p = 0.014; [b] χ^2= 4.997 , df = 1, p =0.046; [c] χ^2= 28.719, df = 1, p <0.001

Table 4.39: Accuracy rates for subpubic concavity by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	-	-	-	-	-	-	-	-	-
	Pre-puberty	10	9	90	-	-	-	10	9	90
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	3	60	2	0	0	3	3	100
	Deceleration	14	13	92.9[a]	5	5	100	9	8	88.9
	Maturation	19	18	94.7[b]	6	5	83.3	13	13	100
	Post-puberty	45	41	91.1[c]	18	14	77.8	27	27	100

[a] χ^2= 10.370, df = 1, p = 0.003; [b] χ^2= 14.702, df = 1, p = 0.001; [c] χ^2= 30.484, df = 1, p<0.001

Table 4.40: Accuracy rates for ventral arc by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	12	10	83.3	2	0	0	10	10	100
	Pre-puberty	14	4	28.6	10	0	0	4	4	100
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	3	60	2	0	0	3	3	100
	Deceleration	12	10	83.3	3	1	33.3	9	9	100
	Maturation	17	15	88.2[a]	4	2	50	13	13	100
	Post-puberty	43	36	83.7[b]	17	10	58.8	26	26	100

[a] χ^2= 7.367, df = 1, p = 0.044; [b] χ^2= 19.929, df = 1, p<0.001

Ischiopubic ramus ridge, greater sciatic notch size & shape, and sacrum shape, had overall accuracy that met or exceeded 80% in both the maturation and post-puberty stages of pubertal growth. The ischiopubic ramus ridge performed better on males

compared to females in both the maturation and post-puberty stages, with only the maturation stage having an accuracy exceeding 80% for both sexes (Table 4.41). The greater sciatic notch size & shape almost consistently performed substantially better for females than males, with the exception of the around PHV and maturation stages (Table 4.42). The post-puberty stage had both females and males exceeding 80% accuracy for greater sciatic notch size & shape, whereas females fell well-below 80% for the maturation stage. Despite overall accuracies exceeding 80% in both the maturation and post-puberty stages for sacrum shape, the trait performed very well in males but failed to meet 80% in females (Table 4.43). Fisher's exact test showed that the high accuracies seen in the maturation and post-puberty stages for ischiopubic ramus and greater sciatic notch size & shape were statistically significant, however only the post-puberty stage for sacrum shape was significant (Tables 4.41-4.43).

Table 4.41: Accuracy rates for ischiopubic ramus ridge by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	11	7	63.6	2	0	0	9	7	77.8
	Pre-puberty	17	5	29.4	12	0	0	5	5	100
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	3	60	2	0	0	3	3	100
	Deceleration	13	10	76.9	5	3	60	8	7	87.5
	Maturation	18	16	88.9[a]	6	5	83.3	12	11	91.7
	Post-puberty	44	35	79.5[b]	18	11	61.1	26	24	92.3

[a] χ^2= 10.125, df = 1, p = 0.004; [b] χ^2= 17.789, df = 1, p <0.001

Table 4.42: Accuracy rates for greater sciatic notch size & shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	18	7	38.9	3	3	100	15	4	26.7
	Pre-puberty	17	13	76.5	12	12	100	5	1	20
	Acceleration	5	3	60	2	2	100	3	1	33.3
	Around PHV	5	3	60	2	1	50	3	2	66.7
	Deceleration	14	10	71.4	5	4	80	9	6	66.7
	Maturation	19	17	89.5[a]	6	4	66.7	13	13	100
	Post-puberty	45	40	88.9[b]	18	17	94.4	27	23	85.2

[a] χ^2= 10.978, df = 1, p = 0.004; [b] χ^2= 27.515, df = 1, p<0.001

Table 4.43: Accuracy rates for sacrum shape by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	1	1	100	-	-	-	1	1	100
	Pre-puberty	-	-	-	-	-	-	-	-	-
	Acceleration	3	2	66.7	1	0	0	2	2	100
	Around PHV	3	3	100	1	1	100	2	2	100
	Deceleration	14	9	64.3	5	2	40	9	7	77.8
	Maturation	17	14	82.4	5	3	60	12	11	91.7
	Post-puberty	44	37	84.1[a]	17	13	76.5	27	24	88.9

[a] χ^2= 19.258 , df = 1, p<0.001; maturation: χ^2= 5.236, df = 1, p = 0.53

Three pelvic traits (dorsal pubic pitting, iliac crest curvature, and development of muscle markings) had only one pubertal stage with an overall sex classification accuracy exceeding a minimum of 80%. While the infant category for dorsal pubic pitting had an overall accuracy of approximately 92% (Table 4.44), the high accuracy was not statistically significant. Iliac crest curvature had 100% accuracy for individuals around PHV (Table 4.45), however this high accuracy was not statistically significant. While iliac crest curvature for post-puberty individuals did not meet the 80% accuracy threshold required for this study, a Fisher's exact test showed that the accuracy observed in this pubertal stage (~76%) was statistically significant (χ^2= 12.465, df = 1, p=0.001). The post-puberty stage was the only pubertal stage group that exceeded 80% accuracy for development of muscle markings with an accuracy of approximately 87% (Table 4.46). A Pearson's chi-square test showed that the sex classification accuracy for this pubertal stage was statistically significant (χ^2= 28.902, df = 1, p<0.001). While marginally failing to achieve an 80% minimum, a Pearson's chi-square showed that the sex classification for the maturation stage when using development of muscle markings (79%) was statistically significant (χ^2= 7.892, df = 1, p = 0.005).

Table 4.44: Accuracy rates for dorsal pubic pitting by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	**Infant**	12	11	91.7	1	0	0	11	11	100
	Pre-puberty	18	6	33.3	12	0	0	6	6	100
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	3	60	2	0	0	3	3	100
	Deceleration	14	8	57.1	5	0	0	9	8	88.9
	Maturation	19	14	73.7	6	1	16.7	13	13	100
	Post-puberty	44	27	61.4	17	1	5.9	27	26	96.3

Table 4.45: Accuracy rates for iliac crest curvature by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; * indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	17	3	17.6	3	3	100	14	0	0
	Pre-puberty	17	13	76.5	12	12	100	5	1	20
	Acceleration	5	3	60	2	1	50	3	2	66.7
	Around PHV	5	5	100	2	2	100	3	3	100
	Deceleration	14	9	64.3	5	3	60	9	7	77.8
	Maturation	19	14	73.7	6	4	66.7	13	10	76.9
	Post-puberty	45	34	75.6*	18	15	83.3	27	19	70.4

Table 4.46: Accuracy rates for development of muscle markings by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; * indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	-	-	-	-	-	-	-	-	-
	Pre-puberty	9	5	55.6	6	5	83.3	3	0	0
	Acceleration	3	1	33.3	1	1	100	2	0	0
	Around PHV	5	2	40	2	2	100	3	0	0
	Deceleration	14	8	57.1	5	5	100	9	3	33.3
	Maturation	19	15	79*	6	6	100	13	9	69.2
	Post-puberty	45	39	86.7*	18	17	94.4	27	22	81.5

Arch criterion was the only pelvic trait that did not have a pubertal stage that met or exceeded an 80% minimum accuracy for combined sexes (Table 4.47). The arch criterion, however, consistently performed well on females (accuracy between 80-100%) but very poorly on males (accuracy between 20-74%). A Pearson's chi-square test showed that the sex classification accuracy for the post-puberty stage (~78%) was

statistically significant (χ^2= 16.273, df = 1, p < 0.001) despite not meeting the 80% threshold.

Table 4.47: Accuracy rates for arch criterion by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Asterisk indicates statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	17	6	35.3	3	3	100	14	3	21.4
	Pre-puberty	17	13	76.5	12	12	100	5	1	20
	Acceleration	5	3	60	2	2	100	3	1	33.3
	Around PHV	5	3	60	2	2	100	3	1	33.3
	Deceleration	14	7	50	5	4	80	9	3	33.3
	Maturation	19	13	68.4	6	5	83.3	13	8	61.5
	Post-puberty	45	35	77.8*	18	15	83.3	27	20	74.1

When examining final sex assessment when all 18 pelvic traits are used, four pubertal stages (around PHV, deceleration, maturation, and post-puberty) had accuracies at or above 80%. This observation was seen both when all traits are weighed equally and when juvenile traits were weighed more heavily. As shown in Table 4.48, when all traits were weighed equally, female accuracy failed to meet a minimum of 80% for the around PHV stage only. Moreover, males had higher final sex assessment accuracy in the deceleration and maturation stages compared to females, but correct sex assessment was comparable for both sexes in the post-puberty stage. Generally, overall accuracy in final sex assessment steadily increased from the around PHV stage to the post-puberty stage. A Pearson's chi-square test showed that the accuracy obtained for the deceleration (χ^2= 10.516, df = 1, p = 0.005), maturation (χ^2= 19.000, df = 1, p<0.001), and post-puberty (χ^2= 45.000, df = 1, p<0.001) stages were statistically significant, but not the accuracy achieved for the around PHV stage (Table 4.48). When juvenile traits were weighed more when conducting a final sex assessment, an increase in overall accuracy was seen for the around PHV and deceleration stages (Table 4.49). A minimum of 80% accuracy was exceeded for both females and males in all four pubertal stages. A Pearson chi-square showed that the sex assessment accuracy achieved in the deceleration (χ^2= 10.370, df = 1, p=0.001), maturation (χ^2= 19.000, df = 1, p<0.001), and post-puberty (χ^2= 45.000, df = 1, p<0.001) stages are statistically significant (Table 4.49).

Table 4.48: Accuracy of final sex assessment (all traits weighed equally) by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; bolded accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	18	12	66.7	3	2	66.7	15	10	66.7
	Pre-puberty	18	5	27.8	12	0	0	6	5	83.3
	Acceleration	5	3	60	2	0	0	3	3	100
	Around PHV	5	4	80	2	1	50	3	3	100
	Deceleration	14	12	**85.7**	5	4	80	9	8	88.9
	Maturation	19	18	**94.7**	6	5	83.3	13	13	100
	Post-puberty	45	44	**97.8**	18	18	100	27	26	96.3

Table 4.49: Accuracy of final sex assessment (juvenile traits weighed more) by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; bolded accuracies indicate statistical significance.

Category		N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Pubertal Stage	Infant	18	5	27.8	3	3	100	15	2	13.3
	Pre-puberty	18	14	77.8	12	12	100	6	2	33.3
	Acceleration	5	3	60	2	2	100	3	1	33.3
	Around PHV	5	5	100	2	2	100	3	3	100
	Deceleration	14	13	**92.9**	5	5	100	9	8	88.9
	Maturation	19	18	**94.7**	6	5	83.3	13	13	100
	Post-puberty	45	44	**97.8**	18	18	100	27	26	96.3

Generally, when examining overall accuracy of the logistic regression models proposed by Albanese (2003), no clear patterns were observed with the progression of puberty stages. Accuracy tended to fluctuate as pubertal stages progress throughout growth. Two models (Model 3 and 8), however, had accuracies exceeding 80% in four pubertal stages: around PHV, deceleration, maturation, and post-puberty (Table 4.50). When examined closer, Model 3 and 8 performed comparably well for males and females in the around PHV and deceleration stages but these two models performed substantially better on males than on females in the maturation and post-puberty stages. Despite the male biases observed in the maturation and post-puberty stages, Fisher's exact tests showed that the accuracies observed for Model 3 in the deceleration, maturation, and post-puberty stages were statistically significant (Table 4.50). For Model 8, a Fisher's exact test showed that only the results for the deceleration and post-puberty stages were significant.

Table 4.50: Accuracy of logistical regression models 3 & 8 by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Logit. Model	Age Category	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Model 3	**Around PHV**	2	2	100	1	1	100	1	1	100
	Deceleration	13	11	84.6[a]	5	4	80	8	7	87.5
	Maturation	15	14	93.3[b]	4	3	75	11	11	100
	Post-puberty	41	35	85.4[c]	16	11	68.8	25	24	96
Model 8	**Around PHV**	2	2	100	1	1	100	1	1	100
	Deceleration	13	11	84.6[d]	5	4	80	8	7	87.5
	Maturation	17	14	82.4	5	2	40	12	12	100
	Post-puberty	41	36	87.8[e]	16	12	75	25	24	96

[a] χ^2= 5.923, df = 1, p = 0.32; [b] χ^2= 10.313, df = 1, p = 0.009; [c] χ^2= 19.758, df = 1, p<0.001; [d] χ^2= 5.923, df = 1, p = 0.032; [e] χ^2= 22.712, df = 1, p<0.001

Five logistic models had three pubertal stages with an accuracy exceeding the 80% minimum required for this research. Models 20, 4, and 10 had accuracies above 80% for the deceleration, maturation, and post-puberty categories. Models 20 and 4 performed comparably well for males and females in both the deceleration and post-puberty stages but performed markedly better in males in the maturation stage, where female accuracy fell well below 80%. Model 10 performed markedly better on females in the deceleration stage with males failing to achieve the 80% threshold, while the opposite was true for the maturation and post-puberty stages. Fisher's exact test showed that the results achieved in models 20 & 4 for the deceleration, maturation, and post-puberty stages are statistically significant, while model 10 only has statistically significant results for the deceleration and post-puberty stages (Table 4.51).

Models 1 and 11, had sex classification accuracies above 80% for the around PHV, maturation, and post-puberty stages. Both models performed comparably well for males and females in the around PHV stage, however both Model 1 & 11 performed substantially better on males than females in the maturation and post-puberty stages. In both pubertal stages for the two logistic regression models, female accuracy fell well below the required threshold. A Fisher's exact test showed that the results obtained using Model 1 for the post-puberty stage are statistically significant, while the results for

the maturation and post-puberty stages obtained for Model 11 are significant (Table 4.51).

Logistic regression Model 2 only had two pubertal stages, maturation and post-puberty, with an overall sex assessment accuracy surpassing 80%. In both pubertal categories, model 2 performed substantially better on males compared to females, where female accuracy failed to meet 80%. Despite the disparate efficacy between the sexes, a Fisher's exact test showed that the results for the maturation and post-puberty stages using model 2 are statistically significant (Table 4.52). Lastly, only the maturation stage had an overall accuracy surpassing 80% when using Model 5, although the post-puberty stage only marginally failed to meet the 80% minimum. Model 5 performed better for males compared to females in the maturation stage, but the accuracy for females still met the 80% threshold. A Fisher's exact test showed that the results obtained for the maturation stage is statistically significant, and despite failing to meet 80% accuracy, the results for the post-puberty stage is also statistically significant (Table 4.52).

Table 4.51: Accuracy of logistical regression models 1, 4, 10, 11, and 20 for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Logit. Model	Age Category	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Model 20	Around PHV	2	1	50	1	1	100	1	0	0
	Deceleration	13	12	92.3[a]	5	5	100	8	7	87.5
	Maturation	17	15	88.2[b]	6	4	66.7	11	11	100
	Post-puberty	43	42	97.7[c]	18	17	94.4	25	25	100
Model 4	Around PHV	3	2	66.7	1	1	100	2	1	50
	Deceleration	13	11	84.6[d]	5	4	80	8	7	87.5
	Maturation	17	15	88.2[e]	5	3	60	12	12	100
	Post-puberty	41	35	85.4[f]	16	13	81.3	25	22	88
Model 10	Around PHV	3	2	66.7	1	1	100	2	1	50
	Deceleration	13	11	84.6[g]	5	5	100	8	6	75
	Maturation	16	13	81.3	5	2	40	11	11	100
	Post-puberty	41	35	85.4[h]	16	12	75	25	23	92
Model 1	**Around PHV**	2	2	100	1	1	100	1	1	100
	Deceleration	13	10	76.9	5	4	80	8	6	75
	Maturation	16	13	81.3	5	2	40	11	11	100
	Post-puberty	41	36	87.8[i]	16	12	75	25	24	96
Model 11	**Around PHV**	3	3	100	1	1	100	2	2	100
	Deceleration	13	10	76.9	5	4	80	8	6	75
	Maturation	16	14	87.5[j]	5	3	60	11	11	100
	Post-puberty	41	35	85.4[k]	16	11	68.8	25	23	92

[a] χ^2= 9.479, df = 1, p = 0.005; [b] χ^2= 9.590, df = 1, p = 0.006; [c] χ^2= 39.049, df = 1, p<0.001; [d] χ^2= 5.923, df = 1, p =0.032; [e] χ^2= 8.743, df = 1, p = 0.015; [f] χ^2= 19.662, df = 1, p<0.001; [g] χ^2= 6.964, df = 1, p = 0.016; [h] χ^2= 19.476, df = 1, p<0.001; [i] χ^2= 22.712, df = 1, p<0.001; [j] χ^2= 8.123, df = 1, p = 0.018; [k] χ^2= 16.628, df = 1, p<0.001

Table 4.52: Accuracy of logistical regression models 2 & 5 by pubertal stage for combined and separate sexes. N_F = total females, n_F = correctly identified females, N_M = total males, n_M = correctly identified males. Bolded pubertal stages indicate those with accuracy at or above 80%; superscript accuracies indicate statistical significance.

Logit. Model	Age Category	N	Correct Class.	%	N_F	n_F	%	N_M	n_M	%
Model 2	Pre-puberty	2	1	50	2	1	50	-	-	-
	Around PHV	3	2	66.7	1	0	0	2	2	100
	Deceleration	13	10	76.9	5	3	60	8	7	87.5
	Maturation	16	14	87.5[a]	5	3	60	11	11	100
	Post-puberty	40	35	87.5[b]	15	11	73.3	25	24	96
Model 5	Around PHV	3	2	66.7	1	1	100	2	1	50
	Deceleration	13	9	69.2	5	4	80	8	5	62.5
	Maturation	17	16	94.1[c]	5	4	80	12	12	100
	Post-puberty	41	32	78.1[d]	16	12	75	25	20	80

[a] χ^2= 8.123, df=1, p = 0.018; [b] χ^2= 21.460, df=1, p<0.001; [c] χ^2=12.554, df=1, p = 0.002; [d] χ^2= 12.159, df=1, p = 0.001

Chapter 5 - Discussion

5.1 Sexual Dimorphism in Subadults

The results of this study provide evidence that six pelvic traits (pelvic inlet shape, true pelvis size & shape, acetabulum size & orientation, pubic bone shape, ilium shape, and subpubic concavity) can obtain an accuracy of 80% and above for determining the sex of subadults (in its broadest definition). These six traits also produced accuracies higher than those obtained when using the complete list of 18 traits (79.7%). Four of these traits (pelvic inlet shape, true pelvis size & shape, acetabulum size & orientation, and subpubic concavity), however, require either the full fusion of the pelvic bones or the fusion of the pubis to the ischium. These traits with high accuracy, therefore, can only be applied to subadults once fusion of the pelvic elements has occurred, which typically occurs at 5-8 years for the fusion of the ischiopubic ramus and the subsequent fusion of the ilium at the acetabulum, which occurs at 11-16 years in females and 14-17 years in males (Scheuer and Black 2000). While an attempt was made to articulate the remains of younger individuals to assess pelvic inlet shape, true pelvis size & shape, acetabulum size & orientation, and subpubic concavity, the true morphology would be distorted since the articulation of the skeletal elements alone would not account for the space occupied by cartilage during life.

Two traits (pelvic inlet shape and true pelvis size & shape) are directly related to the birth canal and have the highest overall accuracies (92.6% and 91.5% respectively) of all 18 pelvic features. Pelvic inlet shape and true pelvis size & shape also demonstrate high accuracy in females and males, 93.3% and 92.2% respectively for pelvic inlet shape and 87.1% and 94.1% for true pelvic size and shape. LaVelle (1995) found that, between 8 and 18 years of age, the female pelvis expands more in the dimensions of the true pelvis than does the male pelvis. The results of this study, therefore, provide evidence that supports LaVelle's observation that pelvic morphology related to the birth canal becomes dimorphic prior to adulthood, as defined by skeletal maturity. The accuracies achieved in this study for both traits are higher than those achieved in adults by Rogers and Saunders (1994), which may contradict the notion that sexual dimorphism should be greater in adults than subadults. Huseynov and

colleagues (2016), however, have shown that the dimensions of the true pelvis expand from puberty until about 25-30 years of age, but that after 40 years of age those dimensions diminish in magnitude. Since the age distribution of Rogers and Saunders' (1994) sample was not disclosed, it might be possible that the lower accuracy in their study may be the result of older females in the sample with a reduced level of true pelvis dimensions. Although pelvic inlet shape and true pelvis size & shape have the highest accuracies in this study, problems with these traits arise with the level of observer error. Both traits have intra-observer errors greater than 10%, where true pelvis size & shape shows error of 10.7% and pelvic inlet shape 12.5%. The level of agreement between observations, however, is substantial for true pelvis size & shape and pelvic inlet shape (κ_c =0.768 and 0.741, respectively). Thus, while there was some error in repeated scores between assessments, the scores that were similar were not the result of chance. The higher level of observer error for pelvic inlet shape compared to true pelvis size & shape mimics the pattern of observer error found by Rogers and Saunders (1994). As a result, the pelvis inlet shape and true pelvis size & shape can be cautiously applied to subadults since these traits show high accuracy but questionable precision.

Acetabulum size and orientation has the third highest accuracy of all pelvic traits (90.3%) and is comparably effective for both female and male sex classification, 88.9% and 91.2% respectively. The accuracy achieved in this study is also quite comparable to the accuracy achieved for adults by Rogers and Saunders (1994) [91.7%]. Thus, this study provides evidence that the acetabulum becomes dimorphic prior to adulthood and achieves a level of dimorphism that is quite comparable to that of adults. Although providing promising results in terms of accuracy, acetabulum size and shape fails to provide precise assessments, with an intra-observer error of 13.8%. The agreement that is observed for this trait (κ_c= 0.699), however, is considered substantial, albeit at the lower end of the range (0.61-0.80) as described by Landis and Koch (1977). Therefore, while there was error in the ability to repeat scores between assessments, the scores that were repeatable were not the result of chance. Resultantly, the high level of error observed might offset the high accuracy of this trait, thereby making the acetabulum size and orientation less than ideal for sex determination of all subadults if used alone.

Subpubic concavity is the last trait dependent on fusion of skeletal elements that met the threshold of accuracy set out in this study. While meeting the minimum accuracy of 80%, subpubic concavity performs substantially better for subadult males (98.6%) than females (54.2%). The male bias this trait demonstrates renders this trait's efficacy on subadults questionable. Moreover, subpubic concavity fails to meet the acceptable level of observer error, albeit not by a substantial amount (1.2%). The combination of substantial male bias and problematic observer error would suggest that this trait, alone, might not be the most adequate for subadult sex assessment. To my knowledge, only one study has examined the use of the subpubic concavity for subadult sex determination. Klales and Burns (2017) examined the application of this trait to subadults using a modified scoring system that used an ordinal scale of pronounced concavity (1 or females), straight (2 or indeterminate), and convex (3 or male). In their study, the authors achieved an overall accuracy of 71.95% for subadults between 1 and 20.5 years of age, where accuracy increased with age (Klales and Burns 2017). A higher accuracy was achieved for subadults in the present study using the binary approach of scoring set forth by Phenice (1969) than the accuracy achieved by Klales and Burns' (2017) ordinal approach. Two factors could explain the difference in accuracy achieved in this study compared to Klales and Burns (2017). The first is that Klales and Burns (2017) tested their ordinal scale of assessing the subpubic concavity using radiographs. The use of radiographs, however, is not necessarily comparable to observations on dry bones (DiGangi and Moore 2013: 107). Given that this study was done on dry bone, it may explain the difference in overall accuracy between the two studies. Second, the use of ordinal scales may prove to be problematic for the purpose of sex determination. While the use of an ordinal scale includes the full range of variation of a trait's expression (Klales et al. 2012), which is important, when a binary method of scoring is used, investigators must choose which of the two forms of expression best resembles the pelvis under analysis. Having to choose one of two options for a trait, while not incorporating the range of human variation, may result in higher sex classification and minimizes the number of indeterminate assessments.

Despite pubic bone shape achieving an overall correct sex classification accuracy (82.8%) that is higher than the accuracy produced using all 18 traits, this trait

should be used with caution. The overall accuracy achieved when pubic bone shape is applied to subadults is 3.4% lower than the accuracy achieved when this trait is applied to adults, as per Rogers and Saunders (1994). In addition to the promising overall accuracy, the pubic bone shape is a trait that does not require the fusion to any other pelvic element and can therefore be assessed when found unfused. This trait also shows an acceptable level of precision, as demonstrated by its low intra-observer error (6.5%) and the high level of agreement that is "near perfect" (κ_c = 0.828). While the high overall accuracy and low level of observer error indicates that the pubic bone shape may be a promising trait for subadult sex assessment, this trait is much more effective for male sex classification (97.3%) than female sex classification (61.2%). Due to the substantial male bias exhibited by the pubic bone shape, when applied to unidentified skeletal remains, there is a strong possibility of misclassifying a subadult female given the poor accuracy of correct female classification.

There has been substantial research on the sexual dimorphism of separate features on the ilium, such as the auricular surface, iliac crest curvature, and great sciatic notch (Weaver 1980; Schutkowski 1991; Mittler and Sheridan 1992; Sutter 2003; Vlak et al. 2008; Wilson et al. 2011; Wilson et al. 2015; Luna et al. 2017), but the general shape of the ilium has not been examined. In this study, the ilium shape provided an overall accuracy of 80.8%, thereby meeting the minimum threshold set out for this research. The overall accuracy achieved for subadults is only approximately 3% lower than what is found for adults (Rogers and Saunders 1994). There is, however, a slight female bias in which females have a higher accuracy of correct sex classification (84%) compared to males (78.7%). The male accuracy, however, only marginally fails to meet the 80% minimum. Of greater concern is the level of intra-observer error determined for this trait. In subadults, ilium shape has a high level of intra-observer error (24.7%) and the intra-observer agreement observed is moderate at best (κ_c =0.505). The level of error found in this study contrasts substantially to the level of error found in adults, which was determined to be approximately 6.4% (Rogers and Saunders 1994). Thus, while ilium shape is a relatively accurate trait, the ability to consistently obtain the same sex assessment is quite poor for subadults and this trait may not be applicable, on its own, to all subadults.

Dorsal pubic pitting was determined to be the least accurate trait for subadult sex assessment, which is not unexpected. Dorsal pubic pitting is known to be an unreliable trait and not accurate, even for adult sex determination. In their study, Rogers and Saunders (1994) found that dorsal pubic pitting had the second lowest correct sex classification, with an accuracy of only 35.7%. A more recent meta-analysis on four studies conducted by McFadden and Oxenham (2018) showed that, when pubic pitting was used as a predictor of sex, accuracy ranged from only 67-71%. Despite the low accuracy of dorsal pubic pitting, this study showed that intra-observer error is low (2.2%) but the degree of observer agreement achieved is marginally substantive (κ_c = 0.655). These results generally mirror those obtained by Rogers and Saunders (1994) who determined that there was no intra-observer error for this trait. Thus, while results of this study and Rogers and Saunders (1994) suggest that dorsal pubic pitting is a precise trait, it fails to provide an accurate sex classification in both subadults and adults. As a result, dorsal pubic pitting should not be used for skeletal sex determination. Given the mounting evidence that dorsal pubic pitting is a poor predictor of sex, even for adults, this trait will not be included heavily in further discussion.

While sex assessment using the complete list of 18 morphological traits just fails to meet the minimum accuracy required for this study (80%), when more weight is placed on pelvic traits developed specifically for subadult sex assessment, the overall accuracy of correct sex classification increases by 2.3 percent, from 79.7% to 82%. The increase in overall accuracy, in addition to a decrease in the number of indeterminate cases (8.6% to 3.9%), points to a promising trend but this trend occurs at the detriment of correct sex classification for males. The increase in overall accuracy of sex classification, however, is attributed to a marked increase in correct sex classification of females of approximately 33% (from 64.7% to 98%) but a decrease in accuracy of male sex classification of 18.2% (from 89.6% to 71.4%). Thus, both methods of assessing final sex produced a substantial sex bias, which exceeded 20%. When examining the accuracy of traits proposed specifically for subadult sex assessment, four traits (greater sciatic notch angle, greater sciatic notch shape, iliac crest curvature, and arc criterion) perform substantially better on females than males. The female bias observed in this study for subadult pelvic traits would explain the substantial increase in overall female

sex assessment when more weight is placed on subadult specific traits. The pattern of higher sex classification for females found in this study, however, contrasts the numerous studies that have shown that the methods proposed for subadult sex assessment are more successful at correctly identifying males than females (Weaver 1980; Hunt 1990; Mittler and Sheridan 1992; Schutkowski 1993; Sutter 2003; Vlak et al. 2008). In this study, the auricular surface height is the only trait that follows the trend of male bias in accurate sex classification observed in previous studies. The higher accuracy of female sex determination found in this study may be the result of an unconscious bias towards identifying the female expression of traits, which is one of the most problematic areas of subadult sex assessment given that there is a pattern of lower accuracy for the identification of females (Wilson and Humphrey 2017).

Five pelvic traits (auricular surface height, dorsal pubic pitting, ventral arc, pubic bone shape, and greater sciatic notch size and shape) show an intra-observer error that meets the allowable maximum error of 10% (Table 4.3). Auricular surface and dorsal pubic pitting showed the lowest levels of error (2.2%). The level of agreement achieved with the auricular surface was the strongest of all traits, with near perfect agreement (κ_c = 0.931). While pubic pitting had the lowest accuracy of all traits, auricular surface failed to meet the 80% accuracy threshold by almost 2%. Despite the seeming potential for auricular surface height, there is a strong male bias, where accuracy for males (95.9%) far exceeds that of females (50%). The male bias for the auricular surface found in this study mirrors the patterns found by Weaver (1980) and Mittler and Sheridan (1992). These studies, however, do not examine levels of error. While Luna and colleagues (2017) examine the auricular surface and show low observer error with promising accuracy, the authors examine features of the auricular surface other than elevation. Although there is no subadult comparison for observer error, in adults the auricular surface elevation shows error levels that exceed 10% (11.3%) and poor accuracy (73.5%) as determined by Rogers and Saunders (1994).

The level of error obtained with the ventral arc met the acceptable level (6.4%) and the level of agreement that was observed was substantial (κ_c =0.778). The ventral arc, however, failed to meet the required minimum for accuracy and only achieved an accuracy of 75.5%. Failure to meet the minimum accuracy set out for this research is

not unexpected given that the ventral arc, as described by Phenice (1969), does not appear until 20-23 years of age (Phenice 1969; Sutherland and Suchey 1991). The late appearance of the Phenice's ventral arc would explain the difference in error between subadults and adults, where this trait in adults does not exhibit intra-observer error (Rogers and Saunders 1994). Moreover, given that this study also used the precursor arc, as described by Sutherland and Suchey (1991), as a proxy for the ventral arc, an accuracy of 75.5% for all subadults is also not unexpected. Sutherland and Suchey (1991) showed that the precursor arc appears in females at approximately 14 years of age, and becomes the most common form of expression in females by 20 years. It is possible, then, that the level of inaccuracy of this trait for subadults, as a broad group, can be attributed to the lack of expression of the precursor arc in individuals under 14 years of age.

The greater sciatic notch size and shape is the last trait with an observer error under 10% (pubic bone shape is discussed above). The greater sciatic notch size & shape has the highest level of error of the five traits, just meeting the threshold with an error of 9.4%. The agreement that is observed with this trait, however, is "near perfect" ($\kappa_c = 0.812$). While the level of error found in this study is slightly higher than a 7% error found by Olivares and Aguilera (2016), the level of agreement is quite comparable ($\kappa_c = 0.87$) between the two studies. Despite the strong level of agreement and an acceptable level of error, the accuracy of this trait is adequate at best (76.4%). The greater sciatic notch size & shape performed substantially better on females (90.2%) than males (67.1%) in this study. The greater sciatic notch has received considerable attention for subadult sex determination (Schutkowski 1993; Sutter 2003; Vlak et al. 2008; Olivares and Aguilera 2016). The overall accuracy achieved in this study for this trait is lower than those achieved by Schutkowski (1993) and Sutter (2003), 79.6% and 81.5% respectively, but substantially higher than the accuracy achieved by Vlak and colleagues (2008) who achieved an accuracy of only 59%. In the present study, the greater sciatic notch size & shape shows a female bias, where accuracy for female sex classification (90.2%) far exceeds that of male classification (67.1%). The female bias found in this study is similar to that found by Sutter (2003) but contrasts a male bias reported by Schutkowski (1993), Vlak and colleagues (2008), and Olivares and Aguilera

(2016). The variability in accuracy and patterns between all studies may be the result of the general discrepancies that exist between investigators when methods are tested for validation, a problem that most noticeably affects subadult sex assessment (Vlak et al. 2008; Lesciotto and Doershuk, 2017; Wilson and Humphrey 2017; Horbaly et al. 2019). While it is possible observer subjectivity in scoring the greater sciatic notch size & shape may contribute to some of the variability seen between all studies, the low level of intra-observer error and high-level of agreement found in this study and by Olivares and Aguilera (2016) would suggest that observer subjectivity does not completely explain the variability in accuracy. On the contrary, the low level of intra-observer error and high level of agreement would suggest that the descriptions for trait expressions are clear enough to produce comparable results between assessments. The variability in accuracy and sex bias patterns seen between studies may be more influenced by the different age ranges included in the different studies. Schutkowski (1993), Sutter (2003), and Olivares and Aguilera (2016) tested the sciatic notch size & shape on samples with an age range of 0-5 years, while Vlak and colleagues (2008) included a wider age range of 0-15 years. None of the aforementioned studies include an age-range as expansive as the current study. It is possible, then, that alternations to the morphology of the sciatic notch that occur during ontogeny (Vlak et al. 2008) may, in part, explain the level of variation in accuracy between the various studies.

The ranking of pelvic traits based on accuracy and precision produced in this study (Table 4.4) shows that the top 5 pelvic traits showing promise for subadult sex determination, and thus sufficiently dimorphic prior to adulthood, include (in descending order): pubic bone shape, true pelvis size & shape, auricular surface height, pelvic inlet shape, and preauricular sulcus. Pubic bone shape and preauricular sulcus are the only two traits that scored the same for accuracy and precision (4th and 7th, respectively). True pelvis size & shape and pelvic inlet shape both ranked high on accuracy (ranked second and first) but less so for precision (6th and 11th, respectively). Conversely, auricular surface height ranked mid-range for accuracy (9th) but high for precision (1st). Despite the discrepancy in ranking of accuracy and precision, true pelvis size & shape scored an overall rank of 1 with pubic bone shape. Since there has not yet been a study assessing accuracy and precision to rank the efficacy of a comprehensive list of pelvic

traits for subadults, the only comparison available is a ranking of pelvic traits for adults. Only one trait, true pelvis size & shape, is within the top 5 traits in both subadults and adults. While this trait received an overall rank of 1 in subadults, it scored a rank of 3 in adults (Rogers and Saunders 1994). Moreover, the opposite pattern is seen in adults in terms of ranking for accuracy and precision, where true pelvis size & shape scored higher on precision (1^{st}) than accuracy (7^{th}) (Rogers and Saunders 1994). The discrepancy in ranking of pelvic traits between subadult and adults may indicate an increase in dimorphism of certain traits, compared to others, well into adulthood. This would certainly be true for the ventral arc, which ranked 1 in adults, since it does not develop or become prominent until after 20-23 years of age (Phenice 1969; Sutherland and Suchey 1991).

In terms of measurable differences in the pelvis between males and females, all nine logistic regression models performed at or above 80% accuracy. While the overall accuracies for the nine logistic regression models seem promising, only model 20 produced an accuracy of 90% or above for both females and males. The remaining eight models performed substantially better for male sex classification than female classification. The results observed in this study differ from the patterns obtained by Albanese (2003), where a decrease in overall accuracy of approximately 10-15% is seen for eight logistic regression models when they are applied to subadults. Moreover, the regression models performed comparably well for males and females in Albanese's (2003) study with accuracy surpassing 90% for both sexes in most cases. The best performing logistic regression equation in this study (Model 20) only saw a decrease of 2.4% in overall accuracy compared to Albanese (2003). Model 20 is also the only equation tested in this study that did not require a measurement for the superior pubis ramus length (SPRL). All other measurements (hip bone height, iliac breadth, acetabular-ischium length, maximum diameter or femur head, and epicondylar breadth) were required for this regression model. It is possible, then, that the superior pubis ramus length is a measurement that is not applicable to subadults as it is to adults. This observation may provide evidence of the fact that significant growth (medio-lateral elongation) of the pubis continues after the age of 18 in females but ends at approximately 18 years in males (Tague 1989; Budinoff and Tague 1990). The

prolonged growth of the female pubis in adulthood could likely explain why the logistic regression equations that include SPRL show a substantially higher accuracy, both overall and for females, in Albanese's (2003) study than what was found here.

5.2 Age Related Trends

When examining the ages of trait appearance and stabilization between males and females, three general patterns emerge. The majority of traits (seven traits) appear to have a male "default" morphology where a divergence in female dimorphism develops later in life, with an average and minimum age of trait stabilization in female expression occurring in the mid- to late teenage years. The second pattern, occurring in only two traits, shows a female "default" morphology with subsequent divergence of the male form occurring in childhood (iliac crest curvature) or in the later teen years (development of muscle markings). These first two patterns observed of a male or female "default" trait expression may provide evidence of divergent ontogenetic trajectories. Divergent trajectories suggest that shape changes are confined to the later stages of postnatal growth where no difference in shape between the sexes would be present in the youngest stages of growth (Wilson et al. 2015). The last pattern, observed in four traits, shows female and male trait expressions first appearing by birth, with subsequent average age of stabilization occurring in the teenage years for both sexes. Five traits have comparable ages of first expression for males and females in the early teen years (13 and 12 years, respectively), which is dependent on the fusion of pelvic or sacral elements.

Traits that show a male "default" morphology at birth include: subpubic concavity, ischiopubic ramus, ventral arc, pubic bone shape, dorsal pubic pitting, auricular surface height, and preauricular sulcus. The subpubic concavity showed high accuracy for the infant age category (90% accuracy) but only 1 female was represented in this age cohort for the trait and was erroneously classified. Thus, the overall 90% accuracy found in the infant category only represents the ability to use this trait on subadult males. One issue with applying this trait to younger individuals was that an estimated articulation had to be used; therefore any promising accuracy for individuals with an unfused pubis and ischium should be questionable at best. A decrease in accuracy

between the infant and child cohorts is observed, but overall accuracy gradually increases between the child and late adolescent groups. The late adolescent cohort exhibited the highest overall accuracy at 93.2%. In this age group, both males and females surpassed 80% accuracy but a male bias is seen (98% for males, 82.6% for females). This is not surprising given that this research found that the average age of trait stabilization in females occurs at 17.8 years, with a first appearance and minimum age of trait stabilization for females occurring at 14 years of age. Klales and Burns (2017) also found that their late adolescent group (15.6-20.5 years old) showed the highest overall accuracy (97.2%) and the lowest level of sex bias (0.7%). Contrary to Klales and Burns (2017), however, the current study shows that the subpubic concavity is not reliable for young adolescents, but there is agreement between both studies that suggests the subpubic concavity is not suitable for subadults aged 1-12.5 years. Both this study, and Klales and Burns (2017) shows a substantial decrease in sex bias among the late adolescent group, which is likely the result of the divergence in development of this trait in both sexes. In fact, Coleman (1969) found that there is different directional growth that occurs at the middle of the ischiopubic ramus between males and females, and LaVelle (1995) further showed that, during the adolescent growth spurt, females showed increased growth in regions of the pubis and ischium resulting in, among other things, an obtuse subpubic angle. Moreover, Phenice (1969), who first examined the subpubic concavity, suggested that this trait was not well developed until 20 years of age in females and, therefore, could not be used until then. The results of this study, then, provide evidence that the subpubic concavity can be used for subadults 17-20 years of age and, given the consistent high accuracy of male classification, potentially as early as 14 years of age based on the minimum age of female trait stabilization.

In this study, the ventral arc was primarily expressed as the precursor arc. The precursor arc first appeared in the study sample at 14 years, with subsequent stabilization at approximately 19 years of age. The age of appearance and stabilization of the precursor arc mirror the patterns obtained by Sutherland and Suchey (1991). The ventral (precursor) arc demonstrated the highest accuracy in the late adolescent cohort (89.7%), but a strong male bias was observed in this age group (100% accuracy for

males, 63% for females). In this study, however, when the precursor arc was present, it was only expressed in females. No males exhibited the precursor arc, and were not erroneously scored as female. As a result, it appears that the precursor arc is exclusively seen in females and the *presence* of this trait can be used to provide, at least, a minimum number of females if used on an archaeological assemblage. The absence of the ventral (precursor) arc is more complicated, particularly within subadults between the ages of 14 and 20 years, since an absence of the trait does not exclusively suggest a male individual. The arc's exclusive presence in females provides evidence that, while females may often be misidentified, classification of female tends to be associated with fewer false positives and thus, greater certainty of correct identification (Wilson and Humphrey 2017). While the precursor arc, as described by Sutherland and Suchey (1991), was first observed in 14-year-olds, 8 individuals between four and 10 years of age appear to exhibit a "twisting" of the symphyseal face, where the billowed texture of the pubic symphysis seems to gradually curve towards the anterior (ventral) surface of the pubis inferiorly. This "twisting" was only observed; the eight individuals were not scored as females since the "twisting" was only identified during the second round of sex assessments. Seven of the individuals exhibiting the "symphyseal twisting" are females and only one 6-year-old is a male. The observation on the male, however, was made with the caveat that there was some damage to the pubic symphysis, which may have resulted in an erroneous inference of the twisting being present. It may be possible that, similar to the precursor arc, the "symphyseal twisting" may be exclusively present in female children. Further examination of this trait expression, however, is required on a larger sample of subadults between 4 and 10 years of age to determine if the presence of the "symphyseal twisting" can help identify younger female subadults.

The female expression of the ischiopubic ramus ridge first appears at 14 years of age in this study and stabilizes, on average, at 19 years with a minimum age of stabilization occurring one year earlier. While the overall accuracy of this trait increases between the child and late adolescent age cohorts, a male bias in correct sex classification is observed. The male bias is particularly strong in the infant, child, and adolescent cohorts where male classification accuracy is approximately 67-100% higher than female accuracy (Table 4.18). The magnitude of bias is less in late adolescents (a

difference of 18.8%) and this age category has an overall accuracy meeting the threshold for this study, but female accuracy fell below 80% (72.7%). The consistent low accuracy for this trait throughout the age cohorts is likely the result of the late epiphyseal development of the ischio-pubic ramus. The ischio-pubic ramus ridge assesses both the presence of a sharp ridge and the breadth of the ischio-pubic surface (Phenice 1969). Females tend to have a narrower ischiopubic ramus, which may result from medial-lateral lengthening of the pubic bone that occurs during adolescent growth (Coleman 1969; Tague 1989; Anderson 1990; Budinoff and Tague 1990; Klales and Burns 2017). Males, on the other hand, tend to have broader surface of the ischio-pubic ramus. Thus, the late growth of the pubic bone in females, which subsequently impacts the morphology of the ischio-pubic ramus, likely reflects the substantial male bias throughout all age cohorts. As for the presence of the sharp ridge, the ischial epiphysis does not completely fuse until 20-23 years of age, and does not extend halfway along the ramus until 19-20 years (Scheuer and Black 2000). Although this trait is the least understood trait proposed by Phenice (Klales and Burns 2017), if the presence of a "sharp ridge" is the result of muscular forces acting on the ischiopubic ramus, it would not be surprising that differences in musculature impacts the observed male bias. Although this trait seems appropriate for late adolescents given its substantial overall accuracy (85.5%), the substantial male bias seen even in the late adolescent cohort suggests that the ischiopubic ramus alone may not be ideal for subadult sex assessment.

The auricular surface height first showed a distinct trait expression for females at 8 years of age. The average age of stabilization for the female expression occurred at approximately 18 years, but the earliest stabilization can be seen at 16 years. This trait had a high overall accuracy for the infant category but was completely male dependent, as males had 100% accuracy while females had 0% accuracy. A substantial decrease is observed in the overall accuracies between the infant and child categories, but a consistent increase in accuracy is later seen between the child and late adolescent categories. The highest accuracy for the auricular surface height is seen in the late adolescent cohort (approximately 89%). Mittler and Sheridan (1992) also found that the effectiveness of the auricular surface height increased with age in both sexes. In their

study, the authors noted that females 10-18 years of age had 66.7% accuracy with higher accuracy (71.4%) occurring in females 14-18 years of age (Mittler and Sheridan 1992). The present study has shown that there is a substantial male bias seen throughout all age cohorts and despite the overall accuracy for the late adolescent cohort being approximately 89%, female accuracy does not meet 80% accuracy by 2.7%. The results obtained in this study show similar patterns observed by Mittler and Sheridan (1992), where accuracy for females in the older cohorts remains problematic. Moreover, both studies show that the auricular surface height was least effective in younger children, from birth to 9 years old (Mittler and Sheridan 1992). Unfortunately, results of the present study cannot be directly compared to the results obtained by Weaver (1980) or Hunt (1990). Weaver (1980) limited his study to fetal and infant remains, while Hunt (1990) conducted his study on an archaeological sample with unknown sex and relied on a hypothetical 1:1 sex distribution to examine the efficacy of the auricular surface height. Due to female accuracy just failing to meet 80% accuracy, the auricular surface height can be cautiously applied to subadults 17-20 years of age.

The female expression of the preauricular sulcus is first observed at 6 years of age and exhibited an average age of stabilization at 17.5 years, with minimum age of stabilization at age 16. There is a gradual increase in accuracy from the child cohort to late adolescent cohort where the late adolescent group achieved the highest overall accuracy (approximately 92%). A male bias exists, however, in all age groups and while the bias gradually decreases between subsequent age cohorts, it still remains in late adolescent cohort (approximately 20% difference in accuracy). Female accuracy in this age cohort fails to meet 80% accuracy by only 1.7%. The preauricular sulcus (or groove) is a trait that has been commonly suggested to be a sign of parity, where a “groove of ligament” is said to occur in males and non-parous women and a “groove of pregnancy” is only observed in parous women (McFadden and Oxenham 2018). McFadden and Oxenham’s (2018: 201) meta-analysis has suggested, however, that the preauricular sulcus is a better predictor of sex than parity. As a predictor of sex, the preauricular sulcus has shown accuracies as low as 49% but as high as 85%, while as a predictor of parity this trait has correct classification that ranges from 42-66% (McFadden and Oxenham 2018). Although the preauricular sulcus has a wider range of

accuracy as a predictor of sex, its predictive power is moderate at best (66%). The results achieved in the present study for the late adolescent group, however, closely resemble the accuracy achieved by Rogers and Saunders (1994) who found the preauricular sulcus had an accuracy of 91.6% in adults. It is possible, then, that the preauricular sulcus can be cautiously applied, with some success, to individuals 17-20 years of age.

The pubic bone shape first showed the female form in 12-year-old individuals. On average, the female form of expression stabilized at approximately 18 years, with the earliest age of stabilization occurring at 16 years of age. With the exception of the infant age category, there is a general increase in overall accuracy with age that peaks in the late adolescent cohort at 97% accuracy. The infant, child, and adolescent cohorts display a substantial male bias, where male accuracy exceeds that of females by 50-100%. The accuracy for both male and female sex classification is most comparable in the late adolescent cohort, where correct female classification achieved 100% accuracy and male classification has approximately 96% accuracy. As discussed above, the pubic bone shows substantial growth after 18 years of age in females, but completes growth in males by that age (Tague 1989; Budinoff and Tague 1990). While the results of the pubic bone shape does not necessarily dispute that substantial growth of the pubis occurs after 18 years of age, the results do suggest that shape differences in the pubic bone between the sexes are dimorphic enough, and can be visually assessed with confidence, in individuals 17-20 years of age, and possibly as early as 16 years.

Two traits demonstrate the second pattern showing a sort of female “default” expression at birth with later emergence of the male expression. Male expressions for development of muscle markings and iliac crest curvature have an average age of stabilization occurring at the tail end of adolescence (19.4 and 19.7 years, respectively). The minimum age of stabilization occurs only a few months to a year before. Both traits demonstrate an increase in overall accuracy with increasing age. The development of muscle markings exhibits a pronounced female bias throughout all age cohorts (32-92% difference). Iliac crest curvature exhibits a female bias in all age cohorts with the exception of the adolescent group. The level of sex bias is markedly reduced in the late adolescent cohort (6.3%). None of the age cohorts, however, have an overall accuracy

meeting 80% and while the late adolescent cohort for the development of muscle markings misses the accuracy threshold set out in this study by only 1.8%, there is pronounced female bias even in this age group (32% difference). The lack of efficacy of both development of muscle markings and iliac crest curvature for subadult sex assessment does not come as a surprise. Development of muscle markings is, hypothetically, dependent on repetitive use and strain of muscles that usually comes with age, and relies on the robustness of muscles since males are generally believed to be more robust (in musculature) than females (Plavcan 2012; Stulp and Barrett 2016; Dunsworth 2020). Rogers and Saunders (1994), however, determined that this trait is a poor predictor of sex as it performs marginally better than chance (approximately 57% accuracy) in adults. The development of muscle markings on the ilium is likely more dependent on activity than sexual dimorphism. The mechanisms and physiological rationale for the iliac crest curvature was never explained in Schutkowski's (1993) study. Subsequent tests of this trait by Sutter (2003) and Olivares and Aguilera (2016) do not address the physiological rationale that may account for dimorphism in the curvature of the iliac crest. It is possible that, like the development of muscle markings, the iliac crest curvature may be influenced in part by musculature. The gluteus medius muscle attaches posterior-laterally (to the back and side) of the ilium, while the iliacus muscle attaches to the ilium medially (or internally) (Scheuer and Black 2000). Sex differences in the robustness of these muscles, should they exist, may influence the purported differences in iliac crest curvature between males and females. Similar to the development of muscle markings, if the gluteus medius and iliacus muscles impact the level of "dimorphism" in iliac crest curvature, this trait may be more influenced by activity and body size as opposed to dimorphism. Ultimately, however, due to the results obtained in this study, development of muscle markings and iliac crest curvature are pelvic traits that should not be used for subadult sex determination at any age.

The third general pattern emerging from this study in terms of the appearance and stabilization of pelvic traits suggests that traits, such as the greater sciatic notch size & shape, greater sciatic notch angle, arc criterion, and ilium shape, show both male and female forms of expression by birth. This pattern could provide evidence, on a macro scale, of parallel ontogenetic trajectories in which females and males exhibit

different shapes, in this instance of a morphological trait, at the onset of ontogeny and growth trajectories point in the same direction. The shape differences between the sexes will never resemble one another at any point (Wilson et al. 2015). All four traits show an earlier average and minimum age of stabilization of female expression than male expression (Table 4.4). Average age of stabilization for the female expression of the four traits occurs anywhere between 2.5-4.5 years before males, while the minimum age of trait expression shows the greatest difference between the sexes (between 5-11 years). Generally, the greater sciatic notch size & shape, greater sciatic notch angle, arc criterion, and ilium shape exhibit a gradual increase in overall accuracy with increasing age. Although both male and female forms of the four traits are observed by birth, they are all poor predictors of sex for the infant age category given the exceptionally low overall accuracies obtained range between 35% (greater sciatic notch angle and arc criterion) and 44% (ilium shape). By childhood, the greater sciatic notch size & shape and greater sciatic notch angle become moderate predictors of sex (76%) but the arc criterion and ilium shape remain problematic predictors of sex (71 and 62% accuracy, respectively). The greater sciatic notch size & shape, greater sciatic notch angle, and arc criterion demonstrate a female bias throughout all age categories but only the former two traits exhibited a substantial decrease in the level of female bias in the late adolescent category. The greater sciatic notch size & shape has a female bias of only 3% in the late adolescent cohort, whereas the greater sciatic notch angle has a bias of 8.7%. The ilium shape shows low levels of sex bias in both the adolescent and late adolescent cohorts (8.3 and 0.7%, respectively). The greater sciatic notch size & shape and the greater sciatic notch angle both show overall accuracies surpassing 80% in the late adolescent categories (85% and 84%, respectively), while ilium shape has overall accuracies above 90% for the adolescent and late adolescent categories (93 and 92% respectively).

Reynolds (1947) and Boucher (1957) argued that sex differences in fetal and infant remains exist because the level of hormones released have the potential to impact the skeleton. These early studies highlighted the importance of androgens and testosterone in fetal sex differentiation. Androgens were argued to be responsible for the development of male sex characteristics since fetal testosterone peaks at 15 weeks,

resulting in the onset of major sexual differentiation (Boucher 1957; Weaver 1980). Recently, the importance of estrogen has also been introduced in the discussion of sexual dimorphism in humans. Estrogen is produced in greater amounts in females compared to males and the amount of estrogen produced changes across the life course (Greulich and Thoms 1944; Dunsworth 2020). During the last few weeks of fetal development, estrogen levels are at their highest in gestation (Dunsworth 2020). It is clear that the high levels of testosterone in males and high levels of estrogen in females during fetal development contribute to sex differentiation. The results obtained in this study do not necessarily negate that sex differences may be extended to the skeleton, particularly the pelvis, as demonstrated by a level of dimorphism present. However, the degree of sexual dimorphism is certainly not at a level that can be used for skeletal sex determination of infants and children. Hormones continue to be released, although at lower levels, until about 2-3 years of age (Bogin 1999) and may be influencing the substantial increase in accuracy between the infant and child cohorts, in addition to general growth, seen in this study. Hormonal release in adolescents, coupled with ongoing growth, might be what causes the continued increase in accuracy of these traits, peaking in late adolescents. The greater sciatic notch size & shape, greater sciatic notch angle, and arc criterion were all proposed as traits for subadult sex assessment for infants and children, but in this study, those age categories show the lowest levels of accuracy. This pattern is not unexpected since Cardoso and Saunders (2008) found comparable poor accuracy for the arc criterion in infant individuals, while Vlak and colleagues (2008) obtained poor accuracies for the greater sciatic notch shape and angle, albeit higher than the accuracy obtained here.

Due to the high accuracy and low sex bias observed for the greater sciatic notch size & shape in late adolescents, this trait is applicable to individuals 17-20 years of age, and possibly as early as 15 or 16. Despite the greater sciatic notch angle demonstrating low sex bias and accuracy over 80% in the late adolescent category, this trait does show substantial intra-observer error when applied to all subadults, late adolescents included (approximately 20% error). The ilium shape also exhibits the same issue with an error level of approximately 25% for all age groups combined. The error levels of these two traits is problematic, but these traits may be cautiously used for

individuals 17-20 years of age, and even 12-16.9 years for the ilium shape, although these traits should not be used alone. While the arc criterion has an overall accuracy above 80% in the late adolescent category, a substantial female bias remains even in this age cohort (25% difference in accuracy between the sexes). Moreover, the arc criterion demonstrates the second highest level of error, at 26%, when it is applied to all age categories. Due to the high level of sex bias and high level of intra-observer error, the arc criterion is not an effective trait for subadults at any age.

The remaining five traits (pelvic inlet shape, true pelvis size & shape, sacrum shape, obturator foramen shape, and acetabulum size & orientation) are dependent on fusion of pelvic elements but have comparable age of appearance and stabilization of female and male expression. Pelvic inlet shape, true pelvis size & shape, and sacrum shape are all features related to the birth canal. The female expression of all three traits is first seen in 12-year-olds, while the male expression if first seen in 13 year-old-males. The average age of female trait stabilization for all three traits is approximately 18 years of age, with minimum age of stabilization of female expression occurring at 14-16 years for pelvic inlet shape, 17 years for true pelvis size & shape, and 18 years for sacrum shape. The average age of stabilization for the male expression of these three traits occurs a year after females (19 years) and have a minimum age of stabilization at 17 years. Pelvic inlet shape has accuracies at or exceeding 90% in the adolescent and late adolescent categories. While the pelvic inlet shape seems promising for the adolescent category, only one male individual was represented in the age category. The correct classification of one male, then, cannot be interpreted as being successful for male individuals in general. The pelvic inlet shape must be examined on a sample with more males between 12 and 16.9 years of age to determine if this trait can be successfully applied to female *and* male adolescents within that age range. True pelvis size & shape and the sacrum shape both exhibit a marked increase between the adolescent and late adolescent cohorts. While both traits have overall accuracies that exceed 80% in the late adolescent group, only true pelvis size & shape demonstrates high accuracy for both males and females with minimal female bias (6.2%). Correct female classification in the late adolescent cohort using the sacrum shape is moderate at best (67%) and there is a high correct male classification bias in this age group (21%).

Sexual dimorphism related to the pelvic cavity, such as pelvic inlet shape and true pelvis size & shape, has been stated to develop during adolescent growth, particularly during late adolescence, when male and female pelves experience remodeling of the pelvic cavity with lateral displacement of the ishium and acetabulum in both males and females (Coleman 1969; LaVelle 1995). Studies have shown that among 18 year olds, females have larger birth canal and pelvic cavity dimensions (Wood and Chamberlain 1986; Arsuaga and Carretero 1994; LaVelle 1995). In fact, LaVelle (1995) found that between 8 and 18 years of age, female pelves expand more in the dimensions of the true pelvis than males'. Additionally, Huseynov and colleagues (2016) found that the dimensions of the true pelvis expand from puberty until about 25-30 years of age. Obstetric selection pressure has traditionally been considered the main influence in female pelvic morphology where features, such as a round pelvic inlet and larger subpubic angle, are associated with adequate space for childbirth (Tague 1989). Recent literature has suggested that estrogen may also play a vital role in the emergence of sexual dimorphism in humans (Dunsworth 2020). Muscles of the pelvic floor and the round muscles and ligaments of the uterus contain estrogen receptors. These receptors are absent in other skeletal muscles, such as the rectus abdominis, which suggests that the pelvic muscles are under special hormonal control (Smith et al. 1990; Dunsworth 2020). The uterus, for example, exhibits rapid growth around the age of 10, coinciding with the onset of puberty and increasing levels of hormones (such as leuteinizing hormone, follicle stimulating hormone, estradiol) (Dunsworth 2020). Thus, it is possible that the release of estrogen may influence the morphology of estrogen sensitive muscles of the pelvis. Additionally, female reproductive organs take up more volume and are placed within a more skeletally constrained region of the pelvis compared to male reproductive organs (Dunsworth 2020). Dunsworth (2020) has, therefore, argued that the pelvis might exhibit plastic accommodation for developing soft tissue, particularly in females, in a similar way that the skull grows to accommodate a growing brain or the bony orbits to accommodate the eyes. Whether it is obstetric pressures, estrogen levels and sensitivity, or a combination of the two, this research has shown that at least two features of the pelvic cavity (pelvic inlet shape and true pelvis size & shape) can be confidently used for sex determination in individuals between 17

and 20 years of age. Further research incorporating a higher number of male individuals between 12-16.9 years of age is required to assess whether or not the pelvic inlet shape can be extended to that age group. The sacrum shape should likely be avoided, even for late adolescents, given its persistent and substantial male bias.

Obturator foramen shape and acetabulum size & orientation are the last two traits that require the fusion of pelvic elements. Both traits first show the female expression in 12-year-olds, while the male expression is first seen in 13-year-olds. The average age of female trait stabilization occurs at approximately 18 years in females, with a minimum age of stabilization seen at 16 years. Male trait stabilizations is seen, on average, in 19 year olds for both traits, but the minimum age of stabilization for acetabulum size & orientation is seen in 17-year-olds whereas obturator foramen shape stabilizes in 15 year old males. Both traits exhibit an increase in accuracy with age, both achieving their highest accuracy in the late adolescent group (obturator foramen 88% and acetabulum 94.5%). Obturator foramen shape exhibits a slight male bias in the late adolescent group (7.4%), while sex classification using the acetabulum size & orientation is comparable for both sexes (95.7% females, 94% males). Studies have shown that among 18 year olds, males tend to have larger acetabulum size (Wood and Chamberlain 1986; Arsuaga and Carretero 1994; LaVelle 1995). This pattern of larger acetabulum size among males 18 year of year is mirrored in the present study given that the average age of trait stabilization for acetabulum size & orientation occurs at 19 years of age in males and 17.9 years of age in females (Table 4.5). Morphologically, acetabulum size in males is described as “large”, whereas for females the trait is described as “small”. Intra-observer error exceeds 10% when assessed across all age groups for the obturator foramen shape and acetabulum size & orientation. It is, therefore, possible that the traits can be cautiously applied to 17-20 year olds as there is high accuracy and low sex bias regardless of the high error for the entire sample as the error for this group is low.

An important consideration to take into account is that the majority of traits examined in this study exhibit average ages of trait stabilization in the later teen years (between 17 and 20 years of age). The late average age of trait stabilization is likely the result of the disparate age distribution in the study sample. More than half of the entire

study sample (n=73) is between the ages of 17 and 20 years, most of which are 19 and 20-years-olds (n=46). This skew in age distribution, then, likely increases the average age of trait stabilization observed in this study given the large quantity of individuals that are 19 and 20 years of age. Additionally, there are no subadults that are 2 or 9 years of age in the study sample and only females are represented in five ages (5-, 7-, 12-, 14-, and 16-year-olds) and only males are represented in three ages (3-, 13-, and 15-year-olds). The unequal representation of sexes at these specific ages impact clear age related trends, particularly during the early to mid-teenage years. For example, since there are no 15-year old females, it remains unclear if traits that show the first female expression at 16 years might in fact occur a year earlier. Furthermore, the minimum age of stabilization of a trait is likely also impacted and it is possible that the first signs of trait stabilization may, in fact, occur earlier than what is reported here. The skewed age distribution may also impact the patterns observed with respect to age categories. The infant and adolescent categories consistently have lower sample sizes compared to the child and late adolescent categories. Additionally, the infant, child, and adolescent categories all have substantially lower sample sizes than the late adolescent category, between 52-58 less individuals. As a result, it is possible that the accuracies obtained for the infant, child, and adolescent age categories are not reflective of what they would be if the sample sizes were more comparable to the late adolescent group. As Wilson and Humphrey (2017) note, unequal sampling of ontogeny is an extrinsic limitation to the study of subadult remains and is primarily dictated by the demographic composition of skeletal collections available for study. However, despite the disparate age distribution in the sample, age related patterns still emerge and the patterns associated with some age categories, namely the adolescent and late adolescent cohorts, are statistically significant. Future research including more individuals aged 14-16 years is required to examine whether or not average ages of stabilization closely align with the minimum age of stabilization found in this study (Table 4.5).

When all 18 traits are used to conduct a final sex assessment, there is a general pattern of increasing accuracy with age. When all traits are weighted equally, only the late adolescent group produces promising results, where overall accuracy is approximately 96% with little difference between the sexes (Table 4.24). When more

weight is placed on the traits proposed specifically for subadult individuals, the child and adolescent cohorts exhibit a marked increase in sex assessment accuracy (approx. 32 and 26% difference, respectively). The accuracy for the late adolescent group remains virtually unchanged. While the marked increase for the adolescent group seems quite promising (100%), the sex distribution in this age category is quite disparate since there is four times the number of females than males (n= 12 and 3, respectively). If all 18 pelvic traits are available for examination, a sex assessment can be conducted, with confidence, on individuals 17-20 years of age. Based on the results of this study, I argue that a sex assessment using all pelvic traits can also be *cautiously* conducted on individuals aged 12-16.9 years, with more weight given to subadult traits. However, further research is required using a sample with a more equal ratio of adolescent males and females to be certain.

Albanese's (2003) metric method of sex assessment was only applicable to the adolescent and late adolescent groups give that measurements, superior pubis ramus length in particular, required the fusion of the acetabulum. There is a general increase in accuracy observed between the two age groups, with the exception of model 8, which demonstrates a 2% decrease in overall accuracy with age. The late adolescent cohort for model 20 exhibits the highest overall accuracy (94%) with female and male accuracy exceeding 80%. The late adolescent groups for models 4, 5, 8, 10, and 11 achieve an overall accuracy that meet or exceed 80%, but these models also exhibit male bias in the late adolescent cohorts where female accuracy fails to achieve 80% and ranges from 65-76%. The adolescent groups of the aforementioned models do achieve accuracies of 80% and above, but have small sample sizes (n= 9 or 10, depending on the model) and exhibit either substantial male (model 11) or female (model 4, 5, and 10) bias. The adolescent cohort of model 8 shows one of the highest accuracies among that age group, but only one male is represented in the age group and therefore overwhelmingly reflects the accuracy for female sex classification. Research on a large sample of adolescent individuals (12-16.9 years) will help determine whether the high accuracy for model 8 does actually show promise for sex determination. The overall accuracies of models 1, 2, and 3 surpass 80% but only for the late adolescent age group. Despite the high overall accuracy, substantial male bias is observed in this age

group where correct classification of males surpasses that of females by 25.5 - 27.1%. Additionally, the closest female accuracy to 80% between the three regression models occurs in model 3, where female accuracy is only 70%. A male bias is consistently observed in the late adolescent categories for all nine logistic regression models examined in this study. The lowest level of male bias observed is found in the only regression model not requiring a measurement for the superior pubic ramus length (model 20). The persistent male bias seen in the late adolescent groups for the logistic regression models requiring measurements of the pubis may be the result of significant growth of the pubis continuing after the age of 18 in females but ending at approximately 18 years in males (Tague 1989; Budinoff and Tague 1990). Prolonged growth of the pubis in females is further exemplified by the lack of sex bias found by Albanese (2003) in adults. The results of this study, then, suggest that model 20 can be confidently applied to individuals between 17 and 20 years of age for sex determination.

5.3 Puberty and Sexual Dimorphism

An important requisite often identified for sexual dimorphism to be observed in the skeleton is the cessation of puberty. The indicators of pubertal stage progressed in the individuals observed in this study in the way proposed by Lewis and colleagues (2016), where full development of the hamate hook occurs first, followed by fusion of the distal phalangeal epiphyses, fusion of the distal radius, and ending with the fusion of the iliac crest (Lewis et al. 2016). The pubertal chart produced in this study (Figure 4.14) demonstrates that females were maturing, on average, earlier than males. This pattern is not unexpected since it has been well established that females tend to mature faster than males (Bogin, 1999; Rogol et al. 2000). The pubertal stage, peak height velocity (PHV) is of particular interest as it corresponds to overt soft tissue changes in females and males that, in the past, would mark a child's transition into adulthood (Bogin, 1999; Rogol et al. 2000; Shapland and Lewis 2013, 2014; Arthur et al. 2016; Lewis et al. 2016). PHV occurred, on average, at 12 years for females and 14 years for males. The age difference observed between the sexes, however, is likely the result of a low number of females in the PHV stage (n=2). This is supported by the fact that the difference in mean age of PHV between the sexes is not statistically significant. In girls,

menarche occurs after PHV is attained and during the deceleration stage of the pubertal growth curve (Tanner, 1973; Sinclair 1989; Bogin 1999). The mean age of attainment of the deceleration stage for females in this study was 15.2 years. Females in this sample, therefore, exhibit a delay in menarche of approximately three years compared to modern standards (Chumlea et al. 2003; Aksglaede 2008). The females in this study also show a delay in menarche compared to the U.S. average from 1950-1970 (13 years) but comparable to the average age of menarche in the U.S. in 1877 (14.75 years) (Kaplowitz, 2006). External factors, such as malnutrition, low body weight, and socio-economic status, have the potential to delay menarche (Sinclair 1989; Lewis et al. 2016). The subadults of the Hamann-Todd and Terry collections lived during periods of changing social and economic conditions that likely had an impact on their health. The Great Migration, in particular, resulted in an influx of Black labourers from Southern US to the North. This led to overcrowded and unsanitary living conditions in urban neighbours, such as Cleveland and St. Louis (Kusmer 1978; Phillips 1999). Additionally, despite Northern US urban centers providing job opportunities that were not available in the South, they were not always high paying jobs, particularly in relation to the cost of living in an urban center (Phillips 1996). As a result, the precarious economic situation may have led to episodes of food insecurity, resulting in a level of malnutrition in children while the overcrowded living conditions may have increased the disease burden for Black labourers and their families living in urban centers. The delay in age of menarche seen in this study, then, likely reflects the nature of those who comprise the skeletal collections used in this study, namely unclaimed bodies from morgues and hospitals from the Mid-West who likely lived in less than optimal conditions. To my knowledge, Henderson and Padez (2017) are the only authors who have examined pubertal stages on a skeletal collection of documented age-at-death. Similar patterns are observed between the two studies where acceleration occurs around the age of 10-11 and menarche is attained by approximately 15 years of age (Henderson and Padez 2017). Given that the skeletal collection Henderson and Padez (2017) examined, the Coimbra Identified Collection, is temporally contemporaneous to the collections examined in this study, it is possible that geographical differences in the attainment of

menarche are quite minimal and that temporal trends (i.e. skeletal assemblages from different time periods) may be of more importance (Kaplowitz 2006).

The relationship between pubertal stages and pelvic sexual dimorphism explored in this study suggests that puberty does not necessarily have to be complete in order to apply dimorphic traits of the pelvis for skeletal sex determination. What seems to be more important, however, is achievement of peak height velocity (PHV). Most traits produce high enough accuracy (>80%) for the sex determination of individuals in the deceleration and maturation pubertal stages. Both these stages precede the post-pubertal stage, which suggest that skeletal sex determination using the pelvis is possible with confidence before the previously thought post-pubertal period. Following Lewis and colleagues' (2016a) pubertal stage designations (see Table 3.6), the maturation stage is not considered post-pubescent in this study since individuals in this stage do not have a completely fused distal radius and iliac crest. The pubertal growth spurt is considered complete with the full fusion of long bone epiphyses and pelvic epiphyses (Marshall and Tanner 1986). Subpubic concavity, ilium shape, pelvic inlet shape, and acetabulum size & orientation are traits that show a high degree of dimorphism beginning at the deceleration stage. Both overall and sex specific accuracies for the four traits are at or substantially above 80%. The acceleration stage, or beginning of puberty, occurs on average at 10 years of age in females and 12 years in males and the deceleration stage is reached, on average, at 15.2 years in females and 17.9 in males. Thus, it takes about five to six years from the time puberty starts for these traits to become sufficiently dimorphic for skeletal sex determination. The ilium shape and pelvic inlet shape may show substantial promise in individuals at the PHV stage (100% accuracy for both traits), but a low sample size of individuals represented in this pubertal stage for each trait (n= 5 and 2, respectively) precludes any definitive conclusions. Further research is required to determine if the high accuracy of the ilium shape and pelvic inlet shape in the PHV stage hold true. The ilium shape further shows some promise for individuals in an even earlier stage of pubertal growth (acceleration stage) since only one male was erroneously identified. Low sample size is also seen in the acceleration stage (n=5) thereby making it difficult to present any definitive conclusions. The pelvic inlet shape and acetabulum size & orientation exhibit a general

increase in overall accuracy as pubertal stages progress, and although the subpubic concavity and ilium shape show fluctuating accuracies, they remain above 85%. Thus, if pubertal stages can be assigned to adolescent skeletons, the subpubic concavity, ilium shape, pelvic inlet shape, and acetabulum size & orientation can be used to confidently determine the sex of individuals from the deceleration to post-puberty stages.

Four traits (pubic bone shape, greater sciatic notch angle, true pelvis size & shape, and obturator foramen shape) provide high sex classification accuracy from the maturation stage and onward. The overall accuracies for these traits range from 84-100% in the maturation stage, while sex specific classification accuracies range from around 80-100%. The maturation stage is reached, on average, at 18 years in females and 19 years in males. Given that in this sample the acceleration stage of puberty occurred at an average age of 10 and 12.3 years for females and males, respectively, it takes approximately 7-8 years from the time of pubertal on-set for these four pelvic traits to become sufficiently dimorphic for accurate sex assessment. While the overall accuracies of the true pelvis size & shape and obturator foramen shape show promising results for individuals in the PHV stage, their use for individuals in the PHV stage cannot be ascertained because of low sample sizes seen in this stage for each trait (n = 2 and 5, respectively). The ischiopubic ramus ridge exhibits high accuracy (89%) for the maturation stage and appears to only be successful on individuals in this pubertal stage. Despite achieving an overall accuracy of approximately 80% (79.5%) in the post-puberty stage, female accuracy for this trait is only 61.1% whereas male accuracy is 92.3%. Thus, the ischiopubic ramus ridge, if applied, should be applied cautiously only to individuals that are within the maturation stage of pubertal growth. Pubic bone shape, greater sciatic notch angle, true pelvis size & shape, and obturator foramen shape, however, can be applied for sex determination of individuals in the maturation and post-puberty stages.

The results of this research suggest that five pelvic traits do, in fact, become sexually dimorphic enough for sex differentiation after puberty. The preauricular sulcus, greater sciatic notch size & shape, development of muscle markings, sacrum shape, and auricular surface height showed overall and sex specific accuracies surpassing 80% in the post-puberty stage. Similar to the traits sufficiently dimorphic in the

maturation stage, the traits in the post-pubertal stage are sexually dimorphic enough (>80% accuracy) 7-8 years from the time of pubertal onset. The difference between traits in the maturation and post-puberty stages is that the preauricular sulcus, greater sciatic notch size & shape, development of muscle markings, sacrum shape, and auricular surface height require the full fusion of skeletal indicators of puberty for accurate sex determination. Although the sacrum shape achieves an overall accuracy of 84%, female sex classification is moderate at best (76.5%) and should therefore be used cautiously and possibly avoided if it is the only trait available for sex assessment. Preauricular sulcus, greater sciatic notch size & shape, development of muscle markings, and auricular surface height, on the other hand, can all be used for sex determination for individuals that are post-pubertal.

Four pelvic traits (dorsal pubic pitting, ventral arc, iliac crest curvature, and arch criterion) proved to be ineffective in all pubertal stages. Three of these traits failed to achieve a minimum overall accuracy of 80%. Arch criterion showed some promise for the post-puberty stage, but correct male sex classification was moderate (74%). Despite overall accuracies for the ventral arc surpassing 80% in the deceleration, maturation, and post-puberty stages, female accuracy was poor throughout with the highest accuracy being only slightly better than chance (59%). All four traits outlined here should certainly not be used if they are the only traits available for assessment. As noted above, the ventral or precursor arc may show promise if it is used in an archaeological assemblage to provide a minimum number of females as represented by the presence of the precursor arc.

When all 18 pelvic traits examined in this study are used in unison, sex determination can be conducted with a high level of accuracy on individuals in the deceleration stage and onward, with increasing accuracy as pubertal stages progress. The pattern of increasing accuracy with the progression of puberty and high accuracy beginning at the deceleration stage are seen when all 18 traits are weighted equally and when subadult traits are weighted more. When the latter is applied, the deceleration stage exhibits an increase in accuracy of approximately 7%. There is also promising results for the PHV stage when all 18 traits are used for skeletal sex assessment, but low sample size (n=5) prevents any confident conclusion.

Low accuracy of correctly identifying female skeletons has been attributed to pro-male bias prior to puberty given that most traits, particularly of the pelvis, are based on secondary sex characteristics (Rogers 2009; Wilson and Humphrey 2017). The results of this research have proven that, as a whole, this pattern holds true since high accuracy in morphological traits and overall sex determination is seen, at the earliest, in the deceleration stage. While some pelvic traits show promise of being sufficiently dimorphic by PHV, there are no pelvic traits with high accuracy prior to the onset of puberty. This research, however, has shown that the completion of puberty is not required for all pelvic traits to express full dimorphism. Instead, passing peak height velocity might, in fact, be more important that being post-pubertal. Differences in female pelvic growth at puberty are believed to be restricted to regions directly involved with the development of the birth canal (Wood and Chamberlain 1986; Scheuer and Black 2000) and developmental differences occurring at the time of puberty results in the differences in the subpubic concavity between males and females (Coleman 1969). This, in part, can be seen with the high level of dimorphism in the pelvic inlet shape and subpubic concavity during the deceleration stage. True pelvis size & shape and sacrum shape, two other traits related to the pelvic cavity, show high dimorphism at the later stages of puberty (maturation and post-puberty, respectively). Ilium shape and acetabulum size & orientation, however, are not necessarily directly related to the birth canal but exhibit a high degree of dimorphism in the deceleration stage. Thus, this study provides evidence that pelvic growth is not necessarily *restricted* to regions of the birth canal. Scheuer and Black (2000) have suggested that starting around mid-puberty, sexual dimorphism in the pelvis is probably advanced enough in the females to allow reliable assessment of sex. If acceleration marks the beginning of puberty and maturation marks the end, puberty (in this sample) lasts approximately 8 years in females. Since sex determination using all 18 morphological traits exhibits high accuracy starting at the deceleration stage, or five years after the onset of puberty, this research has proved that Scheuer and Black (2000) were, approximately, correct.

With respect to metric methods for sex determination, Albanese's (2003) logistic regression models 3 and 8 show high overall accuracy for the PHV, deceleration, maturation, and post-puberty stages. Although both logistic regression models show

promising results for the around PHV stage, their efficacy on subadults in this stage cannot be ascertained due to the low sample size in this pubertal stage for both models (n = 2). While these models also seem promising for the deceleration to post-puberty stages, a decrease in correct female classification is observed with the progression of puberty. This may be the result of a higher number of females represented in each progressive pubertal stage, thereby providing a clearer pattern of male and female sex classification. Logistic regression models 1, 4, 10, 11, and 20 show high overall accuracy for deceleration, maturation, and post-puberty. Models 20 and 4 have high male and female classification accuracies in the deceleration and post-puberty stages, but a marked decrease in female accuracy is seen in the maturation category. This pattern does seem confounding and while it could be explained by a higher number of females in the maturation stage than the deceleration stage, it would not account for the subsequent increase in accuracy in the post-puberty stage, which has more females compared to the maturation stage. The remaining three regression models (models 1, 10, and 11) show problematic accuracies for at least one sex in all pubertal stages, where sex classification is moderate at best (75%). Due to the problematic sex specific accuracies, models 1, 10, and 11 should not be used for subadult sex determination. Although models 11 and 1 have perfect accuracy for the around PHV stage, the sample size for this pubertal stage is too low for both regression models (n=3 and 2, respectively) to provide any conclusions. Model 5 shows some promise for accurate sex determination for individuals only in the maturation stage given that overall and sex specific accuracies meet and surpass 80%. Finally, logistic regression model 2 shows the least promise despite high overall accuracy in the maturation and post-puberty stages, given that the highest female classification accuracy of all pubertal stages is 73%. This model should, therefore, not be used for subadult sex determination.

5.4 Puberty Stage Versus Age Groups

Overall, the efficacy of most traits (n=8) exhibits substantial overlap between the age range represented by both “traditional” age groupings and pubertal stages. The auricular surface height, pubic bone shape, preauricular sulcus, greater sciatic notch size & shape, greater sciatic notch angle, true pelvis size & shape, and obturator

foramen all show high overall and sex specific accuracies in the late adolescent cohort and either the maturation and/or post-puberty stages. Thus, there is agreement in age and developmental categories that the aforementioned traits are most successful when applied to individuals 17-20 years of age. Although the age range for post-puberty females is 14-20 years of age, only one female is 14-years-old and two females are 16-years-old. Thus, most females in the post-puberty stage are, in fact, 17-20 year olds. Ilium shape exhibits high accuracy (overall and sex specific) in the adolescent and late adolescent cohorts and in the deceleration to post-puberty stages. There is overlap in ages, then, that suggests the ilium shape can be used in adolescents 14-20 years of age. The age distribution for the adolescent cohort, however, extends to individuals that are 12 years of age, which would suggest that using age groups might be better for scoring the ilium shape as it would include younger individuals compared to those included in the deceleration pubertal stage (14-19 year olds). If future research is conducted on a sample with a high number of individuals in the PHV stage for the ilium shape, which includes 12-year-old females and 13-15 year old males, and yield promising results, then the ilium shape could, in fact, be useful for individuals between 12-20 years of age, regardless of age categorization.

When comparing age groups and puberty stages for three traits (subpubic concavity, pelvic inlet shape, and acetabulum size & orientation), the results of this study show that these traits can be applied to younger individuals when separated by pubertal stage compared to age groups (Appendix A.1). Subpubic concavity, pelvic inlet shape, and acetabulum size & orientation exhibit their highest overall and sex specific accuracies in the late adolescent group (17-20 years). These traits, however, also show high accuracy in the deceleration stage. In females, this pubertal stage corresponds to 14-16 years of age. This is particularly important for the subpubic concavity given that this traits minimum age of stabilization in females is at the age of 14. Thus, if it is possible to assess the pubertal stage of an adolescent and they are in the deceleration stage, it is possible to assess sex with high accuracy of individuals who may in fact be younger than what is suggested by the age component of this research.

Development of muscle markings and sacrum shape provided poor-moderate accuracies when using traditional age categories, so much so that they were not

recommend for sex determination. Development of muscle markings, however, appears to be effective on individuals in the post-puberty stage (87% accuracy). Sacrum shape, on the other hand, can be cautiously applied to post-pubertal individuals since the female accuracy just fails to meet 80% minimum (approximately 77%). This suggests that puberty, and the social responsibilities that come with it, may impact these two traits more so than just chronological age. The arch criterion, iliac crest curvature, and dorsal pubic pitting show agreement with age categories and pubertal stages that, due to low overall accuracies in all groupings, they are not suitable for subadult sex determination, especially on their own. When all 18 morphological traits are available for sex determination, there appears to be a general agreement that sex can be determined with high accuracy (over 90%) among 17-20 year olds, which corresponds to the late adolescent group and the maturation and post-puberty stages (Appendix A, Table A-19). However, in order to confidently conduct sex assessments on individuals younger than 17 years of age, particularly females, it appears that pubertal stages may be a better age classification method to use compared to age categories given the high overall and sex specific accuracies (80% and higher) achieved in the deceleration stages, which includes 14-16 year old females (Table A-19, Appendix A).

Differences in pubertal stage and age groups appear to be more complex for metric methods of sex determination. Model 20 shows promise for age groupings, being most effective for the late adolescent group (17-20 years of age). Despite having similar age ranges, Model 20 is not effective for individuals in the maturation stage because of a substantial male bias, where female accuracy fails to achieve 70%. Model 8, on the other hand, shows some consistency since this model appears to be effective on adolescent individuals (12-16.9 years) and individuals in the deceleration stage only (14-16 in females, 17-19 males). The discrepancy seen between pubertal stages and age groupings for Model 20 may be the result of a reallocation of late adolescent females (17-20 year old) into two categories (maturation and post-puberty) for the pubertal stage analysis. Consequently, the redistribution of late adolescent females resulted in a lower number of females in the maturation and post-puberty stages (n= 6 and 18, respectively) compared to the late adolescent age category (n=23) [see Appendix A for comparative tables]. Moreover, of the three females 17-20 years of age

with incorrect sex classification, two are in the maturation stage, which has a low sample size (n=6), and one is in the post-puberty stage. Despite only two incorrect sex classifications for females in the maturation stage, with a sample size of six, the accuracy for females in that puberty stage appears markedly lower than its corresponding age category. A sample with a large number of females in the maturation stage could help determine if the low sex classification for this pubertal stage seen in this study is, in fact, the result of low sample size. Conversely, it is also possible that measureable size differences between the sexes are more closely related to growth as opposed to development. Therefore, age categories may be better equipped to capture measureable size related patterns (growth) of sexual dimorphism compared to developmental stages (puberty stages). Thus, until further research can be conducted that examines the applicability of Albanese's (2003) method to adolescents it is likely best to conduct sex assessments using morphological traits than metric measurements.

SUMMARY

The appearance and stabilization of sexual dimorphism in the growing pelvis is, unsurprisingly, a complex process. This research has shown that male and female expressions do not appear and stabilize at the same time for all pelvic traits, thereby suggesting that age-related trends in sexually dimorphic pelvic traits do exist. While the applicability of most pelvic traits (n=8) shows a concordance between chronological age groupings and pubertal stages, three traits can be successfully used in younger individuals, particularly females, if pubertal stages are used rather than chronological age. Moreover, two traits show promising results only when pubertal stages are applied, but are not applicable at all when assessed with respect to chronological age. Albanese's (2003) metric method of sex determination shows less agreement between age groupings and pubertal stages, complicating its use for subadult sex determination. Due to the complexity of choosing which pelvic traits are most effective when applying age categories versus pubertal stages, a summary of best practices for morphological and metric subadult sex assessment using the pelvis is presented in Figure 5.1.

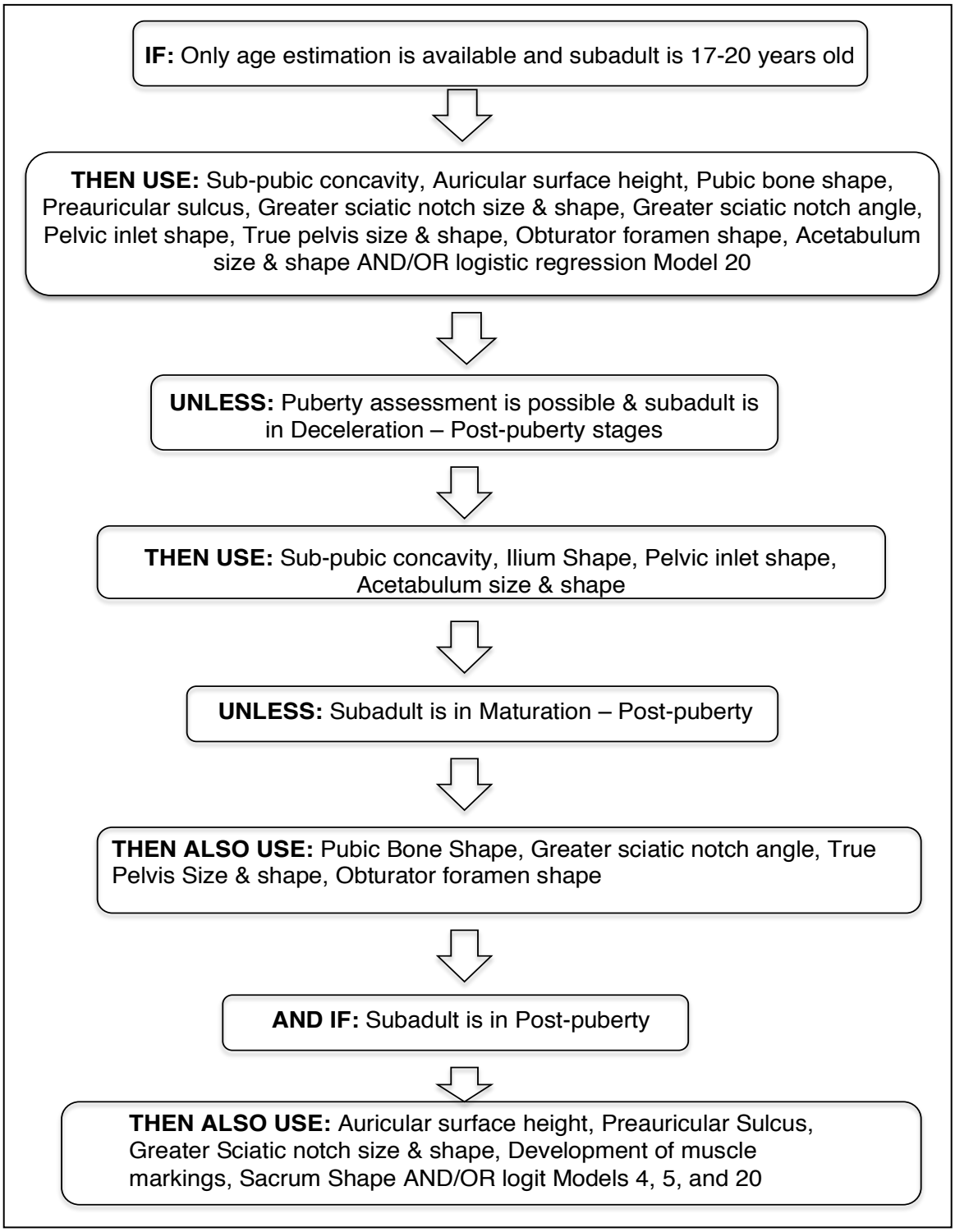

Figure 5.1: Summary of best practices when conducting sex assessments on subadults using pelvic traits

Chapter 6 - Conclusion

The purpose of this research was to explore one potential confounding factor impacting subadult sex determination, namely age-related variation in the expression of sexual dimorphism in the pelvis. Sex determination in subadult skeletons has undoubtedly remained the most problematic area in juvenile osteology given the variability in success when testing purported sexually dimorphic pelvic traits between investigators (Lewis 2018). The development of sexual dimorphism in the human skeleton is complex and the ability to determine the sex of subadults is likely influenced by factors such as health and nutrition, differences in age related patterns of development, population differences, and differences in the duration of skeletal growth (Wilson and Humphrey 2017). Through the use of two skeletal collections with documented age-at-death and sex, this research tracked the appearance and stabilization of a comprehensive list of 18 sexually dimorphic morphological pelvic traits, as well as nine logistic regression equations for adult sex determination. Thus, this research was the first to examine age related trends of sexual dimorphism in the pelvis and the efficacy of these morphological traits and logistic regression equations in subadults.

An oft-cited prerequisite for the full expression of sexual dimorphism in the pelvis is the cessation of puberty (Lewis, 2007; Rogers, 2009; Klales and Burns, 2017), which was also explored in this research. This study is the first to examine the relationship between osteological indicators of pubertal stages and sexual dimorphism. By incorporating pubertal stage analysis, this research provided a means to assess whether “traditional” age or developmental groupings encompass a wider age range that allows for successful sex determination in subadult individuals.

6.1 Revisiting Research Objectives & Hypotheses

The primary and overarching objective of this research was to identify the ages of appearance and stabilization of sexually dimorphic traits in the pelvis. There has been minimal research (Sutherland and Suchey 1993; Rogers 2009; Klales and Burns 2017) examining the earliest age at which sexually dimorphic traits used in adults can be applied to subadults that has yielded promising results. This research, therefore,

contributes a novel strategy to addressing this issue. The analysis and results provided in this research project found that three general patterns of pelvic sex trait development emerge. First, pelvic traits can demonstrate a male “default” expression at birth with female expression appearing later on in growth, particularly in the teenage years. The second pattern observed shows a female “default” in trait expression, with male expression appearing in the teenage years. These first two patterns of appearance and stabilization of trait expression correspond to divergent ontogenetic trajectories outlined by Wilson and colleagues (2015). Third, this research shows traits can display female and male expression by birth, with sex differences becoming more pronounced with age, thereby exemplifying parallel ontogenetic trajectories (Wilson et al. 2015). Similarly, this research complements previous studies, which have argued that some sex differences exist in traits that require the fusion of pelvic elements and can be seen by the time fusion has occurred. This research has partially substantiated Wilson and Humphrey’s (2017) suggestion that applying the same age groups *a priori* to multiple sexually dimorphic traits may not be ideal given that the onset of dimorphism is not necessarily the same across all pelvic traits, which is observed in this study. Future testing of these pelvic traits should therefore consider applying trait-dependent age groupings to minimize the level of bias in successful sex classification. The presence of the three general patterns in the appearance and stabilization of sexually dimorphic traits in the pelvis made it possible to reject the null hypothesis proposed for this research objective, that age related variation does not exist in the expression of sexual dimorphism. The first null hypothesis was further rejected because many morphological pelvic traits, in addition to the metric method of sex determination, also exhibited an increase in accuracy as age progressed.

The second objective set out in this research was to examine the relationship that exists between the efficacy of sexually dimorphic traits of the pelvis at predicting sex and pubertal stages. Substantial previous research has concluded that sexual dimorphism in the human skeleton can only be seen in adulthood, after puberty has been attained (Rissech and Malgosa 2005; Rogers 2009; Klales and Burns 2017; Stull et al. 2017). The research presented here has shown that this is not necessarily true for all traits in the human pelvis. The subpubic concavity, ilium shape, pelvic inlet shape,

and acetabulum size & orientation are all traits that show high sex classification accuracy in the deceleration stage of the pubertal growth curve. An additional five traits show high sex classification accuracy in the maturation stage, or tail end, of puberty. This research, however, does not completely discredit the argument that full expression of dimorphism occurs in adulthood since five traits show high accuracy only in the post-pubertal stage. However, when the 18 traits examined in this study are used together, accurate sex determination can be conducted on individuals as early as the deceleration stage. Anderson (1990) had boldly claimed that, prior to the onset of puberty, the pelvic bones of males and females are “virtually indistinguishable”. This study offers a partial challenge to Anderson’s (1990) claim by identifying some pelvic differences between the sexes before puberty, but accepting that those differences are not nearly strong enough to determine skeletal sex with confidence. The second null hypothesis tested in this study, that sexual dimorphism in the pelvis cannot be seen until adulthood (or after puberty) can be rejected given the presence of high sex classification accuracy in the deceleration stage for four pelvic traits and for overall sex determination. This study has shown that surpassing peak height velocity (PHV) appears to be more important than the post-pubertal period for the pelvis to display sufficient dimorphism to distinguish between the sexes. This is further supported by the fact that six traits (ilium shape, pelvic inlet shape, pubic bone shape, greater sciatic notch angle, true pelvis size & shape, and obturator foramen) show preliminary promise of displaying substantial dimorphism in the PHV stage. Overall sex determination using all 18 traits also shows preliminary promise in the peak height velocity stage. The low sample sizes in this pubertal stage (highest n=5), however, prevent any definitive conclusions. Patterns between metric sex determination and pubertal stage are less clear, thus when pubertal analysis is possible morphological sex determination seems to be the better option compared to the metric method tested in this research.

In summary, this research has demonstrated that:

- Age-related trends in the appearance and expression of sexually dimorphic pelvic traits exist
    - Three general patterns emerge:
        1. Male default expression with subsequent female expression
        2. Female default expression with subsequent male expression

3. Male and female expression by birth or pelvic fusion with subsequent increase in dimorphism

- The post-pubertal period is not required for the full expression of sexual dimorphism of all pelvic traits
    - Surpassing peak height velocity is more important; high sex classification accuracy can be observed in the deceleration stage

- Albanese's (2003) metric method can be applied for subadult sex determination to some degree but no clear relationship is observed with puberty

6.2 Limitations of this Research

While this study has provided promising results, particularly for sex determination of adolescent skeletal remains, this research does have a couple of limitations. Overall, there are an unequal number of males and females in the sample analyzed. Of the 128 subadults included in this study, 51 are female and 77 are male. The uneven sex distribution in the study sample is more problematic in the infant and child age categories than the adolescent and late adolescent groups. Seemingly promising results in the younger age categories are, more often than not, the result of high sex classification of male individuals only. Females are either not represented in these age groups or are represented by one or two individuals whose sex was erroneously identified. The issue of uneven sex, however, is ameliorated in the adolescent and late adolescent categories. Although males continue to outnumber females in these age groups, the statistical testing conducted (chi-square) in this study ensures that the patterns observed in the later age groups are not the result of chance and, thus, minimizes the impact of uneven sex distribution.

A limitation that impacts this research to a greater extent is an unequal sampling of ontogeny. Two ages (2 and 9 years) are not represented in the study sample used in this research. Additionally, only females are represented in five ages (5-, 7-, 12-, 14-, and 16-year-olds) and only males are represented in three ages (3-, 13-, and 15-year-olds). The unequal representation of sexes at these specific ages impact clear age related trends, particularly during the early to mid-teenage years. It is possible that the minimum age of trait stabilization occurs earlier than what is reported here but the unequal representation of sexes at certain ages obscures those patterns. Moreover, the skew in age distribution towards older adolescents likely increases the average age of

trait stabilization reported, making it appear that there is a substantial delay between the minimum and average age of trait stabilization for certain traits. Unequal sampling is also observed in the pubertal stages, where the acceleration and around PHV stages are both represented by only 5 individuals. In contrast, the sample sizes in the other pubertal stages are at least 3 times greater or more than the PHV stage. Unequal sampling of ontogeny is a limitation that is commonly seen in juvenile osteology (Wilson and Humphrey 2017) and needs to be recognized. Despite the unequal sampling, however, the trends observed for successful sex classification (particularly in relation to pubertal stages) are still found to be statistically significant. Given that patterns still emerge regardless of the limitations outlined above, if these limitations can be minimized in future research, the patterns observed in this study might be stronger and the ages of stabilization may occur earlier than what is observed in this study.

6.3 Future Research

This research has provided a number of avenues for further research to clarify the patterns observed. Further research is required to determine whether the age related, and developmental, patterns observed in this study are also found in other populations. This research could be expanded to other skeletal collections of documented age-at-death and sex that contain a substantial subadult sample. Such documented skeletal collections include: the Luis Lopes Collection and the Coimbra Identified Skeletal Collection in Portugal, the Certosa Bologna Collection and the CAL Milano Cemetery Skeletal Collection in Italy, and the Grenada Osteological Collection in Spain. Examining the 18 morphological traits used in this study, and their relationship with puberty stages, on collections like these can help determine if the patterns observed in this research can be detected in other skeletal samples with large subadult sample sizes. Expanding this research to include other samples could help better understand the applicability of the trends observed in this study to other groups. Given that the majority of the collections identified above are contemporaneous to the Hamann-Todd and Terry collections (early twentieth century), it may be possible to combine the samples, ensuring date of birth and/or death overlap between all individuals, to observe patterns of appearance and stabilization of sexually dimorphic traits that are population independent. Conducting similar research on multiple skeletal

collections will also clarify whether developmental stages or age groupings are best for subadult sex determination.

Continued work on understanding the development of sexual dimorphism in the subadult skeleton is required in order to identify which regions of the subadult skeleton provide the most promise for sex determination. While ordinal scales to describe the expression of sexually dimorphic traits may prove to be somewhat problematic for subadult sex determination because it has the potential of increasing the number of "indeterminate" cases, the use of this type of scoring system for sex assessment may prove useful to understand the development of sexual dimorphism. The use of an ordinal scale for trait expression, like that used by Klales and Burns (2017), has the potential to clearly identify the age, or stage of development, at which sexual dimorphism becomes most pronounced. This could be done particularly through the use of a three-point scale of female expression, intermediate expression, and male expression. The inclusion of an intermediate expression may help identify the critical time at which this form of expression is significantly reduced and subsequently followed by an increase in the divergence of the male and female forms. This could then clearly identify the nuances in the appearance of sexual dimorphism in the subadult skeleton throughout growth and development.

One fortunate outcome of this research was the identification of a potential trait, identified and referred to here as "symphyseal twisting", for female sex identification. The presence of the "symphyseal twisting" observed in this research, however, must be examined in other documented skeletal collections to determine whether this trait is expressed in other populations or is specific to the females of the Hamann-Todd collection. As discussed in section 5.2, the presence or absence of the precursor arc and/or symphyseal twisting does not appear to be an adequate indicator of sex in subadult individuals. However, the exclusive presence of the precursor arc and symphyseal twisting in females may prove to be beneficial in archaeological assemblages and be used to identify a minimum number of subadult females in an assemblage. Should similar patterns for this trait be seen in other collections, the symphyseal twisting and precursor arc could have the potential to identify the minimum number of females between 4 and 20 years of age in archaeological contexts.

One final avenue for future research this project generated is the exploration of another potential confounding factor of subadult sex determination, particularly health and nutrition (Cardoso and Saunders 2008; Wilson and Humphrey 2017). Poor nutrition and disease have the potential to impact the timing of puberty and the successful progression through pubertal stages (Proos and Gustafsson 2012; Lewis et al. 2016). Chronic disease, in particular, has been shown to delay pubertal attainment in modern populations (Lewis et al. 2016). Given the delay in menarche found in this study, compared to both modern and temporally comparable standards, examining the health of the subadults in the Hamann-Todd and Terry collections may provide evidence confirming the impact health has on pubertal stage attainment. Moreover, given the strong relationship between the deceleration stage of puberty, as well as the potential of the PHV stage, and sexual dimorphism found in this study, there exists the possibility that adverse health may also impact the appearance and stabilization of sexually dimorphic traits in the pelvis. In turn, the impact of stress may be a further consideration that needs to be taken into account when attempting to conduct and explore other avenues of subadult sex determination. The current research, then, represents only the start of our understanding of the complex nature of sexual dimorphism in the human skeleton.

References

Aksglaede L, Olsen LW, Sorensen TIA, Juul A. 2008. Forty years trends in timing of pubertal growth spurt in 157,000 Danish school children. PLoS ONE 3(7): e2728.

Albanese J. 2003. A metric method for sex determination using the hipbone and femur. Journal of Forensic Sciences 48: 263-273.

Anderson BE. 1990. Ventral arc of the os pubis: anatomical and developmental considerations. American Journal of Physical Anthropology 83: 449-458.

Arsuaga JL, Carretero JM. 1994. Multivariate analysis of the sexual dimorphism of the hip bone in a modern human population and in early homonids. American Journal of Physical Anthropology 93: 241-257.

Arthur NA, Gowland RL, Redfern RC. 2016. Coming of age in Roman Britain: osteological evidence of pubertal timing. American Journal of Physical Anthropology 159: 689-713.

Baker BJ, Dupras TL, Tocheri MW. 2005. The Osteology of Infants and Children. College Station, TX: Texas A&M University Press.

Baxter JE. 2005. The Archaeology of Childhood: Children, Gender, and Material Culture. Walnut Creek : AltaMira Press.

Baxter JE. 2008. The archaeology of childhood. Annual Review of Anthropology 37: 159–175.

Blake KAS. 2011. An investigation of sex determination from the subadult pelvis: a morphometric analysis. Unpublished Doctoral Dissertation. University of Pittsburg

Bogin B. 1999. Patterns of human growth. Cambridge UK: Cambridge University Press.

Boucher BJ. 1955. Sex differences in the foetal sciatic notch. Journal of Forensic Medicine 2:51–54.

Boucher BJ. 1957. Sex differences in the foetal pelvis. American Journal of Physical Anthropology 15:581–600.

Budinoff LC, Tague RG. 1990. Anatomical and developmental bases for the ventral arc of the human pubis. American Journal of Physical Anthropology 82: 73-79.

Buikstra JE, Ubelaker DL. 1994. Standards for Data Collection from Human Skeletal Remains. Arkansas Archaeological Survey research series no. 44. Fayetteville: Arkansas Archaeological Survey.

Caffey J. 1993. Caffey's Pediatric X-ray Diagnosis, 9th edition. St. Louis MO: Mosby Yearbook.

Camacho FF, Pellico L, Rodriguez R. 1993. Osteometry of the human iliac crest: patterns of normality and its utility in sexing human remains. Journal of Forensic Sciences 38(4): 779-787.

Cardoso HFV, Saunders SR. 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: 24-29.

Carlson BM. 2013. Human Embryology and Developmental Biology. Amsterdam: Elsevier.

Chertkow S. 1980. Tooth mineralization as an indicator of the pubertal growth spurt. American Journal of Orthodontics 77: 79-91.

Chumlea WC, Schubert CM, Roche AF, Kulin HE, Lee PA, Himes JH, Sun SS. 2013. Age at menarche and racial comparisons in US girls. Pediatrics 111: 844-850.

Cole HL, Ohanian LE. 2007. A second look at the US Great Depression from a neoclassical perspective. Great Depression of the Twentieth Century. Minneapolis: Federal Reserve Bank of Minneapolis, p. 21-58.

Coleman WH. 1969. Sex differences in the growth of the human bony pelvis. American Journal of Physical Anthropology 31: 125–152.

Cutler GB. 1997. The role of estrogen in bone growth and maturation during childhood and adolescence. Journal of Steroid Biochemistry 61: 141-144.

De Vito C, Saunders SA. 1990. A discriminant function analysis of deciduous teeth to determine sex. Journal of Forensic Science 35: 845–848.

DiGangi EA, Moore MK. 2013. Research Methods in Human Skeletal Biology. Oxford: Elsevier Academic Press.

Doe DM, Rascon Perez J, Cambra-Moo O, Campo Martin M, Gonzalez Martin A. 2017. Assessing pubertal stage in adolescent remains: an investigation of the San Nicolas Maqbara burial site (Murica, Spain). Archaeological and Anthropological Sciences 11(2): 541-554.

Drennan RD. 2010. Statistics for Archaeologists: A Common Sense Approach. 2nd Edition. Springer Trade.

Dunsworth HM. 2020. Expanding the evolutionary explanations for sex differences in the human skeleton. Evolutionary Anthropology: 1-9.

Falys CG, Schutkowski H, Weston DA. 2005. The distal humerus – A blind test of Rogers' sexing technique using a documented skeletal collection. Journal of Forensic Sciences 50(6): 1-5.

Fazekas IG and Kosa F. 1978. Forensic Fetal Osteology. Budapest: Akademiai Kiado.

Fehling H. 1876. Die Form des Beckens beim Fötus und Neugeborenen und ihre Bezierhung zu der beim Erwachsenen. Arch Gynäk 10:1–80.

Flores-Mir C, Burgess CA, Champney M, Jensen RJ, Pitcher MR, Major PW. 2006. Correlation of skeletal maturation stages determined by cervical vertebrae and hand-wrist evaluations. Angle Orthodontist 76: 1-5.

Garvin HM. 2012. Adult sex determination: methods and application. In: DC Dirkmaat (ed) A Companion to Forensic Anthropology. Massachusetts: Wiley-Blackwell Publishing Ltd

Garvin HM, Sholts SB, Mosca LA. 2014. Sexual dimorphism in human cranial trait scores: effects of population, age, and body size. American Journal of Physical Anthropology 154: 259-269.

Genovés S. 1957. L'estimation des différences sexuelles dans l'os coxal; différences métriques et différences morphologiques, Bull. Mem. Soc. Anthropol. Paris 10 (1959) 3–95.

Granados JAT, Roux AVD. 2009. Life and death during the Great Depression. Proceedings of the National Academy of Sciences 106(41): 17290-17295.

Grave KC, Brown T. 1976. Skeletal ossification and the adolescent growth spurt. American Journal of American Orthodontics 69: 611-619.

Gravetter FJ, Wallnau LB. 2004. Statistics for Behavioral Sciences. Belmont CA: Thomson Wadsworth.

Greulich WW, Pyle SI. 1959. Radiographic atlas of skeletal development of the hand and wrist. Stanford: Stanford University Press.

Greulich WW, Thoms H. 1944. The growth and development of the pelvis of individual girls before, during, and after puberty. Yale Journal of Biology and Medicine 17(1): 91-97.

Griffin WW. (2005). African Americans and the color line in Ohio, 1915-1930. Columbus: Ohio State University Press.

Grumbach MM. 2000. Estrogen, bone, growth and sex: a sea change in conventional wisdom. Journal of Pediatric Endocrinology and Metabolism 13(Suppl 6): 1439-1455.

Hägg U, Taranger J. 1982. Maturation indicators and the pubertal growth spurt. American Journal of Orthodontics 82: 299-309.

Halcrow SE, Tayles N. 2008. The bioarchaeological investigation of childhood and social age: problems and prospects. Journal of Archaeological Method and Theory 15: 190-215.

Hassell B, Farmer AG. 1995. Skeletal maturation evaluation using cervical vertebrae. American Journal of Orthodontics and Dentofacial Orthopedics 107: 58-66.

Henderson CY, Padez C. 2017. Testting times: identifying puberty in an identified skeletal sample. Annals of Human Biology 44: 332-337.

Hewitt D, Acheson R. 1961. Some aspects of skeletal development through adolescence. I. Variation in the rate and pattern of skeletal maturation at puberty. American Journal of Physical Anthropology 19: 321-331.

Holcomb SMC, Konigsberg LW. 1995. Statistical study of sexual dimorphism in the human fetal sciatic notch. American Journal of Physical Anthropology 97:113–125.

Horbaly HE, Kenyhercz MW, Hubbe M, Steadman DW. 2019. The influence of body size on the expression of sexually dimorphic traits. Journal of Forensic Sciences 64: 52-57.

Houston W. 1980. Relationship between skeletal maturity estimated from hand-wrist radiographs and the timing of the adolescent growth spurt. European Journal of Orthodontics 2: 81-93.

Humphrey LT. 1998. Growth patterns in the modern human skeleton. American Journal of Physical Anthropology 105: 57-72.

Hunt DR. 1990. Sex determination in the subadult ilia: an indirect test of Weaver's non-metric sexing method. Journal of Forensic Science 35:881–885.

Hunt DR, Albanese J. 2005. History and demographic composition of the Robert J. Terry anatomical collection. American Journal of Physical Anthropology 127: 406-417.

Hunt EE, Gleiser I. 1955. The Estimation of age and sex of preadolescent children from bones and teeth. American Journal of Physical Anthropology 13: 479-487.

Huseynov A, Zollikofer CPE, Coudyzer W, Gascho D, Kellenberger C, Hinzpeter R, Ponce de Leon MS. 2016. Developmental evidence for obstetric adaptation of the human female pelvis. Proceedings of the National Academy of Sciences 113(19): 5227-5232.

Iscan MY, Steyn M. 2013. The Human Skeleton in Forensic Medicine. Illinois: Charles C Thomas Publisher Ltd.

Kaplowitz P. 2006. Pubertal development in girls: secular trends. Current Opinion in Obstetrics and Gynecology 18: 487-491.

Karapanou O, Papadimitriou A. 2010. Determinants of menarche. Reproductive Biology and Endocrinology 8(1): 115.

Kirchengast S. 2014. Human sexual dimorphism – a sex and gender perspective. Anthropologischer Anzeiger 71: 123-133.

Klales AR. 2020. Practitioner preferences for sex estimation from human skeletal remains. In: Klales AR (ed), Sex Estimation of the Human Skeleton: History, Method, and Emerging Techniques. San Diego: Academic Press.

Klales AR, Burns TL. 2017. Adapting and applying the Phenice (1969) adult morphological sex estimation techniques to subadults. Journal of Forensic Sciences 62: 748-752.

Klales AR, Ousley SD, Vollner JM. 2012. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. American Journal of Physical Anthropology 149: 104-114.

Knickmeyer RC, Baron-Cohen S. 2006. Fetal testosterone and sex differences. Early Human Development 82: 755-760.

Kousser JM. 1980. Separate but not equal: the Supreme Court's first decision on racial discrimination in schools. The Journal of Southern History 46(1): 17-44.

Krogman WM. 1962. The Human Skeleton in Forensic Medicine. CC Thomas.

Kusmer KL. 1978. A Ghetto Takes Shape: Black Cleveland, 1870-1930. Urbana: University of Illinois Press.

Lamer M, Spake L, Cardoso HFV. 2021. Testing the cross-applicability of juvenile sex estimation from diaphyseal dimensions. Forensic Science International 321: 110739.

Lamparski D. 1972. Skeletal age assessment utilizing cervical vertebrae. Pittsburgh: University of Pittsburgh.

Landis R, Koch G. 1977. The measurement of observer agreement from categorical data. Biometrics 33(1): 159-174.

LaVelle M. 1995. Natural selection and developmental sexual variation in the human pelvis. American Journal of Physical Anthropology 98: 59-72.

Lesciotto KM, Doershuk LJ. 2017. Accuracy and reliability of the Klales et al. (2012) morphoscopic pelvic sexing method. Journal of Forensic Sciences 63(1): 214-220.

Lewis M. 2007. The Bioarchaeology of Children: Perspectives from Biological and Forensic Anthropology. Cambridge: Cambridge University Press.

Lewis ME. 2018. Children in bioarchaeology: methods and interpretations. In: A Katzenberg and A Grauer (eds), Biological Anthropology of the Human Skeleton. New York: Academic Press pp. 119-143.

Lewis ME, Shapland F, Watts R. 2016a. On the threshold of adulthood: a new approach for the use of maturation indicators to assess puberty in adolescents from medieval England. American Journal of Human Biology 28: 48-56.

Lewis ME, Shapland F, Watts R. 2016b. The influence of chronic conditions and environment on pubertal development. An example from medieval England. International Journal of Paleopathology 12: 1-10.

Lillehammer G. 1989. A child is born. The child's world in an archaeological perspective. Norwegian Archaeological Review 22 (2): 89-105.

Liversidge HM, Smith BH, Maber M. 2010. Bias and accuracy of age estimation using developing teeth in 946 children. American Journal of Physical Anthropology 143: 545-554.

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

Louis GMB, Gray LE, Marcus M, Ojeda SR, Pescovitz OH, Witchel SF, Sippell W, Abbott SH, Soto A, Tyl RW. 2008. Environmental factors and puberty timing: expert panel research needs. Pediatrics 121 (Suppl 3): S192-S207.

Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ. 1985. Multifactorial determination of skeletal age at death: a method and blind test of its accuracy. American Journal of Physical Anthropology 68: 1-14.

Luna LH, Aranda CM, Santos AL. 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: 898-911.

Marks C. 1989. Farewell – We're Good and Gone: The Great Black Migration. Bloomington: Indiana University Press.

Marshall W, Tanner J. 1986. Puberty. In: Falkner F, Tanner J, (eds). Human Growth. A Comprehensive Treatise. New York: Plenum Press. p. 171-209.

Massey DS, Rothwell J, Domina T. 2009. The changing bases of segregation in the United States. The Annals of the American Academy of Political and Social Sciences 626(1): 74-90.

McFadden C, Oxenham MF. 2018. Sex, parity, and scars: a meta-analytic review. Journal of Forensic Sciences 63(1): 201-206.

Meindl RS, Lovejoy CO, Mensforth RP, Carlos LD. Accuracy and direction of error in the sexing of the skeleton: implications for paleodemography. American Journal of Physical Anthropology 68: 79-85.

Miles AEW, Bulman JS. 1995. Growth curves of immature bones from a Scottish island population of sixteenth to mid-nineteenth century: Shoulder girdle, ilium, pubis and ischium. International Journal of Osteoarchaeology. 5(1): 15-27.

Mittler DM, Sheridan SG. 1992. Sex determination in subadults using auricular surface morphology: a forensic science perspective. Journal of Forensic Sciences 37: 1068-1075.

Molleson T, Cruse K, Mays S. 1998. Some sexually dimorphic feature of the human juvenile skull and their value in sex determination in immature skeletal remains. Journal of Archaeological Science 25: 719–728.

Moore DS, McCabe GP, Craig BA. 2009. Introduction to the Practice of Statistics. New York: W.H Freeman and Company.

Muller JL, Pearlstein KE, de la Cova C. 2017. Dissection and documented skeletal collections: embodiments of legalized inequality. In: Nystrom KC (ed) The Bioarchaeology of Dissection and Autopsy in the United States. Switzerland: Springer International Publishing.

Nettle D. 2002. Women's height, reproductive success and the evolution of sexual dimorphism in modern humans. Proceedings of the Royal Society 269: 1919-1923.

Nichol CR, Turner CG. 1986. Intra- and Interobserver concordance in classifying dental morphology. American Journal of Physical Anthropology 69(3): 299-315.

Olivares JI, Aguilera IA. 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 ways of analyzing and interpreting the results. International Journal of Legal Medicine 130: 1623-1632.

Ousley SD. 1995. Should we estimate biological or forensic stature? Journal of Forensic Sciences 40(5): 768-773.

Özer T, Kama JD, Özer SD. 2006. A practical method for determining pubertal growth spurt. American Journal of Orthodontics and Dentofacial Orthopedics 130: 131-136.

Parker GJ, Yip JM, Eerkens JW, Salemi M, Durbin-Johnson B, Kiesow C, Haas R, Buikstrak JE, Klaus H, Regan LA, Rocke DM, Phinney BS. 2019. Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. Journal of Archaeological Science 101: 169-180.

Phenice TW. 1969. A newly developed visual method of sexing the os pubis. American Journal of Physical Anthropology 30: 297-301.

Phillips KL. 1996. “But it is a fine place to make money”: migration and African-American families in Cleveland, 1915-1929. Journal of Social History 30: 393-413.

Phillips KL. 1999. Alabama North: African-American migrants, community, and working class activism in Cleveland, 1915-1945. Urbana: University of Illinois Press.

Plavcan JM. 2012. Body size, size variation, and sexual size dimorphism in early Homo. Current Anthropology 44: 25-53.

Proos L, Gustafsson J. 2012. Is early puberty trigged by catch-up growth following undernutrition? International Journal of Environmental Research and Public Health 9(5): 1791-1809.

Quigley C. 2001. Skulls and Skeletons: Human Bone Collections and Accumulations. Jefferson NC: McFarland & Company Inc.

Ranganathan P, Pramesh CS, Aggarwal R. 2017. Common pitfalls in statistical analysis: measures of agreement. Perspectives in Clinical Research 8(4): 187-191.

Reynolds EL. 1947. The bony pelvis in prepubertal childhood. American Journal of Physical Anthropology 5: 165-200.

Rissech C, Malgosa A. 2005. Ilium growth study: applicability in sex and age diagnosis. Forensic Science International 147: 165-174.

Rogers TL. 1999. The Attribution of Ancestry for European and Indian (South Asian) Individuals within a Forensic Context. Unpublished doctoral thesis. Simon Fraser University.

Rogers TL. 2009. Sex determination of adolescent skeletons using the distal humerus. American Journal of Physical Anthropology 140: 143-148.

Rogers TL, Saunders SR. 1994. Accuracy of sex determination using morphological traits of the human pelvis. Journal of Forensic Sciences 39 (4): 1047-1056.

Rogol AD, Clark PA, Roemmich JN. 2000. Growth and pubertal development in children and adolescents: effects of diet and physical activity. American Journal of Clinical Nutrition 72(2): 521S-528S.

Scheuer L. 2002. A blind test of mandibular morphology for sexing mandibles in the first few years of life. American Journal of Physical Anthropology 119: 189-191.

Scheuer L, Black S. 2000. Developmental Juvenile Osteology. New York: Elsevier Academic Press.

Schutkowski H. 1993. Sex determination of infant and juvenile skeletons. I. morphognostic features. American Journal of Physical Anthropology 90: 199-205.

Shapland F, Lewis ME. 2013. Brief communication: a proposed osteological method for the estimation of pubertal stage in human skeletal remains. American Journal of Physical Anthropology 151: 302-310.

Shapland F, Lewis ME. 2014. Brief communication: a proposed method for the assessment of pubertal stage in human skeletal remains using cervical vertebrae maturation. American Journal of Physical Anthropology 153: 144-153.

Shennan S.1997. Quantifying Archaeology. Iowa City: University of Iowa Press.

Silva Braz V. 2009. Anthropological estimation of sex. In: Blau S and Ubelaker DH editors. Handbook of Forensic Anthropology and Archaeology. California: Left Coast Press Inc.

Sinclair D. 1989. Human Growth After Birth. New York: Oxford University Press.

Smith P, Heimer G, Norgren A, Ulmsten U. 1990. Steroid hormone receptors in pelvic muscles and ligaments in females. Gynecologic and Obstetric Investigation 30: 27-30.

Soegiharto B, Cunningham S, Moles D. 2008. Skeletal maturation in Indonesian and white children assessed with hand-wrist and cervical vertebrae methods. Journal of Orthodontics and Dentofacial Orthopedics 134: 217-226.

Stanojevich V. 2012. The role of a forensic anthropologist in a death investigation. Journal of Forensic Research 3: 154-155.

Stewart NA, Gerlach RF, Gowland RL, Gron KJ, Montgomery J. 2017. Sex determination of human remains from peptides in tooth enamel. Proceedings of the National Academy of Sciences 114: 13649-13654.

Stock MK. 2018. A preliminary analysis of the age of full expression of sexually dimorphic cranial traits. Journal of Forensic Sciences 63(6): 1802-1808.

Stuckler D, Meissner C, Fishback P, Basu S, McKee M. 2011. Banking crisis and mortality during the Great Depression: evidence from US urban populations, 1929-1937. Journal of Epidemiology and Community Health 66: 410-419.

Stulp G, Barrett L. 2016. Evolutionary perspectives on human height variation. Biological Review 91: 206-234.

Stull KE, L'Abbé EN, Ousley SD. 2017. Subadult sex estimation from diaphyseal dimensions. American Journal of Physical Anthropology 163: 64-74.

Sundick RI. 1977. Age and sex determination of subadult skeletons. Journal of Forensic Science 22:141–144.

Sutherland LD, Suchey JM. 1991. Use of the ventral arc in pubic sex determination. Journal of Forensic Sciences 36(2): 501-511.

Sutter RC. 2003. Nonmetric subadult skeletal sexing traits: I. a blind test of the accuracy of eight previously proposed methods using prehistoric known-sex mummies from Northern Chile. Journal of Forensic Sciences 48: 927-935.

Tague RG. 1989. Variation in pelvic size between males and females. American Journal of Physical Anthropology 80(1): 59-71.

Tague RG. 1994. Maternal mortality or prolonged growth: age at death and pelvic size in three prehistoric Amerindian populations. American Journal of Physical Anthropology 95(1): 27-40.

Tanner JM. 1973. Growth at Adolescence. Oxford: Blackwell.

Tanner JM, Healy MJR, Goldstein H, Cameron N. 2001. Assessment of skeletal maturity and prediction of adult height (TW3 method). London: Harcourt Publishers.

Tapia Granados JA, Diez Roux AV. 2009. Life and death during the Great Depression. PNAS 106(41): 17290-17295.

Thompson JL, Crandall JJ, Alfonso-Durruty MP. 2014. Children in prehistory: Now seen, now heard. In: Thompson JL, Crandall JJ, Alfonso-Durruty MP editors. Tracing Childhood: Bioarchaeological Investigations of Early Lives in Antiquity. Florida: University of Florida Press p. 1-14.

Tolnay SE. 2003. The African American “great migration” and beyond. Annual Review of Sociology 29: 209-232.

Tolnay SE, Beck EM. 1990. Black flight: lethal violence and the great migration, 1900-1930. Social Science History 14(3): 347-370.

Tolnay SE, Beck EM. 1992. Racial violence and black migration in the American South, 1910 to 1930. American Sociological Review 57(1): 103-116.

Trotter M. 1981. Robert James Terry, 1871-1966. American Journal of Physical Anthropology 6: 489-501.

Vance V, Steyn M, L’Abbe EN. 2008. Nonmetric sex determination from the distal and posterior humerus in black and white South Africans. Journal of Forensic sciences 56: 710-714.

Veroni A, Nikitovic D, Schillaci MA. 2010. Brief communication: Sexual dimorphism of the juvenile basicranium. American Journal of Physical Anthropology 141: 147-151.

Vlak D, Roksandic M, Schillaci MA. 2008. Greater sciatic notch as a sex indicator in juveniles. American Journal of Physical Anthropology 137: 309-315.

Walker PL. 2005. Greater sciatic notch morphology: sex, age, and population differences. American Journal of Physical Anthropology 127: 385-391.

Washburn SL. 1948. Sex differences in the pubic bone. American Journal of Physical Anthropology 6(2): 199-208.

Weaver DS. 1980. Sex differences in the ilia of a known sex and age sample of fetal and infant skeletons. American Journal of Physical Anthropology 52:191–195.

Wilkerson I. 2010. The Warmth of Other Suns: The Epic Story of America's Great Migration. New York: Random House.

Williams BA, Rogers TL. 2006. Evaluating the accuracy and precision of cranial morphological traits for sex determination. Journal of Forensic Sciences 51(4): 729-735.

Wilson LA, MacLeod N, Humphrey LT. 2008. Morphometic criteria for sexing juveniles human skeletons using the ilium. Journal of Forensic Sciences 53(2): 269-278.

Wilson LAB, Ives R, Cardoso HFV, Humphrey LT. 2015. Shape, size, maturity trajectories of the human ilium. American Journal of Physical Anthropology 156: 19-34.

Wilson LAB, Humphrey LT. 2017. Voyaging into 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: 32-46.

Wood CE. 2012. The Influence of Growth And Development in the Expression of Human Morphological Variation. Unpublished doctoral dissertation. University of Toronto.

Wood BA, Chamberlain AT. 1986. The primate pelvis: allometry or sexual dimorphism? Journal of Human Evolution 15(4): 257-263.

Żądzińska E, Karasinska M, Jedrychowska-Danska K, Watala C, Witas HW. 2008. Sex diagnosis of subadult specimens from medieval Polish archaeological sites: metric analysis of deciduous dentition. HOMO 59(3): 175-187.

Zhang SY, Liu G, Liu LJ, Ma ZG, Han YS, Shen XZ, Xu RL. 2008. Relationship of certain skeletal maturity indicators of hand and wrist with adolescent group spurt. Zhonghua Yi XueZaZhi 88: 2198-2200.

Appendix

Appendix A: Age Category and Puberty Stage Comparison Tables

A.1 - Morphological traits

Table A-1: Comparison of the overall and sex specific accuracies for the ilium shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	16	43.8	3	33.3	13	46.2	Infant	1-2.9 F+M	16	43.8	3	33.3	13	46.2
Child	3 – 11.9	21	61.9	12	75	9	44.4	Prepub	3-13 F 3-11 M	17	58.8	11	72.7	6	33.3
								Accel	8-12 F 10-17 M	5	80	2	100	3	66.7
Adol.	12-16.9	15	93.4	12	91.7	3	100	PHV	12 F 13-15 M	5	100	2	100	3	100
								Decel	14-16 F 17-19 M	14	85.7	5	100	9	77.8
Late Adol.	17-20	73	91.8	23	91.3	50	92	Matura	17-19 F 17-20 M	19	94.7	6	83.3	13	100
								Post-puberty	14-20 F 18-20 M	45	93.3	18	88.9	27	96.3

Table A-2: Comparison of the overall and sex specific accuracies for the pelvic inlet shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	1	0	-	-	1	0	Infant	1-2.9 F+M	-	-	-	-	-	-
Child	3 – 11.9	1	100	-	-	1	100	Prepub	3-13 F 3-11 M	-	-	-	-	-	-
								Accel	8-12 F 10-17 M	-	-	-	-	-	-
Adol.	12-16.9	10	90	9	88.9	1	100	PHV	12 F 13-15 M	2	100	1	100	1	100
								Decel	14-16 F 17-19 M	13	84.6	5	80	8	87.5
Late Adol.	17-20	69	94.2	21	95.2	48	93.8	Matura	17-19 F 17-20 M	18	88.9	5	80	13	92.3
								Post-puberty	14-20 F 18-20 M	44	97.7	17	100	27	96.3

Table A-3: Comparison of the overall and sex specific accuracies for the pubic bone shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	14	92.9	2	50	12	100	Infant	1-2.9 F+M	14	92.9	2	50	12	100
Child	3 – 11.9	21	42.9	12	0	9	100	Prepub	3-13 F 3-11 M	18	33.3	12	0	6	100
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	60	12	50	3	100	PHV	12 F 13-15 M	5	80	2	50	3	100
								Decel	14-16 F 17-19 M	14	78.6	5	60	9	88.9
Late Adol.	17-20	72	97.2	23	100	49	95.9	Matura	17-19 F 17-20 M	19	100	6	100	13	100
								Post-puberty	14-20 F 18-20 M	44	95.5	18	94.4	26	96.2

Table A-4: Comparison of the overall and sex specific accuracies for the greater sciatic notch angle between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	17	35.3	3	100	14	21.4	Infant	1-2.9 F+M	17	35.3	3	100	14	21.4
Child	3 – 11.9	21	76.2	13	92.3	8	50	Prepub	3-13 F 3-11 M	17	70.6	12	91.7	5	20
								Accel	8-12 F 10-17 M	5	80	2	100	3	66.7
Adol.	12-16.9	15	93.3	12	100	3	66.7	PHV	12 F 13-15 M	5	80	2	100	3	66.7
								Decel	14-16 F 17-19 M	14	71.4	5	100	9	55.6
Late Adol.	17-20	73	83.6	23	91.3	50	80	Matura	17-19 F 17-20 M	19	84.2	6	83.3	13	84.6
								Post-puberty	14-20 F 18-20 M	45	91.1	18	94.4	27	88.9

Table A-5: Comparison of the overall and sex specific accuracies for true pelvis size & shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	1	100	-	-	1	100	Infant	1-2.9 F+M	-	-	-	-	-	-
Child	3 – 11.9	1	100	-	-	1	100	Prepub	3-13 F 3-11 M	-	-	-	-	-	-
								Accel	8-12 F 10-17 M	-	-	-	-	-	-
Adol.	12-16.9	10	60	9	55.6	1	100	PHV	12 F 13-15 M	2	100	1	100	1	100
								Decel	14-16 F 17-19 M	13	61.5	5	40	8	75
Late Adol.	17-20	70	95.7	22	100	48	93.8	Matura	17-19 F 17-20 M	18	100	5	100	13	100
								Post-puberty	14-20 F 18-20 M	45	95.6	18	94.4	27	96.3

Table A-6: Comparison of the overall and sex specific accuracies for the obturator foramen shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Child	3 – 11.9	21	42.9	12	0	9	100	Prepub	3-13 F 3-11 M	18	33.3	12	0	6	100
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	66.7	12	58.3	3	100	PHV	12 F 13-15 M	5	100	2	100	3	100
								Decel	14-16 F 17-19 M	14	71.4	5	60	9	77.8
Late Adol.	17-20	73	87.7	23	82.6	50	90	Matura	17-19 F 17-20 M	19	89.5	6	83.3	13	92.3
								Post-puberty	14-20 F 18-20 M	45	86.7	18	77.8	27	92.6

Table A-7:Comparison of the overall and sex specific accuracies for the preauricular sulcus between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	13	84.6	1	0	12	91.7	Infant	1-2.9 F+M	13	84.6	1	0	12	91.7
Child	3 – 11.9	21	47.6	13	15.4	8	100	Prepub	3-13 F 3-11 M	17	41.2	12	16.7	5	100
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	53.3	12	41.7	3	100	PHV	12 F 13-15 M	5	80	2	50	3	100
								Decel	14-16 F 17-19 M	14	71.4	5	20	9	100
Late Adol.	17-20	73	91.8	23	78.3	50	98	Matura	17-19 F 17-20 M	19	89.5[a]	6	66.7	13	100
								Post-puberty	14-20 F 18-20 M	45	91.1[b]	18	83.3	27	96.3

Table A-8: Comparison of the overall and sex specific accuracies for the acetabulum size & orientation between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Child	3 – 11.9	4	75	1	100	3	66.7	Prepub	3-13 F 3-11 M	3	66.7	1	100	2	50
								Accel	8-12 F 10-17 M	4	50	2	50	2	50
Adol.	12-16.9	15	73.3	12	75	3	66.7	PHV	12 F 13-15 M	5	60	2	50	3	66.7
								Decel	14-16 F 17-19 M	14	85.7[a]	5	80	9	88.9
Late Adol.	17-20	73	94.5	23	95.7	50	94	Matura	17-19 F 17-20 M	19	94.7[b]	6	83.3	13	100
								Post-puberty	14-20 F 18-20 M	45	97.8[c]	18	100	27	96.3

Table A-9: Comparison of the overall and sex specific accuracies for the auricular surface height between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	14	85.7	2	0	12	100	Infant	1-2.9 F+M	14	85.7	2	0	12	100
Child	3 – 11.9	21	42.9	13	7.7	8	100	Prepub	3-13 F 3-11 M	17	29.4	12	0	5	100
								Accel	8-12 F 10-17 M	5	80	2	50	3	100
Adol.	12-16.9	12	66.7	9	55.6	3	100	PHV	12 F 13-15 M	4	75	1	0	3	100
								Decel	14-16 F 17-19 M	13	92.3[a]	4	75	9	100
Late Adol.	17-20	72	88.9	22	77.3	50	94	Matura	17-19 F 17-20 M	19	79[b]	6	66.7	13	84.6
								Post-puberty	14-20 F 18-20 M	44	90.9[c]	17	82.4	27	96.3

Table A-10: Comparison of the overall and sex specific accuracies for the subpubic concavity between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	10	90	1	0	9	100	Infant	1-2.9 F+M	-	-	-	-	-	-
Child	3 – 11.9	21	42.9	12	0	9	100	Prepub	3-13 F 3-11 M	10	90	-	-	10	90
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	66.7	12	58.3	3	100	PHV	12 F 13-15 M	5	60	2	0	3	100
								Decel	14-16 F 17-19 M	14	92.9[a]	5	100	9	88.9
Late Adol.	17-20	73	93.2	23	82.6	50	98	Matura	17-19 F 17-20 M	19	94.7[b]	6	83.3	13	100
								Post-puberty	14-20 F 18-20 M	45	91.1[c]	18	77.8	27	100

Table A-11: Comparison of the overall and sex specific accuracies for the ventral arc between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	12	83.3	2	0	10	100	Infant	1-2.9 F+M	12	83.3	2	0	10	100
Child	3 – 11.9	17	41.2	10	0	7	100	Prepub	3-13 F 3-11 M	14	28.6	10	0	4	100
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	13	45.5	10	20	3	100	PHV	12 F 13-15 M	5	60	2	0	3	100
								Decel	14-16 F 17-19 M	12	83.3	3	33.3	9	100
Late Adol.	17-20	68	89.7	19	63.2	49	100	Matura	17-19 F 17-20 M	17	88.2[a]	4	50	13	100
								Post-puberty	14-20 F 18-20 M	43	83.7[b]	17	58.8	26	100

Table A-12: Comparison of the overall and sex specific accuracies for the ischiopubic ramus ridge between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	11	63.6	2	0	9	77.8	Infant	1-2.9 F+M	11	63.6	2	0	9	77.8
Child	3 – 11.9	20	40	12	0	8	100	Prepub	3-13 F 3-11 M	17	29.4	12	0	5	100
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	46.6	12	33.3	3	100	PHV	12 F 13-15 M	5	60	2	0	3	100
								Decel	14-16 F 17-19 M	13	76.9	5	60	8	87.5
Late Adol.	17-20	69	85.5	22	72.7	47	91.5	Matura	17-19 F 17-20 M	18	88.9[a]	6	83.3	12	91.7
								Post-puberty	14-20 F 18-20 M	44	79.5[b]	18	61.1	26	92.3

Table A-13: Comparison of the overall and sex specific accuracies for the greater sciatic notch size & shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	18	38.9	3	100	15	26.7	Infant	1-2.9 F+M	18	38.9	3	100	15	26.7
Child	3 – 11.9	21	76.2	13	100	8	37.5	Prepub	3-13 F 3-11 M	17	76.5	12	100	5	20
								Accel	8-12 F 10-17 M	5	60	2	100	3	33.3
Adol.	12-16.9	15	80	12	83.3	3	66.7	PHV	12 F 13-15 M	5	60	2	50	3	66.7
								Decel	14-16 F 17-19 M	14	71.4	5	80	9	66.7
Late Adol.	17-20	73	84.9	23	87	50	84	Matura	17-19 F 17-20 M	19	89.5[a]	6	66.7	13	100
								Post-puberty	14-20 F 18-20 M	45	88.9[b]	18	94.4	27	85.2

Table A-14: Comparison of the overall and sex specific accuracies for the sacrum shape between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	1	100	-	-	1	100	Infant	1-2.9 F+M	1	100	-	-	1	100
Child	3 – 11.9	1	100	-	-	1	100	Prepub	3-13 F 3-11 M	-	-	-	-	-	-
								Accel	8-12 F 10-17 M	3	66.7	1	0	2	100
Adol.	12-16.9	12	58.3	10	50	2	100	PHV	12 F 13-15 M	3	100	1	100	2	100
								Decel	14-16 F 17-19 M	14	64.3	5	40	9	77.8
Late Adol.	17-20	70	81.4	21	66.7	49	87.8	Matura	17-19 F 17-20 M	17	82.4	5	60	12	91.7
								Post-puberty	14-20 F 18-20 M	44	84.1[a]	17	76.5	27	88.9

Table A-15: Comparison of the overall and sex specific accuracies for the dorsal pubic pitting between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	12	91.7	1	0	11	100	Infant	1-2.9 F+M	12	91.7	1	0	11	100
Child	3 – 11.9	21	42.9	12	0	9	100	Prepub	3-13 F 3-11 M	18	33.3	12	0	6	100
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	20	12	0	3	100	PHV	12 F 13-15 M	5	60	2	0	3	100
								Decel	14-16 F 17-19 M	14	57.1	5	0	9	88.9
Late Adol.	17-20	72	69.4	22	9.1	50	96	Matura	17-19 F 17-20 M	19	73.7	6	16.7	13	100
								Post-puberty	14-20 F 18-20 M	44	61.4	17	5.9	27	96.3

Table A-16: Comparison of the overall and sex specific accuracies for the iliac crest curvature between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	17	35.3	3	100	14	0	Infant	1-2.9 F+M	17	17.6	3	100	14	0
Child	3 – 11.9	21	71.4	13	100	8	37.5	Prepub	3-13 F 3-11 M	17	76.5	12	100	5	20
								Accel	8-12 F 10-17 M	5	60	2	50	3	66.7
Adol.	12-16.9	15	73.3	12	66.7	3	100	PHV	12 F 13-15 M	5	100	2	100	3	100
								Decel	14-16 F 17-19 M	14	64.3	5	60	9	77.8
Late Adol.	17-20	73	74	23	78.3	50	72	Matura	17-19 F 17-20 M	19	73.7	6	66.7	13	76.9
								Post-puberty	14-20 F 18-20 M	45	75.6*	18	83.3	27	70.4

Table A-17: Comparison of the overall and sex specific accuracies for the development of muscle markings between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Child	3 – 11.9	10	40	5	80	5	0	Prepub	3-13 F 3-11 M	9	55.6	6	83.3	3	0
								Accel	8-12 F 10-17 M	3	33.3	1	100	2	0
Adol.	12-16.9	15	73.4	12	91.7	3	0	PHV	12 F 13-15 M	5	40	2	100	3	0
								Decel	14-16 F 17-19 M	14	57.1	5	100	9	33.3
Late Adol.	17-20	73	78.1	23	100	50	68	Matura	17-19 F 17-20 M	19	79*	6	100	13	69.2
								Post-puberty	14-20 F 18-20 M	45	86.7*	18	94.4	27	81.5

Table A-18: Comparison of the overall and sex specific accuracies for the arch criterion between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	17	35.3	3	100	14	21.4	Infant	1-2.9 F+M	17	35.3	3	100	14	21.4
Child	3 – 11.9	21	71.4	13	100	8	25	Prepub	3-13 F 3-11 M	17	76.5	12	100	5	20
								Accel	8-12 F 10-17 M	5	60	2	100	3	33.3
Adol.	12-16.9	15	73.3	12	83.3	3	33.3	PHV	12 F 13-15 M	5	60	2	100	3	33.3
								Decel	14-16 F 17-19 M	14	50	5	80	9	33.3
Late Adol.	17-20	73	83.6	23	87	50	62	Matura	17-19 F 17-20 M	19	68.4	6	83.3	13	61.5
								Post-puberty	14-20 F 18-20 M	45	77.8*	18	83.3	27	74.1

Table A-19: Comparison of the overall and sex specific accuracies for final sex estimaes (all traits weighted equally) between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	18	66.7	3	66.7	15	66.7	Infant	1-2.9 F+M	18	66.7	3	66.7	15	66.7
Child	3 – 11.9	22	40.9	13	7.7	9	88.9	Prepub	3-13 F 3-11 M	18	27.8	12	0	6	83.3
								Accel	8-12 F 10-17 M	5	60	2	0	3	100
Adol.	12-16.9	15	73.3	12	66.7	3	100	PHV	12 F 13-15 M	5	80	2	50	3	100
								Decel	14-16 F 17-19 M	14	85.7	5	80	9	88.9
Late Adol.	17-20	73	95.9	23	95.7	50	96	Matura	17-19 F 17-20 M	19	94.7	6	83.3	13	100
								Post-puberty	14-20 F 18-20 M	45	97.8	18	100	27	96.3

Table A-20: Comparison of the overall and sex specific accuracies for final sex estimate (juvenile traits weighted more) between age categories and puberty stage

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Infant	0-2.9	18	27.8	3	100	15	13.3	Infant	1-2.9 F+M	18	27.8	3	100	15	13.3
Child	3 – 11.9	22	72.7	13	100	9	33.3	Prepub	3-13 F 3-11 M	18	77.8	12	100	6	33.3
								Accel	8-12 F 10-17 M	5	60	2	100	3	33.3
Adol.	12-16.9	15	100	12	100	3	100	PHV	12 F 13-15 M	5	100	2	100	3	100
								Decel	14-16 F 17-19 M	14	92.9	5	100	9	88.9
Late Adol.	17-20	73	94.5	23	95.7	50	94	Matura	17-19 F 17-20 M	19	94.7	6	83.3	13	100
								Post-puberty	14-20 F 18-20 M	45	97.8	18	100	27	96.3

A.2 – Metric Logistic Regression Models

Table A-21: Comparison of the overall and sex specific accuracies for Model 3 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	9	77.8	8	75	1	100	PHV	12 F 13-15 M	2	100	1	100	1	100
								Decel	14-16 F 17-19 M	13	84.6[a]	5	80	8	87.5
Late Adol.	17-20	#	87.5	20	70	44	95.5	Matura	17-19 F 17-20 M	15	93.3[b]	4	75	11	100
								Post-puberty	14-20 F 18-20 M	41	85.4[c]	16	68.8	25	96

Table A-22: Comparison of the overall and sex specific accuracies for Model 8 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	9	88.9	8	97.5	1	100	PHV	12 F 13-15 M	2	100	1	100	1	100
								Decel	14-16 F 17-19 M	13	84.6[d]	5	80	8	87.5
Late Adol.	17-20	#	86.2	20	65	45	95.6	Matura	17-19 F 17-20 M	17	82.4	5	40	12	100
								Post-puberty	14-20 F 18-20 M	41	87.8[e]	16	75	25	96

Table A-23: Comparison of the overall and sex specific accuracies for Model 20 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	9	88.9	9	88.9	-	-	PHV	12 F 13-15 M	2	50	1	100	1	0
								Decel	14-16 F 17-19 M	13	92.3[a]	5	100	8	87.5
Late Adol.	17-20	#	93.9	23	87	44	97.7	Matura	17-19 F 17-20 M	17	88.2[b]	6	66.7	11	100
								Post-puberty	14-20 F 18-20 M	43	97.7[c]	18	94.4	25	100

Table A-24: Comparison of the overall and sex specific accuracies for Model 4 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	#	80	8	87.5	2	50	PHV	12 F 13-15 M	3	66.7	1	100	2	50
								Decel	14-16 F 17-19 M	13	84.6[d]	5	80	8	87.5
Late Adol.	17-20	#	86.4	21	76.2	45	91.1	Matura	17-19 F 17-20 M	17	88.2[e]	5	60	12	100
								Post-puberty	14-20 F 18-20 M	41	85.4[f]	16	81.3	25	88

Table A-25: Comparison of the overall and sex specific accuracies for Model 10 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	#	80	8	87.5	2	50	PHV	12 F 13-15 M	3	66.7	1	100	2	50
								Decel	14-16 F 17-19 M	13	84.6[g]	5	100	8	75
Late Adol.	17-20	#	83.1	21	66.7	44	90.9	Matura	17-19 F 17-20 M	16	81.3	5	40	11	100
								Post-puberty	14-20 F 18-20 M	41	85.4[h]	16	75	25	92

Table A-26: Comparison of the overall and sex specific accuracies for Model 1 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	9	77.8	8	75	1	100	PHV	12 F 13-15 M	2	100	1	100	1	100
								Decel	14-16 F 17-19 M	13	76.9	5	80	8	75
Late Adol.	17-20	#	84.6	21	66.7	44	93.2	Matura	17-19 F 17-20 M	16	81.3	5	40	11	100
								Post-puberty	14-20 F 18-20 M	41	87.8[i]	16	75	25	96

Table A-27: Comparison of the overall and sex specific accuracies for Model 11 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	#	80	8	75	2	100	PHV	12 F 13-15 M	3	100	1	100	2	100
								Decel	14-16 F 17-19 M	13	76.9	5	80	8	75
Late Adol.	17-20	#	83.1	21	66.7	44	90.9	Matura	17-19 F 17-20 M	16	87.5[j]	5	60	11	100
								Post-puberty	14-20 F 18-20 M	41	85.4[k]	16	68.8	25	92

Table A-28: Comparison of the overall and sex specific accuracies for Model 2 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Child	3 – 11.9	-	-	-	-	-	-	Prepub	3-13 F 3-11 M	2	50	2	50	-	-
								Accel	8-12 F 10-17 M	-	-	-	-	-	-
Adol.	12-16.9	#	72.7	9	66.7	2	100	PHV	12 F 13-15 M	3	66.7	1	0	2	100
								Decel	14-16 F 17-19 M	13	76.9	5	60	8	87.5
Late Adol.	17-20	#	87.3	19	68.4	44	95.5	Matura	17-19 F 17-20 M	16	87.5[a]	5	60	11	100
								Post-puberty	14-20 F 18-20 M	40	87.5[b]	15	73.3	25	96

Table A-29: Comparison of the overall and sex specific accuracies for Model 5 between age categories and puberty stages

Age Cat.	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)	Puberty Stage	Age Range	N	Accuracy (%)	N_F	Female Acc. (%)	N_M	Male Acc. (%)
Adol.	12-16.9	#	80	8	87.5	2	50	PHV	12 F 13-15 M	3	66.7	1	100	2	50
								Decel	14-16 F 17-19 M	13	69.2	5	80	8	62.5
Late Adol.	17-20	#	80.3	21	76.2	45	82.2	Matura	17-19 F 17-20 M	17	94.1[c]	5	80	12	100
								Post-puberty	14-20 F 18-20 M	41	78.1[d]	16	75	25	80