

https://dspace.org/

An evaluation of the Acsádi-Nemeskéri Complex Method for adult age estimation on a black South african sample

Mokoena, Amanda Valeria

2024

University of the Witwatersrand, Johannesburg

https://hdl.handle.net/10539/46876

Mokoena, Amanda Valeria . (2024). An evaluation of the Acsádi-Nemeskéri Complex Method for adult age estimation on a black South african sample [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/46876

Downloaded from a DSpace repository. https://dspace.org/

AN EVALUATION OF THE ACSÁDI-NEMESKÉRI COMPLEX METHOD FOR ADULT AGE ESTIMATION ON A BLACK SOUTH AFRICAN SAMPLE

by

Amanda Valeria Mokoena

1624784

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine.

Supervisor:

Dr A. Meyer

February 2024 in Johannesburg

Declaration

I Amanda Valeria Mokoena declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



11 day of October 2024 at Johannesburg

Abstract

The Acsádi and Nemeskéri Complex Method is a multifactorial method that utilises four skeletal indicators and a weighting system to assess adult skeletal age. It uses four skeletal indicators, including pubic symphysis, endocranial sutures, and internal structural changes in the proximal femur and humerus. The Complex Method has been criticized for its small sample size, statistical validity, and constant under- and overestimation of age for older and younger individuals, respectively. This study tested the Complex Method's performance on a black South African sample (n = 103), and Bayesian statistics was employed to see whether a more statistically robust approach would improve accuracy. Results somewhat support the criticism of the Complex Method's tendency to under- and overestimate age. In contrast to Acsádi and Nemeskéri's findings, combinations utilising two or three skeletal indicators generally outperformed combinations using all four skeletal indicators. Given the method's poor performance, population- and sex-specific skeletal indicator combinations and weighting systems must be developed before future use in this sample. Bayesian statistics provided some improvements, decreasing over- and underestimations, especially in males. Some benefits of using the weighting system were observed, as the Complex Method outperformed the Bayesian Method in females. Bayesian statistics in age estimation has some merit but requires large sample sizes with a good spread of age-related changes.

Acknowledgements

I would like to extend my deepest gratitude to the individuals who donated their bodies to the Raymond A. Dart Collection of Modern Human Skeletons housed at the University of the Witwatersrand, without them this project would not be possible. I am extremely thankful to my supervisor, Dr Anja Meyer, for her unwavering support, insightful suggestions, continuous patience and profound belief in my work and abilities. Many thanks to Dr Nicholas Bacci, Mr Mashudu Mulaudzi and Dr Brendon Billings for their assistance with data collection in the Raymond A. Dart Collection of Modern Human Skeletons. Thank you also to Ms Amelia Rammekwa of the Wits Research Animal Facility, for her assistance with the radiographs. Special thanks to Ms Claudia Landsman, for her help with the repeatability scores. Many thanks to Mrs Natasha Loubser for always asking how I'm doing. I cannot thank Ms Ozyl Marufu, Ms Khanyisa Madubane, Ms Mokgatjane Gaffane and Ms Nelly Matine enough for constantly listening to my worries and frustrations and offering me words of encouragement. I am extremely grateful to my parents, Prof Sello Mokoena and Mrs Monica Mokoena, for their support. Finally, I would like to acknowledge the financial support from the National Research Foundation, South Africa, the Smieszek J.J.J. Fellowship, the Faculty Research Committee, and the Postgraduate Merit Award, for this project.

CONTENTS

CHAPTER 1: INTRODUCTION 1

1.1. Aims and Objectives 4

CHAPTER 2: LITERATURE REVIEW 6

2.1. Multifactorial approaches 7

2.1.1. Cranial Sutures 8

2.1.2. Pubic symphysis 12

2.1.3. Proximal humerus and femur 15

2.1.4. Multifactorial approaches 18

2.2. Statistical analyses 21

2.2.1. Bayesian statistical approach in age estimation 23

CHAPTER 3: MATERIALS AND METHODS 25

3.1. Materials 25

3.2. Acsádi & Nemeskéri Complex Method 26

3.2.1. Cranial sutures 26

3.2.2. Pubic Symphysis 30

3.2.3. Proximal humerus and proximal femur 33

3.3. Testing the Complex Method 42

CHAPTER 4: RESULTS 48

4.1. Complex Method 48

4.1.1. Descriptive Statistics and Normality 48

4.1.2. Intra- and interobserver repeatability 49

4.1.3. Differences between female and male true ages per phase for each of the four skeletal indicators....51

4.1.4. Correlation between increasing true age and phase of each of the four skeletal indicators. 51

4.1.5. Correlation between true and estimated age using different skeletal indicator combinations. 53

4.1.6. Accuracy of the Complex Method....55

4.1.7. Inaccuracy and bias 61

4.2. Bayesian Methods....67

4.2.1. Development of the Bayesian Method 67

4.2.1.1. Informative prior....67

4.2.1.2. Likelihoods using transition analysis 68

4.2.1.3. Highest posterior density (HPD) and highest posterior density regions (HPDR) for each of the skeletal indicators....70

4.2.2. Testing the Bayesian Method....72

4.2.2.1. Correlation between true and Bayesian estimated age using different skeletal indicator combinations....72

4.2.2.2. Comparison of age using the Complex Method and Bayesian Method....75

4.2.2.3. Accuracy using the Bayesian Method compared to the Complex Method. 77

4.2.3. Collapsed Bayesian Method....80

4.2.3.1. Likelihoods using transition analysis. 85

4.2.3.2. Highest posterior density (HPD) and highest posterior density regions (HPDR) for each of the skeletal indicators. 87

4.2.3.3. Correlation between true and Collapsed Bayesian estimated age using different skeletal combinations. 89

4.2.3.4. Comparison of age using the Complex Method, Bayesian and Collapsed Bayesian Methods. 92

4.2.3.5. Accuracy using the Collapsed Bayesian Method compared to the Bayesian and Complex Methods. 94

4.2.4. Pooled Bayesian Method 98

4.2.4.1. Informative prior 98

4.2.4.2. Likelihood using transition analysis. 102

4.2.4.3. Highest posterior density (HPD) and highest posterior density regions (HPDR) for each of the skeletal indicators 104

4.2.4.4. Correlation between true and Pooled Bayesian estimated age using different skeletal combinations 104

4.2.4.5. Comparison of age using the Complex Method, Bayesian, and Pooled Bayesian Methods. 108

4.2.4.6. Accuracy using the Pooled Bayesian Method compared to the Bayesian and Complex Methods. 109

CHAPTER 5: DISCUSSION 117

5.1. The performance of the Complex Method 118

5.1.1. Intra-and inter-observer repeatability 118

5.1.2. Descriptive statistics 119

5.1.3. Skeletal indicator correlations with age 121

5.1.4. Correlations between true and estimated ages. 124

5.1.5. Accuracy, inaccuracy, and bias.....125

5.2. The performance of the Complex Method scoring criteria using Bayesian statistics 126

CHAPTER 6: CONCLUSION 135

REFERENCES 138

APPENDIX A 152

APPENDIX B 164

ETHICS CLEARANCE CERTIFICATE 174

PLAGIARISM DECLARATION 175

TURNITIN REPORT 176

LIST OF FIGURES

Figure 1: Number of individuals in the study. 25

Figure 2: Five stages of endocranial suture closure (Reproduced from Figure 3 of Perizonius, 1984: 204). 27

Figure 3: Acsádi & Nemeskéri (1970) endocranial suture stages as observed in the black South African population. 29

Figure 4: Acsádi & Nemeskéri (1970) pubic symphysis phases as observed in the black South African population. 32

Figure 5: Acsádi & Nemeskéri (1970) proximal humerus phases as observed in the black South African population. 35

Figure 6: Acsádi & Nemeskéri (1970) proximal femur phases as observed in the black South African population. 38

Figure 7: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for pooled sexes. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 58

Figure 8: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 59

Figure 9: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 60

Figure 10: True versus estimated age using Acsádi & Nemeskéri Complex Method and Bayesian Method for females. 76

Figure 11: True versus estimated age using Acsádi & Nemeskéri Complex Method and Bayesian Method for males. 77

Figure 12: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for the pooled sexes using the Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur.....82
Figure 13: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females using the Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 83
Figure 14: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males using the Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 84
Figure 15: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Collapsed Bayesian Method for females.....93
Figure 16: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Collapsed Bayesian Method for males. 94
Figure 17: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for the pooled sexes using the Collapsed Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur.....99
Figure 18: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females using the Collapsed Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur.....100
Figure 19: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males using the Collapsed Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur.....101
Figure 20: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Pooled Bayesian Method for females.....109

Figure 21: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Pooled Bayesian Method for males. 110
Figure 22: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for the pooled sexes using the Pooled Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 114
Figure 23: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females using the Pooled Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur 115
Figure 24: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males using the Pooled Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal femur 116

LIST OF TABLES

Table 1: Endocranial suture stages with associated age ranges (Reproduced from Table 32 of Acsádi & Nemeskéri, 1970: 121). 28

Table 2: Pubic symphysis stages and associated age ranges (Reproduced from Table 34 of Acsádi & Nemeskéri, 1970: 126). 31

Table 3: Proximal humerus stages and associated age ranges (Reproduced from Table 33 of Acsádi & Nemeskéri, 1970: 126). 34

Table 4: Proximal femur stages and associated age ranges (Reproduced from Table 35 of Acsádi & Nemeskéri, 1970: 128). 37

Table 5: Limit ranges provided by Acsádi & Nemeskéri (1970: 131). 40

Table 6: Example of calculating the age estimate using the Complex Method. 40

Table 7: Single skeletal indicator phases and associated mean ages (Reproduced from Acsádi & Nemeskéri, 1970: 135). 42

Table 8: Interpretation of Cohen's weighted Kappa coefficient 44

Table 9: Interpretation of Spearman's rank correlation coefficient. 44

Table 10: Descriptive statistics comparing true age and estimated age. 48

Table 11: Test of normality results. 49

Table 12: Weighted Cohen's Kappa values to assess intra-observer repeatability for each skeletal indicator. 50

Table 13: Weighted Cohen's Kappa values to assess interobserver repeatability for each skeletal indicator. 50

Table 14: Differences in true age per phase between females and males. 52

Table 15: Spearman rank correlation coefficients of increasing true age for each skeletal indicator. 53

Table 16: Spearman's rank correlation coefficient between true and estimated age. 54

Table 17: Inaccuracy and bias using the Complex Method. 62
Table 18: Inaccuracy and bias per age cohort for females and males. 64
Table 19: Gompertz hazard parameters for females and males. 68
Table 20: Log transformed and raw ages-at-transition for females and males. 69
Table 21: HPD and HPDR per phase calculated for females and males. 71
Table 22: Spearman's rank correlation coefficient between true and Bayesian estimated age. 74
Table 23: Log transformed and raw ages-at-transition for Collapsed Bayesian Method females and males. 86
Table 24: HPD and HPDR per phase calculated for females. 88
Table 25: HPD and HPDR per phase calculated for males. 89
Table 26: Spearman's rank correlation coefficient between true and Collapsed Bayesian estimated age. 91
Table 27: Gompertz hazard parameters for the pooled sexes. 102
Table 28: Log transformed and raw ages-at-transition for pooled sexes. 103
Table 29: HPD and HPDR per phase calculated for pooled sexes. 105
Table 30: Spearman's rank correlation coefficient between true and Pooled Bayesian estimated age. 106

CHAPTER 1: INTRODUCTION

The identification of unknown individuals begins with the development of a biological profile consisting of age, sex, stature, and ancestry (Tersigni-Tarrant & Shirley, 2013; Garvin *et al*., 2012; Christensen *et al*., 2014). Although sex and ancestry are often the first factors established, an estimation of age is essential as it eliminates more potential matches from a missing person's list than any other demographic indicator, better aiding in establishing a positive identification (Tersigni-Tarrant & Shirley, 2013). Therefore, developing accurate age estimation techniques is a necessary requirement in forensic anthropology.

Forensic anthropologists estimate chronological age, or the actual number of years lived, from biological age, which reflects the physiological condition of an individual's skeleton (İşcan & Steyn, 2013; Kemkes-Grottenthaler, 2002; Fays & Lewis, 2011). Biological age is influenced by extrinsic and intrinsic factors such as environmental conditions, like the amount of ultraviolet B exposure, activity levels, genetics, nutrition, and general physiological health of an individual throughout their life, amongst others (Garvin *et al*., 2012). These factors have been shown to affect bone homeostasis and, subsequently, skeletal ageing rates (Kemkes-Grottenthaler, 2002; Hartnett-McCann *et al*., 2018). This is especially true for adults as age-related changes mainly reflect degeneration of the skeletal system, as opposed to age estimation methods in children which are directed by the more predictable and genetically driven patterns of development and growth (Tersigni-Tarrant & Shirley, 2013; Christensen *et al*., 2014). As an adult age, they are exposed to more extrinsic factors, such as activity levels, diet, and exposure to diseases, which affect biological ageing and often result in varying or very wide age range estimates (Garvin *et al*., 2012). Extrinsic and intrinsic factors also affect the presence and degree of degenerative changes at different skeletal regions, such as the pubic symphysis and proximal head of the femur, of an individual at different rates. Associated with this is a constant

overestimation of age for younger individuals and an underestimation of age for older individuals, a phenomenon Masset (1989:81) referred to as "attraction to the middle," which has been linked to the age distribution of individuals in the reference sample used to develop an age estimation technique.

Although wide age ranges are considered accurate, they are not precise (Christensen *et al*., 2014). That is to say, they encompass the age of an unidentified individual but do not give the exact age, making them less effective in eliminating excessive matches from a missing person's list in forensic instances (Christensen *et al*., 2014). An estimated age range should be as narrow as possible while taking into consideration possible variations associated with adult ageing (Christensen *et al*., 2014). Studies have shown that the timing and degree of age-related senescent changes used in adult age estimation techniques differ among populations (Katz & Suchey, 1989; Hoppa, 2000; Schmitt *et al*., 2002; Kemkes-Grottenthaler, 2002; Jooste *et al*., 2016). Skeletal age progression has also been shown to vary between the left and right sides of an individual's body, as well as between the sexes (Kemkes-Grottenthaler, 2002; Garvin *et al*., 2012). Therefore, variations such as population differences and inter-and intrapersonal variability should be considered when developing ageing methods (Kemkes-Grottenthaler, 2002). It has, therefore, been argued that the study of intra- and inter-population and -individual changes is the key to understanding age-related skeletal changes and refining adult age estimation methods (Kemkes-Grottenthaler, 2002).

To overcome the effects of population variability and the inherent biases that may be found when using a single skeletal indicator, many have argued for the use of multifactorial age estimation methods (Acsádi & Nemeskéri, 1970; Lovejoy *et al*., 1985; Aykroyd *et al*., 1999; Boldsen *et al*., 2002; Kemkes-Grottenthaler, 2002; Tersigni-Tarrant & Shirley, 2013). Acsádi & Nemeskéri (1970) were among the first to develop a multifactorial method, termed the 'Complex Method', that utilises four skeletal age indicators in assessing age. It makes use of

the pubic symphysis, endocranial sutures, and internal structural changes in the proximal femur and humerus (observable through cross-sectioning of bone or radiographs). Acsádi & Nemeskéri (1970) reported a confidence level of 80-85% with a margin of error of ±2.5 years. Although other multifactorial age estimation methods have been developed through the years (e.g. Lovejoy *et al*., 1985a; Saunders *et al*., 1992; Boldsen *et al*., 2002; Uhl & Nawrocki, 2010), the Complex Method remains promising as it is easy to apply and allows for use of one or any combination of the four skeletal indicators. However, the Complex Method has been criticised for its small sample size, which is not large enough to accurately represent inter-individual differences and its constant under- and overestimation of age for older and younger individuals, respectively (Saunders *et al*., 1992; Molleson *et al*., 1993; Meyer *et al*., 2021).

An ageing method's under- and overestimation of age has been linked to the use of least square regression analyses to develop mean age estimates in traditional methods (Tersigni-Tarrant & Shirley, 2013). Skeletal features are regressed onto age and inverse calibration is then utilised to obtain mean age estimates. (Aykroyd, 1997; Tersigni-Tarrant & Shirley, 2013; Hartnett-McCann *et al*., 2018). As inverse calibration involves regression towards the mean, the resulting age estimates will be biased in the direction of the mean of the reference sample (Tersigni-Tarrant & Shirley, 2013). This is the "attraction to the middle" phenomenon by Masset (1989:81), which has been linked to the consequences of age mimicry by Masset & Bocquet-Appel (1982). Age-mimicry is the tendency of the target sample's estimated age-at-death distribution to mimic the reference sample's age-at-death distribution (Tersigni-Tarrant & Shirley, 2013). This phenomenon described as "age mimicry" led to the development of the Rostock Manifesto, which ultimately outlines the solution to age mimicry (Boldsen *et al*., 2002). The solution was to use Bayesian statistics, which uses prior knowledge of the age-at-death distribution of the reference sample to estimate age (Hoppa and Vaupel, 2002). Many studies (e.g. Godde & Hens, 2012; Hens & Godde, 2016; Nikita *et al*., 2018) have begun

focusing on using Bayesian statistics to counter the effects of bias often seen in adult age estimation methods.

1.1. Aims and Objectives

This study was comprised of two aims.

The first aim of this study was to test the performance of the Acsádi-Nemeskéri Complex Method on a black South African sample. This was achieved by addressing the following objectives:

- To assess the intra- and inter-observer error of the Acsádi-Nemeskéri (1970) scoring criteria.
- To assess the correlation of each skeletal indicator with increasing true age to test whether the use of the pubic symphysis, failing that of the proximal femur, as a weighting factor is warranted in this sample.
- To assess the correlation between estimated and true age using the Acsádi-Nemeskéri (1970) Complex Method.
- To assess whether correlations between estimated and true ages differ for males and females.
- To determine the level of bias and inaccuracy between the true and estimated ages for males and females.

The second aim of this study was to use Bayesian statistics to test whether it is more appropriate than the original Acsádi & Nemeskéri (1970) averaging method for the black South African sample. This was achieved by addressing the following objectives:

- Compute transition curves, highest posterior density functions and highest posterior density regions, using informed priors, for the black South African sample.
- To compare the accuracy of the estimated ages obtained using the Acsádi & Nemeskéri (1970) Complex Method versus the Bayesian Method for the black South African sample.

CHAPTER 2: LITERATURE REVIEW

The first step in identifying an unknown individual is creating a biological profile that includes details on their stature, age, sex, and ancestry (Tersigni-Tarrant & Shirley, 2013; Garvin *et al*., 2012; Christensen *et al*., 2014). Even though sex and ancestry are usually the first factors evaluated, estimating age is important because, more than any other factor, it eliminates more probable matches from a missing people list, enhancing the chance of identification. (Tersigni-Tarrant & Shirley, 2013). As a result, accurate age estimate methods are needed. Most adult age estimates, such as the Acsádi-Nemeskéri Complex Method (Acsádi & Nemeskéri, 1970), were developed using European populations. To enable reliable age calculation of unknown or unidentified persons; however, population-specificity in ageing has been proven repeatedly to be an important component that has to be researched (Kemkes-Grottenthaler, 2002).

When estimating chronological age, or the number of years lived, forensic anthropologists consider the skeleton's physiological state or biological age of an individual (İşcan & Steyn, 2013). The latter is more difficult to estimate as it is influenced by extrinsic and intrinsic factors such as environmental conditions, like the amount of ultraviolet B exposure, activity levels, genetics, diet and an individual's overall health (Garvin *et al*., 2012). This is further compounded by factors such as population differences and inter- and intrapersonal variability (Kemkes-Grottenthaler, 2002). Multifactorial techniques have been advocated by many as a way to eliminate biases, such as those in predicting biological age from skeletal indicators (Acsádi and Nemeskéri, 1970; Lovejoy *et al*., 1985; Aykroyd *et al*., 1999; Boldsen *et al*., 2002; Kemkes-Grottenthaler, 2002). Most researchers agree that using multiple indicators in age estimation far outperforms using a single age indicator (Garvin *et al*., 2012).

2.1. Multifactorial approaches

Age estimation techniques that make use of multiple traits have been proposed as they can present a more comprehensive picture of the ageing process by taking into account all possible morphological variations in the determination of age (Tersigni-Tarrant & Shirley, 2013). The use of multifactorial age estimation methods has also been recommended to account for population variability in ageing onset and progression, adjust for inherent bias that may be found when using single skeletal indicators, and determine whether and/or how the senescent process varies across the individual skeleton or skeletal assemblages (Acsádi & Nemeskéri, 1970; Lovejoy *et al*., 1985; Aykroyd *et al*., 1999; Boldsen *et al*., 2002; Kemkes-Grottenthaler, 2002; Tersigni-Tarrant & Shirley, 2013).

Acsádi & Nemeskéri (1970) were among the first to develop a multifactorial method, termed the 'Complex Method', that utilises four skeletal age indicators to assess age. The Workshop of European Anthropologists recommended its use in 1980 (Maat, 1987). This method was developed using 105 autopsied Hungarian individuals, looking at age-related changes in the pubic symphysis, endocranial sutures, and internal structural changes in the proximal femur and humerus (observable through cross-sectioning of bone or radiographs). Acsádi & Nemeskéri (1970) developed scoring criteria for each skeletal age indicators using a phase system whereby each phase is correlated to an age range and mean age. The authors found that the pubic symphysis was more strongly correlated to age, followed by the proximal femur. Based on these findings they proposed using a weighting system in the averaging procedure used to estimate the final age estimate. This requires the weighing of the age indicators based on the phase score of the pubic symphysis or the proximal femur in its absence. Acsádi & Nemeskéri (1970) acknowledged that not all skeletal remains will necessarily be preserved, especially in archaeological and fragmentary forensic cases. As a result, they made provision for their multifactorial method to make use of all four of the skeletal age indicators, or any

combination thereof (depending on the skeletal elements present). When testing the method's accuracy, Acsádi & Nemeskéri (1970) reported a confidence level of 80-85% with a margin of error of ±2.5 years.

The following sections will discuss in more detail research relating to the cranial sutures, the pubic symphysis, the proximal humerus, and the proximal femur, as these are the components that make up the Acsádi & Nemeskéri (1970) Complex Method, which is dealt with in this study. Thereafter, will continue to discuss other multifactorial methods and different statistical approaches.

2.1.1. Cranial Sutures

In the 1920s, Todd & Lyon published several articles of their research examining endocranial and ectocranial sutures on North American White and Black male crania of known age (Todd & Lyon, 1924; 1925a-c). Their studies provided several significant findings regarding the variation in the timing of cranial suture closure, the differences between endocranial and ectocranial sutures, and the progress of suture closure (Ruengdit *et al.*, 2020). However, they cautioned against the sole use of cranial sutures as an age indicator, stating it should be used in conjunction with other skeletal age indicators. After their works, several researchers followed, investigating the use of cranial suture closure as an age indicator (e.g., Krogman, 1949; Singer, 1953; Brooks, 1955; McKern & Stewart, 1957).

Acsádi & Nemeskéri (1970) developed a method utilising 352 crania of a Hungarian sample. They examined the degree of ossification, both endocranially and ectocranially, in three main sutures: coronal, sagittal, and lambdoid. Each suture was divided into sections, a total of sixteen (six coronal, four sagittal, and six lambdoid), and scored individually according to the degree of suture closure, using Martin's five-point scale (0-4) (Martin & Saller, 1957), after which the

overall mean suture closure was calculated. Results from their study showed that ossification began endocranially and progressed outwards and that sex does not affect the process of cranial suture closure (Acsádi & Nemeskéri, 1970). Their findings allowed them to correlate particular phases of suture closure to age ranges. Based on their findings, they warned against the sole use of cranial sutures in age estimation.

Years later, Perizonious (1984) tested the accuracy of the Acsádi & Nemeskéri (1970) method on a Dutch sample from Amsterdam. He found no significant sex differences; however, he also found that suture closure was only correlated to chronological age in individuals under the age of 50 years. Therefore, Perizonious (1984) suggested that unknown individuals should first be aged using another skeletal indicator to determine whether they are below or above 50 years, thereafter, cranial sutures should be used for younger individuals. Perizonious (1984) proposed that a modified ageing method was required due to the correlation differences between younger (50 years) and older (>50 years) individuals. In this modified method, unknown individuals would first be evaluated to determine if they are younger or older than 50 years. Then, specific suture sections (endocranial and ectocranial) would be used to determine the age. Perizonious (1984) proposed using ten endocranial suture sections for younger individuals and seven endocranial and ectocranial (three and four respectively) suture sections for older individuals. Using this method, he reported stronger correlations between composite cranial suture scores and age.

Proposed modified and improved methods mostly involve macroscopically scoring the degree of suture closure from open to completely obliterated, to get a composite score, which is compared to an age range to obtain an age estimate. This technique has been greatly questioned and is generally not very accurate; however, it is easy to understand and use, especially where other skeletal indicators may be absent due to incomplete preservation (İşcan & Steyn, 2013). Authors who proposed improved methods include Masset (1989) who modified the Acsádi &

Nemeskéri (1970) method to use only ten sections (three coronal, four sagittal and three lambdoid) observed endocranially. Baker (1984) who proposed a method using five sections (left coronal, right coronal, sagittal, left lambdoid and right lambdoid) which were observed endocranially and scored into three stages (1=open, 2=partially closed and 3=totally closed) and Mann *et al.* (1991) who developed five degrees of obliteration for the closure of palatine sutures.

Utilising the Hamann-Todd Collection, Meindl & Lovejoy (1985) proposed an improved cranial suture method using four instead of five phases. This method involved scoring 1 cm sections at specific sites on various sutures to improve repeatability. They selected ten specific suture sites ectocranially, which were categorised into the "vault system" and the "lateral-anterior system" (Meindl & Lovejoy, 1985). Results showed that the "lateral-anterior system" was more accurate, and no sex or population differences were observed. Although they stated that their method produced smaller standard deviations than those obtained by Acsádi & Nemeskéri (1970), the standard deviations and age ranges obtained using their improved method remain rather large. Buikstra & Ubelaker (1994) combined the methods of Todd & Lyon (1924; 1925a-c), Baker (1984), Meindl & Lovejoy (1985), and Mann *et al.* (1987) to develop a four-stage method, which uses 17 cranial suture sections, namely four palatine sutures, three endocranial sutures, and ten ectocranial sutures proposed by Meindl & Lovejoy (1985).

Key *et al.* (1994) tested the accuracy of the methods proposed by Acsádi & Nemeskéri (1970), Perizonious (1984) and Meindl & Lovejoy (1985) on a Spitalfields sample. They reported that the Acsádi & Nemeskéri (1970) method performed better than the other two, suggesting that this is because the method uses endocranial sutures while the Perizonious (1984) and Meindl & Lovejoy (1985) methods used ectocranial sutures, which are more susceptible to interpopulation differences in the dimorphism of suture closure. Following these results, Key

et al. (1994) proposed a method that utilises a three-phase system and the pre-assignment of individuals into age groups of either younger or older than 50 years. Key *et al*. (1994) reported improved results when using their improved method; however, they recommended using cranial suture closure to sort individuals into age ranges, after which more precise age estimation techniques should be used to obtain narrower age estimates.

Using the same three-phase system, Nawrocki (1998) proposed a technique using 27 cranial suture landmarks (16 ectocranial, seven endocranial, and four palatal). He developed linear regression formulae, for the whole sample (males and females combined), and males and females separately, which used 14 different combinations of landmarks. He provided each equation's adjusted r-square values, inaccuracy, bias, and standard error (Nawrocki, 1998). Zambrano (2005) tested the method and found that the whole sample equation performed better than the others.

To improve age estimation results, researchers proposed other techniques that do not require scoring of landmarks or sections of a suture line. These researchers include Lynnerup & Jacobson (2003) who suggested using the entire suture length, by tracing the sagittal and coronal suture lines and analysing them using Sigmascan 3.0 morphometrical software, and Obert *et al*. (2010) who proposed the use of high-resolution coronal, flat-panel volumetric computed tomography images of the sagittal suture and analysis of those images using an algorithm that identifies whether the suture is open or closed based on its internal three-dimensional structure. These approaches of age estimation were further explored by other authors (e.g., Chiba *et al*., 2013; Boyd *et al*., 2015; Nikolova *et al*., 2019), who found that there is a slight improvement in accuracy results when using these techniques.

Due to their poor correlation with chronological age, the use of cranial sutures for age estimation has been criticised for many years (McKern & Stewart, 1957; Christensen *et al*., 2014; Hartnett-McCann *et al*., 2018).

2.1.2. Pubic symphysis

The pubic symphysis has long been regarded as the most reliable skeletal indicator and seems to remain one of the most popular skeletal age estimators for adult age estimation (Garvin *et al*., 2012.). Using the pubic symphyses of 306 white males, aged between 18 years and 50+ years, from the Western Reserve Collection, Todd (1920; 1921a-c; 1923; 1930) was the first to systematically describe changes seen on the symphyseal face and their relation to age (Hartnett-McCann *et al*., 2018). He documented the morphological changes in sections, including the surface, the ventral (anterior) border or rampart, the dorsal (posterior) border or plateau, the superior extremity, and the inferior extremity (Todd, 1920). Additionally, he mentioned certain surface characteristics that may be seen, such as the existence of "ridges," "grooves," "ossific nodules," and a "granular" or "smooth" appearance (Todd, 1920). From these observations, Todd (1920) developed a ten-phase scoring system with corresponding age ranges and descriptions of the morphological state of the symphyseal face at that stage (Hartnett-McCann *et al*., 2018). As this method was developed from a white male sample, Todd (1921a-c) tested his ten-phase system on white females, black females, and black males, reporting no significant differences between sexes and population groups.

Todd's ten-phase method was tested by Hanihara (1952) on a Japanese male sample and Brooks (1955) on a Native American sample, both found the method tended to overestimate age by three years. Brooks (1955) also found that the symphyseal face showed a stronger correlation to age in males than in females, suggesting that this may be due to sex differences.

Furthermore, Brooks (1955) noted that an excessively narrow pubic symphysis in females provides the impression of increased age.

McKern & Stewart (1957) were the first to suggest that symphyseal face age-related changes may be influenced by childbearing. Examining the skeletal ageing of 450 Americans war dead from the Korean War, they modified Todd's method, creating a three-component method, consisting of the symphyseal rim, ventral rampart, and dorsal plateau, each with five developmental stages. They state that this method is easier to use and less restrictive than the phase method. Stewart (1957) reported on sex differences in the morphology of the pubic symphysis observed in Eskimo skeletons, such as pitting and irregularities in the dorsal edge of the symphyseal surface only seen in females. Due to this apparent sexual dimorphism of the pubic symphysis, Gilbert & McKern (1973) developed female standards for the three-component method described in McKern & Stewart (1957).

Several studies such as Acsádi & Nemeskéri (1970), Hanihara & Suzuki (1978), Snow (1983), Jackes (1985), Meindl *et al*. (1985), and Pfeiffer (1985), have attempted to modify and improve on the work done by Todd (1920) and McKern & Stewart (1957). Based on their observations in a Hungarian sample, Nemeskéri *et al*. (1960) and Acsádi & Nemeskéri (1970) chose to modify Todd's (1920) phase-based method by reducing the number of phases and adjusting the age ranges. They suggested a five-phase system with descriptions of the morphogenesis of the surface, ventral and dorsal borders, and the formation of a rim around the rami.

Hanihari & Suzuki (1978) used multiple regression analyses to estimate age using seven morphological features on the pubic symphysis that Todd (1920) recognised as exhibiting evident age-related variations. Each feature was scored on a scale of 1 to 4, making it easy for beginners. The method was found inapplicable to individuals younger than 18 years or older than 38 years.

Meindl *et al.* (1985) tested the effectiveness of the Todd (1920), McKern & Stewart (1957), Gilbert & McKern (1973), and Hanihari & Suzuki (1978) methods on a sample from the Hamann-Todd Collection. They found that the Todd method performed the best; however, all methods underestimated the age, especially in older individuals. They also reported no significant biases were observed between the sexes or the two population groups (Garvin *et al*., 2012.). They then developed five biological stages that attempted to simplify Todd's methods while accounting for variation.

Katz & Suchey (1986) evaluated Todd's (1920) and McKern & Stewart's (1957) ageing methods using a sample of males aged between 14 and 92 years (Garvin *et al*., 2012). They reported that Todd's method overestimates age and that both methods do not account for age variability in the advanced stages of degeneration in the pubic symphyses. They also criticized the use of three components in McKern & Stewart's method, stating that there is not enough variation among the three components to allow for independent scoring (Garvin *et al*., 2012). Katz & Suchey (1986) recommended using a six-phase system based on the original Todd method, paving the way for the development of the popular Suchey-Brooks method, which is still used today.

Brooks & Suchey (1990) examined 273 female pubic symphyses and 739 male pubic symphyses from the study by Katz & Suchey (1986) to develop their method. To successfully capture a large degree of variability in the symphyseal phase, their sample consisted of individuals aged between 14 and 99 years from different socio-economic backgrounds and different geographical locations within and outside the United States (Brooks & Suchey, 1990). They revised the original Todd method to a six-phase method with descriptions of morphological changes, wide, overlapping age ranges associated with each phase, and standards for both sexes.

Due to its popularity, the Suchey-Brooks method has been tested many times (e.g., Klepinger *et al*., 1992; Schmitt, 2004; Djurić *et al*., 2007; Hens *et al.*, 2008; Joubert *et al*., 2019). Klepinger *et al*. (1992) tested the accuracy of the McKern-Stewart, Gilbert-McKern, and Suchey-Brooks ageing methods on a modern autopsy sample. They found that the Suchey-Brooks method is the most accurate among the three; however, it underestimated the variability for middle-aged and older adults. These results were supported by Schmitt (2004), who tested the Suchey-Brooks method on a largely middle-aged and older Thai sample and found that bias and inaccuracy increased with age and that the method tends to underestimate age, except for younger individuals, who were fairly accurate. Hens *et al.* (2008) tested the method on an Italian sample with similar results, stating that it underestimates true ages for individuals over 40 years.

2.1.3. Proximal humerus and femur

The structural changes that can be observed in spongiosa of the humeral epiphysis were first described by Wachholz (1894). Following his work, several researchers have explored how to evaluate the age-related changes seen in the proximal humerus using radiographs and cross-sections (Schranz, 1959; Ponsold, 1950; Hansen, 1953/54; Nemeskéri, 1960; Acsádi & Nemeskéri, 1970; Bergot & Bocquet, 1976; Bocquet & Bergot, 1977). Wachholz (1894) examined the internal structure of 230 freshly sectioned humeri from individuals aged between eight and 54 years and developed age-related timings and descriptions for both males and females (Schranz, 1959). These descriptions are based on observations of the union of the epiphyses to complete maturing, followed by the structural senescent changes visible in the trabecular pattern and medullary cavity (Schranz, 1959).

In contrast, Hansen (1953/54) examined 250 macerated and sectioned proximal humeri and developed a chronological sequence for ageing. This sequence describes the ages at which epiphyseal union would begin and end, the transition from radial to columnar trabecular structure, the timing of lacunae formation and enlargement, and the increase in overall transparency. After observing the age-related changes in the proximal humerus, Hansen (1953/54) also studied the age-related changes in the proximal femur, developing an ageing chronological sequence for the proximal femur as well. This sequence describes changes in the location of the medullary cavity's apex, changes in the structural organisation of the trabeculae, the degree of rarefaction, the formation of cavities, and the atrophy on the femoral head's outer surface (Acsádi & Nemeskéri, 1970). Before Hansen's work, von Meyer (1867) was the first to describe the trabecular arrangement of the proximal femur, and the age-related changes occurring in the trabeculae were mentioned by Testut & Laterjut (1948).

Schranz (1959) expanded on Wachholz's (1894) work, criticising Wachholz's use of freshly sectioned bones, arguing that the presence of soft tissue hides many of the subtle internal structural changes. As a result, Schranz (1959) suggested using macerated sectioned proximal humeri, which he claimed offered even more detail than radiological images. He stated that the results of his study validated the research done by Hansen (1953/54). Hansen (1953/1954) and Schranz (1959) noted minor variations in the onset and timing of senescent traits between males and females. According to their theories, females aged faster than males, and this disparity grew larger as they aged, with the two sexes' differences amounting to a difference of seven to ten years.

Nemeskéri *et al*. (1960) and Acsádi & Nemeskéri (1970) examined 105 skeletons of known sex and age to study the age-related changes in the proximal humerus and femur, using cross-sections or radiographs. Based on work done by past researchers, they developed a six-phase scoring system for both the proximal humerus and proximal femur that is not sex-specific as

they did not observe any differences in the timing of changes between males and females. These phases describe the changes in the height of the apex of the medullary cavity, changes in the trabecular pattern, rarefaction, cavity formation and cortical bone thinning (Acsádi & Nemeskéri, 1970). Acsádi & Nemeskéri (1970) discovered that age-related changes occurred quite gradually at a young age and progressively accelerated in old age. They noted that the various phases represented age limits with varying ranges and group intervals, and as such, individual factors should always be considered. As a result, later phases overlap partially or completely (Acsádi & Nemeskéri, 1970). This method offers very wide age ranges in earlier phases, and somewhat narrower age ranges in later phases, making it partially useful in older individuals (İşcan & Steyn, 2013).

Drawing from work done by Hansen (1953/54) and Acsádi & Nemeskéri (1970), other authors proposed modified methods for the proximal humerus and femur. Szilvàssy & Kritscher (1990) proposed a seven-phase scoring system using cross-sections, radiographs, or tracing. Tracings would be used when the observer is unsure about using radiographs and would be created by placing the X-ray plate onto a transparency and tracing the spongiosa structure (Szilvàssy & Kritscher, 1990). Apart from phase one, Szilvàssy & Kritscher's (1990) method provided age ranges with ten-year intervals, which are somewhat narrower than those provided by Acsádi & Nemeskéri (1970), who reported age ranges with a plus/minus 15 years. Using 60 radiographs of living individuals, Lynnerup *et al.* (1990) proposed a five-phase scoring system for the proximal femur. They attributed the number of phases to the fact that soft tissue decreased the detailed resolution of radiographs. Similar to Acsádi & Nemeskéri (1970), they reported very wide age ranges with approximately 20 years between phases and recommended that their method be used along with other age estimation techniques.

For their study, Walker & Lovejoy (1985) examined the proximal femur, humerus, clavicle and calcaneus of 500 individuals from the Hamann-Todd Collection using radiographic

assessment and seriation. Results showed that the clavicle provided the strongest correlation to age, followed by the femur and humerus, with the calcaneus showing the worst correlation. In contrast to Acsádi & Nemeskéri (1970), Walker & Lovejoy (1985) observed sex-specific trends in ageing at these locations. They reported a high rate of bone activity, relating to trabecular arrangement and bone loss, in the proximal femora of females in the third and fourth decade of life, suggesting that its use as an age indicator is unreliable. Both the humerus and femur were found to be equally reliable as age indicators in males, whereas, the humerus performed better in females. Based on their observations, they proposed an improved eight-phase scoring system for the proximal femur and clavicle, with five-year age ranges ranging from 18 to 55+ years (Walker & Lovejoy,1985).

Methods using age-related changes in the proximal humerus and proximal femur, viewed on radiographs, for age estimation have been proposed and modified since 1894. Several other studies have explored alternative approaches to visualise and measure age-related changes in long bones. These include the use of computed tomography (CT), quantitative computed tomography (QCT), multidetector computed tomography (MDCT), microcomputed tomography (micro-CT) and dual-energy X-ray absorptiometry (DXA) (e.g. Tingart *et al*., 2003; Riggs *et al*., 2004; Cooper *et al*., 2007; Cui *et al*., 2008; Lochmüller *et al*., 2008; Bergot *et al*., 2009; Grabherr *et al*., 2009; Ito *et al*., 2011; MacDonald *et al*., 2011; Roosa *et al*., 2012; Chen *et al*., 2013; Botha *et al*., 2019).

2.1.4. Multifactorial approaches

Lovejoy *et al*. (1985a) proposed that a large number of the problems seen in age estimation techniques are because a majority of techniques use one or two age traits at a time. Even with the Acsádi & Nemeskéri (1970) Complex Method, the pubic symphysis outperformed the other

traits. As a result, using a sample from the Hamann-Todd collection, Lovejoy *et al*. (1985a) developed their multifactorial summary age (MSA) method using five indicators, namely the pubic symphysis, auricular surface, proximal femur, dentition (dental wear) and ectocranial suture closure. Maat (1987) compared the practical applicability of the Acsádi & Nemeskéri (1970) Complex Method to the Lovejoy et al. (1985a) MSA approach and discovered that the Complex Method was preferred as it is simpler to apply (Hartnett-McCann et al., 2018).

Other multifactorial methods include those by Boldsen *et al*. (2002), Samworth & Gowland (2007), and Uhl & Nawrocki (2010). Samworth & Gowland (2007) combined the Brooks & Suchey (1990) pubic symphysis method and Lovejoy *et al*. (1985b) auricular surface method to develop a method that makes use of single or multiple indicators using a "lookup table". To develop their method, they used a sample consisting of historical and modern skeletons, which the authors suggested provided a good range of potential variations observed in ageing. Passalacqua (2010) evaluated these "lookup tables" and found that they produced results that were significantly better than those obtained using the Brooks & Suchey (1990) and Lovejoy *et al*. (1985b) methods.

The Acsádi & Nemeskéri (1970) Complex Method has; however, been criticised for its small sample size, which is not large enough to accurately represent inter-individual differences, the statistical validity of the averaging procedure, and its constant under- and overestimation of age for older and younger individuals, respectively (Saunders *et al*., 1992; Molleson *et al*., 1993).

Molleson *et al*. (1993) tested the Complex Method on a sample of 219 individuals of known age recovered from the Christ Church cemetery at Spitalfields. They found that the pubic symphysis performed the best, followed by the femur and humerus; however, they noted for the proximal femur and humerus, it was difficult to distinguish between stages 4 and 5. They

provided several possible explanations for this, including the use of radiographs rather than cross-sections to visualise the internal structure of long bones, post-mortem changes, and pathology, all of which would make it difficult to visualise features that distinguish two phases (Molleson *et al.*, 1993). Overall, they reported a consistent overestimation of age for individuals younger than 40 years and an underestimation of age for individuals older than 70 years, with the method performing slightly better in males than in females (Molleson *et al.*, 1993). They reported that the method performed best in individuals between 50 and 70 years, suggesting that this is because of the interpopulation differences between the Spitalfields and Hungarian samples.

Meyer *et al.* (2021) tested the Complex Method on a modern white South African skeletal sample and also found that it under- and over-estimated age for older and younger individuals, respectively, with only 16.3% of the individuals aged within the reported ±2.5 years of their true ages. Although the results supported the notion that the method tends to under- and overestimate age, they also showed that it performed well for certain age indicator combinations, especially for individuals aged 40-69 years, indicating that the method should not be dismissed (Meyer *et al.*, 2021). Instead, they proposed that the method could be improved by using sex- and population-specific skeletal elements and weighing systems that correlate to the degree of correlation detected for each of the skeletal components included in the method (Meyer *et al.*, 2021). They also suggested using a Bayesian approach to counteract the effects of under- and over-estimation and to make the method more applicable in a forensic context (Meyer *et al.*, 2021).

Most researchers agree that the use of multiple indicators in age estimation far outperforms the use of a single age indicator; however, how to combine information to develop a multifactorial method with the best results is unclear (Garvin *et al.*, 2012). Current multifactorial methods are not seen as statistically reliable, as they were developed using different samples, under a

different set of assumptions, and may even provide different statistical information (Garvin *et al.*, 2012). As a result, new statistical approaches have been proposed, as outlined below.

2.2. Statistical analyses

Traditionally in age estimation techniques, the morphological state of the age indicator is utilised to determine age-at-death (Hartnett-McCann *et al.*, 2018). In the formation of ageing standards, least square regression analyses regress age-related skeletal features onto the age structure of a specific reference population (Aykroyd, 1997; Tersigni-Tarrant & Shirley, 2013). In this sense, the age-at-death is used as the dependent variable in the regression analysis, thereby, utilising inverse calibration to obtain age estimates (Aykroyd, 1997; Tersigni-Tarrant & Shirley, 2013; Hartnett-McCann *et al.*, 2018). Inverse calibration requires regression toward the mean; therefore, using it to develop a method means age estimates obtained using the method will be biased in the direction of the mean of the reference sample or population (Tersigni-Tarrant & Shirley, 2013). This phenomenon is referred to as "age mimicry."

"Age mimicry" was first described by Bocquet-Appel & Masset (1982) in their paper "Farewell to palaeodemography", where they tried to discuss the statistical inadequacies seen in palaeodemographic age estimation studies. This paper inspired many researchers to seek more accurate osteological and mathematical age estimation techniques (Tersigni-Tarrant & Shirley, 2013). An outcome of this was a three-day workshop entitled "Mathematical Modelling for Paleodemography: Coming to Consensus", hosted by the Max Planck Institute for Demographic Research in Rostock, Germany, in June 1999 (Hoppa and Vaupel, 2002). This workshop resulted in the "Rostock Manifesto", which outlines some of the key ideas learned throughout the workshop on how to more skillfully address statistical and sampling challenges in age estimate being accepted (Hoppa and Vaupel, 2002).

The Rostock Manifesto makes several important points, one of which is that methods should be devised to allow for the assessment of Pr(c|a), or the likelihood of observing skeletal characteristic(s) (c), given a known age (a) (Hoppa and Vaupel, 2002). When using age-estimation techniques, one would typically be able to calculate Pr(a|c), which is the likelihood that an individual is that particular age (a), given the presence of one or more skeletal characteristics (c) (Hoppa and Vaupel, 2002). Both probabilities, which are obtained using very different statistical methods, are crucial for age estimates. While Pr(c|a) can be calculated using Bayesian statistics, Pr(a|c) can be acquired from a suitable reference sample. Bayesian statistics require the use of prior distribution of ages-at-death (informed or uniform priors), which assists in countering the effects of "age mimicry" (Hoppa and Vaupel, 2002).

In response to this manifesto, Boldsen *et al*. (2002) developed their multifactorial method known as Transition Analysis (TA) which uses Bayesian statistics. Building on work done by McKern and Stewart (1957), Gilbert and McKern (1973), Lovejoy *et al*. (1985a), and Meindl and Lovejoy (1985), Boldsen *et al*. (2002) proposed a component-based scoring system using the pubic symphysis, auricular surface, and cranial sutures. It is a parametric method that relies on the transition from one phase to a higher, subsequent phase (İşcan & Steyn, 2013). The skeletal features are scored individually according to the descriptions created by Boldsen *et al*. (2002) and these scores are analyzed using the ADBOU computer program. This program was developed by Boldsen *et al*. (2002), and guides the user through data collection on several skeletal features before using transition analysis to determine age-at-death (Garvin *et al.*, 2012). The program includes the use of either an informed or a uniform prior distribution and provides an age estimate with a 95% confidence interval, which addresses the issue of age mimicry (İşcan & Steyn, 2013). Using a sample of modern North American males and females from the Bass Donated Collection and Mercyhurst cases, Milner and Boldsen (2012) conducted a validation study of the method. They reported that TA performed worse than experience-based

age assessments such as the Perizonius (1984) cranial sutures method or the Lovejoy *et al*. (1985b) auricular surface method. They suggested that this was likely because the method is limited to only three skeletal features. Nevertheless, the use of TA and Bayesian statistics has been shown to reduce bias caused by age mimicry successfully (Langley-Shirley & Jantz, 2010; Godde & Hens, 2012; Hens & Godde, 2016).

2.2.1. Bayesian statistical approach in age estimation

Langley-Shirley and Jantz (2010) used Bayesian statistics to study 1289 clavicles sampled from the William F. McCormick Clavicle Collection and the Hamann-Todd Osteological Collection, as well as raw data from McKern & Stewart's (1957) study consisting of Korean War males. Age was determined using McKern & Stewart's (1957) five-phase method and a three-phase scoring system. Using an informative prior, Bayesian analysis provided age ranges with four distinct probability percentages (50%, 75%, 90% and 95%) for each phase of the McKern-Stewart method and the three-phase scoring system. Results showed that epiphyseal union began about four years earlier in modern Americans than in the early 20th century Americans and 3.5 years earlier in Korean War Americans, highlighting the need for modern standards to assess age in modern individuals. They concluded that age ranges developed using Bayesian statistics avoided the problems of age mimicry and developmental outliers.

Godde & Hens (2012) examined the accuracy of age estimations generated by the Suchey-Brooks (1990) method vs Transition analysis combined with Bayesian statistics using a historic archaeological sample from Italy. Additionally, they examined the accuracy using informative and uniform priors. Results showed that Transition analysis coupled with Bayesian statistics increased the accuracy of the Suchey-Brooks method. They noted that the informative and uniform priors performed similarly, suggesting that a uniform prior might be a good substitute

for an informative prior in some cases. Like Langley-Shirley and Jantz (2010), they concluded that Bayesian statistics successfully avoided age mimicry.

Hens & Godde (2016) used Bayesian statistics to compare Lovejoy *et al.* (1985) and Buckberry & Chamberlain (2002) auricular surface ageing methods. They used a sample of Portuguese males as the target sample, Thai males for the informative prior and American white males for the Transition analysis. Results showed that Bayesian statistics improved age estimation accuracy for both methods, with the Lovejoy *et al.* (1985) method significantly better. They also concluded that Bayesian statistics elevated the concerns of bias and age mimicry when using the two methods.

Nikita *et al.* (2018) also tested the implementation of Bayesian statistics to Lovejoy *et al.* (1985) and Buckberry & Chamberlain (2002) iliac auricular surface ageing methods. They explored the performance of the methods using two modern Greek samples (Athenian and Cretan) with various informative priors and transition analysis samples from America, the Balkans and Greece. They found no improvements in age estimates except for the Athenian male sample for the Buckberry & Chamberlain (2002) method. Despite the poor results, Nikita and colleagues (2018) advocated using Bayesian methods, stating that Bayesian statistics can accommodate missing values for individual markers and may combine several skeletal age indicators and exclude markers that show a poor connection with age-at-death.

CHAPTER 3: MATERIALS AND METHODS

3.1. Materials

For this study, a total of 103 black South African individuals (54 males and 49 females), with known ages at death (ranging between 20 and 90 years) were sampled from the Raymond A. Dart Collection of Modern Human Skeletons housed at the University of the Witwatersrand, which houses cadaver-derived human skeletal remains (Dayal *et al*., 2009). The individuals were divided into age cohorts, with ten-year intervals for comparison (e.g., 20-29 years; 40-49 years, etc.). Equal numbers of individuals were sampled for each age cohort as far as possible. Only individuals who presented with a complete cranium and at least one pubic symphysis, proximal humerus, and proximal femur were included in the study sample. Individuals who presented with both the left and right sides of the bilateral skeletal elements, the left side was used for scoring.

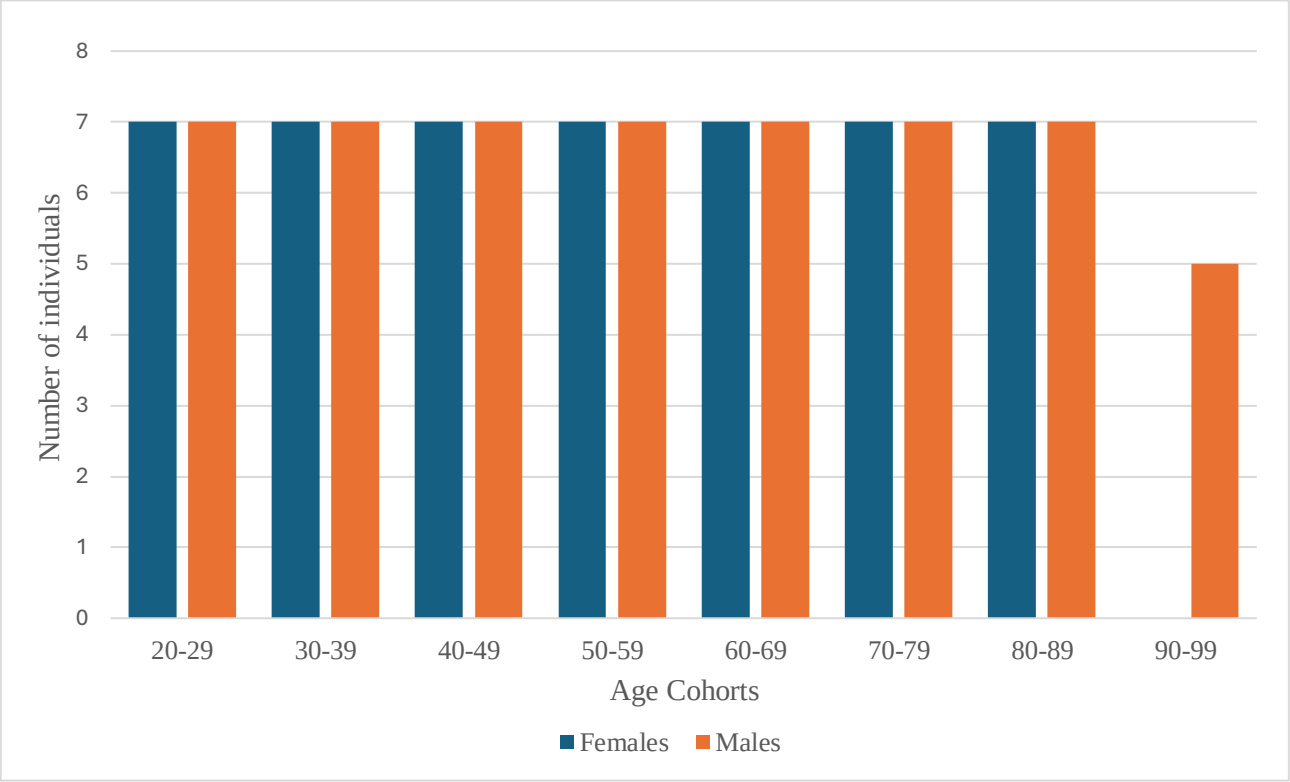


Figure 1: Number of individuals in the study.

3.2. Acsádi & Nemeskéri Complex Method

The Complex Method is a multifactorial age estimation developed by Acsádi & Nemeskéri (1970). It makes use of the endocranial sutures, pubic symphysis and internal structural changes in the proximal femur and humerus (observable through cross-sectioning of bone or radiographs). Each skeletal indicator was analysed based on the descriptions provided by Acsádi & Nemeskéri (1970) and scored using a stage- or phase-based system.

3.2.1. Cranial sutures

The endocranial sutures (coronal, lambdoid and sagittal) were divided into 16 sections of equal length (Perizonius, 1984), the coronal and lambdoid into three sections each (three left and three right), and the sagittal suture into four sections, and scored into five stages of suture closure. When the calotte has not been removed, a probe light was used to view the sutures through the foramen magnum. Each section was individually scored according to its degree of suture closure using the Acsádi & Nemeskéri (1970) and the Perizonius (1984) descriptions. The Acsádi & Nemeskéri (1970) descriptions are as follows: 0 = open suture; 1 = incipient closure; 2 = closure in process; 3 = advanced closure; 4 = complete closure. The classification system used by Acsádi & Nemeskéri (1970) was based on the "Martin's scale" developed by Martin & Saller (1957-1966). During his study into the effectiveness of the Acsádi & Nemeskéri (1970) cranial sutures in a Dutch sample, Perizonius (1984) found that the suture descriptions were too general and simple to be useful. Therefore, after direct communication with Nemeskéri, Perizonius (1984) developed more detailed descriptions with diagrams (Figure 2) for easier scoring. The Perizonius (1984) descriptions are as follows: Stage 0 = Open suture, there is still a little space left between the adjoining bones; Stage 1 = Suture is closed but clearly visible as a continuous, often zigzagged, line; Stage 2 = Suture line becomes thinner,

has fewer zigzags and may be interrupted by complete closure; Stage 3 = Only pits indicate where the suture is located; Stage 4 = Suture completely obliterated, even its location cannot be recognised (Figure 3). After scoring each section of the endocranial sutures, the scores were added and divided by 16 to obtain a mean closure stage, which is converted into a phase associated with a mean age and age range (Table 1). The mean phase was then used in the final age estimation calculations.

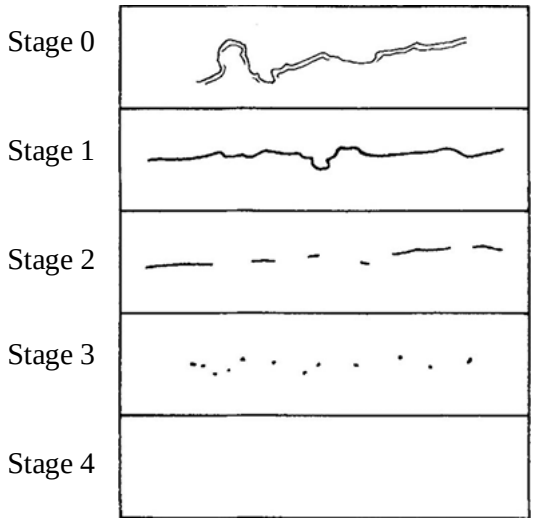


Figure 2: Five stages of endocranial suture closure (Reproduced from Figure 3 of Perizonius, 1984: 204).

Table 1: Endocranial suture stages with associated age ranges (Reproduced from Table 32 of Acsádi & Nemeskéri, 1970: 121).

Mean closure stage	**Phase**	**Mean age**	**Age range**	**Range description***
0.4 – 1.5	**I**	28.6	15 - 40	Juvenile – young adult
1.6 – 2.5	**II**	43.7	30 - 60	Young adult – middle adult
2.6 – 2.9	**III**	49.1	35 – 65	Young adult – middle adult
3.0 – 3.9	**IV**	60.0	45 – 75	Middle adult – old adult
4.0	**V**	65.4	50 - 80	Middle adult – old adult



Stage 0



Stage 1



Stage 2



Stage 3



Stage 4

Figure 3: Acsádi & Nemeskéri (1970) endocranial suture stages as observed in the black South African population.

3.2.2. Pubic Symphysis

Phase descriptions, along with their associated age ranges (Table 2), provided by Acsádi & Nemeskéri (1970) are based on the macroscopic changes to the symphyseal face, ventral and dorsal margins, and rim formation (Figure 4). Each of the five phases is correlated to a mean age and age range, which were used to get an age estimate. The phase descriptions are as follows:

Phase I: The symphyseal face surface is convex, transversed by horizontal ridges and grooves, and a curved transition in the region of the rami.

Phase II: The original structure of the symphyseal surface starts to disappear, with the ridges becoming flatter and the grooves shallower. On the ventral and dorsal margins, rim formation begins.

Phase III: The original symphyseal surface structure is completely destroyed, with some granular remnants remaining. Rim formation continues on the ventral and dorsal margins, with a well-defined border apparent in the rami margin.

Phase IV: The symphyseal face is completely smooth, with a sharp rim along the ventral and dorsal margins. The inferior end of the symphyseal face terminates in a ridge, forming an acute angle.

Phase V: The completely smooth symphyseal face is partly concave with a porous and shrivelled appearance. The fully developed ventral and dorsal rims rise above the face like a crest.

Table 2: Pubic symphysis stages and associated age ranges (Reproduced from Table 34 of Acsádi & Nemeskéri, 1970: 126).

Phase	**Mean age**	**Standard deviation (SD)**	**Age range**	**Calculate range (±3SD)**
I	26.3		18 -45	
II	46.5	1.76	23 – 69	41.2 – 51.7
III	51.1	1.62	25 – 76	45.8 – 56.3
IV	58.1	2.16	24 – 81	51.7 – 64.6
V	68.5	2.53	41 - 86	61.0 – 76.1



Phase I



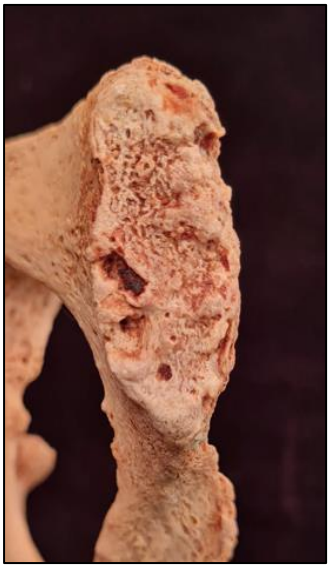
Phase II



Phase III



Phase IV



Phase V

Figure 4: Acsádi & Nemeskéri (1970) pubic symphysis phases as observed in the black South African population.

3.2.3. Proximal humerus and proximal femur

The internal structural changes of the proximal humerus and femur were observed on radiographs. Radiographs were obtained from the University of the Witwatersrand's Central Animal Services Unit using a Shimadzu Mobile Art Evolution MVX 200 mobile X-ray machine. X-rays were taken at between 73 – 80 kilovoltage (kV), 125 – 500 milliamperes (mA) and 10.0 milliamperes-seconds (mAs). All X-rays were exported as DICOM files and viewed using the Microdicom DICOM viewer [Software] version 2022.2.

For the proximal humerus, six characteristic phases along with their associated age ranges, are provided by Acsádi & Nemeskéri (1970). These phase characteristics include an assessment of the height of the apex of the medullary cavity, the transformation and rarefaction of the spongy substance, cavity formation in the major tubercle and thinning of the cortex (Figure 5, Table 3). The individual phase descriptions are as follows:

Phase I: The apex of the medullary cavity is positioned well below the surgical neck of the humerus, with the trabeculae having a radial system.

Phase II: The apex is at the height of the surgical neck or above. The trabecular system appears more fragile, with an ogival (curved and pointed) structure.

Phase III: The apex may have shifted up to reach the epiphyseal line, the trabecular system is ogival and columnar structures appear along the cortex at the border of the diaphysis and epiphysis.

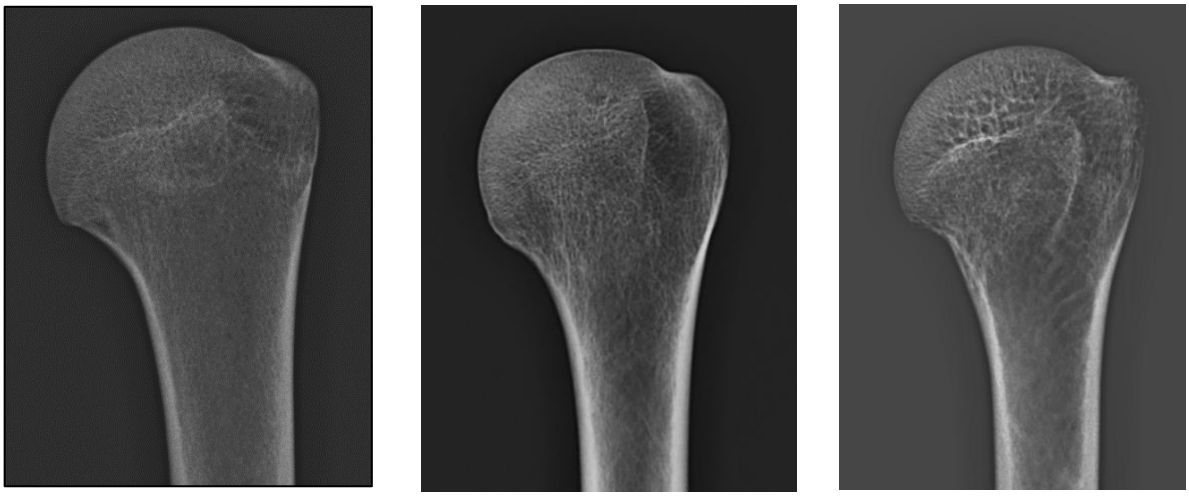
Phase IV: The apex may reach the epiphyseal line or extend above it., with the trabecular system showing gaps in the major tubercle. The columnar structure along both sides of the medullary cavity may occasionally be breached.

Phase V: The apex extends above the epiphyseal line with 2-5 mm lacunae developing in the major tubercle and discontinuous remains of the columnar structure appear on both sides of the medullary cavity.

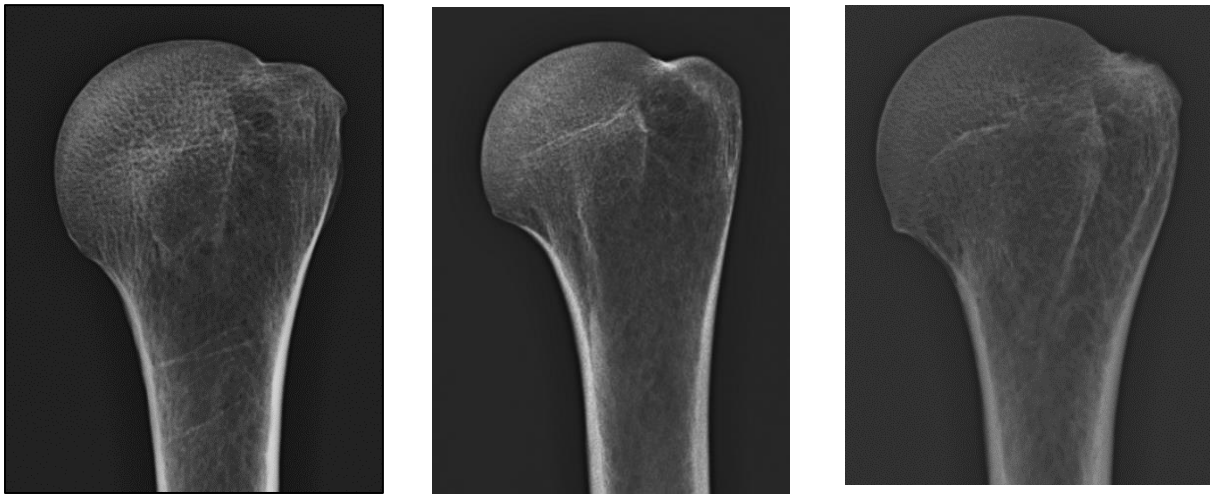
Phase VI: The cavity formed (more than 5 mm) in the major tubercle may reach the cortex. The apex extends upwards and merges with the cavity formed in the major tubercle. The trabecular system in the head of the humerus is intensely rarefied, with trabeculae becoming cobweb-like and torn. The cortex becomes thin and transparent, and only the remains of the spongiosa are left.

Table 3: Proximal humerus stages and associated age ranges (Reproduced from Table 33 of Acsádi & Nemeskéri, 1970: 126).

Phase	**Mean age**	**Standard deviation (SD)**	**Age range**	**Calculate range (±3SD)**
I	41.1	6.60	18 - 68	21.3 – 60.9
II	52.3	2.51	24 – 76	44.8 – 59.8
III	59.8	3.59	37 – 86	49.0 – 70.5
IV	56.0	1.84	19 – 79	50.5 – 61.6
V	61.0	2.05	40 – 84	54.9 – 67.2
VI	61.1	3.39	38 - 84	50.9 – 71.2



Phase I **Phase II** **Phase III**



Phase IV **Phase V** **Phase VI**

Figure 5: Acsádi & Nemeskéri (1970) proximal humerus phases as observed in the black South African population.

Assessment of the proximal femur also includes scoring into one of six phases assessing the position of the apex of the medullary cavity, structure and rarefaction of the trabecular system forming spongiosa, cavitation and atrophy of the outer surface of the femoral head (Figure 6). As with the other skeletal age indicators, the proximal femur phases are also correlated to age ranges provided by Acsádi & Nemeskéri (1970) (Table 4). The individual phase descriptions are as follows:

Phase I: The apex of the medullary cavity is positioned well below the lesser trochanter, with the trabeculae appearing thick and truss-like.

Phase II: The apex reaches or extends past the lower limit of the lesser trochanter at the border of the diaphysis and epiphysis. The trabecular pattern of the fasciculus trochantericus and fasciculus arciformis show rarefaction, with early rarefaction mostly noticeable in the medial part of the femoral neck.

Phase III: The apex reaches the upper limit of the lesser trochanter. Marked rarefaction of the trabecular pattern occurs in the medial part of the neck as well as the greater trochanter.

Phase IV: The apex extends above the upper limit of the lesser trochanter, with a delimited cavity of 5-10 mm in diameter appearing in the medial part of the femoral neck. Distinct rarefaction can be seen in the border of the diaphysis and epiphysis, in the greater trochanter and in the head below the fovea capitis.

Phase V: The apex extends beyond the upper limit of the lesser trochanter and only cellular remnants of the original trabecular system are present in the neck. A delimited cavity of approximately 3-5 mm in diameter appears in the greater trochanter. Cavities form in the head beneath the fovea capitis and at the medial and lateral borders.

Phase VI: Cavities formed in the neck (more than 10 mm) and greater trochanter (more than 5 mm) have enlarged. The cavity formed in the medial part of the femoral neck merges with the

medullary cavity and only remnants of the original trabecular structure remain along the cortex.

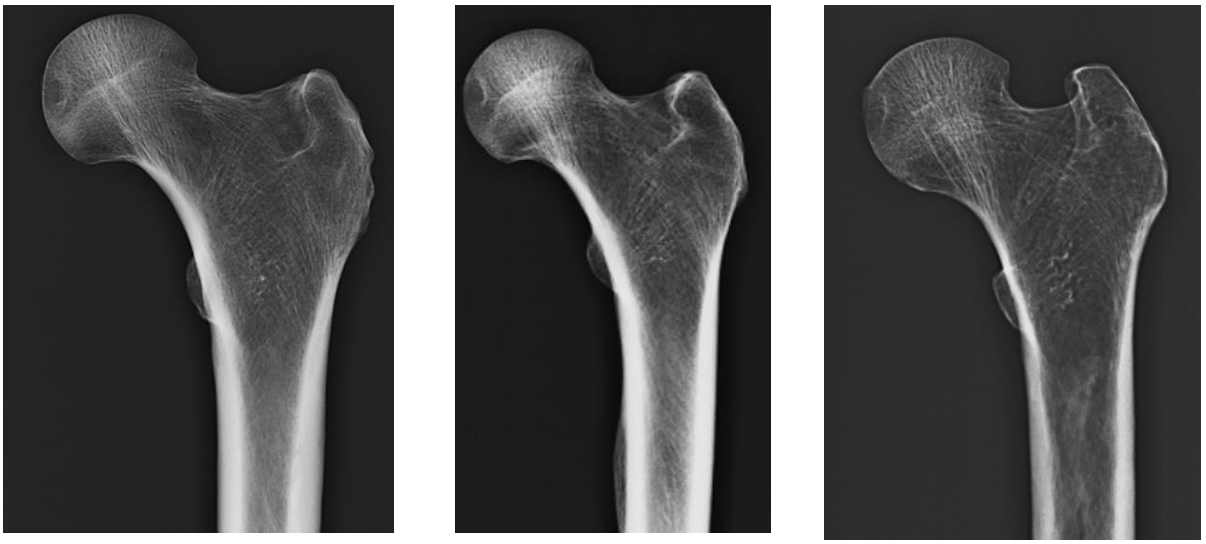
The cortex becomes thin and transparent.

Table 4: Proximal femur stages and associated age ranges (Reproduced from Table 35 of Acsádi & Nemeskéri, 1970: 128).

Phase	**Mean age**	**Standard deviation (SD)**	**Age range**	**Calculate range (±3SD)**
I	31.4		18 – 52	
II	44.0	2.60	19 – 61	36.2 – 51.8
III	52.6	1.86	23 – 72	47.0 – 58.2
IV	56.0	2.32	32 – 86	49.0 – 63.0
V	63.3	2.17	38 – 84	56.8 – 69.9
VI	67.8	3.64	25 - 85	56.9 – 78.7



Phase I **Phase IV** **Phase IV**



Phase IV **Phase IV** **Phase IV**

Figure 6: Acsádi & Nemeskéri (1970) proximal femur phases as observed in the black South African population.

Acsádi & Nemeskéri (1970) acknowledged that not all skeletal remains will necessarily be preserved, especially in archaeological and forensic cases. Therefore, the Complex Method makes provision for missing skeletal indicators. The mean ages from all or a combination of

the four age indicators (endocranial sutures, pubic symphysis, proximal femur and humerus) were used to estimate age in a black South African sample. Age was calculated using the lower-, mean- and upper-limit age ranges (Table 5) provided by Acsádi & Nemeskéri (1970) for each age indicator, starting with the pubic symphysis (as a weighting factor). The phase score of the pubic symphysis determined the age limit. Phases I and II indicate individuals generally younger than 50 years and for those individuals, the lower-limit age ranges were used for each age indicator. In contrast, phase III indicates individuals generally aged about 50 years, and therefore, the mean-limit age ranges were used for each age indicator. Finally, phases IV and V indicate individuals generally older than 50 years, and for these, the upper-limit age ranges were used for each age indicator. The values for all the available skeletal age indicators were added and divided by the total number of age indicators used (maximum 4, minimum 2) to obtain a final age estimate. Acsádi & Nemeskéri (1970) reported a margin of error of ±2.5 years when all four skeletal indicators are used. An example of utilising the limit table to obtain an age estimate is presented in Table 6.

Table 5: Limit ranges provided by Acsádi & Nemeskéri (1970: 131).

	Lower limit range				**Mean**				**Upper limit range**			
Phases	**S**	**PS**	**PF**	**PH**	**S**	**PS**	**PF**	**PH**	**S**	**PS**	**PF**	**PH**
I	23	23	23	23	30	32	33	41	39	40	43	57
II	35	37	35	41	44	44	44	51	52	49	53	61
III	45	46	44	48	53	52	52	57	60	58	59	65
IV	53	54	50	52	60	60	58	59	66	68	66	67
V	58	61	54	54	63	67	63	61	72	75	71	69
VI	-	-	58	55	-	-	67	62	-	-	76	70

S=Suture; PS=pubic symphysis; PF=proximal femur; PH= proximal humerus

Table 6: Example of calculating the age estimate using the Complex Method.

Skeletal region	**Phase**	**Range**	**Mean age**
Pubic symphysis	III	37-49	52
Endocranial sutures	III	45-60	53
Proximal femur	II	35-53	44
Proximal humerus	II	41-61	51
		Total	200 (/4)
		Age estimate	**50 ±2.5 years***

*Margin of error originally provided by Acsádi & Nemeskéri (1970).

When three skeletal indicators were used, the pubic symphysis was used as the weighting factor and the summed value was divided by three. These combinations were reported by Acsádi & Nemeskéri (1970) to have the same margin of error of ±2.5 years. When three skeletal indicators were used, with the pubic symphysis not being one of the indicators, the proximal

femur was used as the weighting factor in the same way as the pubic symphysis. For the proximal femur, phases I and II meant the lower limit ranges should be used; for phase III, the mean limit ranges; and for phases IV, V, and VI, the upper limit ranges. Acsádi & Nemeskéri (1970) reported a margin of error of ±3 years for this combination.

When two skeletal indicators were used, values were averaged. If the skeletal indicators used included the cranial sutures and either the pubic symphysis or proximal humerus, the pubic symphysis was used as a weighting factor. If the cranial sutures and proximal humerus combination is used, the average of the upper and lower limit values of each skeletal indicators' phase should be used. If the skeletal indicators used included the proximal femur and either the cranial sutures or proximal humerus, the proximal femur was used as a weighting factor. Acsádi & Nemeskéri (1970) noted that should the proximal femur phase score differ extremely from the other skeletal indicator, the average of the upper and limit ranges should be used. As Acsádi & Nemeskéri (1970) did not state what constitutes an extreme difference, it was decided that for this study an extreme difference would be a difference of three or more phases. Acsádi & Nemeskéri (1970) reported a margin of error of -4 and +5 years for these combinations,. If the skeletal indicators used included the pubic symphysis and proximal femur, the mean limit range values were averaged and a margin of error of ±4 years (Acsádi & Nemeskéri, 1970). When only one skeletal indicator was available, Acsádi & Nemeskéri (1970) provided phase-specific age ranges (Table 7).

Table 7: Single skeletal indicator phases and associated mean ages (Reproduced from Acsádi & Nemeskéri, 1970: 135).

Cranial sutures			**Pubic symphysis**		
Phase	**Age range**	**Mean age**	**Phase**	**Age range**	**Mean age**
I	23 – 40	32	**I**	23 – 40	32
II – III	30 – 60	45	**II**	35 – 55	45
IV - V	40 - 80	60	**III**	40 – 60	50
			IV	50 – 70	60
			V	60 - 80	70
Proximal humerus			**Proximal femur**		
Phase	**Age range**	**Mean age**	**Phase**	**Age range**	**Mean age**
I	23 – 60	42	**I**	23 – 40	32
II	30 – 60	45	**II**	35 – 55	45
III – IV	40 – 70	55	**III**	40 – 60	50
V - VI	50 – 80	65	**IV**	50 – 70	60
			V	50 - 73	62
			VI	50 - 80	65

3.3. Testing the Complex Method

Age was estimated using blind analysis to avoid researcher bias therefore the true age at death was umknown to the researcher or secondary observer during scoring. The original scoring criteria of Acsádi & Nemeskéri (1970) were used to score each of the four skeletal indicators for each of the 103 individuals. As the Complex Method is a multifactorial method that employs four skeletal indicators that may be combined in multiple ways, age was estimated using various skeletal indicator combinations, totalling 15 combinations for each individual.

These 15 combinations provided by Acsádi & Nemeskéri (1970) included using all four skeletal indicators (1. cranial sutures, pubic symphysis, proximal humerus and proximal femur), using three skeletal indicators (2. cranial sutures, pubic symphysis and proximal humerus; 3. cranial sutures, pubic symphysis and proximal femur; 4. cranial sutures, proximal humerus and proximal femur; 5. pubic symphysis, proximal humerus and proximal femur), using two skeletal indicators (6. cranial sutures and pubic symphysis; 7. cranial sutures and proximal humerus; 8. cranial sutures and proximal femur; 9. pubic symphysis and proximal humerus; 10. pubic symphysis and proximal humerus; 11. proximal humerus and proximal femur) and lastly, using single skeletal indicators (12. cranial sutures; 13. pubic symphysis; 14. proximal humerus; 15. proximal femur) to estimate age. Phase estimates and mean ages were captured in Microsoft Excel and adjusted based on the lower-, mean- and upper-limit age ranges (Table 5) mentioned above to estimate age for each individuals.

The data was tested for normality using the Shapiro-Wilk test (with alpha set at 0.05) and for non-parametric data. The test showed that the data was non-normally distributed for the pooled sexes, females, and males' samples; therefore, non-parametric tests were used for all data analyses.

Descriptive statistics (e.g. mean, median, minimum, maximum, standard deviation and confidence intervals) were calculated to obtain a summary of the spread of the data. This was done for the pooled sexes sample and the females and males sample separately and only for the combination using all four skeletal indicators.

Intra- and inter-observer errors were calculated by re-assessing of all four skeletal indicators by the principal investigator and a secondary observer of a small subsample of the study sample (20 individuals). Levels of repeatability were calculated using Cohen's weighted Kappa coefficient. Table 8, adapted from Altman (1991), was used as a guide to interpret the results.

Table 8: Interpretation of Cohen's weighted Kappa coefficient (Adapted from Altman, 1991).

Value of K	Strength of agreement
< 0.20	Poor
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Good
0.81 – 1.00	Very good

Wilcoxon signed-rank tests (with alpha set at 0.05) were performed to test whether differences exist between age estimates obtained for males and females per phase for each of the four skeletal indicators. Additionally, to test each of the skeletal indicators utilised in the Complex method, Spearman's rank correlation coefficient was used to correlate each of the skeletal indicators' phase scores to increasing age to ascertain whether the use of the pubic symphysis and failing that the proximal femur, is warranted as a weighting indicator in a black South African sample. This was done separately for the pooled sexes, females, and males. Table 9, adapted from Mukaka (2012: 71), was used as a guide to interpret the strength of agreement.

Table 9: Interpretation of Spearman's rank correlation coefficient (Adapted from Mukaka, 2012: 71).

p-value	Strength of agreement
0.90 – 1.00	Very high correlation
0.70 – 0.90	High correlation
0.50 – 0.70	Moderate correlation
0.30 – 0.50	Low correlation
0.00 – 0.30	Very low correlation

Spearman's rank correlation coefficient was also used to determine the level of correlation between the estimated ages obtained using the original Acsádi & Nemeskéri (1970) averaging

method and true age. This was done for each of the 15 skeletal indicator combinations and the pooled sexes, females, and males, separately. Again, Table 9 above was used to interpret the results.

To test for accuracy, the percentage of individuals whose estimated age fell within ±2.5 years, ±5 years, ±10 years, ±15 years and more than 15 years of the true age were calculated. This was done for each of the 15 skeletal indicator combinations and the pooled sexes, females, and males, separately.

Bias and inaccuracy were calculated for the pooled sexes, females and males samples, using all 15 skeletal indicator combinations. Bias refers to the direction of under- or overestimating age, and inaccuracy determines the overall under- and overestimation of age (Meindl *et al.*, 1985). Bias and inaccuracy were calculated per 10-year age cohort as follows:

$$\text{Bias} = \frac{\sum(\text{estimatedage} - \text{ageatdeath})}{n}$$

$$\text{Inaccuracy} = \frac{\sum|\text{estimated age} - \text{age at death}|}{n}$$

All statistical tests were computed using Microsoft Excel version 2203, SPSS Statistics version 25 (IBM, Armonk, NY, USA), and MedCalc for Windows version 18 (MedCalc Software, Ostend, Belgium).

3.2. Testing the Bayesian Methods

In addition to testing the performance of the Acsádi-Nemeskéri Complex Method on a black South African sample, the use of a more statistically robust method was utilised to see whether this would provide more accurate results than that of Acsádi & Nemeskéri's (1970) proposed averaging method. For this, Bayesian statistics, which enable probability estimates, were

explored. Bayes' theorem uses prior probabilities, likelihood, and posterior probabilities of point estimates. The Bayes' theorem is as follows (Tersigni-Tarrant & Shirley, 2013):

$$\Pr(a|c) = \frac{Pr(c|a)\, f(a)}{\int_{shift}^{\infty} Pr(c|x) f(x) dx}$$

First, the prior distribution or prior probability [f(a)], which refers to the probability distribution of ages-at-death in the target sample, was calculated (Tersigni-Tarrant & Shirley, 2013). This was calculated based on the documented ages of all the black individuals housed in the Raymond A. Dart Collection of Modern Human Skeletons using a Gompertz hazard model. The informative priors were calculated for the pooled sexes (n = 2851), females (n = 1114) and males (n = 1737) separately. Next, the likelihood [Pr(c|a)] was determined, where Pr denotes the probability of observing an assemblage of skeletal characteristics (c), given known age-at-death (a) (Tersigni-Tarrant & Shirley, 2013). This was calculated using transition analysis from the black South African sample used in the study. The posterior density function associated with each phase of each skeletal indicator was calculated by combining the prior distributions with the likelihood probabilities obtained. The posterior density function was then used to determine each phase's highest posterior density (HPD) and highest posterior density region (HPDR). The final step was the determination of the posterior probability [Pr(a|c)]; the probability that the skeletal remains are from an individual who died at age (a), given the evidence concerning characteristics of the skeletal region/stage (c). This constituted the final age estimate to which true ages and estimated ages obtained using the Acsádi & Nemeskéri (1970) averaging method were compared.

For this study, the Bayesian statistics described above were used to develop three methods, termed the Bayesian Method, the Collapsed Bayesian Method and the Pooled Bayesian

Method. The Bayesian Method was developed first and during transition analysis, it was noted that the ages-at-transition were very unrealistic as most individuals cannot reach that age. To improve this, the Collapsed Bayesian and Pooled Bayesian Methods were developed. The Collapsed Method was developed with the same sample size; however, with fewer phases to increase the number of individuals per phase. The Pooled Bayesian Method was developed with a bigger sample size and an increased number of individuals per phase. HPDs and HPDRs were calculated using each of the 15 skeletal indicator combinations provided by Acsádi & Nemeskéri (1970).

All Bayesian analyses were run in the R Studio version 4.1.2, using statistical packages, "survival" and "optimx" for the Gompertz hazard models, "VGAM" for the transition analysis and "nleqslv" for the highest posterior density (HPD) and highest posterior density region (HPDR), which were downloaded from CRAN (Comprehensive R Archive Network; https://cran.r-project.org). R code developed by Nikita and colleagues (2018) was used in this study.

Spearman's rank correlation coefficient was calculated to determine the level of correlation between the estimated ages obtained using the Bayesian Method with the true ages. The percentage of individuals whose estimated age fell within ±2.5 years, ±5 years, ±10 years, ±15 years and more than 15 years of the true age were calculated to determine the accuracy of each method. Age estimates obtained using the methods were compared to those originally obtained using the Acsádi-Nemeskéri Complex Method and Bayesian Method (for the Collapsed Bayesian and Pooled Bayesian Methods) to determine whether the former provided more accurate age estimates.

CHAPTER 4: RESULTS

4.1. Complex Method

4.1.1. Descriptive Statistics and Normality

Basic descriptive statistics, including confidence intervals, related to the true ages and the estimated ages for the sample are provided in Table 10. The true age refers to the recorded age-at-death obtained from the Raymond A. Dart Collection of Modern Human Skeletons' catalogue entries reflecting self-reported ages. In contrast, the estimated age refers to the ages obtained using Acsádi & Nemeskéri's Complex Method (1970), derived using all four skeletal indicators. Looking at the mean true age (55.85 years; SD = 20.92) versus the mean estimated age (59.98 years; SD = 8.41) for the pooled sexes sample the estimated age was slightly older. When the sample was separated by sex, the same trend could be seen for both females and males; however, female true versus estimated ages showed a larger discrepancy compared to the male ages (Table 10). The Shapiro-Wilk tests for normality showed significance levels of less than 0.05 for the pooled sexes, females and males, indicating a non-normal distribution for the study sample (Table 11) and that non-parametric statistical tests should be used.

Table 10: Descriptive statistics comparing true age and estimated age.

	Pooled (*n* = 103)		**Female (*n* = 49)**		**Male (*n* = 54)**	
	True age	**Estimated age**	**True age**	**Estimated age**	**True age**	**Estimated age**
Minimum	20	26.5	20	34.5	21	26.5
Maximum	98	71.5	87	70	98	71.5
Mean	55.85	59.98	53.81	58.93	57.70	60.93
Median	56	62.8	56	59.5	58	64.15
SD	20.92	8.41	20.06	7.95	21.69	8.78
CI (95%)	4.09	1.65	5.76	2.28	5.92	2.40

CI = Confidence interval
SD = Standard Deviation

Table 11: Test of normality results.

	Pooled sexes	**Female**	**Male**
n	103	49	53
Skewness	-1.168	-1.264	-1.167
Kurtosis	0.042	0.009	0.028
Statistic	0.957	0.947	0.955
df	103	49	54
Significance	0.002	0.028	0.043

df = Degree of freedom

4.1.2. Intra- and interobserver repeatability

Almost perfect intra-observer agreement was found for all four skeletal indicators (Table 12). The highest intra-observer agreement was obtained for the pooled sexes pubic symphysis (κ = 0.944), followed by the proximal femur (κ = 0.884), cranial sutures (κ = 0.827) and proximal humerus (κ = 0.804). In general, females scored higher compared to males (Table 12).

Fair to substantial interobserver agreement was found for all four skeletal indicator scores (Table 13). The cranial sutures showed the highest overall agreement for pooled sexes (κ = 0.737), females (κ = 0.688) and males (κ = 0.706). This was followed by the pubic symphysis (pooled sexes κ = 0.474), proximal humerus (pooled sexes κ = 0.320) and lastly, the proximal femur (pooled sexes κ = 0.291). The same order was observed for females and males, respectively, with females scoring higher than males for all skeletal indicators, except for the cranial sutures. Males had much lower interobserver agreement scores than females, especially for the proximal femur (females κ = 0.420 and males κ = 0.143).

Table 12: Weighted Cohen's Kappa values to assess intra-observer repeatability for each skeletal indicator.

	Pooled sexes (*n* = 20)	**Females (*n* = 10)**	**Males (*n* = 10)**
Cranial sutures			
Weighted kappa value (κ)	0.827	0.884	0.706
Standard error	0.086	0.109	0.109
95% Confidence Interval	0.659 – 0.994	0.671 – 1.097	0.492 – 0.919
Pubic Symphysis			
Weighted kappa value (κ)	0.944	1.000	0.884
Standard error	0.053	0.000	0.109
95% Confidence Interval	0.840 – 1.049	1.000 – 1.000	0.671 – 1.097
Proximal Humerus			
Weighted kappa value (κ)	0.804	0.836	0.773
Standard error	0.068	0.117	0.078
95% Confidence Interval	0.671 – 0.937	0.606 – 1.066	0.621 – 0.925
Proximal Femur			
Weighted kappa value (κ)	0.884	0.853	0.884
Standard error	0.084	0.145	0.113
95% Confidence Interval	0.720 – 1.049	0.568 – 1.138	0.663 – 1.104

Table 13: Weighted Cohen's Kappa values to assess interobserver repeatability for each skeletal indicator.

	Pooled sexes (*n* = 20)	**Females (*n* = 10)**	**Males (*n* = 10)**
Cranial sutures			
Weighted kappa value (κ)	0.737	0.688	0.706
Standard error	0.095	0.153	0.164
95% Confidence Interval	0.551 – 0.923	0.387 – 0.988	0.385 – 1.027
Pubic Symphysis			
Weighted kappa value (κ)	0.474	0.576	0.318
Standard error	0.146	0.150	0.227
95% Confidence Interval	0.187 – 0.761	0.281 – 0.871	-0.127 – 0.764
Proximal Humerus			
Weighted kappa value (κ)	0.320	0.453	0.235
Standard error	0.124	0.182	0.136
95% Confidence Interval	0.077 – 0.563	0.096 – 0.810	-0.031 – 0.502
Proximal Femur			
Weighted kappa value (κ)	0.291	0.420	0.143
Standard error	0.139	0.178	0.149
95% Confidence Interval	0.019 – 0.563	0.072 – 0.769	-0.148 – 0.434

4.1.3. Differences between female and male true ages per phase for each of the four skeletal indicators

The number of individuals scored into each phase for the four skeletal indicators is shown in Table 14. All the skeletal indicators had the highest number of individuals scored into phase IV for females (cranial sutures $n = 21$; pubic symphysis $n = 20$; proximal humerus $n = 27$; proximal femur $n = 23$) and males (cranial sutures $n = 31$; pubic symphysis $n = 26$; proximal humerus $n = 14$), except for the proximal femur, which had the highest number of males scored into phase III ($n = 21$). No females were scored into phase I for the proximal humerus and phases I and II for the proximal femur. This meant that no significant difference tests could be performed on those phases. The results showing the differences in ageing between females and males per phase associated with each skeletal indicator are provided in Table 14. No statistically significant differences ($p > 0.05$ for all skeletal indicators) between female and male true ages were obtained for any of the phases associated with the pubic symphysis, proximal humerus, or proximal femur. However, a statistically significant difference between the true ages of females and males was found for phase II of the cranial sutures, with females presenting with an older mean rank age ($Z = -2.277$; $p = 0.022$) than males (Table 14).

4.1.4. Correlation between increasing true age and phase of each of the four skeletal indicators.

Spearman's rank correlation coefficient results are provided in Table 15. Overall, all skeletal indicators showed weak to moderate correlations with age. The proximal humerus presented with the lowest correlation with true age (pooled sexes rho = 0.173), followed by the proximal femur (pooled sexes rho = 0.223). The highest correlation with true age was obtained by the pubic symphyses (pooled sexes rho = 0.457). Overall, males presented with stronger correlations for the cranial sutures and pubic symphyses compared to females (Table 15).

However, females presented with stronger correlations for the proximal humerus and proximal femur than males.

Table 14: Differences in true age per phase between females and males.

Cranial sutures						
Phase	**I**	**II**	**III**	**IV**	**V**	
***n* (F)**	8	12	6	21	1	
***n* (M)**	5	9	5	31	4	
Z statistic	0.000	-2.277	-1.278	-1.409	-1.581	
P-value*	1.000	**0.022**	0.247	0.162	0.400	
Pubic symphysis						
Phase	**I**	**II**	**III**	**IV**	**V**	
***n* (F)**	1	3	19	20	6	
***n* (M)**	1	1	14	26	12	
Z statistic	-1.000	-1.342	-0.474	-0.543	-0.094	
P-value*	1.000	0.500	0.647	0.594	0.947	
Proximal humerus						
Phase	**I**	**II**	**III**	**IV**	**V**	**VI**
***n* (F)**	0	3	8	27	9	2
***n* (M)**	4	11	13	14	11	1
Z statistic	-	-0.779	-1.705	-0.908	-0.114	-1.225
P-value*	-	0.473	0.092	0.372	0.941	0.667
Proximal femur						
Phase	**I**	**II**	**III**	**IV**	**V**	**VI**
***n* (F)**	0	0	19	23	5	2
***n* (M)**	1	5	21	20	4	3
Z statistic	-	-	-1.897	-0.085	-0.490	-1.732
P-value*	-	-	0.058	0.938	0.730	0.200

F = Female
M = Male
*Exact significance (two-tailed)
Significant values indicated in **bold**
- Test cannot be performed on empty groups

Table 15: Spearman rank correlation coefficients of increasing true age for each skeletal indicator.

	Pooled sexes	**Female**	**Male**
Cranial sutures			
n	103	49	54
Rho	0.457	0.255	0.631
p-value	<0.001	0.078	<0.001
Pubic symphysis			
n	103	49	54
Rho	0.461	0.396	0.491
p-value	<0.001	0.005	<0.001
Proximal humerus			
n	103	49	54
Rho	0.195	0.211	0.172
p-value	0.048	0.145	0.214
Proximal femur			
n	103	49	54
Rho	0.223	0.387	0.136
p-value	0.024	0.006	0.325

4.1.5. Correlation between true and estimated age using different skeletal indicator combinations.

The highest correlation for the pooled sexes was obtained when a combination of two skeletal indicators was used, specifically, the cranial sutures and pubic symphysis (rho = 0.516), followed by the cranial sutures, pubic symphysis, and proximal femur (rho = 0.505). Correlations for the female sample were stronger than those of males for eight of the skeletal indicator combinations (Table 16). In females, the highest correlation was obtained when using a combination of the pubic symphysis and proximal humerus (rho = 0.508) and in males, when using a combination of the cranial sutures and pubic symphysis (rho = 0.583). Similar to the pooled sexes, the second highest correlations for females (rho = 0.463) and males (rho = 0.534) were obtained when using a combination of the cranial sutures, pubic symphysis, and proximal femur. The combination of all four skeletal indicators was the fourth highest in the pooled sexes

(rho = 0.493), the third highest in females (rho = 0.461) and the fifth highest in males (rho = 0.506).

Table 16: Spearman's rank correlation coefficient between true and estimated age.

	Pooled sexes (*n* = 103)	**Female (*n* = 49)**	**Male (n = 54)**
CS PS PH PF			
Rho	0.493	0.461	0.506
p-value	<0.001	<0.001	<0.001
CS PS PH			
Rho	0.499	0.450	0.528
p-value	<0.001	0.001	<0.001
CS PS PF			
Rho	0.505	0.463	0.534
p-value	<0.001	<0.001	<0.001
CS PH PF			
Rho	0.314	0.340	0.284
p-value	0.001	0.017	0.037
PS PH PF			
Rho	0.449	0.460	0.428
p-value	<0.001	<0.001	0.001
CS PS			
Rho	0.516	0.416	0.583
p-value	<0.001	0.003	<0.001
CS PH			
Rho	0.368	0.321	0.409
p-value	<0.001	0.025	0.002
CS PF			
Rho	0.317	0.351	0.289
p-value	0.001	0.014	0.034
PS PH			
Rho	0.438	0.437	0.421
p-value	<0.001	0.002	0.002
PS PF			
Rho	0.477	0.508	0.448
p-value	<0.001	<0.001	<0.001
PH PF			
Rho	0.213	0.407	0.131
p-value	0.031	0.004	0.346
CS			
Rho	0.377	0.214	0.511
p-value	<0.001	0.140	<0.001
PS			
Rho	0.461	0.396	0.491
p-value	<0.001	0.005	<0.001

Table 16 (continued): Spearman's rank correlation coefficient between true and estimated age.

PH			
Rho	0.137	0.214	0.118
p-value	0.167	0.140	
PF			
Rho	0.223	0.387	0.136
p-value	0.024	0.006	0.325

CS = Cranial sutures
PS = Pubic symphysis
PH = Proximal Humerus
PF = Proximal Femur

4.1.6. Accuracy of the Complex Method

The highest proportion of individuals in the pooled sexes, female and male samples were aged more than 15 years different from their true age for all skeletal indicator combinations used in the Complex Method (Figs. 7 to 9).

Looking at the accuracy for the combination utilising all four skeletal indicators, which was reported with an error margin of ±2.5 years by Acsádi & Nemeskéri (1970), only 10.7% of the pooled sexes, 12.2% of the females and 9.3% of the males were estimated within this error margin of their true age. Individuals whose estimated age fell within ±2.6 to ±5.0 years of their true age had similar percentages of 10.7% for the pooled sexes, 12.2% for females and 9.3% for males. Overall, the combination utilising all four skeletal indicators performed the best in females (24.4% aged with ±5 years; 42.8% aged within ±10 years; 53% aged within ±15 years) compared to the pooled sexes (21.4% aged with ±5 years; 34.0% aged within ±10 years; 49.5% aged within ±15 years) and males (18.6% aged with ±5 years; 26% aged within ±10 years; 46.4% aged within ±15 years).

For the combinations utilising three skeletal indicators, the combination of the cranial sutures, pubic symphysis and proximal femur (pooled sexes: 21.4% aged with ±5 years; 36% aged within ±10 years; 50.6% aged within ±15 years; females: 26.5% aged with ±5 years; 42.8%

aged within ±10 years; 53% aged within ±15 years) as well as the combination of the cranial sutures, pubic symphysis and proximal humerus (pooled sexes: 23.3% aged with ±5 years; 33% aged within ±10 years; 48.5% aged within ±15 years; females: 28.6% aged with ±5 years; 48% aged within ±10 years; 53% aged within ±15 years) provided the best results for the pooled sexes and females. For males, the combination of the cranial sutures, proximal humerus, and proximal femur (18.6% aged with ±5 years; 35.5% aged within ±10 years; 40.9% aged within ±15 years), as well as the combination of the cranial sutures, pubic symphysis, and proximal femur (16.7% aged with ±5 years; 29.7% aged within ±10 years; 48.3% aged within ±15 years), provided the best results. The combination utilising the pubic symphysis, proximal humerus and proximal femur performed the worst in the pooled sexes (21.4% aged with ±5 years; 28.2% aged within ±10 years; 44.7% aged within ±15 years) and females (24.3% aged with ±5 years; 32.7% aged within ±10 years; 47% aged within ±15 years). The combination of the cranial sutures, proximal humerus and proximal femur performed the worst in males (18.6% aged with ±5 years; 35.3% aged within ±10 years; 40.9% aged within ±15 years).

When looking at the combination utilising two skeletal indicators, the combination of the cranial sutures and pubic symphysis performed the best for pooled sexes (27.2% aged with ±5 years; 37.9% aged within ±10 years; 54.4% aged within ±15 years), females (38% aged with ±5 years; 42.8% aged within ±10 years; 59.1% aged within ±15 years) and males (16.7% aged with ±5 years; 33.8% aged within ±10 years; 50.1% aged within ±15 years). The worst performance was by the combination of the pubic symphysis and proximal femur in the pooled sexes (21.3% aged with ±5 years; 33% aged within ±10 years; 44.7% aged within ±15 years) and by the combination of the pubic symphysis and proximal humerus in males (14.9% aged with ±5 years; 26% aged within ±10 years; 42.7% aged within ±15 years). For females, the combination of the cranial sutures and proximal humerus (18.3% aged within ±5 years; 26.5% aged within ±10 years; 46.9% aged within ±15 years) performed the worst.

For the use of a single skeletal indicator, the cranial sutures performed the best in the pooled sexes (17.5% aged with ±5 years; 42.7% aged within ±10 years; 56.3% aged within ±15 years), females (16.4% aged with ±5 years; 42.9% aged within ±10 years; 57.2% aged within ±15 years) and males (18.5% aged with ±5 years; 42.6% aged within ±10 years; 55.6% aged within ±15 years). The worst performance was by the proximal humerus in the pooled sexes (17.5% aged within ±5 years; 32.1% aged within ±10 years; 46.7% aged within ±15 years) and males (18.5% aged with ±5 years; 35.2% aged within ±10 years; 42.6% aged within ±15 years) and by the proximal femur for females (16.3% aged with ±5 years; 36.7% aged within ±10 years; 46.9% aged within ±15 years).

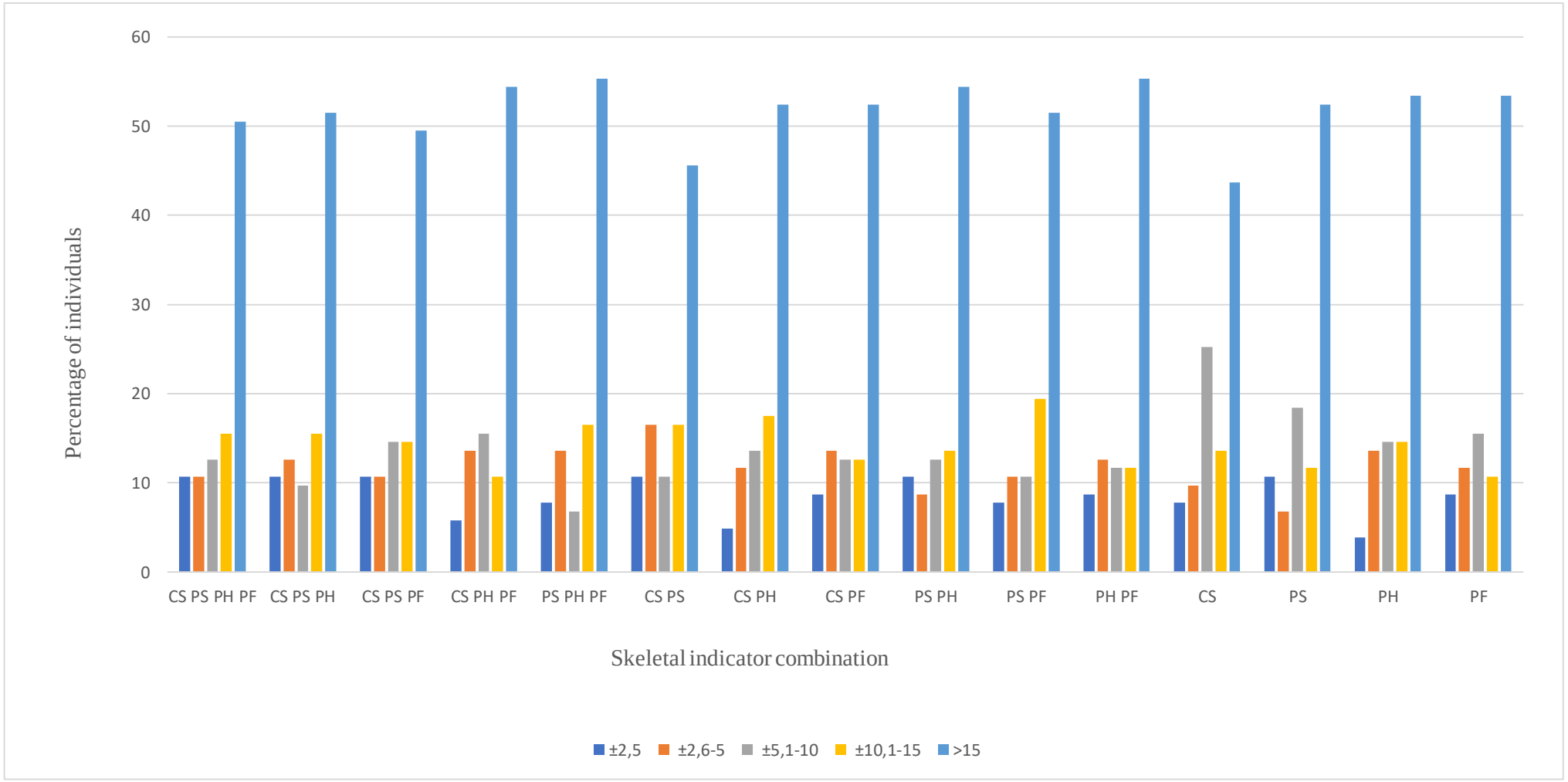


Figure 7: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for pooled sexes. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

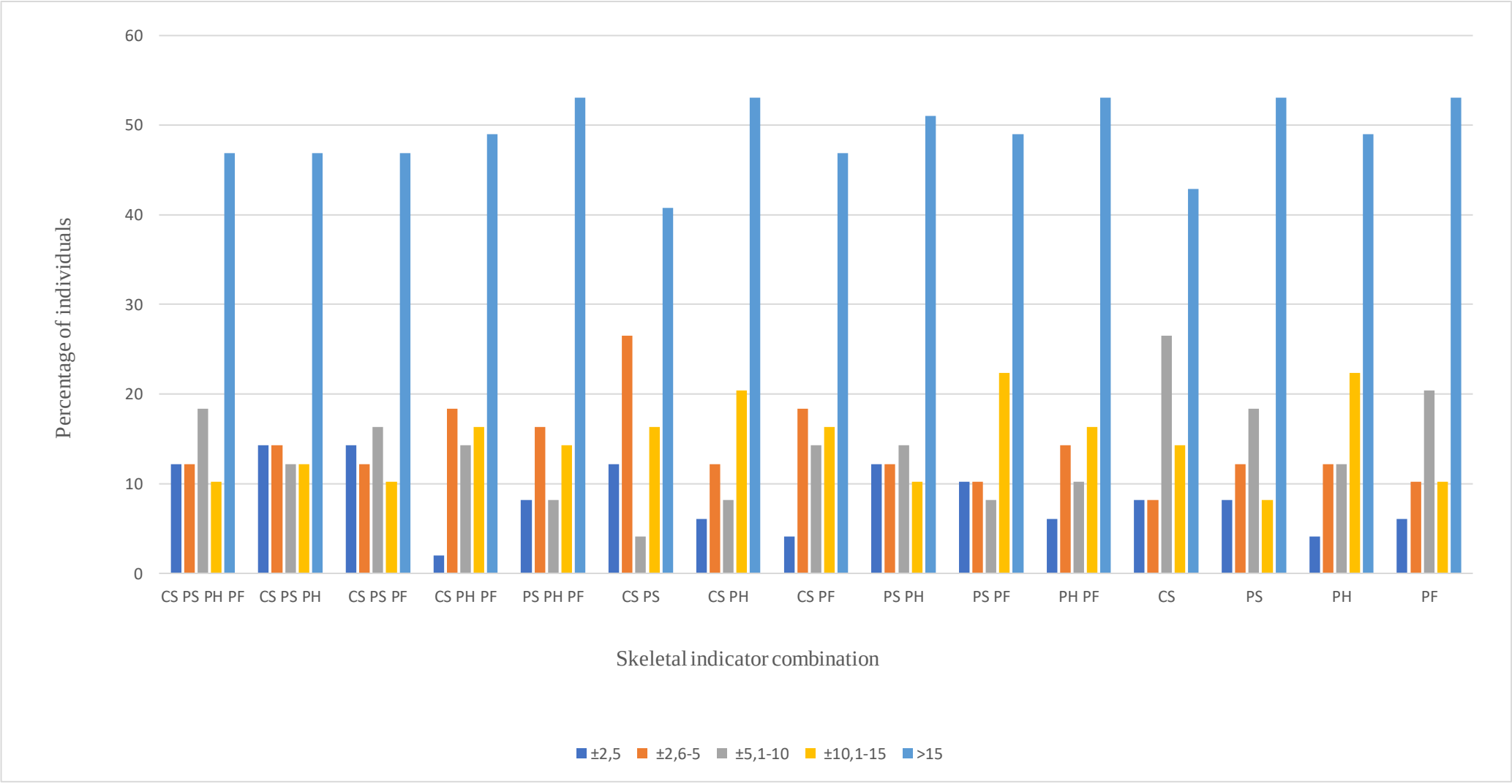


Figure 8: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

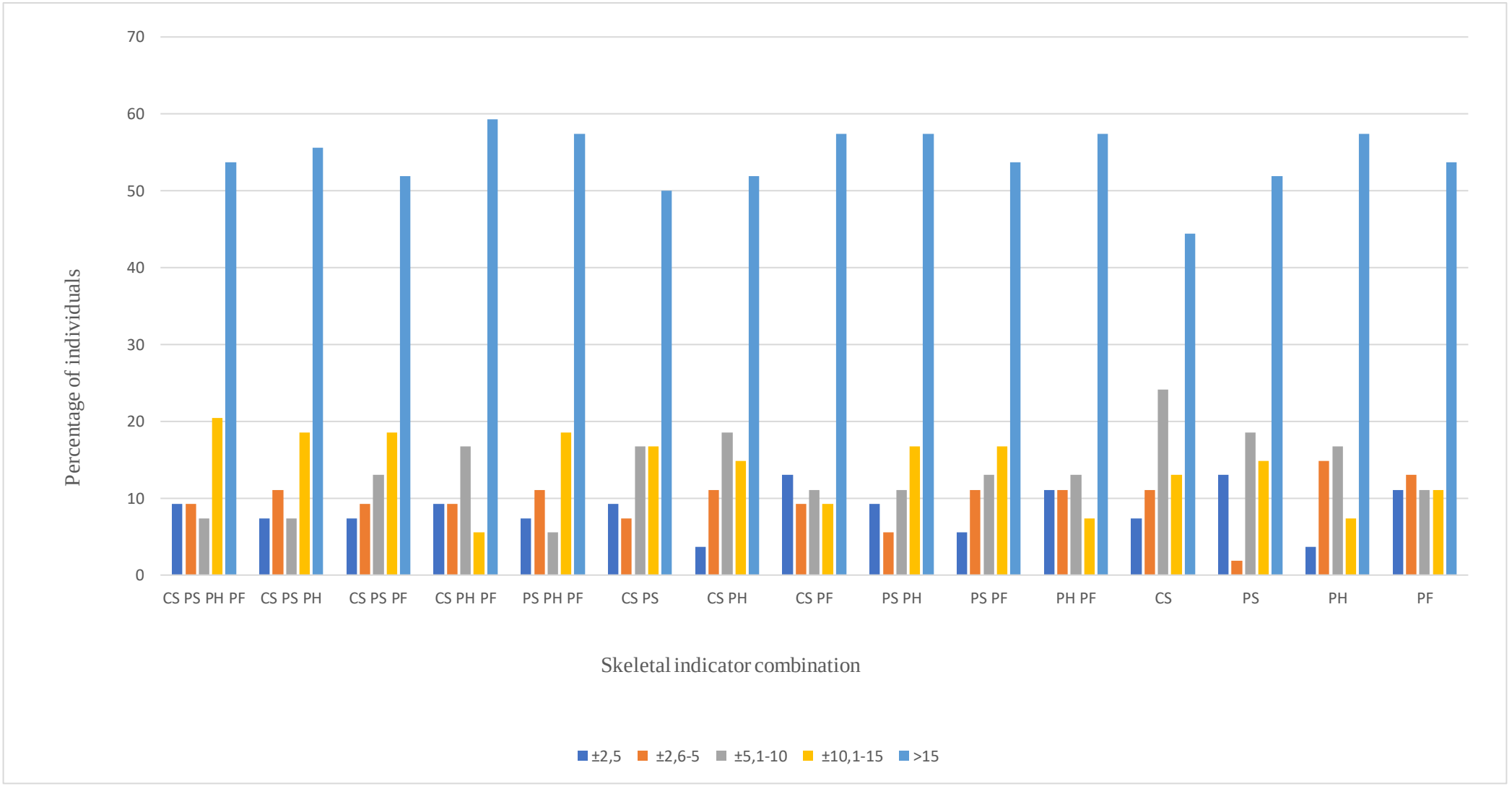


Figure 9: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

4.1.7. Inaccuracy and bias

Table 17 summarises the overall inaccuracy and bias for the black South African sample. The cranial sutures and pubic symphysis skeletal indicator combination provided the lowest overall inaccuracies for the pooled sexes (14.9 years), females (14.2 years) and males (15.6 years). The use of all four skeletal indicators provided the fourth-lowest inaccuracies for the pooled sexes (15.7 years), females (15.1 years) and males (16.2) years. The pubic symphysis as a single skeletal indicator provided the highest inaccuracies for pooled sexes (18.2 years), females (17.6 years) and males (18.7 years). For males, the same high inaccuracy (18.7 years) was obtained using a combination of the proximal humerus and proximal femur.

Using the combination of all four skeletal indicators resulted in an overestimation of age for the pooled sexes (4.1 years), females (5.1 years) and males (3.2 years). The highest bias for the pooled sexes (6.1 years) was obtained when using the combination of the pubic symphysis and proximal humerus, for females (8.7 years) when using the combination of the proximal humerus and proximal femur and for males (19.2 years) when using a combination of the cranial sutures, pubic symphysis, and proximal femur. The lowest bias for the pooled sexes (0.3 years) was obtained when using the proximal humerus as a single indicator, for females (0.7 years) when using the cranial sutures and proximal humerus and for males (0.6 years) when using the proximal humerus and proximal femur. Overall, an overestimation of age was obtained for all the skeletal indicator combinations, except for the combinations using cranial sutures, proximal humerus and proximal femur for males; cranial sutures and proximal humerus for the pooled sexes and males; cranial sutures and proximal femur for males; pubic symphysis and proximal femur for males; proximal humerus a single indicator for males; proximal femur as a single indicator for pooled sexes and males; and cranial sutures as a single indicator for pooled sexes, females and males. These combinations provided estimated ages

that underestimated the true ages. The highest underestimation of age was obtained when using the cranial sutures as a single indicator for the pooled sexes (-4.6 years), females (-4.2 years) and males (-4.9 years). The lowest underestimation of true age for pooled sexes (-0.3 years) was obtained when using the proximal femur as a single indicator and for males (-0.2 years) using a combination of the cranial sutures, proximal humerus, and proximal femur. For females, there was only one underestimation of age, as mentioned above.

Table 17: Inaccuracy and bias using the Complex Method.

	Pooled sexes (n = 103)		**Female (n = 49)**		**Male (n = 54)**	
Indicator combination	**Inaccuracy (years)**	**Bias (years)**	**Inaccuracy (years)**	**Bias (years)**	**Inaccuracy (years)**	**Bias (years)**
CS PS PH PF	15.7	4.1	15.1	5.1	16.2	3.2
CS PS PH	15.5	4.1	14.9	4.7	16.1	3.5
CS PS PF	15.2	3.4	14.5	3.9	15.9	19.2
CS PH PF	16.7	2.4	16.0	5.2	17.4	-0.2
PS PH PF	16.4	5.5	15.8	6.9	16.9	4.2
CS PS	14.9	3.0	14.2	2.8	15.6	3.3
CS PH	16.6	-1.2	16.2	0.7	16.9	-3.0
CS PF	16.5	0.7	15.5	3.2	17.4	-1.6
PS PH	16.5	6.1	16.0	7.2	17.0	5.2
PS PF	16.3	0.5	15.6	2.3	16.9	-1.0
PH PF	17.9	4.4	17.1	8.7	18.7	0.6
CS	16.3	-4.6	16.5	-4.2	16.1	-4.9
PS	15.8	1.5	15.5	1.8	16.1	1.1
PH	18.2	0.3	17.6	3.8	18.7	-2.9
PF	17.1	-0.3	16.1	2.7	18.1	-3.1

CS = Cranial sutures
PS = Pubic symphysis
PH = Proximal Humerus
PF = Proximal Femur

Table 18 shows a full breakdown of the inaccuracy and bias results per age cohort for females and males. When using a combination of all four skeletal indicators, inaccuracies ranged between 4.3 to 24.7 years for females and between 2.8 to 26.4 for males. For females, there was an overestimation of age up until the 50 – 59 age cohort; thereafter, there was an

underestimation of age for individuals aged 60 years and older. For males, there was an overestimation of age up until the 60 – 69 age cohort, thereafter, there was an underestimation of age for individuals aged 70 years and older.

When looking at females, a general overestimation of age was obtained until the 50 – 59 years age cohort, thereafter, an underestimation of age for individuals aged 60 years and older for all indicator combinations, except for the combination of cranial sutures and proximal humerus and single indicator cranial sutures. The use of these combinations resulted in an overestimation of age up until the 40 – 49 years age cohort, thereafter, an underestimation of age for individuals aged 50 years and older was obtained. When using a combination of three skeletal indicators, the lowest overestimation of age (2.8 years) was obtained when using a combination of the cranial sutures, pubic symphysis and proximal femur, and the lowest underestimation of age (-3.2 years) was obtained when using a combination of the cranial sutures, proximal humerus, and proximal femur. Utilising a combination of two skeletal indicators, the lowest overestimation of age (1.5 years) was obtained when using the cranial sutures and proximal femur, and the lowest underestimation of age (-0.4 years) was obtained when using the cranial sutures and proximal humerus. For the single skeletal indicators, the lowest overestimation of age (0.4 years) was obtained when using the pubic symphysis and the lowest underestimation of age (-4.3 years) was obtained when using the cranial sutures.

For males, when utilising a combination of three skeletal indicators, there was a general overestimation of age until the 60 – 69 years age cohort, thereafter, there was an underestimation of age for individuals aged 70 years for all combinations except the cranial sutures, proximal humerus, and proximal femur combination. This combination obtained an overestimation of age until the 50 – 59 years age cohort. The lowest overestimation of age (1.5 years) for the three skeletal indicator combinations was obtained when using a combination of the cranial sutures, pubic symphysis, and proximal humerus as well as a combination of the

cranial sutures, pubic symphysis, and proximal femur. The lowest underestimation of age (-3.8 years) was obtained using a combination of the cranial sutures, proximal humerus, and proximal femur. For the two skeletal indicator combinations, there was a general overestimation of age until the 50 – 59 years age cohort, thereafter, there was an underestimation of age for individuals aged 60 years and older, except for the combinations utilising the cranial sutures and proximal humerus as well as the pubic symphysis and proximal humerus. These combinations obtained an overestimation of age until the 60 – 69 years age cohort. The lowest overestimation of age (1.9 years) for the two indicator combinations was obtained when using the cranial sutures and pubic symphysis, and the lowest underestimation of age (-2.6 years) was obtained when using the proximal humerus and proximal femur. Single skeletal indicators provided an overestimation of age until the 50 – 59 years age cohort, followed by an underestimation of age for individuals aged 60 years and older, except when using the pubic symphysis, which obtained an overestimation of age until the 60 – 69 years age cohort, followed by an underestimation of age for individuals aged 70 years and older. The pubic symphysis provided the lowest overestimation of age (0.4 years), and the proximal femur provided the lowest underestimation of age (-6.7 years).

Table 18: Inaccuracy and bias per age cohort for females and males.

	Inaccuracy (years)		**Bias (years)**	
Age cohorts (years)	**Female (*n* = 49)**	**Male (*n* = 54)**	**Female (*n* = 49)**	**Male (*n* = 54)**
CS PS PH PF				
20-29	22.9	20.2	22.9	20.2
30-39	24.7	23.7	24.7	23.7
40-49	18.1	16.7	18.1	16.7
50-59	4.3	12.4	4.3	12.4
60-69	6.2	2.8	-4.7	1.3
70-79	6.8	10.8	-6.8	-10.8
80-89	22.6	19.8	-22.6	-19.8
90-99	-	26.4	-	-26.4
CS PS PH				
20-29	21.4	19.4	21.4	19.4

Table 18 (continued): Inaccuracy and bias per age cohort for females and males.

30-39	24.5	23.8	24.5	23.8
40-49	18.2	16.4	18.2	16.4
50-59	3.8	13.6	3.8	13.6
60-69	5.7	3.6	-4.6	1.5
70-79	7.1	10.3	-7.1	-10.3
80-89	23.5	18.2	-23.5	-18.2
90-99	-	26.4	-	-26.4
CS PS PF				
20-29	20.2	19.2	20.2	19.2
30-39	23.3	23.1	23.3	23.1
40-49	17.0	15.8	17.0	15.8
50-59	3.3	12.4	2.8	12.4
60-69	6.7	3.0	-5.1	1.5
70-79	6.9	11.3	-6.9	-11.3
80-89	24.0	19.7	-24.0	-19.7
90-99	-	25.9	-	-25.9
CS PH PF				
20-29	29.1	23.7	29.1	23.7
30-39	21.4	21.3	21.4	19.3
40-49	16.0	14.9	16.0	14.9
50-59	7.2	7.0	4.7	7.0
60-69	6.6	5.0	-3.2	-3.8
70-79	8.6	15.0	-8.6	-14.9
80-89	23.1	27.9	-23.1	-27.9
90-99	-	27.3	-	-27.3
PS PH PF				
20-29	26.2	21.7	26.2	21.7
30-39	26.4	25.3	26.4	25.3
40-49	19.9	18.4	19.9	18.4
50-59	5.9	12.7	5.9	12.7
60-69	6.5	3.1	-5.1	2.9
70-79	6.2	10.4	-5.7	-10.1
80-89	19.6	19.4	-19.6	-19.4
90-99	-	26.7	-	-26.7
CS PS				
20-29	17.9	17.6	16.8	17.6
30-39	22.4	22.9	22.4	22.9
40-49	16.7	15.0	16.7	15.0
50-59	2.6	14.1	1.8	14.1
60-69	6.2	4.4	-5.2	1.9
70-79	7.4	10.9	-7.4	-10.8
80-89	25.9	17.3	-25.9	-17.3
90-99	-	25.5	-	-25.5
CS PH				
20-29	24.1	22.2	24.1	22.2
30-39	20.7	16.6	20.7	16.6
40-49	11.8	9.9	11.8	9.9

Table 18 (continued): Inaccuracy and bias per age cohort for females and males.

50-59	5.6	5.0	-0.4	5.0
60-69	6.9	9.0	-6.9	-9.0
70-79	14.8	16.4	-14.8	-16.4
80-89	29.3	27.5	-29.3	-27.5
90-99	-	33.6	-	-33.6
CS PF				
20-29	24.2	21.1	24.2	21.1
30-39	21.2	19.4	21.2	18.6
40-49	14.3	13.4	14.0	13.4
50-59	6.8	6.6	1.5	6.6
60-69	7.1	11.4	-3.4	-9.9
70-79	10.2	16.1	-10.2	-15.9
80-89	24.9	27.6	-24.9	-27.6
90-99	-	26.0	-	-26.0
PS PH				
20-29	25.8	21.4	25.8	21.4
30-39	27.1	26.2	27.1	26.2
40-49	21.1	18.9	21.1	18.9
50-59	6.7	14.6	6.4	14.6
60-69	5.9	4.1	-5.1	4.0
70-79	6.2	9.6	-5.6	-9.0
80-89	19.4	16.9	-19.4	-16.9
90-99	-	26.8	-	-26.8
PS PF				
20-29	24.7	23.5	24.7	23.5
30-39	21.9	20.5	21.9	20.5
40-49	13.6	12.4	13.6	12.4
50-59	2.7	5.0	1.6	5.0
60-69	9.1	4.1	-9.1	-4.1
70-79	12.2	16.3	-12.2	-16.3
80-89	24.6	25.7	-24.6	-25.7
90-99	-	32.3	-	-32.3
PH PF				
20-29	34.8	26.1	34.8	26.1
30-39	27.0	23.6	27.0	21.0
40-49	18.4	17.3	18.4	17.3
50-59	9.3	6.5	7.4	6.5
60-69	6.8	5.9	-3.2	-2.6
70-79	7.4	15.2	-7.4	-14.5
80-89	16.0	29.3	-16.0	-29.3
90-99	-	28.1	-	-28.1
CS				
20-29	14.7	15.7	14.7	15.7
30-39	15.9	15.0	15.9	15.0
40-49	11.7	8.4	8.3	4.7
50-59	9.4	5.6	-4.3	5.6
60-69	9.6	12.1	-9.6	-12.1

Table 18 (continued): Inaccuracy and bias per age cohort for females and males.

70-79	16.7	17.0	-16.7	-17.0
80-89	37.7	27.4	-37.7	-27.4
90-99	-	31.6	-	-31.6
PS				
20-29	22.4	20.9	22.4	20.9
30-39	20.9	21.4	20.9	21.4
40-49	15.1	12.3	15.1	12.3
50-59	4.1	9.9	0.4	9.9
60-69	11.0	3.6	-11.0	0.4
70-79	11.0	14.4	-11.0	-14.4
80-89	24.0	19.1	-24.0	-19.1
90-99	-	31.6	-	-31.6
PH				
20-29	30.0	24.4	30.0	24.4
30-39	24.4	17.4	24.4	17.4
40-49	15.9	15.6	15.9	14.4
50-59	6.9	4.6	2.9	3.4
60-69	10.3	11.7	-10.3	-11.7
70-79	13.1	15.1	-13.1	-15.1
80-89	22.6	28.9	-22.6	-28.9
90-99	-	37.2	-	-37.2
PF				
20-29	29.3	25.1	29.3	25.1
30-39	21.1	19.3	21.1	19.3
40-49	12.3	11.4	12.3	11.4
50-59	4.4	3.4	1.9	0.6
60-69	9.0	6.7	-9.0	-6.7
70-79	13.6	18.7	-13.6	-18.7
80-89	23.0	31.3	-23.0	-31.3
90-99	-	32.8	-	-32.8

CS = Cranial sutures
PS = Pubic symphysis
PH = Proximal Humerus
PF = Proximal Femur
-No individuals sampled for this cohort

4.2. Bayesian Methods

4.2.1. Development of the Bayesian Method

4.2.1.1.Informative prior

The prior probability was calculated using the informative priors based on the documented ages of all the donated skeletal remains from the black South African individuals housed in the

Raymond A. Dart Collection of Modern Human Skeletons. Informative priors were calculated separately for females (n = 1114) and males (n = 1737), given known differences in age progression between sexes and survival curves as this relates to life expectancy. Gompertz hazard parameters calculated for each informative prior are provided in Table 19, and the Kaplan-Meier survival curves with the associated Gompertz hazard curves are provided in Appendix A (Gompertz plots 1 and 2). The R code developed by Nikita *et al.* (2018) was used to obtain the Gompertz parameters; please refer to Appendix B for an example.

Table 19: Gompertz hazard parameters for females and males.

	Female	**Male**
Shift (minimum age)	20	20
Mortality (α)	0.01958575	0.01302971
Senescence (β)	0.03450968	0.04473718

4.2.1.2. Likelihoods using transition analysis

Transition analysis was used to calculate the log-likelihood and mean age-at-transition for each phase related to the cranial sutures, pubic symphysis, proximal humerus, and proximal femur. The standard deviation for all the age-at-transition distributions was also calculated for each skeletal indicator. This was done for both females and males. The log-likelihood age-at-transition and the standard deviation make up the likelihood parameters used in the equation for the Bayesian analyses. Table 20 shows the results of the transition analyses. The transition curves for each of the skeletal indicator phases, for both females and males, are provided in Appendix A (Transition curves 1 to 8). For the R code and an example of the output for the transition analyses, please refer to Appendix B.

The mean ages-at-transition using the cranial sutures ranged from 15 to 401 years for females and 20 to 158 years for males (Table 20). For both females and males, the mean ages-at-transition between Phases IV and V were unrealistic as most individuals can reach that age,

which is likely a factor of the smaller sample size used to create transition curves. Except for the mean age-at-death transition between Phases I and II, males presented with younger mean ages (20 to 158 years) compared to females (15 to 401 years). When using the pubic symphysis, the mean ages-at-transition ranged from 11 to 123 years for females and 11 to 109 years for males. Once again, the mean ages-at-transition between Phases IV and V were somewhat unrealistic for females (123 years) and males (108 years). Males (11 to 108 years) presented with younger mean ages compared to females (11 to 123 years) for all the phases. Using the proximal humerus, mean ages-at-transition ranged from 9 to 647 years for females and 2 to 3834 years for males. However, for the proximal humerus, males transitioned from one phase to the next at older mean ages compared to females. When using the proximal femur, mean ages-at-transition for females ranged from 44 to 252 years and 1 to 3736 years for males. Again, males presented with older mean ages compared to females for all phases. None of the females in this study sample presented with Phase I for the proximal humerus and Phases I and II for the proximal femur. Therefore, the phases were collapsed into five phases for the proximal humerus and four phases for the proximal femur.

Table 20: Log transformed and raw ages-at-transition for females and males.

	Females		**Males**	
Cranial Sutures				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	2.715005	15.10469	3.01008	20.28902
II to III	3.597377	36.50238	3.453715	31.61762
III to IV	4.078305	59.04528	3.6804	39.66225
IV to V	5.993078	400.6461	5.06421	158.2554
SD	0.9785551		0.5098705	
Pubic Symphysis				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	2.414016	11.17877	2.403238	11.05893
II to III	2.953746	19.17766	2.675292	14.51659
III to IV	3.997546	54.46434	3.502094	33.18487
IV to V	4.813809	123.2	4.687145	108.5428
SD	0.6492301		0.6168121	

Proximal Humerus				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	-	-	0.8197417	2.269913
II to III	2.154305	8.621893	2.512687	12.33803
III to IV	3.151156	23.36305	3.853305	47.14865
IV to V	5.129153	168.874	5.466375	236.601
V to VI	6.472971	647.4045	8.251914	3834.961
SD	1.195112		2.147735	
Proximal Femur				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	-	-	0.3147109	1.369863
II to III	-	-	0.876704	2.402966
III to IV	3.784808	44.02723	3.971263	53.05151
IV to V	4.87887	131.4819	6.447292	630.9912
V to VI	5.530404	252.2458	8.225855	3736.314
SD	0.7094909		2.117639	

Log(AAT) = Log-likelihood ages-at-transition
ATT = Age-at-transition in years
SD = Standard deviation
- No individuals in the sample presented with the phase
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

4.2.1.3. Highest posterior density (HPD) and highest posterior density regions (HPDR) for each of the skeletal indicators

Likelihood parameters derived from the transition analyses and the prior probability were used to calculate the posterior density functions associated with each phase of the cranial sutures, pubic symphysis, proximal humerus, and proximal femur. Thereafter, the posterior density functions were utilised to calculate the highest posterior density (HPD) and highest posterior density regions (HPDR) for each phase associated with each skeletal indicator. HPD represented the mean ages per phase, and HPDR represents the respective age ranges. The HPDs and HPDRs are provided in Table 21. For the R code and an example of the generated output of the HPDs and HPDRs, please refer to Appendix B.

In general, there was an increase in HPD with an increase in phase and quite wide HPDRs for all skeletal indicators. For the cranial sutures, HPDs ranged from 20 years for both females and

males to 59 years for females and 67 years for males for Phases I to V. Females presented with slightly younger HDPs compared to males, for all phases, except for Phase I, which were the same. HPDRs were very wide, spanning more than 30 years for all phases for both females and males. Males, however, had slightly narrower HPDRs compared to females. For the pubic symphysis, HPDs and HPDRs ranged from 20 years for both females and males to 60 years for females and 61 years for males for Phases I to V. Phase I and Phase II had the same HPD (20 years) for both females and males. Overall, females had slightly narrower HPDRs compared to males; however, the difference was no more than two years. Again, the HPDRs were quite wide, spanning over 30 years for all phases. The proximal humerus provided HPDs that ranged from 20 years (females' Phase II and males' Phase I) to 57 years (females) and 56 years (males) for Phase VI. Females presented with younger HPDs compared to males across all phases, except for Phase VI. The HPDRs were wide for females and males, spanning more than 40 years for all phases. Females, however, presented narrower HPDRs for all phases, except Phases V and VI. For the proximal femur, HPDs ranged from 20 years (females' Phase III and males' Phase I) to 63 years (females) and 56 years (males). HPDs for females were younger than males for all phases, except Phases V and VI. HPDRs for females were also narrower compared to males for all phases except Phases V and VI.

Table 21: HPD and HPDR per phase calculated for females and males.

	Females		**Males**	
Cranial Sutures				
Phases	**HPD**	**HPDR (95%)**	**HPD**	**HPDR (95%)**
I	20	20 – 63.05	20	20 - 53.55
II	29.38	20 – 69.16	30.97	20 - 63.31
III	39.65	20 – 72.59	40.38	20 - 68.3
IV	48.68	20 – 76.93	53.97	25.84 - 79.66
V	58.92	28.1 – 86.57	66.92	41.88 - 88.24
Pubic Symphysis				
Phases	**HPD**	**HPDR (95%)**	**HPD**	**HPDR (95%)**
I	20	20 – 49.5	20	20 - 51.09
II	20	20 – 57.76	20	20 - 56.97

III	33.97	20 – 68.62	30.31	20 - 65.78
IV	49.97	22.47 – 78.72	50.59	22.29 - 76.52
V	59.49	30.49 – 86.78	61.47	33.51 - 85.01
Proximal Humerus				
Phases	**HPD**	**HPDR (95%)**	**HPD**	**HPDR (95%)**
I	-	-	20	20 - 70.39
II	20	20 – 63.84	41.9	20 - 72.25
III	22.25	20 – 68.61	46.05	20 - 73.52
IV	40.92	20 – 73.71	49.13	20 - 74.69
V	51.23	20.86 – 79.12	52.29	20.15 - 76.26
VI	56.86	25.42 – 84.81	55.7	23.16 - 80.22
Proximal Femur				
Phases	**HPD**	**HPDR (95%)**	**HPD**	**HPDR (95%)**
I	-	-	20	20 - 69.83
II	-	-	35.62	20 - 71.05
III	20	20 – 65.91	44.5	20 - 73.02
IV	47.52	20.33 – 76.06	50.01	20 - 75.07
V	57.57	28.32 – 85.31	53.37	21.07 - 77.61
VI	63.08	34.14 – 89.42	55.82	23.29 - 80.35

HPD = Highest posterior density (mean in years)
HPDR = Highest posterior density range (age range in years)
-No individuals in the sample presented with phase
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

4.2.2. Testing the Bayesian Method

4.2.2.1. Correlation between true and Bayesian estimated age using different skeletal indicator combinations.

The Bayesian Method did not perform well for any of the sub-sample groups. For the pooled sexes, only three of the fifteen skeletal indicator combinations showed an increase in correlation between estimated and true age compared to the results obtained using the Complex Method. When the sexes were separated, slightly better results were obtained, especially for the females, with stronger correlations between estimated and true age in seven of the fifteen skeletal indicator combinations compared to the results obtained for females when using the Complex Method alone. The Bayesian Method only provided stronger correlations in four of the skeletal indicator combinations for males than the results obtained using the Complex

Method. Please refer to Table 16 for the Complex Method and Table 22 for the Bayesian Method results.

The highest correlation for the pooled sexes (rho = 0.505) was obtained using the pubic symphysis as a single skeletal indicator, followed by the combination of the pubic symphysis and proximal femur (rho = 0.415). Correlations for the female sample were higher than those of males for eight of the fifteen skeletal indicator combinations (Table 22). For females, the highest was obtained when using a combination of the cranial sutures and pubic symphysis (rho = 0.516) and for males, when using the pubic symphysis (rho = 0.637). Like when using the Complex Method, the second highest correlation for females (rho = 0.505) was obtained using a combination of the cranial sutures, pubic symphysis, and proximal femur. For males, the combination of the cranial sutures and pubic symphysis (rho = 0.571) provided the second-highest correlation.

The lowest correlations to true age were obtained when using the cranial sutures and proximal humerus combination for the pooled sexes (rho = 0.123) and the proximal humerus as a single indicator for females (rho = 0.137) and males (rho = -0.019). The combination utilising all four skeletal indicators was the fifth highest in the pooled sexes (rho = 0.324), the fourth highest in females (rho = 0.493) and the sixth highest in males (rho = 0.458).

For females, the Bayesian Method provided stronger correlations between estimated and true age when compared to the Complex Method in these seven combinations: cranial sutures, pubic symphysis, proximal humerus, and proximal femur (rho = 0.461 to rho = 0.493); cranial sutures, pubic symphysis and proximal humerus (rho = 0.450 to rho = 0.499); cranial sutures and pubic symphysis (rho = 0.416 to rho = 0.516); cranial sutures and proximal humerus (rho = 0.321 to rho = 0.368); pubic symphysis and proximal humerus (rho = 0.437 to rho = 0.438); cranial sutures (rho = 0.214 to rho = 0.377); pubic symphysis (rho = 0.396 to rho = 0.461). For

males, the Bayesian Method provided stronger correlations between estimated and true age when compared to the Complex Method in these four combinations: cranial sutures and proximal femur (rho = 0.289 to rho = 0.385); pubic symphysis and proximal femur (rho = 0.448 to rho = 0.524); proximal humerus and proximal femur (rho = 0.131 to rho = 0.209); pubic symphysis (rho = 0.491 to rho = 0.637).

Table 22: Spearman's rank correlation coefficient between true and Bayesian estimated age.

	Pooled sexes (*n* = 23)	**Females (*n* = 9)**	**Males (*n* = 14)**
CS PS PH PF			
Rho	0.324	0.493	0.458
p-value	0.131	<0.001	0.099
CS PS PH			
Rho	0.347	0.499	0.520
p-value	0.105	<0.001	0.056
CS PS PF			
Rho	0.382	0.505	0.526
p-value	0.072	<0.001	0.053
CS PH PF			
Rho	0.157	0.314	0.194
p-value	0.474	0.001	0.507
PS PH PF			
Rho	0.295	0.449	0.301
p-value	0.171	<0.001	0.296
CS PS			
Rho	0.290	0.516	0.571
p-value	0.180	<0.001	0.033
CS PH			
Rho	0.123	0.368	0.233
p-value	0.575	<0.001	0.422
CS PF			
Rho	0.212	0.317	0.385
p-value	0.332	0.001	0.174
PS PH			
Rho	0.249	0.438	0.346
p-value	0.252	<0.001	0.225
PS PF			
Rho	0.415	0.477	0.524
p-value	0.049	<0.001	0.055
PH PF			
Rho	0.198	0.213	0.209

Table 22 (continued): Spearman's rank correlation coefficient between true and Bayesian estimated age.

p-value	0.365	0.031	0.474
CS			
Rho	0.204	0.377	0.424
p-value	0.351	<0.001	0.131
PS			
Rho	0.505	0.461	0.637
p-value	0.014	<0.001	0.014
PH			
Rho	0.307	0.137	-0.019
p-value	0.154	0.167	0.948
PF			
Rho	0.257	0.223	0.123
p-value	0.236	0.024	0.675

CS = Cranial sutures
PS = Pubic symphysis
PH = Proximal Humerus
PF = Proximal Femur

4.2.2.2. Comparison of age using the Complex Method and Bayesian Method

Age was estimated for each individual using the highest posterior densities (mean age) and highest posterior densities ratios (age ranges). The mean ages from all or a combination of the skeletal indicators (15 in total) are averaged to obtain an age estimate. The estimated ages obtained using the Acsádi & Nemeskéri Complex Method and the Bayesian Method were compared to the true ages for each individual in the sample. This was done separately for females and males. Figures 10 and 11 show the estimated ages when utilising all four skeletal indicators, using both methods plotted against the true ages of each individual.

In females, the Complex Method overestimated the age for individuals aged between 25 and 56 years, after that, it underestimated the age for older individuals. Meanwhile, the Bayesian Method overestimated age (lowest overestimation = 7.7 years) for individuals aged between 25 and 35 years and underestimated age (lowest underestimation = -2.1 years) for individuals older than 40 years. For individuals younger than 50 years, the Bayesian Method provided

estimated ages closer to the true ages as compared to the estimated ages obtained when using the Complex Method. After the age of 50 years, the Complex Method provided estimated ages that were closer to true ages.

In males, the Complex Method overestimated the age for individuals, between 28 and 62 years, after that, it underestimated the age for older individuals. In comparison, the Bayesian Method overestimated age (lowest overestimation = 0.7 years) for individuals younger than 40 years and underestimated age (lowest underestimation = -3.0 years) for individuals older than 40 years. Overall, the Bayesian Method provided estimated ages closer to the true ages for individuals younger than 60 years and the Complex Method provided estimated ages closer to the true ages for individuals older than 60 years and individuals older than 40 years. Overall, the Bayesian Method provided estimated ages closer to the true ages for individuals younger than 60 years and the Complex Method provided estimated ages closer to the true ages for individuals older than 60 years.

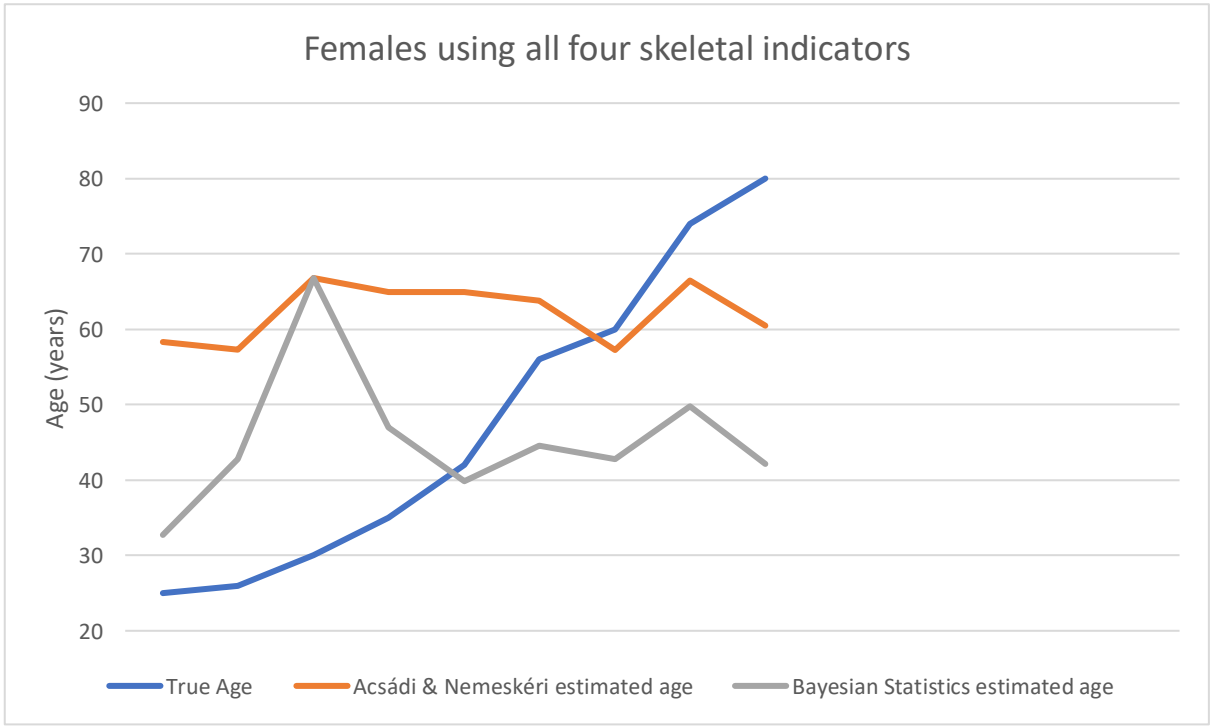


Figure 10: True versus estimated age using Acsádi & Nemeskéri Complex Method and Bayesian Method for females.

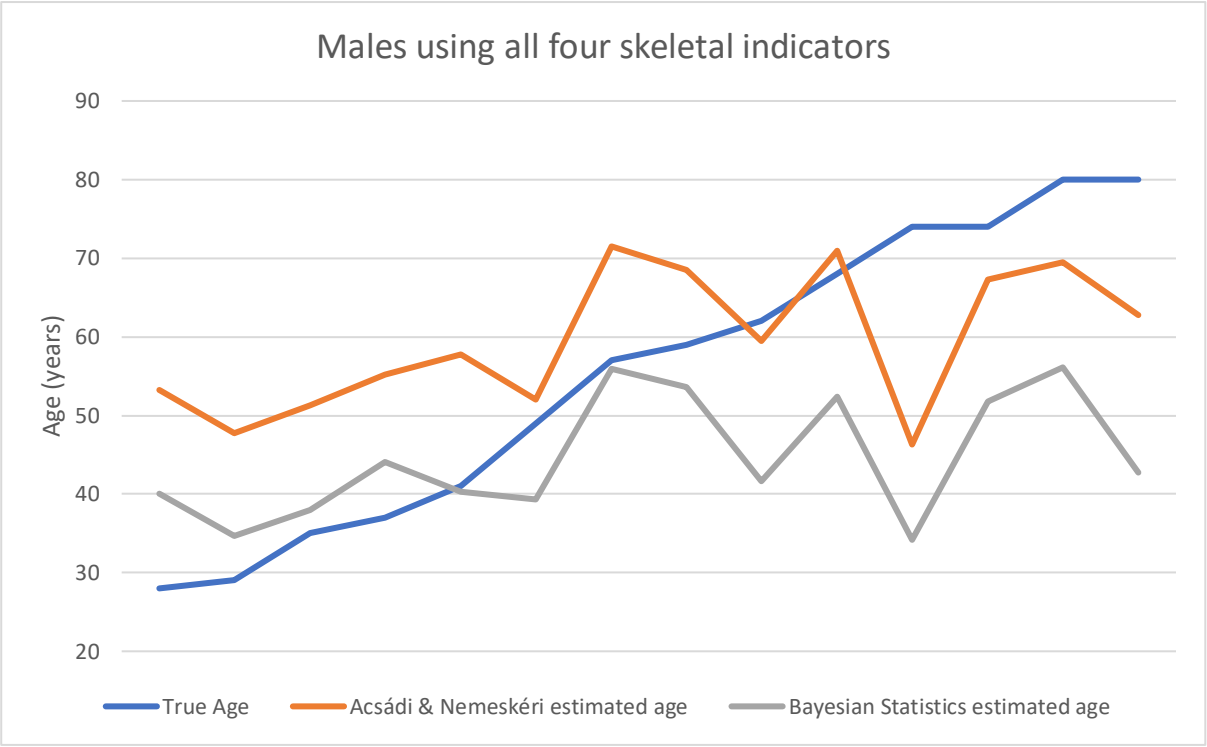


Figure 11: True versus estimated age using Acsádi & Nemeskéri Complex Method and Bayesian Method for males.

4.2.2.3. Accuracy using the Bayesian Method compared to the Complex Method.

The accuracy was obtained by calculating the percentage of individuals whose estimated age fell within ±2.5 years, ±2.6 to ±5 years, ±5.1 to ±10 years, ±10.1 to ±15 years and more than 15 years of true age for each of the 15 skeletal indicator combinations. The highest proportion of individuals in the pooled sexes (ranging from 43.7% to 54.8%), female (ranging from 44.4% to 66.7%) and male (ranging from 28.6% to 50.0%) samples were aged more than 15 years different from their true age for all skeletal indicator combinations used. The only exception was the combination of the pubic symphysis and proximal femur in males, which provided the same proportion (28.6%) of individuals whose estimated age fell within ±10.1 to ±15 years and more than 15 years of true age. Please refer to Figures 7 to 9 for the Complex Method results and Figures 12 to 14 for the Bayesian Method results.

For the combination utilising all four skeletal indicators, only 12.7% of the pooled sexes, 11.1% of the females and 14.3% of the males estimated age differed by ±2.5 years of true age. Compared to the accuracy results obtained when using the Complex Method, the Bayesian Method improved the pooled sexes (Complex Method = 10.7%) and males (Complex Method = 9.3%). However, the Bayesian Method did not improve the results for females (Complex Method = 12.2%). The four skeletal indicator combination performed the best in males (21.4% aged within ±5 years; 50% aged within ±10 years; 57.1% aged within ±15 years) compared to the pooled sexes (16.3% aged within ±5 years; 36.1% aged within ±10 years; 50.8% aged within ±15 years) and females (11.1% aged within ±5 years; 22.2% aged within ±10 years; 44.4% aged within ±15 years).

When looking at the combination utilising three skeletal indicators, the combination of the cranial sutures, pubic symphysis and proximal femur performed the best for the pooled sexes (29.0 % aged within ±5 years; 36.1% aged within ±10 years; 54.4% aged within ±15 years) and males (35.7% aged within ±5 years; 50% aged within ±10 years; 64.3% aged within ±15 years). When using the Complex Method, the combination of the cranial sutures, pubic symphysis and proximal femur performed the best for the pooled sexes (21.4% aged within ±5 years; 36% aged within ±10 years; 50.6% aged within ±15 years) and males (16.7% aged with ±5 years; 29.7% aged within ±10 years; 48.3% aged within ±15 years). When comparing the Complex Method and Bayesian Method results of this combination, cranial sutures, pubic symphysis and proximal femur, the Bayesian Method showed an improvement for the pooled sexes and males. For females, the combination of the cranial sutures, proximal humerus, and proximal femur (11.1% aged within ±5 years; 22.2% aged within ±10 years; 55.6% aged within ±15 years) performed the best. The combination of the pubic symphysis, proximal humerus and proximal femur performed the worst for the pooled sexes (12.7% aged within ±5 years; 32.5% aged within ±10 years; 50.8% aged within ±15 years) and females (11.1% aged with ±5

years; 22.2% aged within ±10 years; 44.4% aged within ±15 years). The same combination of the pubic symphysis, proximal humerus and proximal femur performed the worst in the pooled sexes and females when using the Complex Method. For females, the combination of the cranial sutures, pubic symphysis, and proximal humerus (11.1% aged within ±5 years; 22.2% aged within ±10 years; 44.4% aged within ±15 years) performed equally as poorly as the combination of the pubic symphysis, proximal humerus, and proximal femur. For males, the combination of the cranial sutures, proximal humerus, and proximal femur (14.3% aged within ±5 years; 42.9% aged within ±10 years; 57.2% aged within ±15 years) performed the worst. Again, using the Complex Method, the same combination of the cranial sutures, proximal humerus, and proximal femur performed the worst for males.

For the combinations utilising two skeletal indicators, the combination of the cranial sutures and pubic symphysis performed the best for the pooled sexes (23.4% aged within ±5 years; 41.7% aged within ±10 years; 45.2% aged within ±15 years). This combination performed the best for the pooled sexes (27.2% aged within ±5 years; 37.9% aged within ±10 years; 54.4% aged within ±15 years) when using the Complex Method as well. When comparing the Complex Method results to those obtained using the Bayesian Method, there was an improvement only in the overall proportion of individuals that were aged up to ±10 years of their true ages. The combination of the cranial sutures and proximal femur performed the best for females (11.1% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years) and the combination of the pubic symphysis and proximal femur performed the best for males (35.7% aged within ±5 years; 42.9% aged within ±10 years; 71.4% aged within ±15 years). The worst performance was by the combination of cranial sutures and proximal humerus for the pooled sexes (12.7% aged within ±5 years; 38.1% aged within ±10 years; 45.2% aged within ±15 years) and males (14.3% aged within ±5 years; 42.9% aged within ±10 years; 57.1% aged within ±15 years). For females, the combination of the pubic symphysis and

proximal humerus (11.1% aged within ±5 years; 22.2% aged within ±10 years; 33.3% aged within ±15 years) performed the worst.

Using a single skeletal indicator, the cranial sutures performed the best for the pooled sexes (23.4% aged within ±5 years; 38.1% aged within ±10 years; 50.8% aged within ±15 years). The cranial sutures also performed the best for the pooled sexes (17.5% aged within ±5 years; 42.7% aged within ±10 years; 56.3% aged within ±15 years) when using the Complex Method. There was an improvement in the overall proportion of individuals aged up to ±5 years of their true ages when using the Bayesian Method compared to the Complex Method. The proximal humerus performed the best for females (22.2% aged within ±5 years; 22.2% aged within ±10 years; 44.4% aged within ±15 years), while the pubic symphysis performed the best for males (35.7% aged within ±5 years; 50% aged within ±10 years; 57.1% aged within ±15 years). The worst performance was by the proximal humerus for the pooled sexes (18.3% aged within ±5 years; 29% aged within ±10 years; 47.2% aged within ±15 years) and males (14.3% aged within ±5 years; 35.7% aged within ±10 years; 50% aged within ±15 years). The same skeletal indicator, the proximal humerus, performed the worst for the pooled sexes and males when using the Complex Method. For females, the worst performance was by the pubic symphysis (0% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years).

4.2.3. Collapsed Bayesian Method.

During the development of the Bayesian Method, it was noted that the ages-at-transition were unrealistic as most individuals cannot reach that age. Therefore, to improve this, the collapsed Bayesian Method and pooled Bayesian Method were developed.

To improve the ages-at-transition, phases with fewer individuals were collapsed to increase the number of individuals per phase and to obtain more realistic results. Phases IV and V were

combined to form a single phase for the cranial sutures of both males and females. Phases I and II were combined, and phases IV and V were combined to form new phases for the pubic symphysis of both females and males. For the male's proximal humerus, phases I and II were combined and phases IV, V and VI were combined to form new phases. Phases IV, V, and VI of the female proximal humeri were combined into a single phase. For the male's proximal femur, phases I and II were combined and phases IV, V and VI were combined to form new phases. Phases V and VI were combined to form new phases for the female's proximal femur. The same female and male samples used to develop the Bayesian Method were used; therefore, the Gompertz hazard parameters are the same as those reported in Table 19 above.

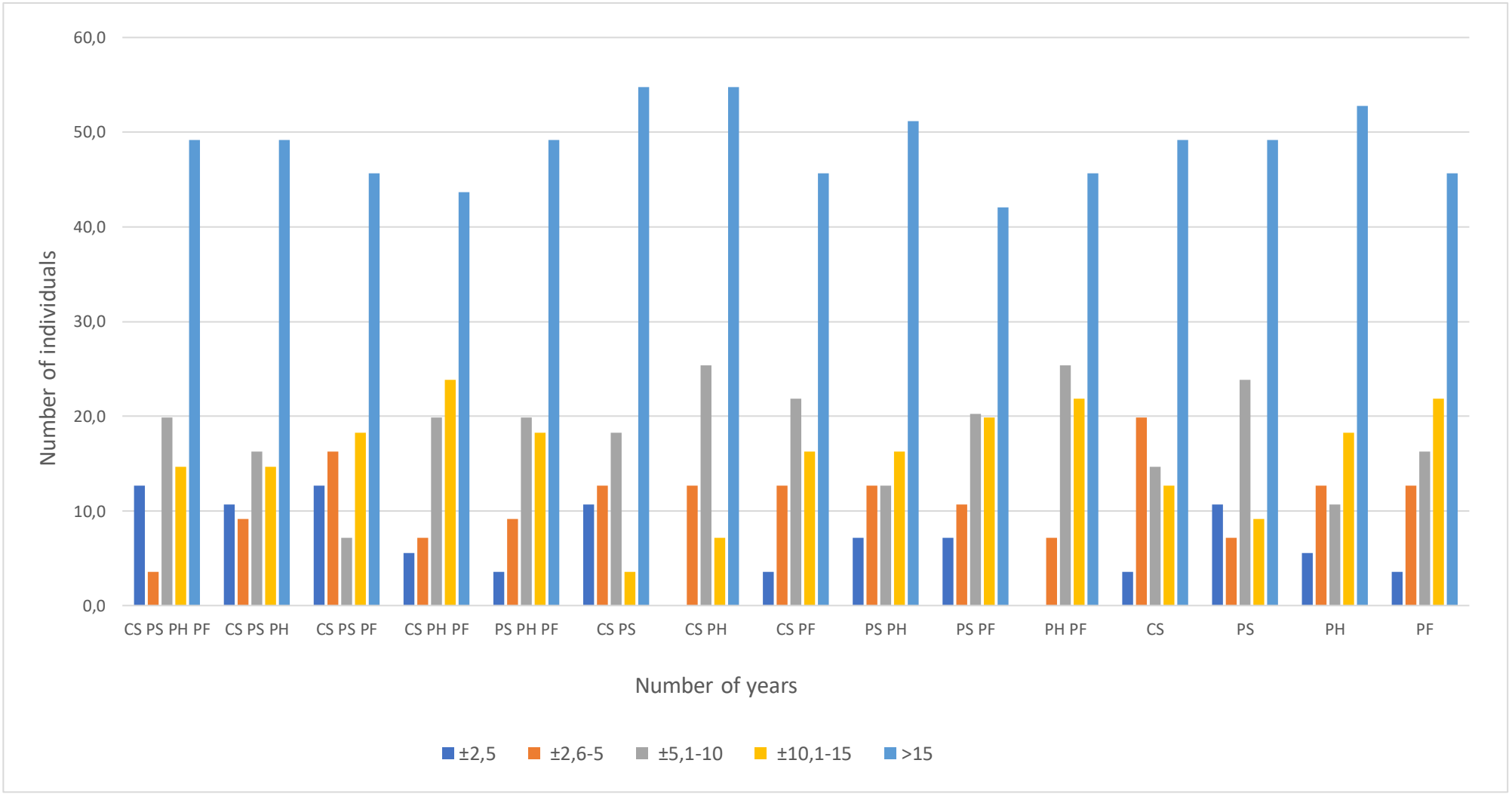


Figure 12: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for the pooled sexes using the Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

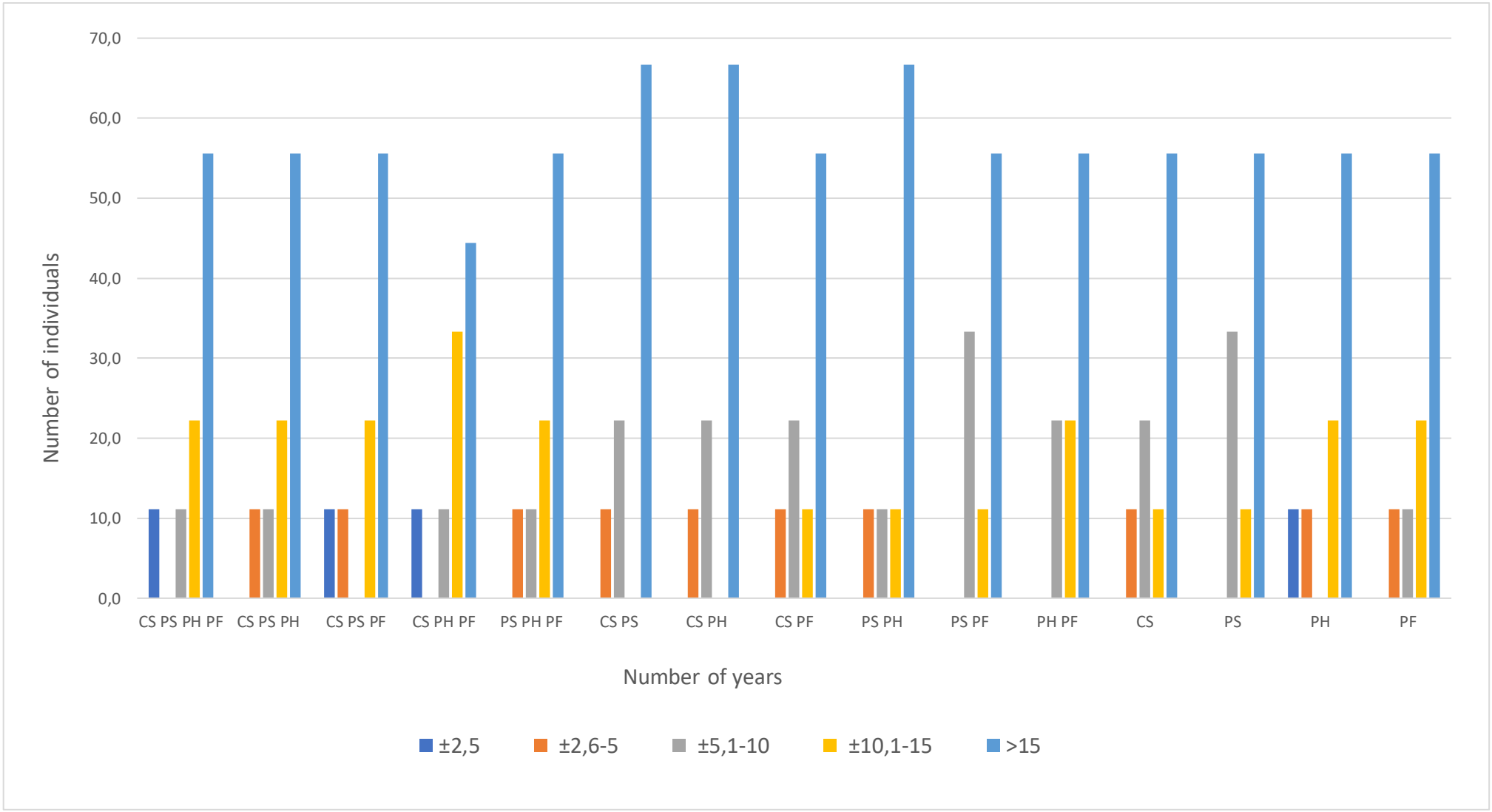


Figure 13: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females using the Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

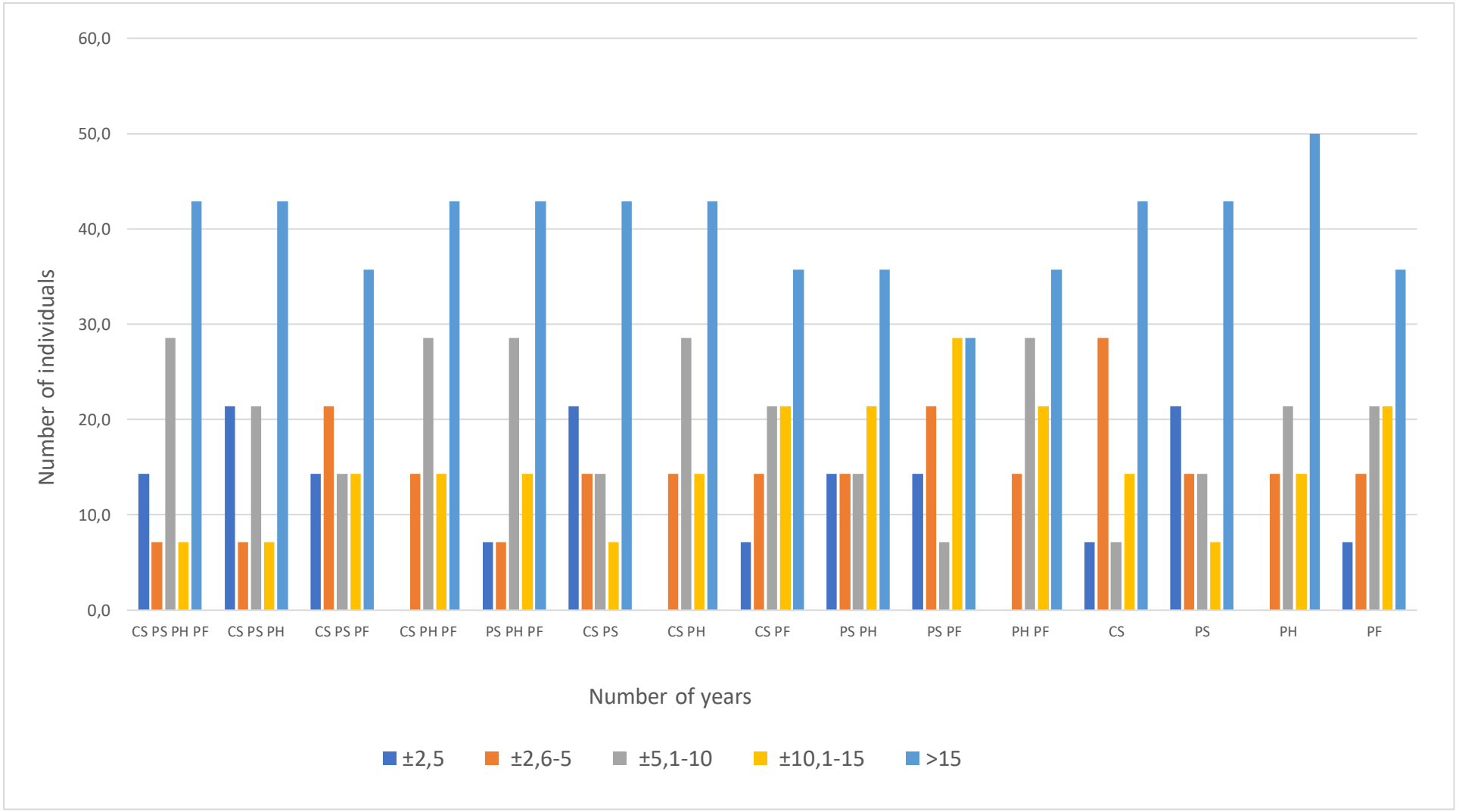


Figure 14: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males using the Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

4.2.3.1. Likelihoods using transition analysis.

Table 23 shows the results of the transition analyses. The transition curves for each of the skeletal indicator phases, for both females and males, are provided in Appendix A (Transition curves 9 to 16). Refer to Table 20 for the Bayesian Method ages-at-transition. For the R code and an example of the output for the transition analyses, refer to Appendix B.

The mean ages-at-transition using the cranial sutures ranged from 15 to 59 years for females and 19 to 39 years for males. There was an improvement from the age-at-transition obtained for the Bayesian Method for females (15 to 401 years) and males (20 to 158 years). When using the pubic symphysis, the mean ages-at-transition ranged from 19 to 55 years for females and 14 to 33 years for males. Again, there was an improvement from the results obtained using the Bayesian Method for both females (11 to 123 years) and males (11 to 109 years). Using the proximal humerus, mean ages-at-transition ranged from 1 to 9 years for females and 12 to 47 years for males. Once again, there was an improvement from the results obtained using the Bayesian Method for both females (9 to 647 years) and males (2 to 3834 years). When using the proximal femur, mean ages-at-transition for females ranged from 43 to 113 years and 3 to 48 years for males. Although the mean ages-at-transition between Phases IV to V for males was still unrealistic, there was an improvement from the results obtained using the Bayesian Method (44 to 252 years). There was also an improvement for females (1 to 3736 years).

None of the females in this study sample presented with phase I for the proximal humerus and Phases I and II for the proximal femur. Therefore, the phases with individuals could not be collapsed further, resulting in the still unrealistic Phases IV to V mean age-at-transition for females.

Table 23: Log transformed and raw ages-at-transition for Collapsed Bayesian Method females and males.

	Females		**Males**	
Cranial Sutures				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	2.697813	14.84723	2.969289	19.47807
II to III	3.593066	36.34534	3.435301	31.04074
III to IV	4.081021	59.20589	3.671523	39.31173
IV to V				
SD	0.9785551		0.5407011	
Pubic Symphysis				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	2.930164	18.7307	2.660069	14.29728
II to III	3.998718	54.52819	3.497407	33.0297
III to IV				
IV to V				
SD	0.6669191		0.6291385	
Proximal Humerus				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	-	-	2.443593	11.51433
II to III	0.1186523	1.125978	3.8501	46.99776
III to IV	2.228583	9.286699		
SD	2.790463		2.252967	
Proximal Femur				
Phases	**Log(AAT)**	**AAT**	**Log(AAT)**	**AAT**
I to II	-	-	1.173995	3.234891
II to III	-	-	3.872415	48.05829
III to IV	3.783109	43.95247		
IV to V	4.891442	133.1454*		
SD	0.7195768		1.878439	

Log(AAT) = Log-likelihood ages-at-transition
ATT = Age-at-transition in years
SD = Standard deviation
- No individuals in the sample presented with the phase
* AAT high because no individuals in phases I and II
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

4.2.3.2. Highest posterior density (HPD) and highest posterior density regions (HPDR) for each of the skeletal indicators.

The HPDs and HPDRs for the Collapsed Bayesian Method are provided in Table 24 for females and Table 25 for males. Table 21 above shows the Bayesian Method results. For the R code and an example of the generated output of the HPDs and HPDRs, please refer to Appendix B.

There was an increase in HPD and a widening of HPDRs with an increase in phase for all skeletal indicators. For the cranial sutures, HPDs ranged from 20 years for both females and males, for Phase I to 49 years (females) and 54 years (males), for Phases IV and V. Females presented with slightly younger HPDs compared to males for all phases. Males presented with slightly narrower HPDRs than females, with HPDRs spanning more than 30 years for males and more than 40 years for females throughout all phases. For the pubic symphysis, HPDs and HPDRs were similar for both females and males, with HPDs ranging from 20 years (females and males) for Phases I and II to 52 years (females) and 53 years (males) for Phase IV and V. HPDRs were very wide, spanning more than 30 years for all phases. Females, however, had slightly younger HPDRs than males, with a difference of no more than 2 years. For females, the proximal humerus provided HPDs that ranged from 20 years for Phase II to 40 years for Phases IV-VI. For males, the proximal humerus provided HPDs that ranged from 40 years for Phases I and II to 51 years for Phases IV-VI. Females presented with younger HPDs compared to males for all phases. Again, HPDRs were very wide for both females and males, spanning more than 50 years for all phases. For the proximal femur, HPDs ranged from 20 years (females' Phase III and males' Phase I and II) to 58 years (females) and 51 years (males) for Phases IV-VI. HPDs for females were younger than males for all phases, except for Phases IV – VI. HPDRs were very wide again for both females and males, spanning more than 40 years for all phases.

In general, the Collapsed Bayesian Method provided younger HPDs for all phases and all skeletal indicators when compared to the Bayesian Method. The difference was the biggest for the last phases, with more than 15 years of difference for the cranial sutures (females and males), pubic symphysis (females and males) and proximal humerus (females).

Table 24: HPD and HPDR per phase calculated for females.

Cranial Sutures		
Phases	**HPD**	**HPDR (95%)**
I	20	20 – 63.22
II	29.31	20 – 69.22
III	39.58	20 – 72.61
IV & V	49.07	20 – 77.23
Pubic Symphysis		
Phases	**HPD**	**HPDR (95%)**
I & II	20	20 – 56.59
III	33.79	20 – 68.73
IV & V	51.92	23.11 – 80.66
Proximal Humerus		
Phases	**HPD**	**HPDR (95%)**
I	-	-
II	20	20 – 69.62
III	26.92	20 – 71.42
IV - VI	39.94	20 – 74.06
Proximal Femur		
Phases	**HPD**	**HPDR (95%)**
I	-	-
II	-	-
III	20	20 – 66.01
IV	47.43	20.33 – 75.93
V & VI	58.37	28.32 – 86.03

HPD = Highest posterior density (mean in years)
HPDR = Highest posterior density range (age range in years)
-No individuals in the sample presented with phase
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

Table 25: HPD and HPDR per phase calculated for males.

Cranial Sutures		
Phases	**HPD**	**HPDR (95%)**
I	20	20 – 54.76
II	30.83	20 – 64.03
III	40.45	20 - 68.73
IV & V	54.29	25.46 – 80.02
Pubic Symphysis		
Phases	**HPD**	**HPDR (95%)**
I & II	20	20 – 55.3
III	30.38	20 – 65.99
IV & V	52.8	23.05 – 78.48
Proximal Humerus		
Phases	**HPD**	**HPDR (95%)**
I & II	39.8	20 – 71.85
III	46.11	20 – 73.55
IV - VI	50.73	20 – 75.41
Proximal Femur		
Phases	**HPD**	**HPDR (95%)**
I & II	20	20 - 69.75
III	48.86	20 – 72.79
IV - VI	51.28	20 – 75.64

HPD = Highest posterior density (mean in years)
HPDR = Highest posterior density range (age range in years)
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

4.2.3.3. Correlation between true and Collapsed Bayesian estimated age using different skeletal combinations.

The Collapsed Bayesian Method did not perform well in any of the sample groups. When compared to Spearman's rank correlation results obtained using the Complex Method, only two of the fifteen skeletal indicator combinations for the pooled sexes showed an increase in correlation between estimated and true age. For males, there was an increase in only one of the fifteen skeletal indicator combinations and none for females. For the proximal femur as a single indicator, the same correlation was observed in females. When compared to Spearman's rank correlation results obtained using the Bayesian Method, only one of the fifteen skeletal indicator combinations for both the pooled sexes and females showed an increase in correlation

between estimated and true age. For males, five of the fifteen skeletal indicator combinations showed an increase in correlation. Please refer to Table 16 for the Complex Method, Table 22 for the Bayesian Method, and Table 26 for the Collapsed Bayesian Method results.

The highest correlation for the pooled sexes (rho = 0.493) was obtained when using the pubic symphysis as a single indicator, followed by the combination of the pubic symphysis and proximal humerus (rho = 0.304). The pubic symphysis also provided the highest correlation results for the pooled sexes when using the Bayesian Method. Correlations for the male sample were higher than those of females for all the skeletal indicator combinations, except the proximal humerus. As all the individuals in the female sample were scored into phases IV to V and these phases were collapsed into one phase, the estimated age was the same. Because the estimated age for each female was constant, it was not possible to compute partial correlations and determine the correlation between the true age and the estimated age using the proximal humerus. For females, the highest correlation was obtained when using the proximal femur (rho =0.387) as a single indicator, followed by the combination of the pubic symphysis and proximal femur (0.254) as well as the combination of the pubic symphysis, proximal humerus, and proximal femur (rho = 0.254). Similar to when using the Bayesian Method, the highest correlation was obtained when using the pubic symphysis (rho = 0.628) as a single indicator, followed by the combination of the cranial sutures and pubic symphysis (rho = 0.578) for males.

The lowest correlation to true age was obtained when using the combination of the cranial sutures, proximal humerus, and proximal femur (rho = 0.069) for the pooled sexes and the combination of the cranial sutures and proximal humerus (rho = -0.225) for females. Like the Bayesian Method, the lowest correlation was obtained when using the proximal humerus (rho = 0.011) as a single indicator for males. The combination utilising all four skeletal indicators

was the ninth highest in the pooled sexes (rho = 0.230), the fifth highest in females (rho = -0.008) and the sixth highest in males (rho = 0.404).

Table 26: Spearman's rank correlation coefficient between true and Collapsed Bayesian estimated age.

	Pooled sexes (*n* = 23)	**Females (*n* = 9)**	**Males (*n* = 14)**
CS PS PH PF			
Rho	0.230	-0.008	0.404
p-value	0.291	0.983	0.152
CS PS PH			
Rho	0.236	-0.214	0.454
p-value	0.279	0.581	0.103
CS PS PF			
Rho	0.292	-0.008	0.529
p-value	0.176	0.983	0.052
CS PH PF			
Rho	0.069	-0.109	0.170
p-value	0.755	0.780	0.560
PS PH PF			
Rho	0.261	0.254	0.258
p-value	0.229	0.509	0.373
CS PS			
Rho	0.289	-0.214	0.578
p-value	0.182	0.581	0.030
CS PH			
Rho	0.073	-0.225	0.290
p-value	0.742	0.560	0.314
CS PF			
Rho	0.108	-0.109	0.244
p-value	0.625	0.780	0.400
PS PH			
Rho	0.304	-0.060	0.410
p-value	0.159	0.879	0.146
PS PF			
Rho	0.269	0.254	0.344
p-value	0.215	0.509	0.229
PH PF			
Rho	0.070	0.158	0.046
p-value	0.749	0.685	0.875
CS			
Rho	0.185	-0.187	0.371
p-value	0.398	0.630	0.192

Table 26 (continued): Spearman's rank correlation coefficient between true and Collapsed Bayesian estimated age.

PS			
Rho	0.493	0.104	0.628
p-value	0.017	0.791	0.016
PH			
Rho	0.076	-	0.011
p-value	0.729	-	0.972
PF			
Rho	0.235	0.387	0.111
p-value	0.281	0.303	0.704

CS = Cranial sutures
PS = Pubic symphysis
PH = Proximal Humerus
PF = Proximal Femur
-Partial correlations cannot be computed

4.2.3.4. Comparison of age using the Complex Method, Bayesian and Collapsed Bayesian Methods.

The estimated ages obtained using the Acsádi & Nemeskéri Complex Method and the Collapsed Bayesian Method were compared to the true ages for each individual in the sample. This was done separately for females and males. Figures 15 and 16 show the estimated ages when utilising all four skeletal indicators, using both methods plotted against the true ages of each individual.

For females, the Complex Method overestimated the age for individuals aged between 25 and 56 years, after that, it underestimated the age for older individuals. The Bayesian Method overestimated age (lowest overestimation = 7.7 years) for individuals aged between 25 and 35 years and underestimated age (lowest underestimation = -2.1 years) for individuals older than 40 years. In comparison, the Collapsed Bayesian Method overestimated age (lowest overestimation = 8.0 years) for individuals younger than 40 years and underestimated age (lowest underestimation = -1.8 years) for individuals older than 40 years. Like the Bayesian

Method, the Collapsed Bayesian Method provided estimated ages for individuals younger than 50 years that were closer to the true ages as compared to the estimated ages obtained when using the Complex Method. After the age of 50 years, the Complex Method provided estimated ages that were closer to true ages.

For males, the Complex Method overestimated the age for individuals aged between 28 and 62 years, thereafter, it underestimated the age for older individuals. In comparison, both the Bayesian Method and the Collapsed Bayesian Method overestimated age for individuals younger than 40 years and overestimated age for individuals older than 40 years. The lowest overestimation of age was 0.9 years, and the underestimation of age was -4.7 years when using the Collapsed Bayesian Method. Again, like the Bayesian Method, the Collapsed Bayesian Method provided estimated ages closer to the true ages for individuals younger than 60 years and the Complex Method provided estimated ages closer to the true ages for individuals older than 60 years.

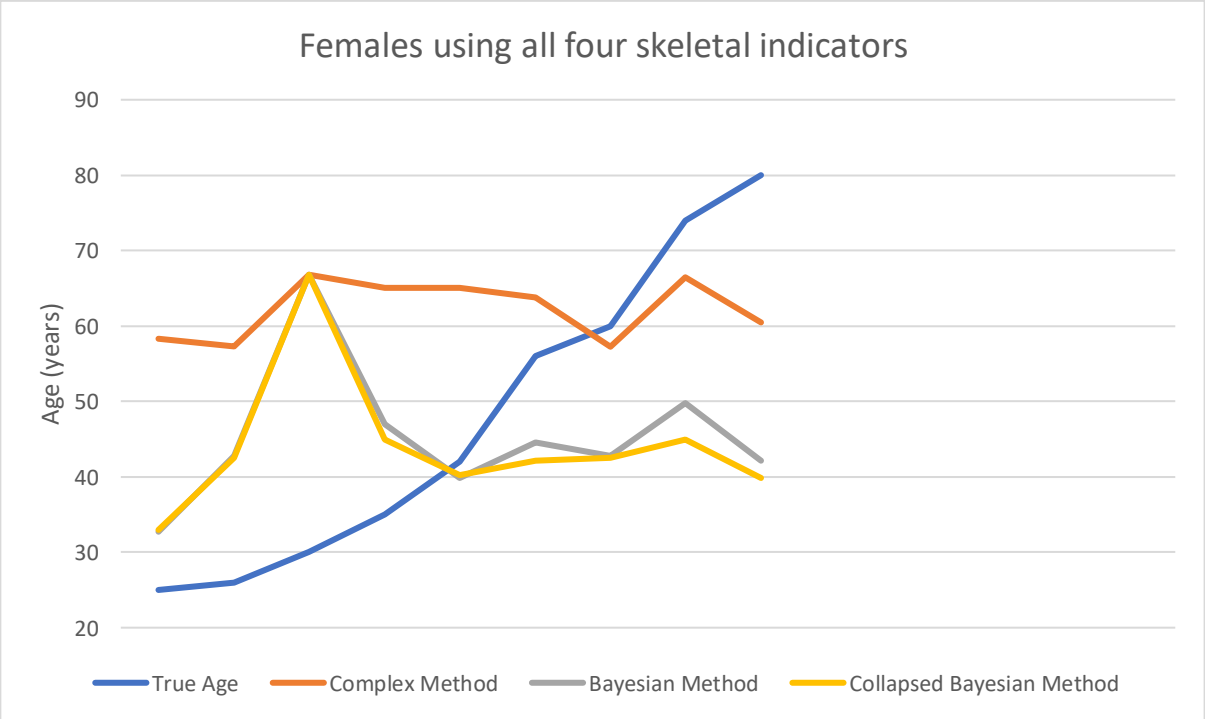


Figure 15: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Collapsed Bayesian Method for females.

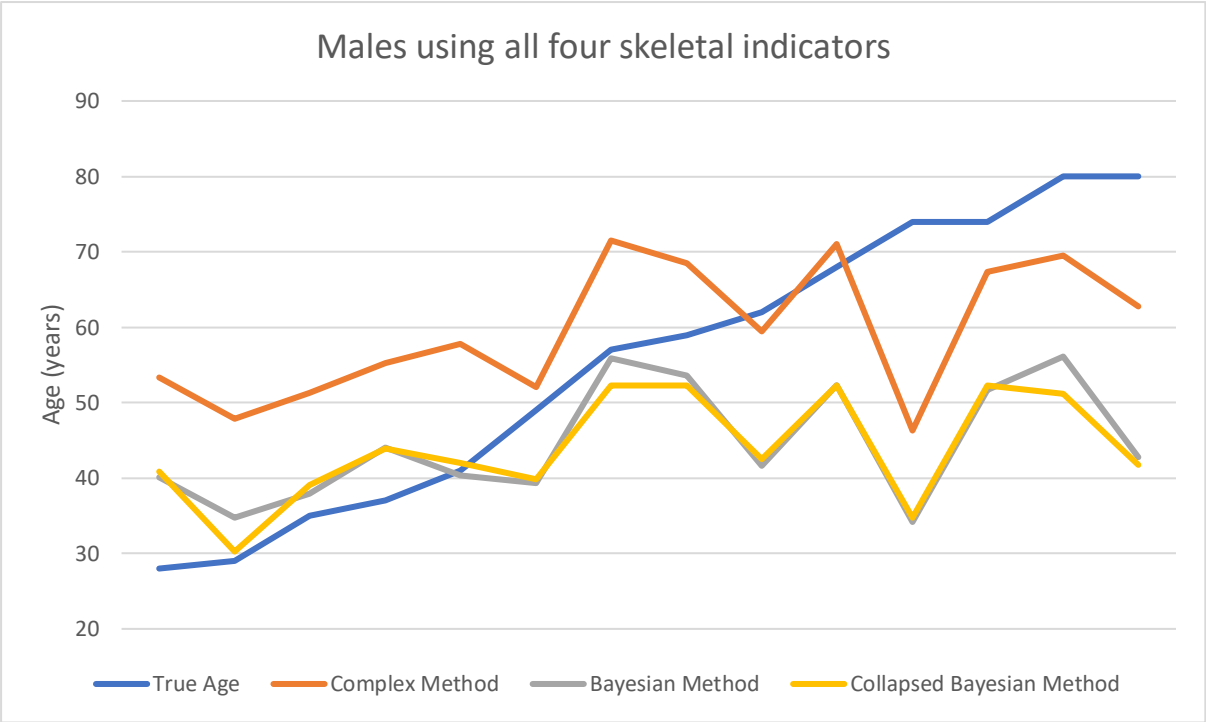


Figure 16: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Collapsed Bayesian Method for males.

4.2.3.5. Accuracy using the Collapsed Bayesian Method compared to the Bayesian and Complex Methods.

The accuracy was obtained by calculating the percentage of individuals whose estimated age fell within ±2.5 years, ±2.6 to ±5 years, ±5.1 to ±10 years, ±10.1 to ±15 years and more than 15 years of true age for each of the 15 skeletal indicator combinations. The highest proportion of individuals in the pooled sexes (ranging from 45.6% to 54.8%), female (ranging from 44.4% to 66.7%) and male (ranging from 35.7% to 50.0%) samples were aged more than 15 years different from their true age for all skeletal indicator combinations used. Please refer to Figures 7 to 9 for the Complex Method results, Figures 12 to 14 for the Bayesian Method results and Figures 17 to 19 for the Collapsed Bayesian Method results.

For the combination utilising all four skeletal indicators, only 12.7% of the pooled sexes, 11.1% of the females and 14.3% of the males estimated age differed by ±2.5 years of true age. The

same results were obtained when using the Bayesian Method. Compared to the accuracy results obtained when using the Complex Method, the Collapsed Bayesian Method improved the pooled sexes (Complex Method = 10.7%) and males (Complex Method = 9.3%). However, the Collapsed Bayesian Method did not improve the results for females (Complex Method = 12.2%). The four skeletal indicator combination performed the best in males (28.6% aged within ±5 years; 50% aged within ±10 years; 57.1% aged within ±15 years) compared to the pooled sexes (19.8% aged within ±5 years; 41.7% aged within ±10 years; 50.8% aged within ±15 years) and females (11.1% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years).

When looking at the combination utilising three skeletal indicators, the combination of the cranial sutures, pubic symphysis and proximal femur performed the best for the pooled sexes (19.8 % aged within ±5 years; 39.7% aged within ±10 years; 50.8% aged within ±15 years). The same combination of the cranial sutures, pubic symphysis and proximal femur performed the best for the pooled sexes when using the Complex Method (21.4% aged within ±5 years; 36% aged within ±10 years; 50.6% aged within ±15 years) and Bayesian Method (29.0% aged within ±5 years; 36.1% aged within ±10 years; 54.4% aged within ±15 years). When comparing the results of the three methods for this combination of the cranial sutures, pubic symphysis and proximal femur, the Collapsed Bayesian Method showed an improvement in the overall proportions of individuals that were aged up to ±10 years and ±15 years of their true ages. For females, the combination of the cranial sutures, proximal humerus, and proximal femur (11.1% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years) performed the best. This combination of the cranial sutures, proximal humerus, and proximal humerus (11.1% aged within ±5 years; 22.2% aged within ±10 years; 55.6% aged within ±15 years) performed the best when using the Bayesian Method for females. When comparing the results obtained when using the Collapsed Bayesian Method and Bayesian Method for this

combination of the cranial sutures, proximal humerus, and proximal femur, the Collapsed Bayesian only showed an improvement in the overall proportions of individuals that were aged up to ±10 years, while the proportions of the individuals that were aged ±5 years remained unchanged. The same combination of the cranial sutures, proximal humerus and proximal femur performed the worst for the pooled sexes (16.3% aged within ±5 years; 38.1% aged within ±10 years; 47.2% aged within ±15 years) and males (21.4% aged with ±5 years; 42.9% aged within ±10 years; 50.0% aged within ±15 years). When using the Complex Method and Bayesian Method, the same combination of the pubic symphysis, proximal humerus, and proximal femur performed the worst in males. For males, the combination of the cranial sutures, pubic symphysis, and proximal humerus (35.7% aged within ±5 years; 50.0% aged within ±10 years; 57.1% aged within ±15 years) performed the best. This combination of the cranial sutures, pubic symphysis, and proximal humerus (11.1% aged within ±5 years; 22.2% aged within ±10 years; 44.4% aged within ±15 years) performed the worst for females. The combination of pubic symphysis, proximal humerus, and proximal femur (11.1% aged within ±5 years; 22.2% aged within ±10 years; 44.4% aged within ±15 years) performed equally bad for females.

For the combinations utilising two skeletal indicators, the combination of the pubic symphysis and proximal femur performed the best for the pooled sexes (14.3% aged within ±5 years; 39.71% aged within ±10 years; 54.4% aged within ±15 years) and males (28.6% aged within ±5 years; 57.1% aged within ±10 years; 64.3% aged within ±15 years). This combination performed the best in males (35.7% aged within ±5 years; 42.9% aged within ±10 years; 71.4% aged within ±15 years) when using the Bayesian Method as well. There was no improvement when comparing the Bayesian Method results to those obtained when using the Collapsed Bayesian Method. When using the Complex Method, the same combination of the pubic symphysis and proximal femur performed the worst for the pooled sexes. For females, the

combination of the cranial sutures and proximal femur (11.1% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years) performed the best. This combination also performed the best for females (11.1% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years) when using the Bayesian Method. The same results were obtained when using either the Bayesian Method or the Collapsed Bayesian Method. The combination of the cranial sutures and proximal humerus (23.8% aged within ±5 years; 41.7% aged within ±10 years; 45.2% aged within ±15 years) performed the worst for the pooled sexes. The worst performance was by the combination of cranial sutures and pubic symphysis in females (0.0% aged within ±5 years; 22.2% aged within ±10 years; 33.3% aged within ±15 years) and males (28.6% aged within ±5 years; 42.9% aged within ±10 years; 50.0% aged within ±15 years).

For the use of a single skeletal indicator, the proximal humerus performed the best for the pooled sexes (14.7% aged within ±5 years; 34.5% aged within ±10 years; 52.8% aged within ±15 years) and females (22.2% aged within ±5 years; 33.3% aged within ±10 years; 55.6% aged within ±15 years). This skeletal indicator also performed the best in females (22.2% aged within ±5 years; 22.2% aged within ±10 years; 44.4% aged within ±15 years) when using the Bayesian Method. When comparing the proximal humerus results obtained when using the two methods, the Collapsed Bayesian Method showed an improvement in the overall proportions of individuals aged up to ±10 years and ±15 years of their true ages. The same skeletal indicator, the proximal humerus, performed the worst (7.1% aged within ±5 years; 35.7% aged within ±10 years; 50% aged within ±15 years) in males. The proximal humerus also performed the worst for the pooled sexes and males when using the Complex Method and Bayesian Method. For males, the pubic symphysis (28.6% aged within ±5 years; 50% aged within ±10 years; 57.1% aged within ±15 years) performed the best. The pubic symphysis also performed the best in males (35.7% aged within ±5 years; 50% aged within ±10 years; 57.1% aged within

±15 years) when using the Bayesian Method. There was no improvement in results when using the Collapsed Bayesian Method. The worst performance was by the proximal femur for the pooled sexes (9.1% aged within ±5 years; 29% aged within ±10 years; 45.2% aged within ±15 years) and females (11.1% aged within ±5 years; 22.2% aged within ±10 years; 33.3% aged within ±15 years). The same skeletal indicator, the proximal femur, performed the worst in females when using the Complex Method.

4.2.4. Pooled Bayesian Method

As mentioned above, the ages-at-transition were found to be very unrealistic during the development of the Bayesian Method because most individuals cannot reach that age. To improve the ages-at-transition, the females and males were combined to increase the number of individuals per phase and to obtain more realistic results.

4.2.4.1. Informative prior

The prior probability was calculated using the informative prior based on the documented ages of all the donated skeletal remains from the black South African individuals (n = 2 851) housed in the Raymond A. Dart Collection of Modern Human Skeletons. Gompertz hazard parameters calculated for the informative prior are provided in Table 27. The Kaplan-Meier survival curve with the associated Gompertz hazard curve is provided in Appendix A (Gompertz plots 3). The R code developed by Nikita *et al*. (2018) was used to obtain the Gompertz parameters; please refer to Appendix B for an example.

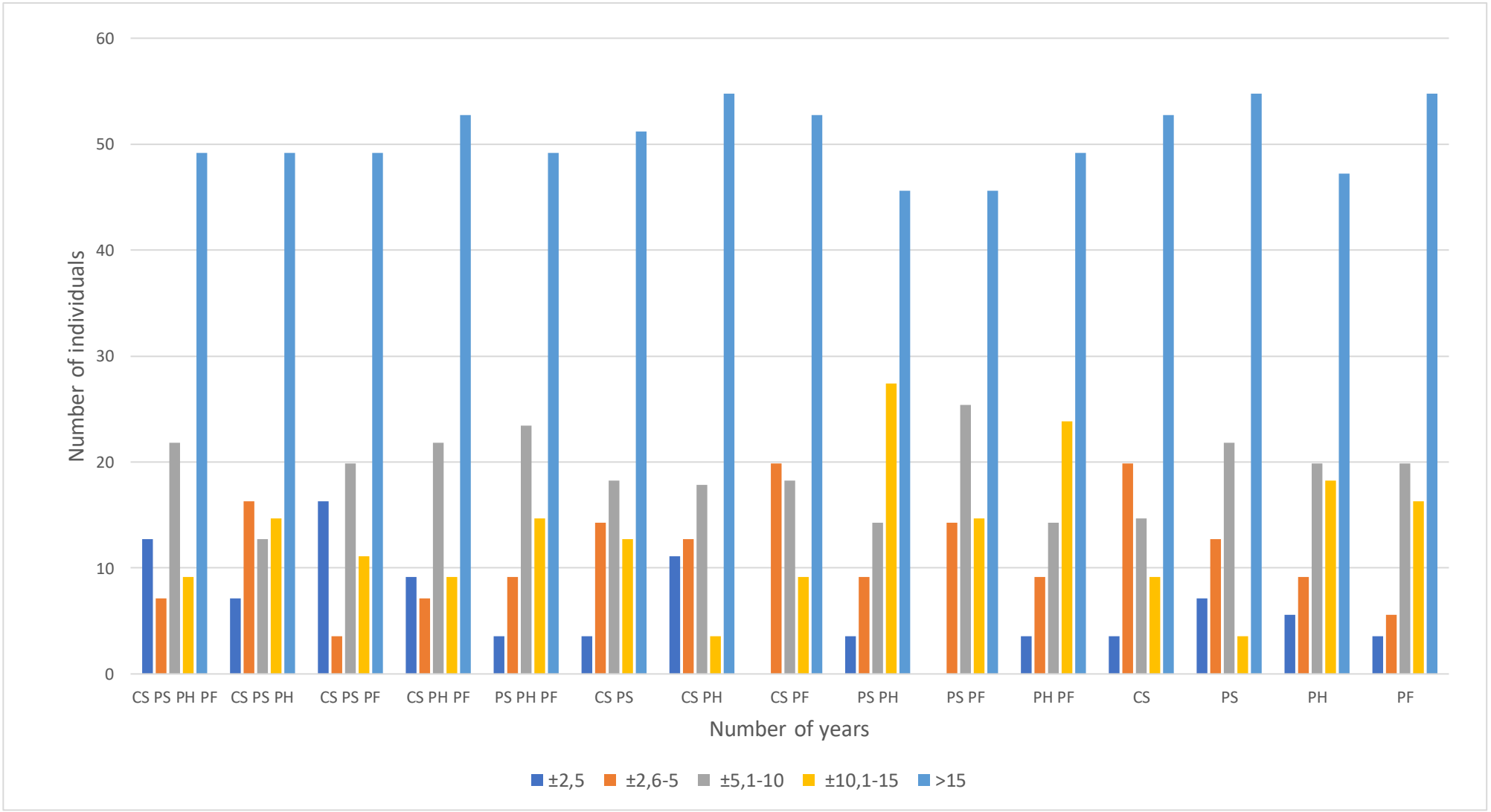


Figure 17: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for the pooled sexes using the Collapsed Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

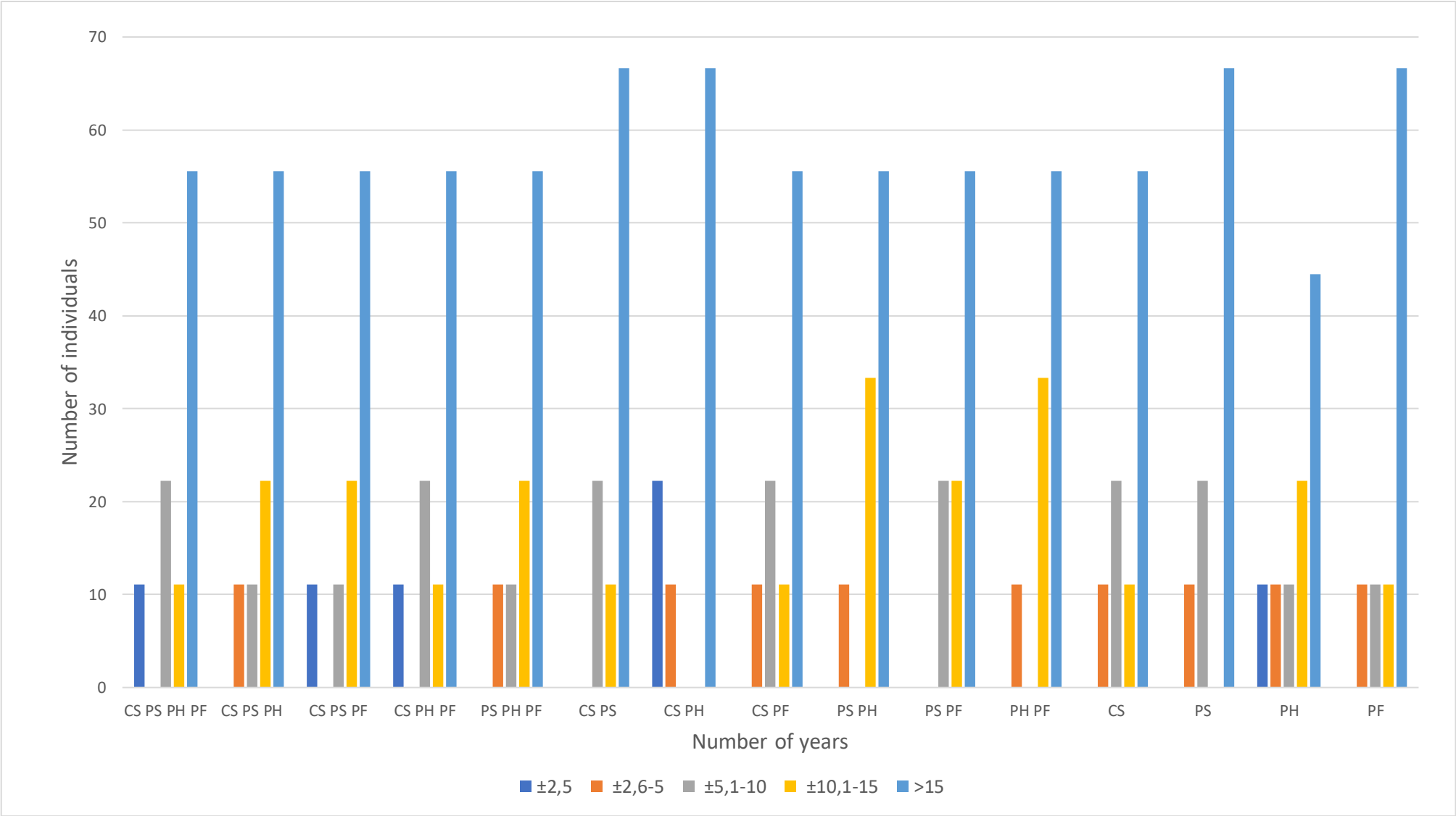


Figure 18: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females using the Collapsed Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

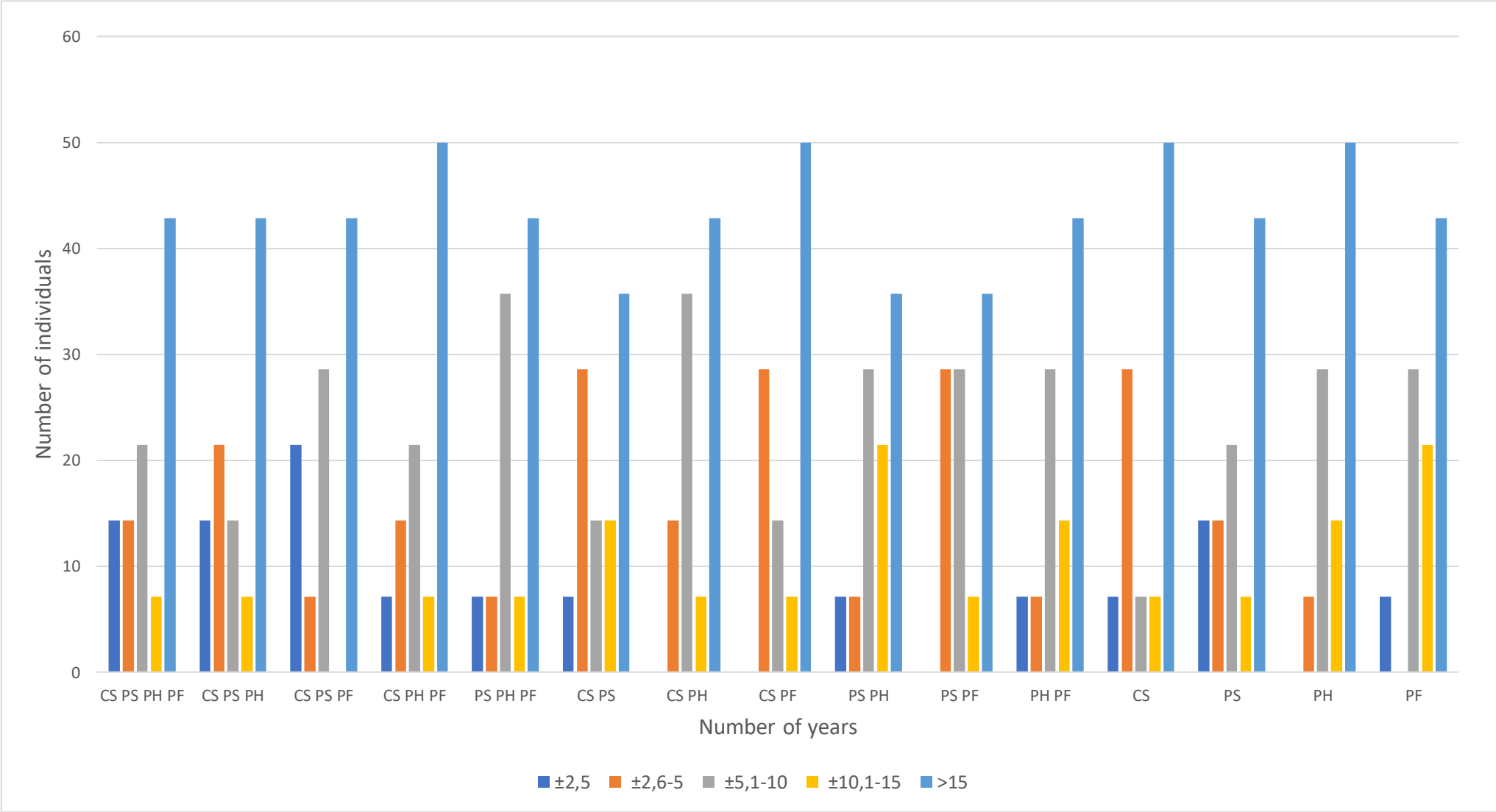


Figure 19: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males using the Collapsed Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

Table 27: Gompertz hazard parameters for the pooled sexes.

Shift (minimum age)	20
Mortality (α)	0.01443632
Senescence (β)	0.04217679

The Kaplan-Meier survival curve, which represents the pooled sexes target sample, is plotted against the Gompertz hazard curve, which represents the informative prior, to determine whether the informative prior is the best fit for the target sample. However, as the target sample and informative prior were taken from the same collection and had similar age-at-death distributions, the survivorship rate plot was not created. Therefore, the informative prior could be used for the Bayesian analyses of the pooled sexes samples.

4.2.4.2. Likelihood using transition analysis.

Table 28 shows the results of the transition analyses. The transition curves for each of the skeletal indicator phases are provided in Appendix A (Transition curves 17 to 20). For the R code and an example of the output for the transition analyses, please refer to Appendix B.

The mean ages-at-transition using the cranial sutures ranged from 18 to 225 years. For the pooled sexes, the mean ages-at-transition between Phases IV to V were still unrealistic, as most individuals cannot reach that age. When using the pubic symphysis, the mean ages-at-transition ranged from 11 to 116 years. Again, the mean ages-at-transition between Phases IV to V were still unrealistic. Using the proximal humerus, mean ages-at-transition ranged from 2 to 1845 years. The mean ages-at-transition between Phases IV to V and Phases V to VI were still unrealistic. When using the proximal femur, mean ages-at-transition ranged from 3 to 518 years. Once again, the mean ages-at-transition between Phases IV to V and Phases V to VI were still unrealistic.

Table 28: Log transformed and raw ages-at-transition for pooled sexes.

Cranial Sutures		
Phases	**Log(AAT)**	**AAT**
I to II	2.898268	18.14269
II to III	3.518172	33.72274
III to IV	3.836072	46.34308
IV to V	5.41659	225.1101
SD	0.7103692	
Pubic Symphysis		
Phases	**Log(AAT)**	**AAT**
I to II	2.402421	11.04989
II to III	2.832983	16.99609
III to IV	3.759358	42.92087
IV to V	4.754884	116.1502
SD	0.6442706	
Proximal Humerus		
Phases	**Log(AAT)**	**AAT**
I to II	0.7750751	2.170755
II to III	2.190054	8.935697
III to IV	3.371184	29.11296
IV to V	5.404084	222.3124
V to VI	7.520432	1845.364
SD	1.756237	
Proximal Femur		
Phases	**Log(AAT)**	**AAT**
I to II	1.261616	3.531124
II to III	1.8562	6.399376
III to IV	3.813177	45.2941
IV to V	5.295953	199.5277
V to VI	6.250211	518.1224
SD	1.091835	

Log(AAT) = Log-likelihood ages-at-transition
ATT = Age-at-transition in years
SD = Standard deviation
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

4.2.4.3. Highest posterior density (HPD) and highest posterior density regions (HPDR) for each of the skeletal indicators

The HPDs and HPDRs for the Pooled Bayesian Method are provided in Table 29. For the R code and an example of the generated output of the HPDs and HPDRs, please refer to Appendix B.

For all skeletal indicators, there was a general increase in HPD with an increase in phase and quite wide HPDRs. For the cranial sutures, HPDs ranged from 20 years for Phase I to 64 years for phase V. The HPDRs were very wide, spanning more than 40 years for all phases. For the pubic symphysis, HPDs ranged from 20 years for Phase I to 61 years for Phase V. Phase I and Phase II had the same HPD of 20 years. Again, the HPDRs were quite wide, spanning more than 30 years for all phases. The proximal humerus provided HPDs that ranged from 20 years for Phase I to 56 years for Phase VI. Once again, the HPDRs were very wide spanning more than 45 years for all phases. For the proximal femur, HPDs ranged from 20 years for Phase I to 60 for Phase VI. The HPDRs were very wide, spanning more than 40 years for all phases.

4.2.4.4. Correlation between true and Pooled Bayesian estimated age using different skeletal combinations.

The Pooled Bayesian Method did not perform well in any of the sample groups (Table 30). When compared to Spearman's rank correlation results obtained using the Complex Method, only one of the fifteen skeletal indicator combinations for both the pooled sexes and males showed an increase in correlation between estimated and true age. For females, there was an increase in only two of the fifteen skeletal indicator combinations.

Table 29: HPD and HPDR per phase calculated for pooled sexes.

Cranial Sutures		
Phases	**HPD**	**HPDR (95%)**
I	20	20 – 60.59
II	33.42	20 – 67.88
III	43.11	20 – 71.61
IV	53.15	23.51 – 79.33
V	63.52	35.43 – 87.17
Pubic Symphysis		
Phases	**HPD**	**HPDR (95%)**
I	20	20 – 51.89
II	20	20 – 59.02
III	35.01	20 – 68.14
IV	51.42	23.06 – 77.92
V	60.91	32.46 – 85.37
Proximal Humerus		
Phases	**HPD**	**HPDR (95%)**
I	20	20 – 68.06
II	30.81	20 – 70.6
III	41.05	20 – 72.41
IV	47.19	20 – 74.71
V	52.16	20.86 – 77.1
VI	56.1	25.42 – 81.44
Proximal Femur		
Phases	**HPD**	**HPDR (95%)**
I	20	20 – 60.69
II	20	20 – 64.39
III	36.37	20 – 70.8
IV	50.09	20 – 75.52
V	56.25	25.07 – 81.85
VI	60.01	29.25 – 84.96

HPD = Highest posterior density (mean in years)
HPDR = Highest posterior density range (age range in years)
Phase scoring based on the Acsádi & Nemeskéri (1970) criteria

When compared to Spearman's rank correlation results obtained using the Bayesian Method, none of the fifteen skeletal indicator combinations for the pooled sexes showed an increase in correlation between estimated and true age. For females, three of the fifteen skeletal indicator combinations showed an increase in correlation and only one for males. For the pubic symphysis as a single indicator, the same correlation was observed in males. Please refer to

Table 16 for the Complex Method, Table 22 for the Bayesian Method, and Table 26 for the Collapsed Bayesian Method results.

The highest correlation for the pooled sexes (rho = 0.467) was obtained when using the pubic symphysis as a single indicator, followed by the combination of the pubic symphysis and proximal femur (rho = 0.239). The pubic symphysis also provided the highest correlation results for the pooled sexes using the Bayesian Method. Correlations for the male sample were higher than those of females for eleven of the fifteen skeletal indicator combinations. The highest correlation was obtained for females when using the proximal humerus (rho =0.634) as a single indicator, followed by the proximal femur (0.422) as a single indicator. Similar to when using the Bayesian Method, the highest correlation was obtained when using the pubic symphysis (rho = 0.637) as a single indicator, followed by the combination of the cranial sutures and pubic symphysis (rho = 0.518) for males.

The lowest correlation to true age was obtained when using the combination of the cranial sutures and pubic symphysis (rho = -0.214) for females. Like the Bayesian Method, the lowest correlation was obtained when using the combination of the cranial sutures and proximal humerus (rho = -0.023) for the pooled sexes and proximal humerus (rho = -0.066) as a single indicator for males. The combination utilising all four skeletal indicators was the eighth highest in the pooled sexes (rho = 0.145), females (rho = 0.075) and males (rho = 0.227).

Table 30: Spearman's rank correlation coefficient between true and Pooled Bayesian estimated age.

	Pooled sexes (*n* = 23)	**Females (*n* = 9)**	**Males (*n* = 14)**
CS PS PH PF			
Rho	0.145	0.075	0.227
p-value	0.509	0.847	0.435
CS PS PH			
Rho	0.167	0.059	0.341

Table 30 (continued): Spearman's rank correlation coefficient between true and Pooled Bayesian estimated age.

p-value	0.447	0.881	0.233
CS PS PF			
Rho	0.219	0.025	0.434
p-value	0.316	0.949	0.121
CS PH PF			
Rho	0.035	-0.092	0.049
p-value	0.873	0.814	0.869
PS PH PF			
Rho	0.147	0.227	0.157
p-value	0.502	0.557	0.591
CS PS			
Rho	0.238	-0.214	0.518
p-value	0.275	0.581	0.058
CS PH			
Rho	-0.023	-0.102	0.008
p-value	0.918	0.794	0.979
CS PF			
Rho	0.031	-0.092	0.116
p-value	0.889	0.814	0.692
PS PH			
Rho	0.174	0.248	0.352
p-value	0.428	0.520	0.217
PS PF			
Rho	0.239	0.186	0.335
p-value	0.272	0.632	0.241
PH PF			
Rho	0.035	0.287	0.000
p-value	0.875	0.454	1.000
CS			
Rho	0.062	-0.187	0.267
p-value	0.779	0.630	0.357
PS			
Rho	0.467	0.104	0.637 * same as BM
p-value	0.025	0.791	0.014
PH			
Rho	0.008	0.634	-0.066
p-value	0.971	0.067	0.822
PF			
Rho	0.127	0.422	0.019
p-value	0.565	0.258	0.948

CS = Cranial sutures
PS = Pubic symphysis
PH = Proximal Humerus
PF = Proximal Femur

4.2.4.5. Comparison of age using the Complex Method, Bayesian, and Pooled Bayesian Methods.

The Acsádi & Nemeskéri Complex Method and the Pooled Bayesian Method estimated ages were compared to the true ages for each individual in the sample. This was done separately for females and males. Figures 20 and 21 show the estimated ages when utilising all four skeletal indicators, using both methods plotted against the true ages of each individual.

In females, the Complex Method overestimated age for individuals aged between 25 and 56 years and underestimated age for older individuals. The Bayesian Method overestimated age for individuals aged between 25 and 35 years and underestimated age for individuals older than 40 years. In comparison, the Pooled Bayesian Method overestimated age (lowest overestimation = 5.0 years) for individuals younger than 46 years and underestimated age (lowest underestimation = -9.2 years) for individuals older than 46 years. For individuals younger than 55 years, the Pooled Bayesian Method provided estimated ages closer to the true ages as compared to the estimated ages obtained when using the Complex Method. After the age of 55 years, the Complex Method provided estimated ages that were closer to true ages.

In males, the Complex Method overestimated the age for individuals aged between 28 and 62 years and underestimated the age for older individuals. In comparison, both the Bayesian Method and the Pooled Bayesian Method overestimated age for individuals younger than 40 years and overestimated age for individuals older than 40 years. The lowest overestimation of age was 1.5 years, and the underestimation of age was -0.4 years when using the Pooled Bayesian Method. Like the Bayesian Method once again, the Pooled Bayesian Method provided estimated ages closer to the true ages for individuals younger than 60 years and the Complex Method provided estimated ages closer to the true ages for individuals older than 60 years.

4.2.4.6. Accuracy using the Pooled Bayesian Method compared to the Bayesian and Complex Methods.

The accuracy was obtained by calculating the percentage of individuals whose estimated age fell within ±2.5 years, ±2.6 to ±5 years, ±5.1 to ±10 years, ±10.1 to ±15 years and more than 15 years of true age for each of the 15 skeletal indicator combinations. The highest proportion of individuals in the pooled sexes (ranging from 43.7% to 61.9%), female (ranging from 44.4% to 66.7%) and male (ranging from 35.7% to 57.1%) samples were aged more than 15 years different from their true age for all skeletal indicator combinations used. Please refer to Figures 12 to 14 for the Complex Method results, Figures 17 to 19 for the Bayesian Method results and Figures 22 to 24 for the Pooled Bayesian Method results.

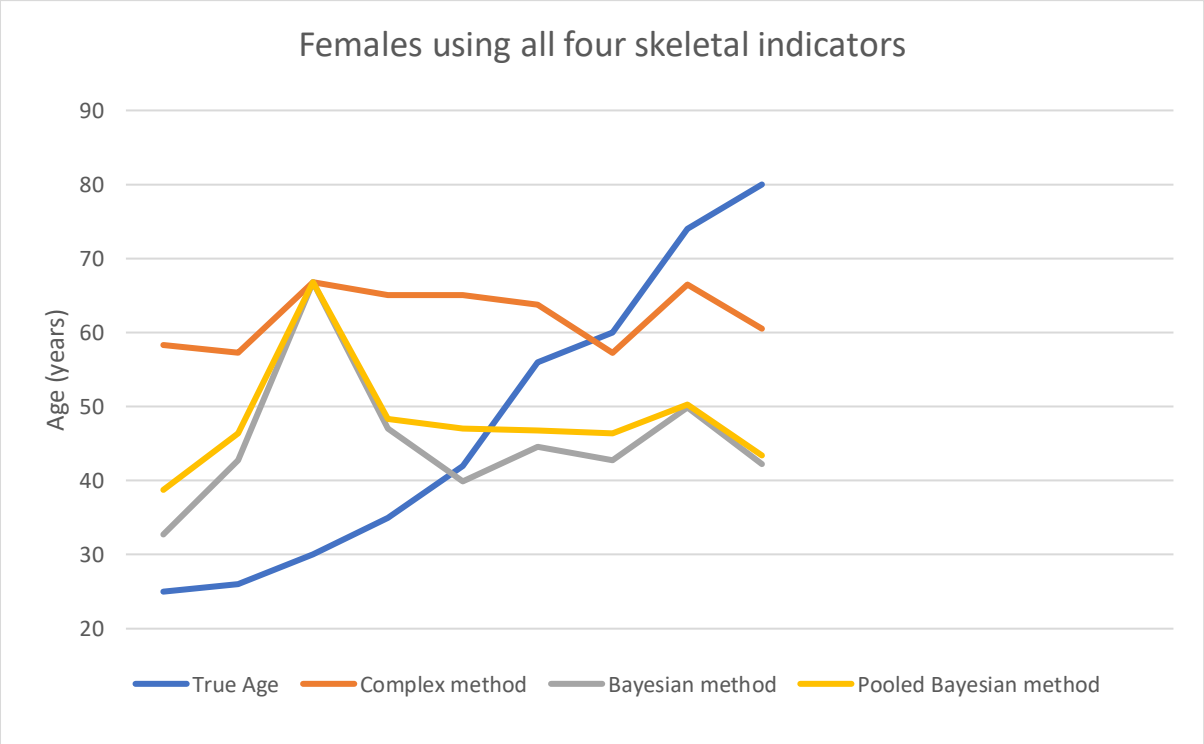


Figure 20: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Pooled Bayesian Method for females.

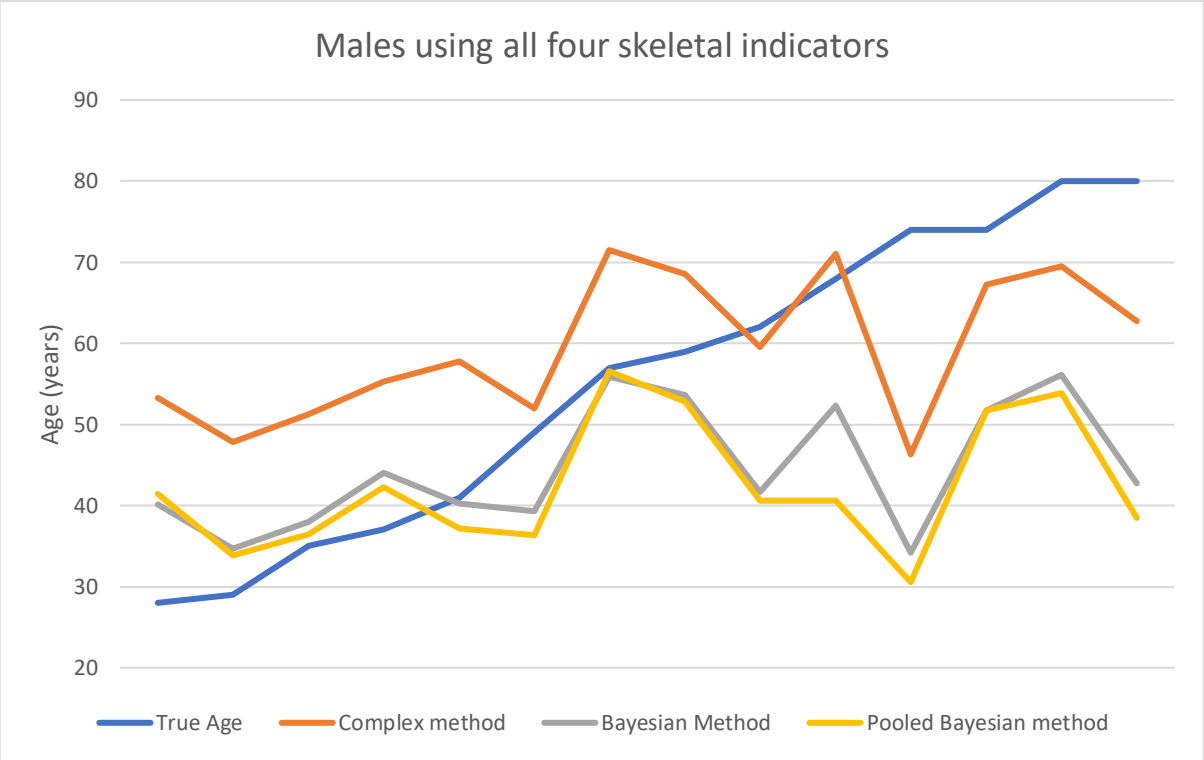


Figure 21: True versus estimated age using Acsádi & Nemeskéri Complex Method, Bayesian Method and Pooled Bayesian Method for males.

For the combination utilising all four skeletal indicators, only 7.1% of the pooled sexes and none of the female's estimated age differed by ±2.5 years of true age. Like the results obtained using the Bayesian Method, 14.3% of the male's estimated age differed by ±2.5 years of true age. Compared to the accuracy results obtained when using the Complex Method, the Pooled Bayesian Method provided an improvement only for the males (Complex Method = 9.3%). The four skeletal indicator combination performed the best in males (28.6% aged within ±5 years; 42.9% aged within ±10 years; 57.1% aged within ±15 years) compared to the pooled sexes (19.8% aged within ±5 years; 32.5% aged within ±10 years; 56.3% aged within ±15 years) and females (11.1% aged within ±5 years; 22.2% aged within ±10 years; 55.6% aged within ±15 years).

For the combination utilising three skeletal indicators, the combination of the cranial sutures, pubic symphysis and proximal humerus performed the best for the pooled sexes (17.9 % aged

within ±5 years; 27.0% aged within ±10 years; 56.3% aged within ±15 years) and males (35.7% aged within ±5 years; 42.9% aged within ±10 years; 57.1% aged within ±15 years). For females, the combination of the cranial sutures, proximal humerus, and proximal femur (11.1% aged within ±5 years; 33.3% aged within ±10 years; 55.6% aged within ±15 years) performed the best. The combination of the cranial sutures, proximal humerus and proximal femur also performed the best for females (11.1% aged within ±5 years; 22.2% aged within ±10 years; 55.6% aged within ±15 years) when using the Bayesian Method. When comparing the Pooled Bayesian Method results to those obtained when using the Bayesian Method, there was an improvement only in the overall proportion of individuals aged up to ±10 years of their true ages. This combination of the cranial sutures, proximal humerus and proximal femur performed the worst for males (21.4% aged within ±5 years; 42.9% aged within ±10 years; 50% aged within ±15 years). The same combination of the cranial sutures, proximal humerus and proximal femur performed the worst in males when using the Complex Method and Bayesian Method. The worst performance was by the combination of the pubic symphysis, proximal humerus, and proximal femur for the pooled sexes (25.4% aged within ±5 years; 32.5% aged within ±10 years; 39.7% aged within ±15 years) and females (22.2% aged within ±5 years; 22.2% aged within ±10 years; 22.2% aged within ±15 years). Again, the same combination of the pubic symphysis, proximal humerus, and proximal femur performed the worst for the pooled sexes and females when using the Complex Method and Bayesian Method.

When looking at the combinations utilising two skeletal indicators, the combination of the cranial sutures and proximal femur performed the best for the pooled sexes (18.3% aged within ±5 years; 40.1% aged within ±10 years; 56.3% aged within ±15 years) and females (22.2% aged within ±5 years; 44.4% aged within ±10 years; 55.6% aged within ±15 years). The combination of the cranial sutures and proximal femur also performed the best for females (11.1% aged within ±5 years; 33.3% aged within ±10 years; 44.4% aged within ±15 years)

when using the Bayesian Method. When comparing the Pooled Bayesian Method results to those obtained using the Bayesian Method, the Pooled Bayesian Method showed an overall improvement in females. However, using the Bayesian Method, this combination of the cranial sutures and proximal femur performed the worst in the pooled sexes. For males, the combination of the pubic symphysis and proximal femur (28.6% aged within ±5 years; 42.9% aged within ±10 years; 64.3% aged within ±15 years) performed the best. This combination of the cranial sutures and proximal femur also performed the best in males (35.7% aged within ±5 years; 42.9% aged within ±10 years; 71.4% aged within ±15 years) when using the Bayesian Method. When comparing the Pooled Bayesian Method results to those obtained when using the Bayesian Method, there was no improvement in any of the overall proportion of individuals. When using the Complex Method, the combination of the pubic symphysis and proximal femur performed the worst for the pooled sexes. The worst performance was by the combination of pubic symphysis and proximal humerus for the pooled sexes (23.4% aged within ±5 years; 32.5% aged within ±10 years; 36.1% aged within ±15 years) and females (11.1% aged within ±5 years; 22.2% aged within ±10 years; 22.2% aged within ±15 years). The combination of the pubic symphysis and proximal femur (11.1% aged within ±5 years; 22.2% aged within ±10 years; 22.2% aged within ±15 years) performed equally bad for females. The combination of the pubic symphysis and proximal humerus also performed the worst in males using the Bayesian Method. For males, the combination of the proximal humerus and proximal femur (35.7% aged within ±5 years; 35.7% aged within ±10 years; 42.9% aged within ±15 years) performed the worst.

For the use of a single skeletal indicator, the proximal femur performed the best for the pooled sexes (10.7% aged within ±5 years; 34.5% aged within ±10 years; 56.3% aged within ±15 years) and females (0.0% aged within ±5 years; 33.3% aged within ±10 years; 55.6% aged within ±15 years). The pubic symphysis performed the best for males (28.6% aged within ±5

years; 42.9% aged within ±10 years; 64.3% aged within ±15 years). When using the Bayesian Method, the pubic symphysis males (35.7% aged within ±5 years; 50% aged within ±10 years; 57.1% aged within ±15 years) also performed the best for males. There was an improvement in the overall proportion of individuals that were aged up to ±15 years of their true ages when using the Pooled Bayesian Method compared to the Bayesian Method. The worst performance was by the proximal humerus for the pooled sexes (16.3% aged within ±5 years; 29.0% aged within ±10 years; 38.1% aged within ±15 years), females (11.1% aged within ±5 years; 22.2% aged within ±10 years; 33.3% aged within ±15 years) and males (21.4% aged within ±5 years; 35.7% aged within ±10 years; 42.9% aged within ±15 years). The same skeletal indicator, the proximal humerus, performed the worst for the pooled sexes and males when using the Complex Method and Bayesian Method.

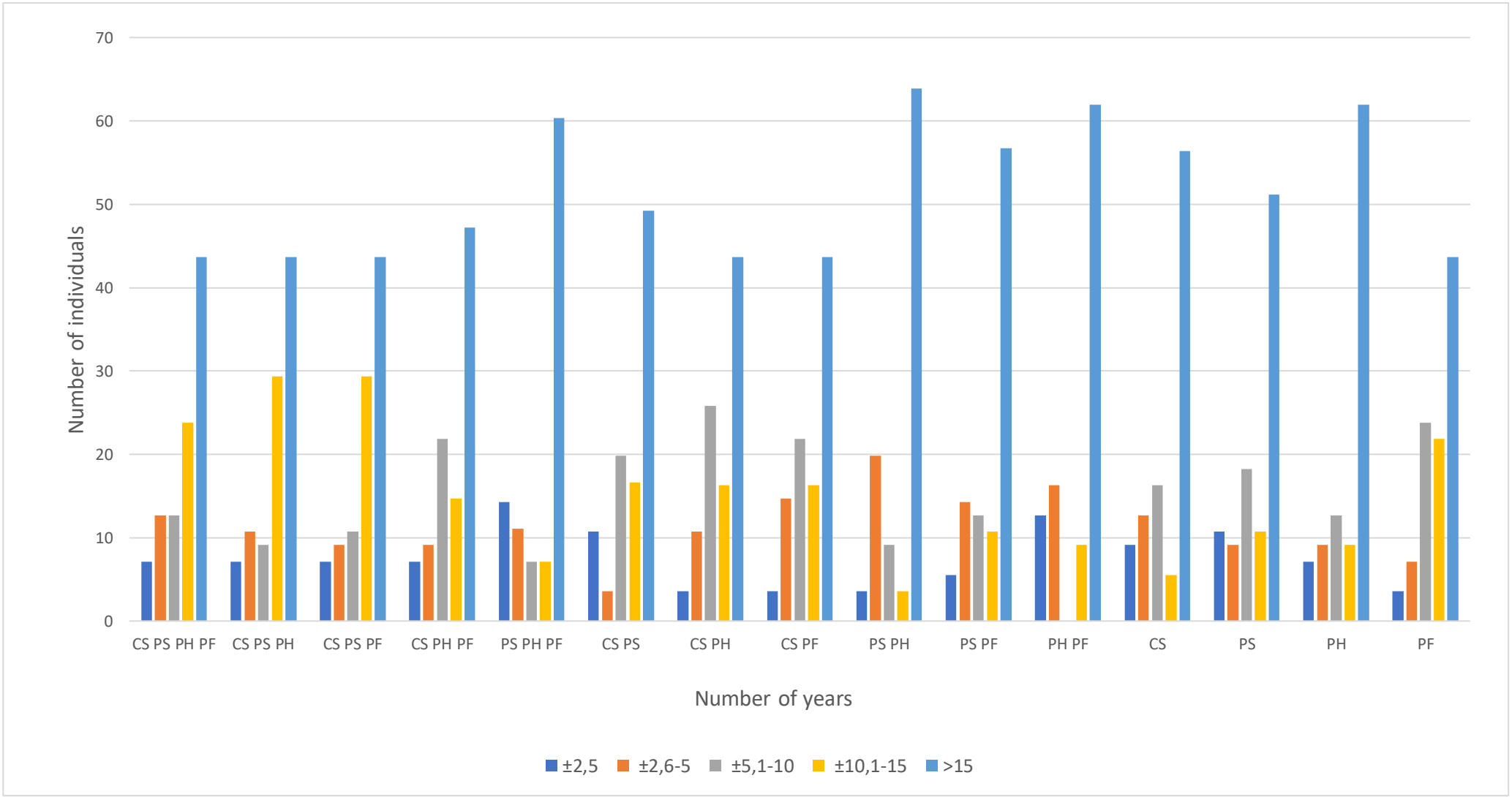


Figure 22: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for the pooled sexes using the Pooled Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

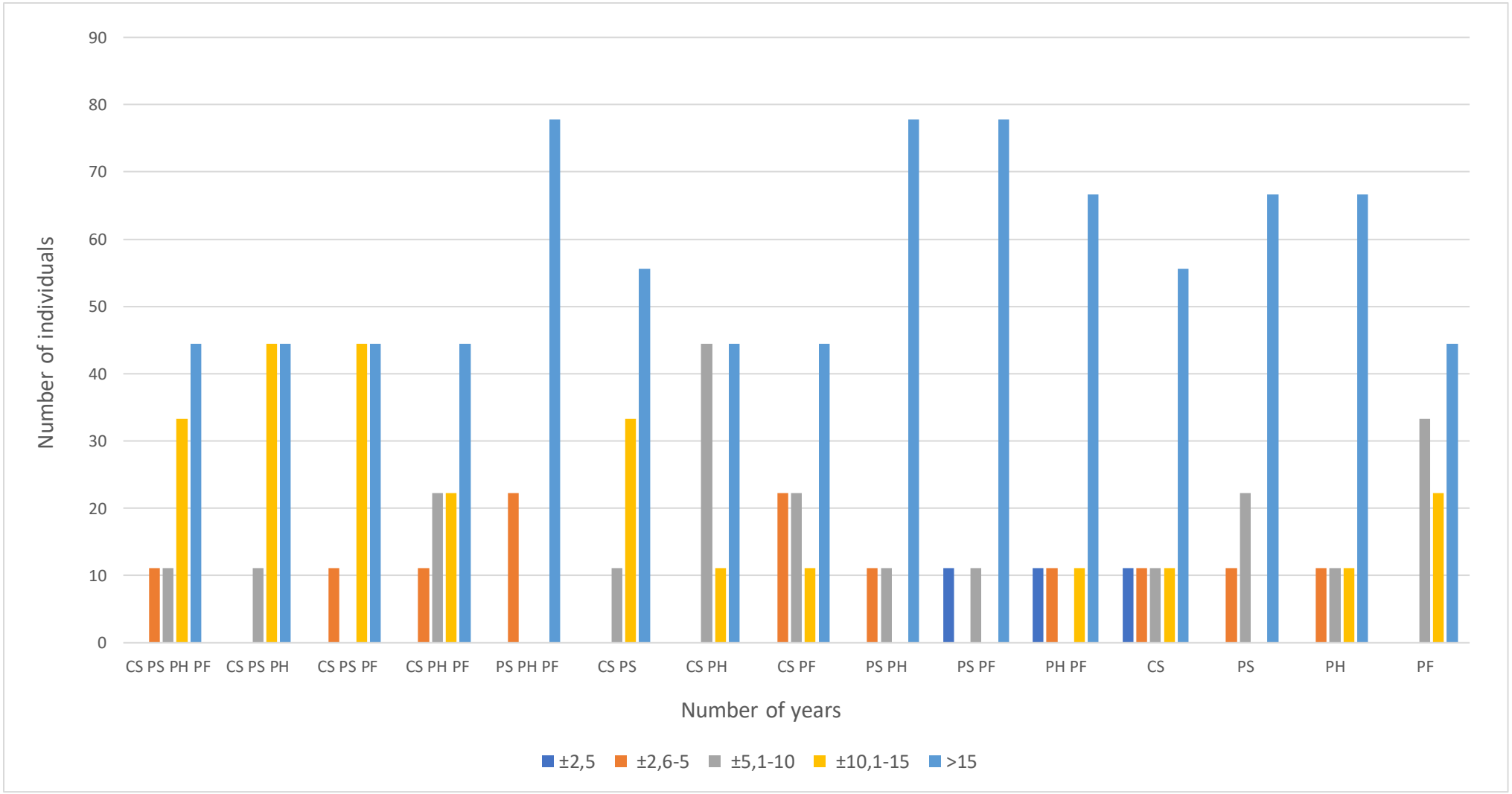


Figure 23: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for females using the Pooled Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal Femur

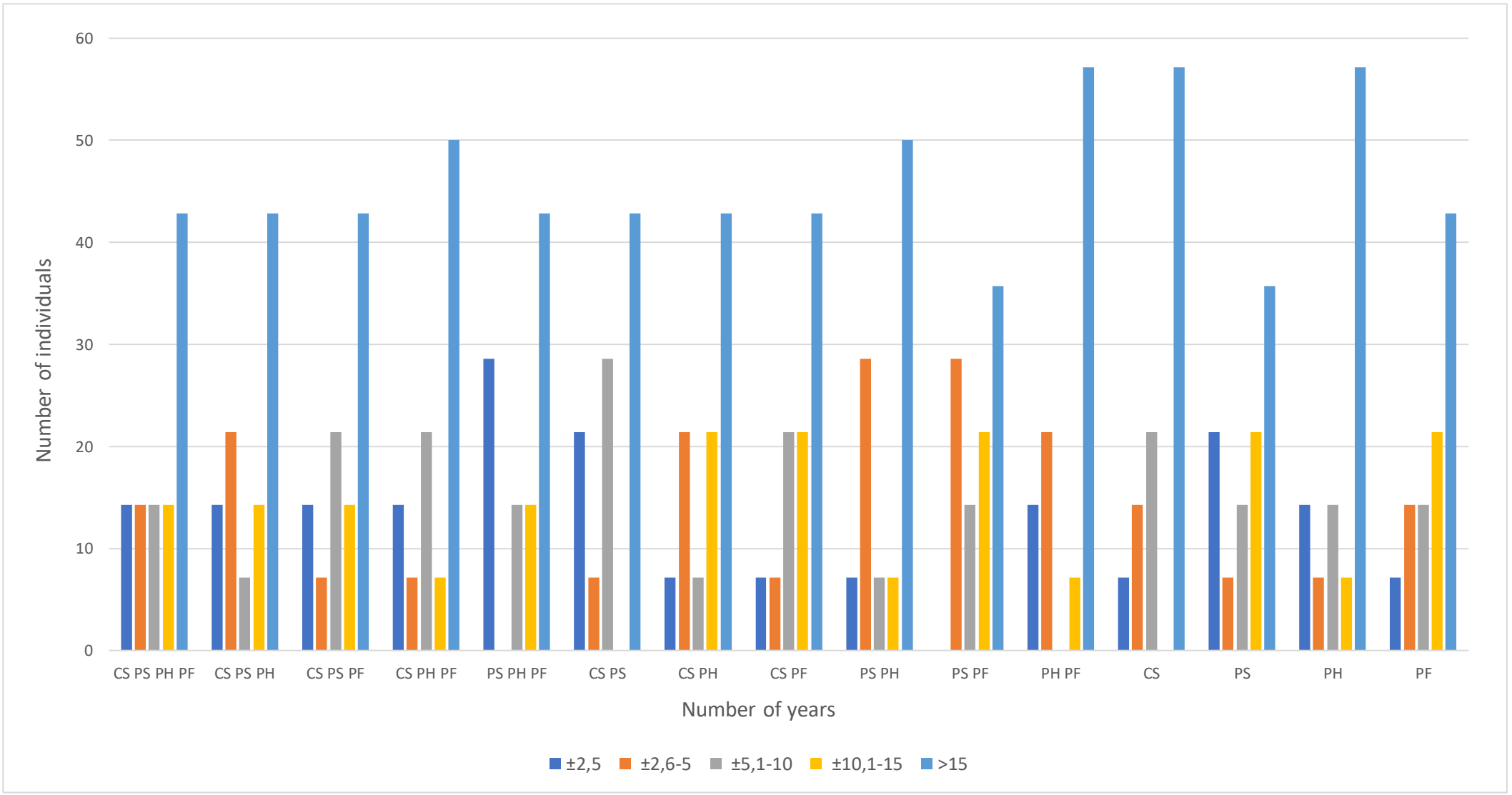


Figure 24: Percentage of individuals whose estimated age fell within ±2.5 and >15 years of their true age for males using the Pooled Bayesian Method. CS = Cranial sutures, PS = Pubic symphysis, PH = Proximal Humerus, PF = Proximal femur

CHAPTER 5: DISCUSSION

Acsádi & Nemeskéri's (1970) Complex Method was one of the first multiple indicator age estimation methods developed, which uses a weighting system to estimate age. It utilises four skeletal age indicators in assessing age: the pubic symphysis, endocranial sutures and internal structural changes in the proximal femur and humerus (observable through cross-sectioning of bone or radiographs). In circumstances when indicators are poorly preserved or missing, such as in forensic and archaeological cases, the method allows for using a single indicator, or any combination of the skeletal indicators mentioned above. Acsádi & Nemeskéri (1970) reported a margin of error of ±2.5 years when all four indicators were used.

Even with such promising results, the Complex Method has been criticised for its small reference sample size which is not large enough to represent inter-individual differences accurately, the statistical validity of the averaging procedure, and its constant under- and overestimation of age for older and younger individuals, respectively (Saunders *et al*., 1992; Molleson *et al*., 1993; Meyer *et al.*, 2021). An ageing method's under- and overestimation of age have. amongst others, been linked to the use of least square regression analyses to develop mean age estimates in traditional methods (Tersigni-Tarrant & Shirley, 2013). Age estimates obtained tend to be biased toward the reference sample used and Bayesian statistics have been proposed to counter this effect (Hoppa and Vaupel, 2002). In this study, the Complex Method was tested on a black South African sample, and Bayesian statistics were used to see whether the results could be improved.

5.1. The performance of the Complex Method

5.1.1. Intra-and inter-observer repeatability

In this study, the intra- and interobserver error results showed that for the most part, the Complex Method was less repeatable in males compared to females (refer to Table 12 and 13 in the results section). The highest interobserver repeatability scores were obtained by the cranial sutures (κ = 0.737), followed by the pubic symphysis (κ = 0.474), then the proximal humerus (κ = 0.320), and the worst by the proximal femur (κ = 0.291). Similar results were obtained by Meyer *et al*. (2021) using a white South African sample, who also found that the method was less repeatable in males for all skeletal indicators, with the cranial sutures performing the best, followed by the pubic symphysis, then the proximal humerus and last the proximal femur. Interobserver error scores obtained in this study using a black South African sample were also worse than those obtained by Meyer *et al*. (2021) using a white South African sample. The high interobserver results for cranial sutures and pubic symphysis and the low interobserver results for the proximal femur and proximal humerus could be due to the scorer's experience level. The second observer was a postgraduate student, and most students are more familiar with age determination using the cranial sutures and pubic symphysis, rather than the proximal femur and proximal humerus through radiographs. This suggests that some experience is required when using the Complex Method, particularly when scoring using radiographs. Meyer *et al.* (2021) also stated that experience is required when scoring changes using radiographs. Molleson *et al*. (1993) noted that radiograph scoring can be challenging because even modest degradation can alter bone appearance and cause confusion for the scorer. The intraobserver results also showed that the method was less repeatable in males than females for all skeletal indicators except the proximal femur (refer to Table 3 in the results section). The highest repeatability for females was obtained by the pubic symphysis (κ = 1.000) and the worst by the proximal humerus (κ = 0.836). For males, the highest repeatability was obtained

by both the pubic symphysis (κ = 0.884) and proximal femur (κ = 0.884) and the worst by the cranial sutures (κ = 0.706). Meyer *et al*. (2021) also found similar results testing the Complex Method on a white South African sample. The high repeatability results for females could be related to more marked changes in females as a result of specific physiological processes such as puberty and menopause (Mays, 2015; Langley-Shirley and Jantz, 2010; San-Millán *et al*., 2016), making age-related skeletal features easier to distinguish and classify. Nevertheless, the intra- and interobserver scores were still within sufficient limits for age estimation methods, showcasing the usefulness of the Complex Method as an adult age estimation method.

5.1.2. Descriptive statistics

When comparing the descriptive statistics for the estimated and true ages, the mean estimated age was older by ±4 years than the mean true age for the pooled sexes, ±5 years for females and ±3 years for males. The estimated minimum ages were older (pooled sexes = 26.5 years) than the true minimum ages (pooled sexes = 20 years), and the estimated maximum ages (pooled sexes = 71.5 years) were younger than the true maximum ages (pooled sexes = 98 years). The older minimum ages and younger maximum ages show an "attraction to the middle" bias, a phenomenon by Masset (1989: 81) and a consequence of age mimicry, a phenomenon by Bocquet-Appel & Masset (1982). Age mimicry is the tendency of the target sample's estimated age-at-death distribution to mimic the reference sample's age-at-death distribution (Tersigni-Tarrant & Shirley, 2013). The reference sample used to develop the Complex Method was skewed towards older individuals, the majority of whom were between 40 and 70 years old. The reference sample used to develop the pubic symphysis, proximal humerus and proximal femur age ranges comprised of 62% pooled sexes, 70% females and 56% males older than 50 years. The reference sample used to develop the cranial sutures age ranges consisted of 71% pooled sexes, 72% females and 71% males older than 50 years. This

means that the age estimates obtained using the Complex Method would be biased towards older individuals, especially for the ages between 40 and 70 years, leading to an overall older mean and median estimated age as was seen in this study. Because of the older reference sample, the overall mean age of the reference sample would be older than that of the sample used in this study, as seen in the results. The estimated mean ages would be biased towards older individuals, resulting in an older estimated minimum age and a younger estimated maximum age.

The reference sample used to develop the Complex Method was primarily composed of older individuals (62% pooled sexes, 70% females and 56% males older than 50 years), but it was also relatively small, which it has been criticised for before (Kemkes-Grottenthaler, 2002; Saunders *et al*., 1992) and consisted of more males (n = 61) than females (n = 44). The small reference sample would affect the formation of phase descriptions, as one would not be able to observe and document the wide variation of age-related changes seen in each skeletal indicator. The phase descriptions formed would not accommodate for some features, especially those seen in younger and older individuals and in females. This would lead to misclassification of individuals, affecting both the accuracy and repeatability of the method. Inherent differences between different populations with different ageing patterns would also affect the accuracy and repeatability of the method. The Complex Method was developed using a European sample. Martins *et al.* (2012), looking at age estimation from the pubic symphysis and auricular surface using European and black South African samples, noted that European and South African populations show distinct morphological variations in ageing, with the European population having a more uniform pattern. The current study tested the Complex Method on a black South African sample, which has a different ageing pattern from the European sample used to develop the method. Another factor that would affect accuracy is secular changes in skeletal maturation. Secular changes are microevolutionary changes resulting from gene flow and environmental

factors (nutrition, health, living conditions, etc.) (Henneberg & van den Berg, 1990; Hawley *et al.*, 2009). These changes can positively or negatively affect the onset of skeletal growth and development. As a developing country, South Africa has experienced substantial socio-economic changes, which may be linked to positive secular changes in skeletal maturity. Hawley *et al.* (2009), for example, found that between 1962 and 2001, black males and females had significant secular increases in skeletal maturity of 9.7 months and 15.8 months, respectively. Such secular changes could also translate into degenerative ageing. The Complex Method was developed using autopsied individuals from the years 1956 to 1957. The sample used in this study consisted of individuals who died between the years 1944 and 2009. Age estimates based on the Complex Method reference sample standards, which include individuals from a period that has undergone significant positive secular changes, may be overestimated for those individuals who died after 1957. However, the most significant factor that may have led to differences in the ageing profiles would still relate to the very different genetic and socio-economic backgrounds of the Complex Method's sample and the current study sample.

5.1.3. Skeletal indicator correlations with age

Looking at the correlation of each skeletal indicator to true age, the pubic symphysis presented the highest correlation for females (rho = 0.396) and the second highest correlation for males (rho = 0.491). Hartnett (2010) and Joubert *et al.* (2019) also found a higher correlation between females and males when using the pubic symphysis when testing the Suchey-Brooks age estimation method. Testing the Complex Method on a white South African sample, Meyer *et al.* (2021) also found a higher correlation in females than males when using the pubic symphysis. The better performance of the pubic symphysis is to be expected as it has long been regarded as one of the most reliable age indicators in adult age estimation methods, keeping in mind population differences (McKern & Stewart, 1973; Acsádi & Nemeskéri,1970; Brooks &

Suchey, 1990; Schmitt, 2004; Hens *et al.*, 2008; Garvin *et al.*, 2012). Acsádi & Nemeskéri (1970) reported that the pubic symphysis also had the strongest correlation to true age and; therefore, used it as the primary weighting factor to obtain the final age estimate. The results obtained for females in this study agree with the findings of Acsádi & Nemeskéri (1970); however, for the male sample, the cranial sutures (rho = 0.631) outperformed the pubic symphysis (rho = 0.491). The cranial sutures performed the best in males (rho = 0.631) and the third best in females (rho = 0.255). The poor performance of the cranial sutures in females is to be expected as several studies have reported the cranial sutures as an unreliable age indicator (Meindl & Lovejoy, 1985; Acsádi & Nemeskéri,1970; Key *et al.*, 1994; Purves *et al.*, 2011; Meyer *et al*., 2021). Because of population- and sex-specific differences in the onset and timing of cranial suture closure, there is still disagreement about their use in age estimation (McKern & Stewart, 1957; Christensen *et al.*, 2014; Hartnett-McCann *et al.*, 2018). In their study testing the Complex Method ecto- and endocranial criteria in a black North American sample, Galera and colleagues (2010) also found a stronger correlation to true age in males than females. They attributed this to sexual differences as well as population differences. Hershkovitz *et al*. (1997) looking at the sagittal suture in a white and black American sample, also found stronger correlations to true age in males compared to females, noting that females had slower suture closure rates compared to males and black American females particularly had slower suture closure rates compared to white American females. Hershkovitz *et al*. (1997) also attributed their findings to sex differences.

The proximal femur and proximal humerus showed low correlations to true age in both females (proximal femur rho = 0.387; proximal humerus rho = 0.211) and males (proximal femur rho = 0.134; proximal humerus rho = 0.172). The proximal femur performed second best in females and worst in males, while the proximal humerus performed worst in females and third in males. The low correlation results of the proximal femur and proximal humerus are not surprising as

black South African individuals have been found to have prolonged bone formation, reaching peak bone mass at a much older age, and higher bone mass density compared to white South African individuals (Daniels *et al*., 1995; Daniels *et al*., 1997; Botha *et al.*, 2020). This means that when internal structural changes in the proximal femur and proximal humerus are used to estimate age, black South African individuals will appear younger than their true age. Botha *et al*. (2020) also found very low correlations between bone mineral density and true age in a black South African sample.

Interestingly, the performance of the pubic symphysis and proximal femur in the female sample followed the order of results obtained by Acsádi & Nemeskéri (1970). That is, the pubic symphysis performed the best, followed by the proximal femur. Acsádi & Nemeskéri (1970) found that the proximal femur had the second strongest correlation to true age and should be used as the weighting factor in the absence of the pubic symphysis. The results of the female sample in this study support the correlation findings and the usage of the pubic symphysis and proximal femur as weighting factors. However, for the male sample, the cranial sutures performed the best, followed by the pubic symphysis, then the proximal humerus and last the proximal femur. This shows that while the weighting system may work for the black South African females, it may not be appropriate for the black South African males. This also suggests that population- and sex-specific weighting systems should be considered. In testing the Complex Method on a white South African sample, Meyer and colleagues (2021) found that the proximal humerus outperformed the proximal femur in females, and they also concluded that a blanket weighting system may not be the most effective option.

5.1.4. Correlations between true and estimated ages.

The degree of correlation between true and estimated age for the fifteen skeletal indicator combinations was also assessed. The highest correlations with true age were obtained when two skeletal indicators were utilised, specifically the cranial sutures and pubic symphysis combination for the pooled sexes (rho = 0.516) and males (rho = 0.583) and the pubic symphysis and proximal humerus combination for females (rho = 0.508). This makes sense given that the pubic symphysis provided the strongest correlation with increasing true age, followed by the cranial sutures for the pooled sexes and male samples. For females, the pubic symphysis provided the strongest correlation with increasing true age, however, the proximal humerus performed the worst. The use of three skeletal indicators, specifically the cranial sutures, pubic symphysis, and proximal femur, performed second highest for the pooled sexes (rho = 0.505), females (rho = 0.463) and males (rho = 0.534). The combination utilising all four skeletal indicators performed fourth in the pooled sexes, third best for males and fifth best for females. These results stand in contrast to Acsádi & Nemeskéri's (1970) findings indicating that all four skeletal indicators provided the best results. Meyer and colleagues (2021) also found the highest correlations with true age when two skeletal indicators were utilized, followed by three skeletal indicators for the pooled sexes, females, and males, when testing the Complex Method on a white South African sample. The findings of this study were rather surprising, as the combinations utilising the proximal humerus performed first and second best in females despite the very low correlation with increasing true age of the proximal humerus. This might be explained by the correlation values of the other two skeletal indicators in the combinations. In addition to the cranial sutures these combinations also make use of the pubic symphysis and proximal femur, both of which had the highest correlations with increasing true age in black South African females. The two combinations would, therefore, still be weighted

by the pubic symphysis, potentially minimising the poor-performing effects of the proximal humerus.

5.1.5. Accuracy, inaccuracy, and bias

The highest accuracy for females was obtained by the combination of the cranial sutures and pubic symphysis, with 59.1% of the individuals aged within ±15 years. This combination also provided the lowest inaccuracy value of 14.2 years; however, the combination of the cranial sutures and proximal humerus provided the lowest bias values of 0.7 years. For males, the cranial sutures provided the highest accuracy, with 55.6% of the individuals aged within ±15 years. The combination of the cranial sutures and pubic symphysis, which had the second highest accuracy (50.1 % aged within ±15 years) provided the lowest inaccuracy (15.6 years). However, the combination of the proximal femur and proximal humerus provided the lowest bias value of 0.6 years. The use of all four skeletal indicators resulted in slightly higher inaccuracy values (pooled sexes = 15.7 years; females = 15.1 years; males = 16.2 years) but rather high bias values (pooled sexes = 4.1 years; females = 5.1 years; males = 3.2 years) compared to the highest combinations mentioned above.

Acsádi & Nemeskéri (1970) reported that when utilising all four skeletal indicators, the Complex Method provides an 80–85% accuracy in estimating age to within ±2.5 years. However, in this study, only 10.7% of the pooled sexes sample, 12.2% of the females' sample and 9.3% of the males' sample fell within 2.5 years. Testing the Complex Method on a white South African sample, Meyer and colleagues (2021) reported only 16.3% of individuals aged within ±2.5 years of true age. These results suggest that the accuracy and error margins provided by Acsádi & Nemeskéri (1970) were extremely optimistic or may reflect a population-specific finding that does not apply to South African populations.

Looking at the overall inaccuracy and bias results obtained in this study, the Complex Method under- and overestimated age for older and younger individuals, respectively, for the black South African sample. This can be explained by Masset's (1989:81) attraction to the middle phenomenon mentioned above. Although the Complex Method under- and overestimated age for the black South African sample, bias values were still low for individuals aged between 50 and 69 years of true age. Molleson *et al*. (1993) and Meyer *et al*. (2021) also found that the Complex Method was most accurate for individuals aged between 50 – 70 years and 40 – 69 years, respectively.

5.2. The performance of the Complex Method scoring criteria using Bayesian statistics

The results of this study show that in a black South African sample, the Acsádi & Nemeskéri (1970) Complex Method underestimates age for older individuals and overestimates age for younger individuals. This demonstrates the "attraction to the middle" bias described by Masset (1989:81), which has been linked to the consequences of age mimicry by Masset & Bocquet-Appel (1982). Age-mimicry is the tendency of the estimated age-at-death distribution of the target sample to mimic the age-at-death distribution of the reference sample (Tersigni-Tarrant & Shirley, 2013). This is because age estimates are obtained through the use of least square regression analyses, which regress the skeletal features onto age and then utilise inverse calibration. (Aykroyd, 1997; Tersigni-Tarrant & Shirley, 2013; Hartnett-McCann *et al*., 2018). Therefore, the age estimates obtained will be biased in the direction of the mean of the reference sample (Tersigni-Tarrant & Shirley, 2013). Bayesian statistics, which use prior knowledge of the age-at-death distribution of the reference sample to estimate age has been proposed to counter these effects of age-mimicry. Therefore, the second part of this study was to determine whether the application of Bayesian statistics to the Complex Method will counter the effects of age-mimicry and increase the accuracy of the method in black South African individuals.

The implementation of Bayesian statistics to the Complex Method scoring criteria resulted in somewhat improved results. Of the fifteen skeletal indicator combinations, the Bayesian Method provided an increase in correlation in three of the combinations for the pooled sexes. For females, there was an increase in seven of the fifteen skeletal indicator combinations and four for males. For females, the Bayesian Method provided stronger correlations between estimated and true age when compared to the Complex Method in these seven combinations: cranial sutures, pubic symphysis, proximal humerus, and proximal femur (rho = 0.461 to rho = 0.493); cranial sutures, pubic symphysis and proximal humerus (rho = 0.450 to rho = 0.499); cranial sutures and pubic symphysis (rho = 0.416 to rho = 0.516); cranial sutures and proximal humerus (rho = 0.321 to rho = 0.368); pubic symphysis and proximal humerus (rho = 0.437 to rho = 0.438); cranial sutures (rho = 0.214 to rho = 0.377); pubic symphysis (rho = 0.396 to rho = 0.461). For males, the Bayesian Method provided stronger correlations between estimated and true age when compared to the Complex Method in these four combinations: cranial sutures and proximal femur (rho = 0.289 to rho = 0.385); pubic symphysis and proximal femur (rho = 0.448 to rho = 0.524); proximal humerus and proximal femur (rho = 0.131 to rho = 0.209); pubic symphysis (rho = 0.491 to rho = 0.637).

When looking at accuracy, there was an overall decrease in the percentage of individuals aged more than 15 years of the true age. For the pooled sexes, the Bayesian Method provided improved results in twelve of the fifteen skeletal indicator combinations, and all skeletal indicator combinations for males, with percentages decreasing from above 50% for all skeletal indicators to below. For males, the Bayesian Method provided the best results for the pubic symphysis and proximal femur combination (53.7% to 28.6%). However, the Bayesian Method did not perform as well for the females, providing better results in only one of the fifteen skeletal indicator combinations, the cranial sutures, proximal humerus, and proximal femur combination (49% to 44%), and an increase in the percentage of individuals aged above

15 years of true age for the other fourteen skeletal indicator combinations compared to when the Complex Method was used.

The Bayesian Method did not fully resolve the over- and underestimation of age problem observed in the Complex Method, but it did show some improvement. There was still an overestimation of age for younger individuals and an underestimation of age for older individuals, however, the number of years for under- and overestimation decreased, with the Bayesian Method providing estimated ages closer to true age compared to the Complex Method. For females younger than 50 years and males younger than 60 years, the Bayesian Method provided estimated ages that were closer to the true ages compared to the Complex Method. These results may be attributed to the prior distribution sample used which included a larger proportion of younger individuals compared to older individuals (Dayal *et al*., 2009).

The differences in results obtained for the female and male samples stress the need for sex-specific ageing standards. The Complex Method performed better in females compared to males, with females having higher correlation scores in eight of the fifteen skeletal indicator combinations. In addition, for thirteen out of the fifteen skeletal indicator combinations, the percentage of males aged above 15 years of true age was higher than the percentage of females. These results make sense as females' skeletal indicator correlation to increasing true age for the pubic symphysis and proximal femur followed the results obtained by Acsádi & Nemeskéri (1970), making the weighting system perform better for the females. In contrast, the Bayesian Method performed better in males compared to females, with males having higher correlation scores in eight of the fifteen skeletal indicator combinations. The absence of a weighting scheme in the Bayesian Method may explain its success among males. The Bayesian Method does not use a weighting system, instead assigning equal weight to each skeletal indicator. In comparison, the Complex Method uses a weighting system that includes the pubic symphysis and proximal femur are used as weighting factors. The poor correlation of the pubic symphysis

and proximal femur to increasing true age in males may have influenced the performance of the weighting system, resulting in the poor performance of the Complex Method in males. Females also had a higher percentage of individuals aged above 15 years of true age than males for all skeletal indicator combinations, when using the Bayesian Method. The poor performance of the Bayesian Method in females might also be because of the lack of individuals scored into phases I of the proximal humerus and phases I and II of the proximal femur. These results again point to the poor performance of the weighting system in males and emphasize the need for sex-specific weighting systems.

The decision to develop the Collapsed Bayesian and Pooled Bayesian Methods was made after it was observed that the ages-at-transition of the Bayesian Method were very unrealistic as most individuals cannot reach that age. The success of Bayesian Statistics is determined by the realisation of various underlying assumptions (Konigsberg and Frankenberg, 1992; Boldsen *et al*., 2002; Nikita *et al*., 2018). One such assumption is during transition analysis (TA) to determine the ages-at-transition, where the accuracy of the acquired results relies on how well the probit regression models describe the TA sample's skeletal ageing characteristic vs. log(age) data (Nikita *et al*., 2018). This assumption is fulfilled if the TA sample is big enough to adequately represent all stages or phases of the skeletal ageing characteristic under investigation (Nikita *et al*., 2018). It was concluded that the unrealistic ages-at-transition obtained for the Bayesian Method was due to the failure to fulfil this assumption. The sample used in the Bayesian Method was rather small (female: n = 40; males: n = 40) and did not have enough individuals per stage or phase, with some having only one individual. Nikita *et al*. (2018) tested the implementation of Bayesian statistics using varying TA sample sizes, one small (n = 81) and another large (n = 372) and found improvement in age estimates only when the larger TA sample was used. Godde & Hens (2012) using a large TA sample size (n = 212 and n = 422), also found increased accuracy when using Bayesian statistics. It was then decided

to try two methods to resolve this matter. The first was by combining various phases or stages to increase the number of individuals and to have roughly equal numbers of individuals per phase, thus forming the Collapsed Bayesian Method. This method however resulted in a very small number of phases, specifically for the pubic symphysis, proximal humerus, and proximal femur, with only three phases. These phases produced very wide age ranges that are not precise and useless in a forensic context as they cannot help narrow down a missing persons list (Garvin *et al*., 2012). The second was to combine the female and male samples, to form a pooled sexes sample with an increased sample size and a number of individuals per phase or stage, thus forming the Pooled Bayesian Method.

Collapsing the phases to increase the number of individuals per phase did result in more realistic ages-at-transition for both female and male samples, except for the proximal femur for the female sample. This was because only two phases of the proximal femur could be compressed as no individuals were scored into phases I and II. However, collapsing the phases for both samples resulted in an increase in mean age estimates and a widening of age ranges for all skeletal indicators. These results were expected as widening of age ranges will result in higher mean ages. These mean age estimates were, however, still younger than those obtained in the Bayesian Method, therefore showing some improvements in the results. The collapsing of phases forming fewer phases and very wide age ranges would affect the method's usefulness. However, these wide age ranges are considered more accurate as they are more likely to encompass true age (Milner & Boldsen, 2012; Buckberry, 2015). For the Pooled Bayesian Method, the increase in TA sample size still resulted in unrealistic ages-at-transition, though the unrealistic ages were younger. One possible explanation for this could be that although the sample size and number of individuals increased for each phase or stage, there was still an unequal number of individuals per phase.

The Collapsed Bayesian and Pooled Bayesian Methods performed similarly for the correlation between estimated and true age. There was an increase in only two (pubic symphysis; proximal femur as single indicators) of the fifteen skeletal combinations for the pooled sexes, in one (pubic symphysis as a single indicator) combination for males and none for females when using the Collapsed Bayesian Method. In comparison, there was an increase in only one (pubic symphysis as a single indicator) of the fifteen skeletal indicator combinations for both the pooled sexes and males and in two (proximal humerus; proximal femur as a single indicator) combinations for females when using the Pooled Bayesian Method. The increase in correlation between true and estimated age in the Collapsed Bayesian and Pooled Bayesian Methods was obtained by skeletal indicators, which showed the strongest correlation to increasing true age.

While they showed some improvement, the overestimation and underestimation of age issues seen in the Complex Method were not entirely resolved by either the Collapsed Bayesian or Pooled Bayesian Methods, similar to the Bayesian Method. There were still instances of overestimation of younger individuals and underestimation of older individuals. However, the number of years for under- and overestimation decreased when the Bayesian Methods were compared to the Complex Method. Comparing the two methods, for females, the Collapsed Bayesian Method performed better than the Pooled Bayesian Method for individuals aged younger than 40 years. After 40 years, the Pooled Bayesian Method performed better than the Collapsed Bayesian Method, providing age estimates closer to true age. In contrast, male performance varied according to age groups. The Collapsed Bayesian Method worked better for individuals in their twenties, the Pooled Bayesian Method worked better for individuals in their thirties, and so on. These results may be attributed to each method's mean ages obtained per phase. The Collapsed Bayesian Method had different mean ages per phase that increased with an increase in phase, while the Pooled Bayesian Method had repeating mean ages for some phases (e.g. proximal femur phases I and II had mean ages of 20 years). When using the

Pooled Bayesian Method, individuals aged into phase II, though meant to be older, would be underestimated. The difference in mean ages per phase could explain the better performance of the Collapsed Bayesian Method in younger individuals. However, the Pooled Bayesian Method had mean ages for later phases that were older (e.g. proximal femur phase VI = 60 years) than the mean ages obtained for the Collapsed Bayesian Method (e.g. proximal femur phases IV-VI = 51 years). The Collapsed Bayesian Method did not accommodate older individuals, resulting in an underestimation of age and poor performance in older individuals compared to the Pooled Bayesian Method. Both methods provided age estimates closer to true age than the Complex Method. For females, the Collapsed Bayesian Method estimated ages closer to the true age for individuals younger than 50 years and younger than 55 years for the Pooled Bayesian Method. Both methods provided close estimated ages for individuals younger than 60 years for males.

Looking at accuracy, both modified Bayesian Methods provided an overall decrease in the percentage of individuals aged more than 15 years of true when compared to the Complex Method. The Collapsed Bayesian Method provided a decrease in percentages of up to 8.8% for the pooled sexes, 4.6% for females and 21.7% for males. The Collapsed Bayesian Method provided a decrease in percentages or improved results in nine of the fifteen skeletal indicator combinations (cranial sutures, pubic symphysis, proximal humerus and proximal femur; cranial sutures, pubic symphysis and proximal humerus; cranial sutures, pubic symphysis and proximal femur; cranial sutures, proximal humerus and proximal femur; pubic symphysis, proximal humerus and proximal femur; pubic symphysis and proximal humerus; pubic symphysis and proximal humerus; proximal humerus and proximal femur; proximal humerus) for the pooled sexes, fourteen skeletal indicator combinations (cranial sutures, pubic symphysis, proximal humerus and proximal femur; cranial sutures, pubic symphysis and proximal humerus; cranial sutures, pubic symphysis and proximal femur; cranial sutures,

proximal humerus and proximal femur; pubic symphysis, proximal humerus and proximal femur; cranial sutures and pubic symphysis; cranial sutures and proximal humerus; cranial sutures and proximal femur; pubic symphysis and proximal humerus; pubic symphysis and proximal humerus; proximal humerus and proximal femur; pubic symphysis; proximal humerus; proximal femur) for males, and one (proximal femur) for females. The Pooled Bayesian Method provided a decrease in percentages of up to 9.7% for the pooled sexes, 8.7% for females and 18% for males. The Pooled Bayesian Method provided a decrease in percentages or improved results in eight of the fifteen skeletal indicator combinations (cranial sutures, pubic symphysis, proximal humerus and proximal femur; cranial sutures, pubic symphysis and proximal humerus; cranial sutures, pubic symphysis and proximal femur; cranial sutures, proximal humerus and proximal femur; cranial sutures and proximal humerus; cranial sutures and proximal femur; pubic symphysis; proximal femur) for the pooled sexes, in fourteen skeletal indicator combinations (cranial sutures, pubic symphysis, proximal humerus and proximal femur; cranial sutures, pubic symphysis and proximal humerus; cranial sutures, pubic symphysis and proximal femur; cranial sutures, proximal humerus and proximal femur; pubic symphysis, proximal humerus and proximal femur; cranial sutures and pubic symphysis; cranial sutures and proximal humerus; cranial sutures and proximal femur; pubic symphysis and proximal humerus; pubic symphysis and proximal humerus; proximal humerus and proximal femur; pubic symphysis; proximal humerus; proximal femur) for males and seven skeletal indicator combinations (cranial sutures, pubic symphysis, proximal humerus and proximal femur; cranial sutures, pubic symphysis and proximal humerus; cranial sutures, pubic symphysis and proximal femur; cranial sutures, proximal humerus and proximal femur; cranial sutures and proximal humerus; cranial sutures and proximal femur; proximal femur) for females. However, both methods yielded worse results in the remaining skeletal indicator combinations. The Collapsed Bayesian Method provided an increase in percentages of up to

5.6% for the pooled sexes, 25.9% for females and 5.6% for males. The Pooled Bayesian Method provided increased percentages of up to 12.6% for the pooled sexes, 24.7% for females and 12.7% for males. The accuracy results may be attributed to the Bayesian Methods' removal of a weighting system. Therefore, males who performed badly in the Complex Method performed better than females in the Bayesian Methods.

Again, there was a clear difference in performance in the female and male samples, with both modified Bayesian Methods performing better in males. Both methods provided somewhat better results than the Complex Method, with the Pooled Bayesian Method performing slightly better than the Collapsed Bayesian Method. Compared to the Collapsed Bayesian Method, the Pooled Bayesian Method provided a slightly larger decrease in percentages or improved results in nine of the fifteen skeletal indicator combinations for the pooled sexes, eight skeletal indicator combinations for females and only two combinations for males. However, the methods provided similar percentages in eight skeletal indicator combinations for males, which were still less than those obtained using the Complex Method. These results show that the increase in prior distribution sample size, TA sample size, and number of individuals per phase, as done for the Pooled Bayesian Method, provides better results. A balance in the number of individuals per phase is also needed for Bayesian statistics to improve results, as seen from the performance of the Collapsed Bayesian Method.

To summarise, the Bayesian Method improved age estimates, especially for males, compared to the Complex Method. The Collapsed and Pooled Bayesian Method showed that sample size and having a good spread of age-related changes in the transition analysis sample is very important. Therefore, in principle, the Bayesian approach to age estimation has merit but requires large sample sizes (Boldsen *et al*., 2002; Langley-Shirley and Jantz, 2010; Godde & Hens, 2012; Hens & Godde, 2016; Nikita *et al*., 2018).

CHAPTER 6: CONCLUSION

Most researchers agree that using multiple indicators in age estimation far outperforms using a single age indicator (Garvin *et al*., 2012). Acsádi & Nemeskéri (1970) were among the first to develop a multifactorial method, termed the 'Complex Method', that utilises four skeletal age indicators to assess age. It makes use of the pubic symphysis, endocranial sutures, and internal structural changes in the proximal femur and humerus (observable through cross-sectioning of bone or radiographs). However, the Complex Method has been criticised for its small sample size, which is not large enough to accurately represent inter-individual differences and its constant under- and overestimation of age for older and younger individuals, respectively (Saunders *et al*., 1992; Molleson *et al*., 1993). An ageing method's under- and overestimation of age has been linked to the effects of age mimicry (Aykroyd, 1997; Boldsen *et al.*, 2002; Hartnett-McCann *et al*., 2018). Bayesian statistics, which uses prior knowledge of the age-at-death distribution of the reference sample to estimate age, has been proposed to counter the effects of age mimicry and other bias often seen in adult age estimation methods (Boldsen *et al.*, 2002; Hens & Godde, 2016; Nikita *et al*., 2018; Hartnett-McCann *et al.*, 2018).

This study set out to assess the performance of the Complex Method in a black South African sample and, secondly, to test whether using Bayesian statistics can improve results. The following results were obtained in this study:

- The intra- and interobserver results were within sufficient limits for age estimation methods; however, the interobserver scores for the proximal femur and proximal humerus indicated that some experience is required when scoring using radiographs.
- The results obtained for the individual skeletal indicators' correlation with true age indicated that the pubic symphysis for females and cranial sutures performed the best.

    - Correlation results for the female sample support Acsádi & Nemeskéri's (1970) findings and the usage of the pubic symphysis and proximal femur as weighting factors.
    - The endocranial sutures performed the best of the four skeletal indicators, followed by the pubic symphysis, proximal humerus and last the proximal femur, in the male sample. This affected the performance of the weighting system in males.
- The weighting system worked better for females than males, suggesting that a sex-specific weighting system should be considered.
- The use of all four skeletal indicators did not provide the best results as reported by Acsádi & Nemeskéri (1970). Combinations utilising two or three skeletal indicators (e.g. the highest correlation was obtained by the cranial sutures and pubic symphysis for females generally performed better than the combination using all four skeletal indicators for females and males.
- The results of this study support the notion that the Complex Method under- and overestimates age for older and younger individuals, respectively. However, bias was relatively low for ages between 50 – 69 years.
- The Complex Method can be improved by developing population- and sex-specific skeletal indicator combinations and weighting systems that consider the degree of correlation found for each skeletal indicator.
- The implementation of Bayesian statistics to the Complex Method did not fully resolve the over- and underestimation of age problem. Still, some improvement in estimated ages and levels of under- and overestimation was observed. The Bayesian Method provided estimated ages closer to true age for females younger than 50 years and males younger than 60 years, compared to the Complex Method. The Bayesian Method also

provided an overall decrease in the percentages of individuals aged more than 15 years of true for all skeletal indicator combinations in males and one skeletal indicator combination (cranial sutures, proximal humerus, and proximal femur) in females.

- The Complex Method outperformed the Bayesian Method for the female sample, suggesting that there is some benefit to using the weighting system.
- From the Pooled Bayesian Method results, we see that increasing the prior distribution sample size, TA sample size and number of individuals per phase provides better results.
- From the Collapsed Bayesian Method results, we see that a balance in the number of individuals per phase is also needed for Bayesian statistics to improve results.
- Future studies should focus on developing population- and sex-specific scoring criteria and weighting systems. The use of Bayesian statistics to improve results should also be explored further, with larger prior distribution sample sizes, larger transition analysis sample sizes, and an equal number of individuals per stage or phase.

REFERENCES

Acsádi, G. Y. and Nemeskéri, J. (1970). *History of human life span and mortality*. Budapest: Akadémiai Kiadó.

Altman, D.G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

Aykroyd, R.G., Lucy, D., Pollard, A.M. and Solheim, T. (1997). Technical Note: Regression analysis in adult age estimation. *American Journal of Physical Anthropology*, 104: 259-265.

Aykroyd, R.G., Lucy, D., Pollard, A.M. and Roberts, C.A. (1999). Nasty, brutish, but not necessarily short: a reconsideration of the statistical methods used to calculate age at death from adult human skeletal and dental age indicators. *American Antiquity*, 64(1): 55-70.

Baker, R. K. (1984). *The Relationship of Cranial Suture Closure and Age Analyzed in a Modern MULTI-rACIAL Sample of Males and Females*. Fullerton, Ann Arbor: California State University, 121-121.

Bergot, C, and Bocquet, J-P. (1976). Etude systématique, en fonction de l'âge, de l'os spongieux et de l'os cortical de l'humérus et du fémur. *Bulletins et Mémoires de la Société d'anthropologie de Paris*, 3(XIII): 215-24.

Bergot, C., Wu, Y., Jolivet, E., Zhou, L.Q., Laredo, J.D. and Bousson, V. (2009). The degree and distribution of cortical bone mineralization in the human femoral shaft change with age and sex in a microradiographic study. *Bone*, 45(3): 435-442.

Bocquet, J. P., & Bergot, C. (1977). Evolution de l'os cortical de l'humérus en fonction de l'âge. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*, 4(4): 359-369.

Bocquet-Appel, J.P. and Masset, C. (1982). Farewell to paleodemography. *Journal of Human Evolution*, 11: 321-333.

Boldsen, J.L., Milner, G.R., Konigsberg, L.W. and Wood, J.W. (2002). Transition analysis: a new method for estimating age from skeletons. In: Hoppa, R.D. and Vaupel, J.W. (ed), *Paleodemography: Age distributions from skeletal samples*. Cambridge: Cambridge University Press, 37-106.

Botha, D., Lynnerup, N. and Steyn, M. (2019). Age estimation using bone mineral density in South Africans. *Forensic Science International*, 297: 307-314.

Boyd, K. L., Villa, C. and Lynnerup, N. (2015) The use of CT scans in estimating age at death by examining the extent of ectocranial suture closure. *Journal of Forensic Sciences*, 60(2): 363–369.

Brooks, S.T. (1955). Skeletal age at death: the reliability of cranial and pubic age indicators. *American Journal of Physical Anthropology*, 13(4): 567-597.

Brooks, S. and Suchey, J.M. (1990). Skeletal age determination based on the os pubis: a comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, 5(3): 227-238.

Buckberry, J. (2015). The (mis) use of adult age estimates in osteology. *Annals of human biology*, *42*(4): 323-331.

Buckberry, J.L. and Chamberlain, A.T. (2002). Age estimation from the auricular surface of the ilium: a revised method. *American Journal of Physical Anthropology*, 119(3): 231-239.

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

Chen, H., Zhou, X., Fujita, H., Onozuka, M. and Kubo, K.Y. (2013). Age-related changes in trabecular and cortical bone microstructure. *International Journal of Endocrinology*, 2013: 1-9.

Chiba, F., Makino, Y., Motomura, A., Inokuchi, G., Torimitsu, S., Ishii, N. and Hayakawa, M. (2013). Age estimation by multidetector CT images of the sagittal suture. *International Journal of Legal Medicine*, 127(5): 1005-1011.

Christensen, A. M., Passalacqua, N. V. and Bartelink, E. J. (2014). *Forensic anthropology: current methods and practice.* Oxford: Academic Press.

Cooper, D.M.L., Thomas, C.D.L., Clement, J.G., Turinsky, A.L., Sensen, C.W. and Hallgrímsson, B. (2007). Age-dependent change in the 3D structure of cortical porosity at the human femoral midshaft. *Bone*, 40(4): 957-965.

Cui, W.Q., Won, Y.Y., Baek, M.H., Lee, D.H., Chung, Y.S., Hur, J.H. and Ma, Y.Z. (2008). Age and region-dependent changes in three-dimensional microstructural properties of proximal femoral trabeculae. *Osteoporosis International*, 19(11): 1579-1587.

Dayal, M.R., Kegley, A., Bidmos, M.A., Strkalj, G. and Kuykendall, K.L. (2009). The history and composition of the Raymond A. Dart collection of human skeletons at the University of the Witwatersrand, Johannesburg, South Africa. *American Journal of Physical Anthropology*, 140(2): 324-335.

Daniels, E.D., Pettifor, J.M., Schnitzler, C.M., Moodley, G.P. and Zachen, D. (1997). Differences in mineral homeostasis, volumetric bone mass and femoral neck axis length in black and white South African women. *Osteoporosis International*, 7(2): 105-112.

Daniels, E.D., Pettifor, J.M., Schnitzler, C.M., Russell, S.W. and Patel, D.N. (1995). Ethnic differences in bone density in female South African nurses. *Journal of Bone and Mineral Research*, 10(3): 359-367.

Djurić, M., Djonić, D., Nikolić, S., Popović, D. and Marinković, J. (2007). Evaluation of the Suchey– Brooks method for aging skeletons in the Balkans. *Journal of Forensic Sciences*, 52(1): 21-23.

Galera, V., Ubelaker, D.H. and Hayek, L.A. (1998). Comparison of macroscopic cranial methods of age estimation applied to skeletons from the Terry Collection. *Journal of Forensic Science*, 43(5): 933-939.

Garvin, H. M., Passalacqua, N. V., Uhl, N. M., Gipson, D. R., Overbury, R. S. and Cabo, L. L. (2012). Adult sex determination: methods and application. In: Dirkmaat, D. C. (ed.), *A companion to forensic anthropology.* West Sussex: Blackwell Publishing Ltd, 202-223.

Gilbert, B.M. and McKern, T.W. (1973). A method for aging the female os pubis. *American Journal of Physical Anthropology*, 38(1): 31-38.

Godde, K. and Hens, S.M. (2012). Age-at-death estimation in an Italian historical sample: a test of the Suchey-Brooks and Transition Analysis methods. *American Journal of Physical Anthropology*, 194: 259-265.

Grabherr, S., Cooper, C., Ulrich-Bochsler, S., Uldin, T., Ross, S., Oesterhelweg, L., and Thali, M. J. (2009). Estimation of sex and age of "virtual skeletons"–a feasibility study. *European Radiology*, 19(2): 419-429.

Hanihara, K. (1952). On the age changes in the male Japanese pubic bone. *Journal of the Anthropological Society of Nippon*, 62: 245-260.

Hanihara, K. and Suzuki, T. (1978). Estimation of age from the pubic symphysis by means of multiple regression analysis. *American Journal of Physical Anthropology*, 48(2): 233-239.

Hansen, G. (1953-54). *Die Altersbestimmung am proximalen Humerus- und Femurende im Rahmen der Identifizierung menschlicher Skelettreste.* Wiss. Z. Humboldt-Universitat Berlin, Mat.-Natunv. Reihe 3t1-73.

Hartnett, K.M. (2010). Analysis of Age-at-Death Estimation Using Data from a New, Modern Autopsy Sample—Part I: Pubic Bone. *Journal of Forensic Sciences*, 55(5), 1145–1151.

Hartnett-McCann, K., Fulginiti, L.C. and Seidel, A.C. (2018). Adult Age-at-Death Estimation in Unknown Decedents: New Perspectives on an Old Problem. In: Latham, K.E., Bartelink, E.J. and Finnegan, M. (ed.), *New Perspectives in Forensic Human Skeletal Identification.* London: Academic Press, 65-85.

Hawley, N.L., Rousham, E.K., Norris, S.A., Pettifor, J.M. and Cameron, N. (2009). Secular trends in skeletal maturity in South Africa: 1962–2001. Annals of Human Biology, 36(5), 584-594.

Henneberg, M. and Van den Berg, E.R. (1990). Test of socioeconomic causation of secular trend: stature changes among favored and oppressed South Africans are parallel. American Journal of Physical Anthropology, 83(4), 459-465.

Hens, S.M. and Godde, K. (2016). Auricular surface aging: comparing two methods that assess morphological change in the ilium with Bayesian analyses. *Journal of Forensic Science*, 61: S30 S38.

Hens, S. M., Rastelli, E. and Belcastro, G. (2008). Age estimation from the human os coxa: a test on a documented Italian collection. *Journal of forensic sciences*, 53(5): 1040-1043.

Hershkovitz, I., Latimer, B., Dutour, O., Jellema, L.M., Wish-Baratz, S., Rothchild, C and Rothchild, B.M. (1997). Why do we fail in aging the skull from the sagittal suture? *American Journal of Physical Anthropology*, 103: 393-399.

Hoppa, R.D. (2000). Population variation in osteological aging criteria: An example from the pubic symphysis. *American journal of physical anthropology*, 111(2): 185–191.

Hoppa, R.D. and Vaupel, J.W. (2002). The Rostock Manifesto for paleodemography: the way from stage to age. In: Hoppa, R.D. and Vaupel, J.W. (ed.), *Paleodemography: Age Distributions from Skeletal Samples*. Cambridge: Cambridge University Press, 1-9.

İşcan, M. Y. and Steyn, M. (2013). *The Human Skeleton in Forensic Medicine*. (3). Illinois: Charles C Thomas Publisher Ltd.

Ito, M., Nakata, T., Nishida, A. and Uetani, M. (2011). Age-related changes in bone density, geometry and biomechanical properties of the proximal femur: CT-based 3D hip structure analysis in normal postmenopausal women. *Bone*, 48(3): 627-630.

Jooste, N., L'Abbé, E.N., Pretorius, S. and Steyn, M. (2016). Validation of transition analysis as a method of adult age estimation in a modern South African sample. *Forensic Science International*. 266, 580.e1–580.e7.

Joubert, L.C., Briers, N. and Meyer, A. (2019). Evaluation of the Suchey–Brooks age estimation method in a white South African population. *Forensic Science International*, 301: e14-e19.

Katz, D. and Suchey, J.M. (1986). Age determination of the male os pubis. *American Journal of Physical Anthropology*, 69(4): 427-435.

Kemkes-Grottenhaler, A. (2002). Historical perspectives on age indicator methods. In: Hoppa, R. D. and Vaupel, J. W. (ed.), *Paleodemography: Age Distributions from Skeletal Samples*. Cambridge: Cambridge University Press, 48-72.

Key, C.A., Aiello, L.C. and Molleson, T. (1994). Cranial suture closure and its implication for age estimation. *International Journal of Osteoarchaeology*, 4: 193-207.

Klepinger, L.L., Katz, D., Micozzi, M.S. and Carroll, L. (1992). Evaluation of cast methods for estimating age from the os pubis. *Journal of Forensic Science*, 37(3): 763-770.

Konigsberg, L.W. and Frankenberg, S.R. (1992). Estimation of age structure in anthropological demography. *American Journal of Physical Anthropology*, 89(2): 235-256.

Krogman, W.M. (1949). The human skeleton in legal medicine: Medical aspects. In: *Symposium on medicolegal problems*. Philadelphia: Lippincott, 1-92.

Langley-Shirley, N. and Jantz, R.L. (2010). A Bayesian approach to age estimation in modern Americans from the clavicle. *Journal of Forensic Science*, 55(3): 571-583.

Lochmüller, E.M., Matsuura, M., Bauer, J., Hitzl, W., Link, T.M., Müller, R. and Eckstein, F. (2008). Site-specific deterioration of trabecular bone architecture in men and women with advancing age. *Journal of Bone and Mineral Research*, 23(12): 1964-1973.

Lovejoy, C.O., Meindl, R.S., Mensforth, R.P. and Barton T.J. (1985a). Multifactorial age determination of skeletal age at death: a method and blind test of its accuracy. *American Journal of Physical Anthropology*, 68: 1-14.

Lovejoy, C.O., Meindl, R.S., Pryzbeck, T.R. and Mensforth, R.P. (1985b). Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. *American Journal of Physical Anthropology*, 68(1): 15-28.

Lynnerup, N. and Jacobsen, J.C.B. (2003). Brief communication: age and fractal dimensions of human sagittal and coronal sutures. *American Journal of Physical Anthropology*, 121(4): 332-336.

Lynnerup, N., Thomsen, I. and Frohlich, B. (1990). A non-invasive technique for age at death determination. *Medicine, Science and the Law*, 30(4): 317-320.

Maat, G.J.R. (1987). Practising methods of age determination. Comments on methods combining multiple age indicators. *International Journal of Anthropology*, 2(4): 293-299.

Macchiarelli, R. and Bondioli, L. (1994). Linear densitometry and digital image processing of proximal femur radiographs: implications for archaeological and forensic anthropology. *American Journal of Physical Anthropology*, 93(1): 109-122.

Macdonald, H.M., Nishiyama, K.K., Hanley, D.A. and Boyd, S.K. (2011). Changes in trabecular and cortical bone microarchitecture at peripheral sites associated with 18 months of teriparatide therapy in postmenopausal women with osteoporosis. *Osteoporosis International*, 22(1): 357-362.

Mann, R. W., Jantz, R. L., Bass, W. M. and Willey, P. S. (1991). Maxillary suture obliteration: A visual method for estimating skeletal age. *Journal of forensic sciences,* 36: 781-791.

Martin, R. and Saller, K. (1957-1966). *Lehrbuch der Anthropologie in systematischer Darstellung*. Fischer: Stuttgart.

Martins, R., Paulo Eduardo Oliveira and Schmitt, A. (2012). Estimation of age at death from the pubic symphysis and the auricular surface of the ilium using a smoothing procedure. *Forensic Science International*, 219(1-3), 287.e1–287.e7.

Masset, C. (1989). Age estimation on the basis of cranial sutures. In: İşcan, M. Y. (ed). *Age Markers in the Human Skeleton*. Springfield Illinois: CC Thomas, 71-103.

Mays, S. (2015). The effect of factors other than age upon skeletal age indicators in the adult. *Annals of Human Biology*, 42(4): 332–341.

Meyer, A., van der Merwe, A. and Steyn, M. (2021). An evaluation of the Acsádi and Nemeskéri Complex Method of adult age estimation in a modern South African skeletal sample. *Forensic Science International*, 321: 110740.

McKern, T.W. and Stewart, T.D. (1957). *Skeletal age changes in young American males: analysed from the standpoint of age identification* (Vol. 45). Headquarters, Quartermaster Research & Development Command.

Meindl, R.S. and Lovejoy, C.O. (1985). Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral anterior suture. *American Journal of Physical Anthropology*, 68: 57-66.

Meindl, R.S., Lovejoy, C.O., Mensforth, R.P. and Walker, R.A. (1985). A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. *American Journal of Physical Anthropology*, 68(1): 29-45.

Milner, G. R. and Boldsen, J. L. (2012). Skeletal age estimation: Where we are and where we should go. In: Dirkmaat, D. C. (ed.), *A companion to forensic anthropology*. West Sussex: Blackwell Publishing Ltd, 224-238.

Molleson, T., Cox, M., Waldron, A. H. and Whittaker, D.K. (1993). *The Spitalfields Report. Volume 2. The Anthropology: The Middling Sort.* CBA Research Report 86. Council for British Archaeology, York.

Mukaka, M.M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3): 69-71.

Nawrocki, S.P. (1998). Regression formulae for estimating age at death from cranial suture closure. *Forensic Osteology: Advances in the Identification of Human Remains*. Springfield. IL: Charles C Thomas.

Nemeskéri, J., Harsányi, L. and Acsádi, G.Y. (1960). Methoden zur diagnose des lebensalters von skellett funden. *Anthropologischer Anzeiger*, 24: 70-95.

Nikita, E., and Nikitas, P. (2019). Skeletal age-at-death estimation: Bayesian versus regression methods. *Forensic Science International*, 297: 56–64.

Nikita, E., Xanthopoulou, P. and Kranioti, E. (2018). An evaluation of Bayesian age estimation using the auricular surface in modern Greek material. *Forensic Science International: Genetics*, 291: 1–11.

Nikolova, S., Toneva, D., Georgiev, I. and Lazarov, N. (2019). Sagittal suture maturation: Morphological reorganization, relation to aging, and reliability as an age-at-death indicator. *American Journal of Physical Anthropology*, 169(1): 78-92.

Obert, M., Schulte-Geers, C., Schilling, R.L., Harth, S., Kläver, M., Traupe, H. and Verhoff, M.A. (2010). High-resolution flat-panel volumetric CT images show no correlation between human age and sagittal suture obliteration—independent of sex. *Forensic Science International*, 200(1-3): 180.e1–180.e12.

Passalacqua, N.V. (2010). The utility of the Samworth and Gowland age-at-death "look-up" tables in forensic anthropology. *Journal of Forensic Sciences*, 55(2): 482-487.

Ponsold, A. (1950). *Lehrbuch der Gerichtlichen Medizin*. Stuttgart: Thieme.

Purves, S., Woodley, L. and Hackman, L., 2011. Age determination in the adult. *Forensic Anthropology—2000 to 2010*, 29-59.

Riggs, B.L., Melton III, L.J., Robb, R.A., Camp, J.J., Atkinson, E.J., Peterson, J.M. and Khosla, S. (2004). Population-based study of age and sex differences in bone volumetric density, size, geometry, and structure at different skeletal sites. *Journal of Bone and Mineral Research*, 19(12): 1945-1954.

Roosa, S.M., Hurd, A.L., Xu, H., Fuchs, R.K. and Warden, S.J. (2012). Age-related changes in proximal humerus bone health in healthy, white males. *Osteoporosis International*, 23(12): 2775- 2783.

Samworth, R. and Gowland, R. (2007). Estimation of adult skeletal age-at-death: statistical assumptions and applications. *International Journal of Osteoarchaeology*, 17(2): 174-188.

San-Millán, M., Carme Rissech and Turbón, D. (2016). New approach to age estimation of male and female adult skeletons based on the morphological characteristics of the acetabulum. *International Journal of Legal Medicine*, 131(2), 501–525.

Saunders, S.R., Fitzgerald, C., Roger, T., Dudar C. and McKillop, H. (1992). Test of several methods of skeletal age estimation using a documented archaeological sample. *Canadian Society of Forensic Science Journal*, 25: 97-118.

Schmitt, A. (2004) Age-at-death assessment using the os pubis and the auricular surface of the illum: a test on an identified Asian sample. *International Journal of Osteoarchaeology*, 14: 1-6.

Schmitt, A., Murali, P., Cunha, E. and Rouge, D. (2002). Variability of the pattern of aging on the human skeleton: Evidence from bone indicators and implications on age at death estimation. *Journal of forensic science,* 47(6): 1203-1209.

Schranz, D. (1959). Age determination from the internal structure of the humerus. *American Journal of Physical Anthropology*, 17: 273-277.

Singer, R. (1953). Estimation of age from cranial suture closure. *Journal of Forensic Medicine*, 1: 52-59.

Szilvàssy, J. and Kritscher, H. (1990). Estimation of chronological age in man based on the spongy structure of long bones. *Anthropologischer Anzeiger*: 289-298.

Tersigni-Tarrant, M. A. and Shirley, N. R. (2013). *Forensic Anthropology An Introduction.* New York: CRC Press.

Testut, L. and Latarjet, A., 1948. *Traité d'anatomie humaineé*. G. Doin et Cie.

Tingart, M.J., Apreleva, M., von Stechow, D., Zurakowski, D. and Warner, J.J.P. (2003). The cortical thickness of the proximal humeral diaphysis predicts bone mineral density of the proximal humerus. *The Journal of Bone and Joint Surgery. British volume*, 85(4): 611-617.

Todd, T.W. (1920) Age changes in the pubic bone. I. The male white pubis. *American Journal of Physical Anthropology*, 3: 285-339.

Todd, T.W. (1921a-c). Age changes in the pubic bone. *American Journal of Physical Anthropology*, 4(1): 1-70.

Todd, T.W. (1923). Age changes in the pubic symphysis: VII. The anthropoid strain in human pubic symphyses of the third decade. *Journal of Anatomy*, 57(Pt 3): 274.

Todd, T.W. (1930). Age changes in the pubic bone. VIII. Roentgenographic differentiation. *American Journal of Physical Anthropology*, 14(2): 255-271.

Todd, T.W. and Lyon, D.W. (1924). Endocranial suture closure. Its progress and age relationship. Part I.—Adult males of white stock. *American Journal of Physical Anthropology*, 7(3): 325-384.

Todd, T.W. and Lyon, D.W. (1925a). Cranial suture closure. Its progress and age relationship. Part II. —Ectocranial closure in adult males of white stock. *American Journal of Physical Anthropology*, 8(1): 23-45.

Todd, T.W. and Lyon, D.W. (1925b). Cranial suture closure. Its progress and age relationship. Part III. —Endocranial closure in adult males of Negro stock. *American Journal of Physical Anthropology*, 8(1): 47-71.

Todd, T.W. and Lyon, D.W. (1925c). Suture closure—Its progress and age relationship. Part IV. —Ectocranial closure in adult males of Negro stock. *American Journal of Physical Anthropology*, 8(2):149-168.

Uhl, N.M. and Nawrocki, S.P. (2010). Multifactorial estimation of age at death from the human skeleton. In: Latham, K.E. and Finnegan, M. (ed). *Age estimation of the human skeleton*. Springfield: Charles C. Thomas, pp. 243-261.

Von Meyer, G.V. (1867). Die architektur der spongiosa. Archisfur Anatomie, Physiologie und wissenschaftliche Medizin. *Reichert und DuBois-Reymonds Archiv*, 34: 615-628.

Wachholz, L. (1894). O oznaczaniu wieku ze zwłok na podstawie kostnienia główki kości ramieniowej. *Akademia Umiejętności*.

Walker, R.A. and Lovejoy, C.O. (1985). Radiographic changes in the clavicle and proximal femur and their use in the determination of skeletal age at death. *American Journal of Physical Anthropology*, 68(1): 67-78.

Zambrano, C.J., 2005. *Evaluation of regression equations used to estimate age at death from cranial suture closure*. MS thesis, University of Indianapolis, Indianapolis, IN.

APPENDIX A



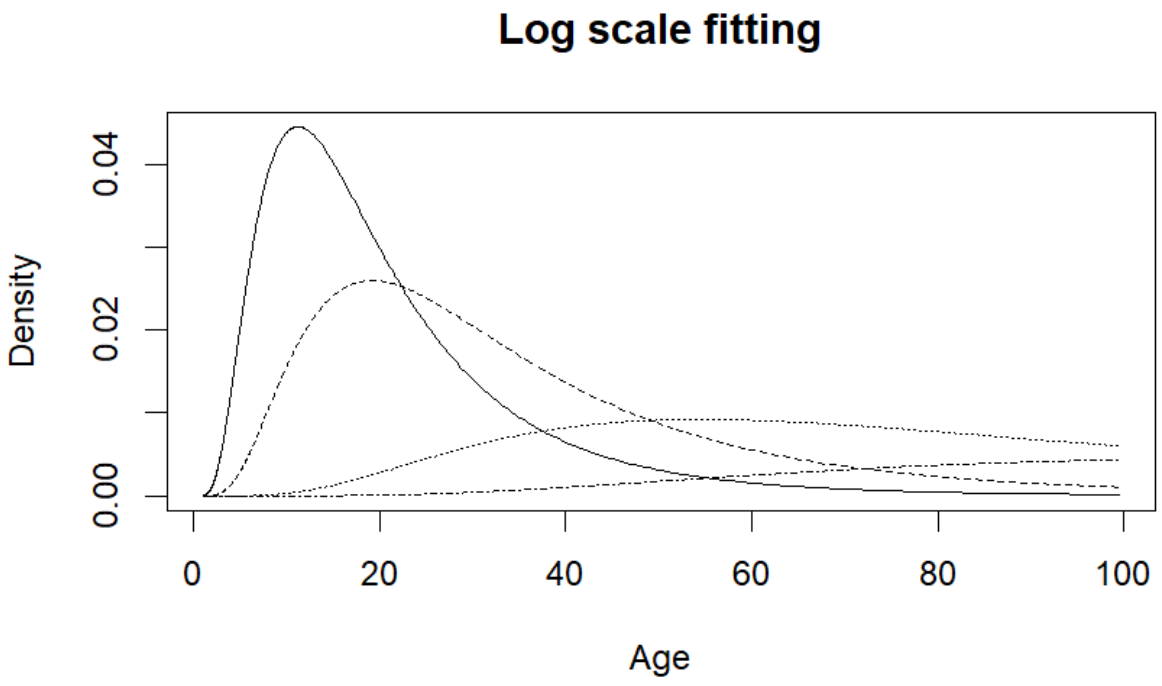
Gompertz plot 1: Female survival curve (dashed lines) with Gompertz hazard curve (solid line).



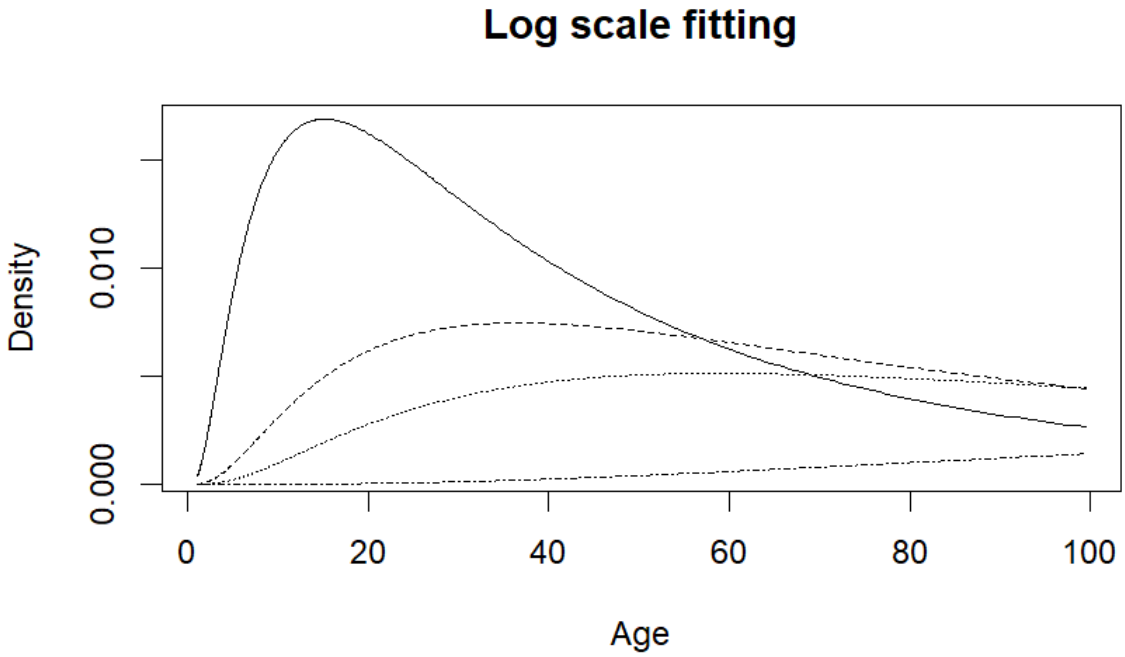
Gompertz plot 2: Male survival curve (dashed lines) with Gompertz hazard curve (solid line).



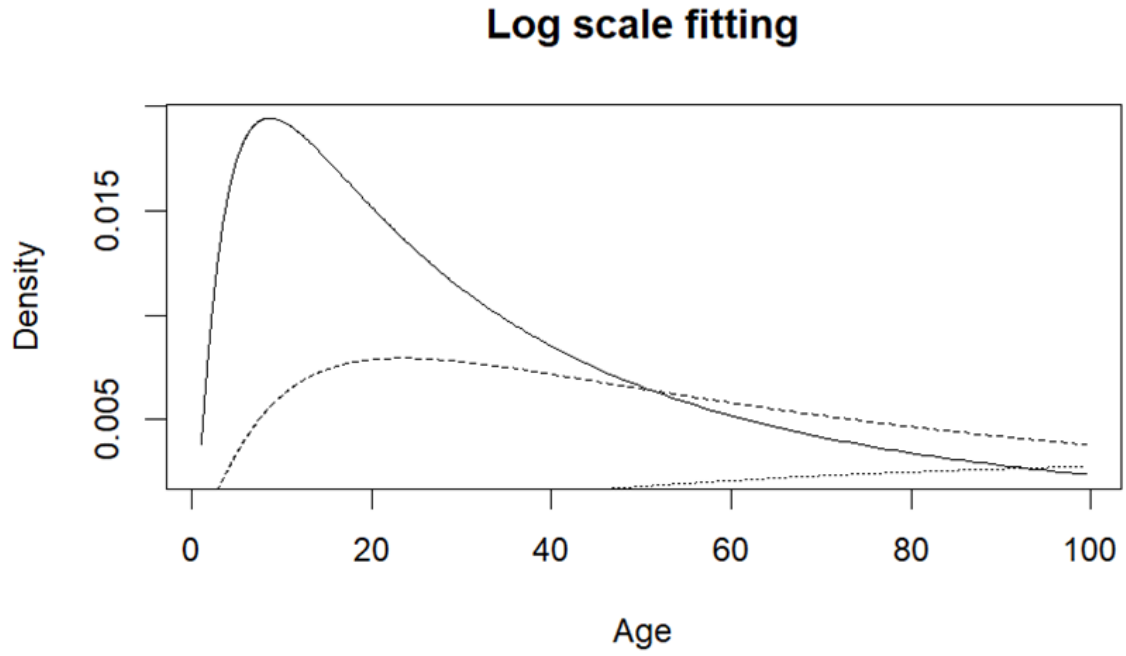
Gompertz plot 3: Pooled sexes survival curve (dashed lines) with Gompertz hazard curve (solid line).



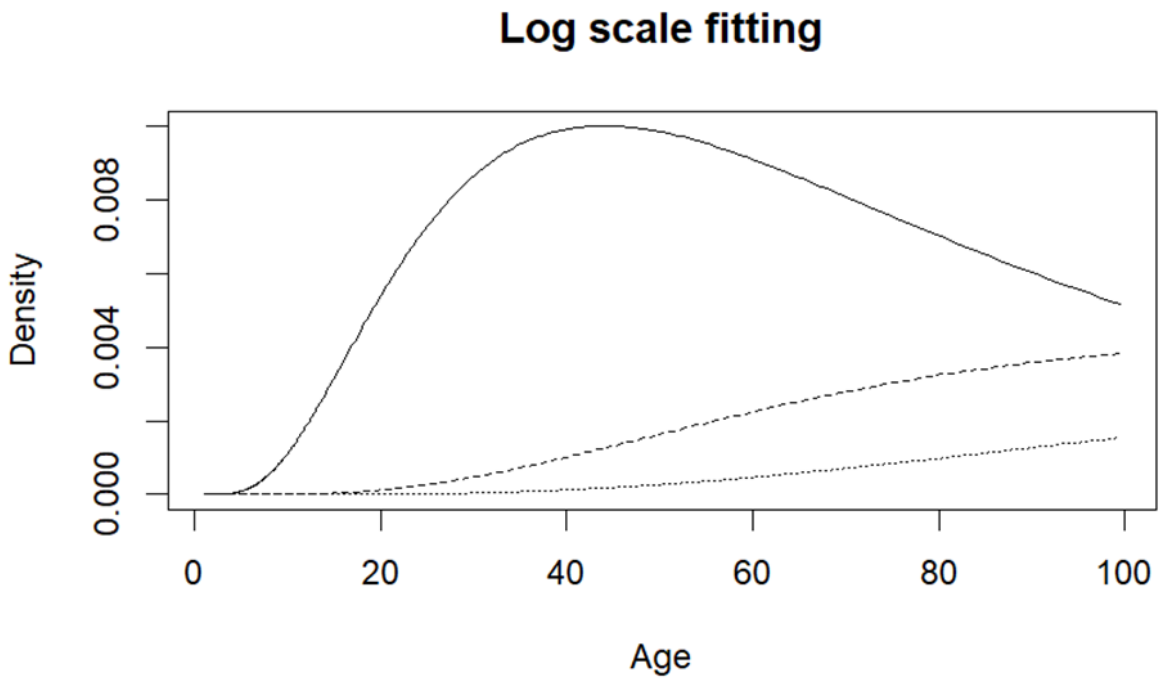
Transition curve 1: Female cranial suture.



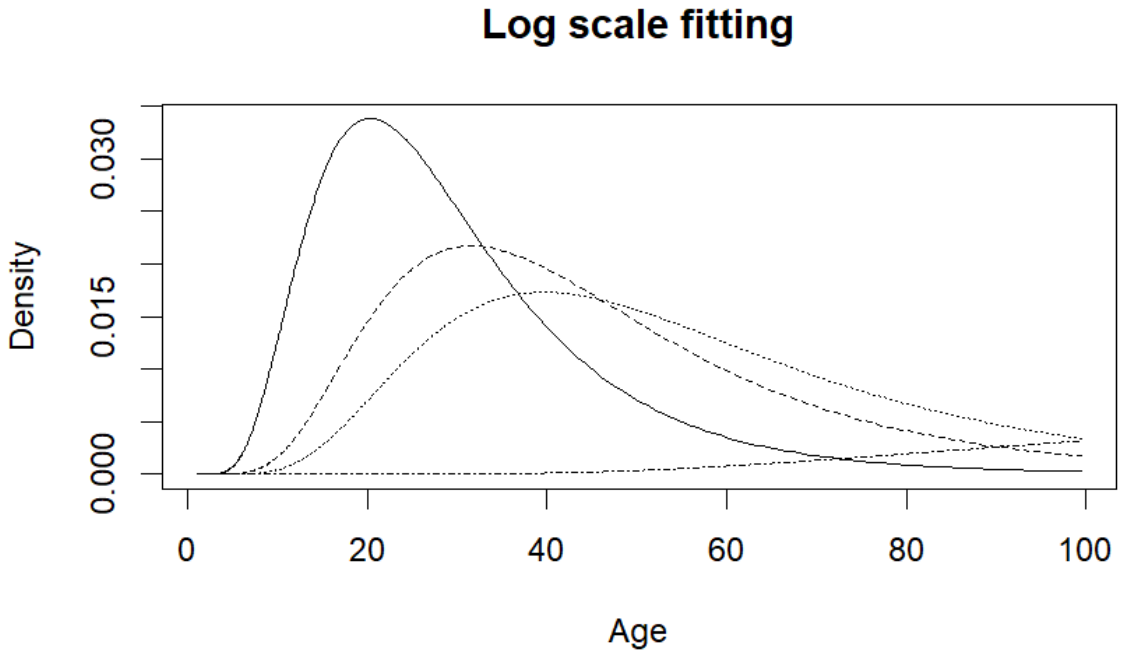
Transition curve 2: Female pubic symphysis.



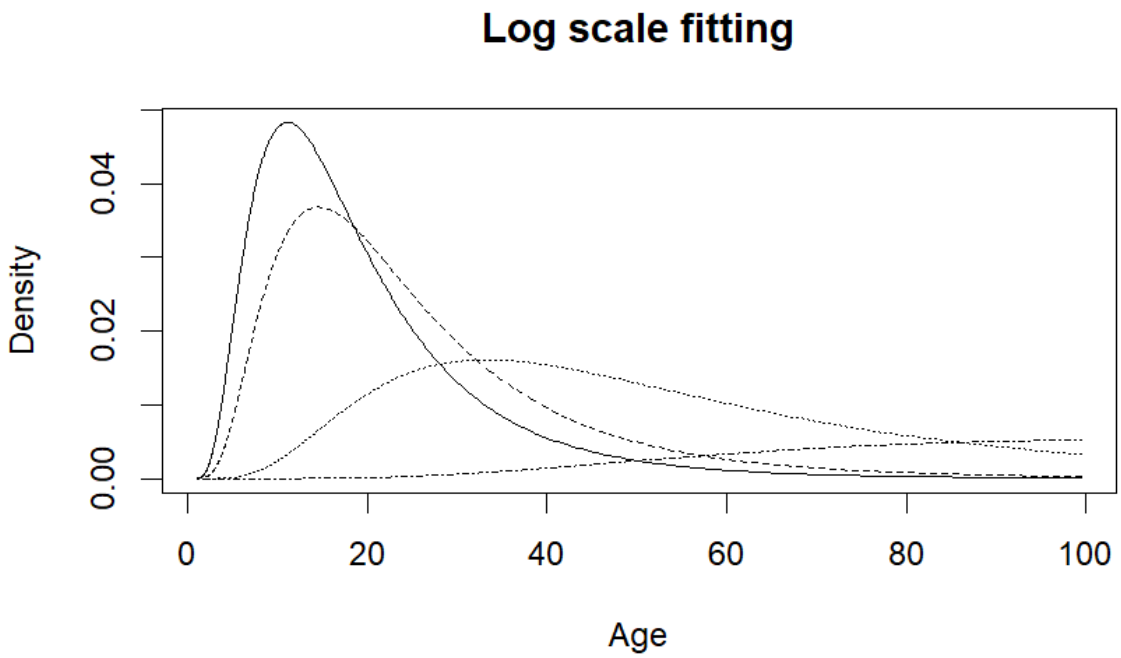
Transition 3: Female proximal humerus.



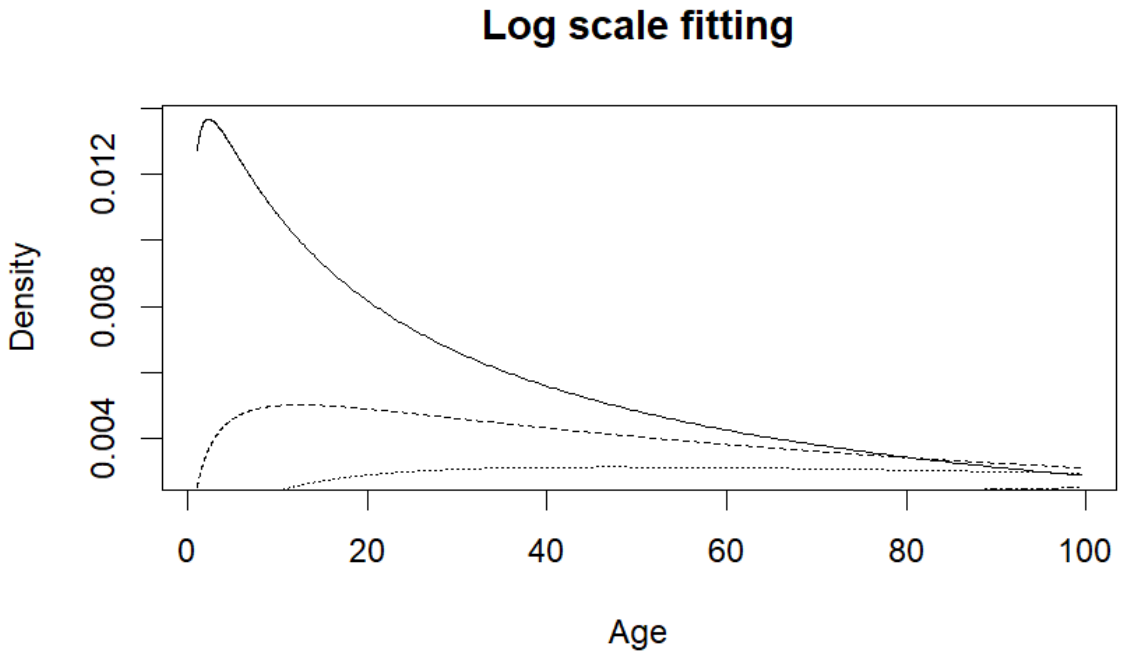
Transition curve 4: Female proximal femur.



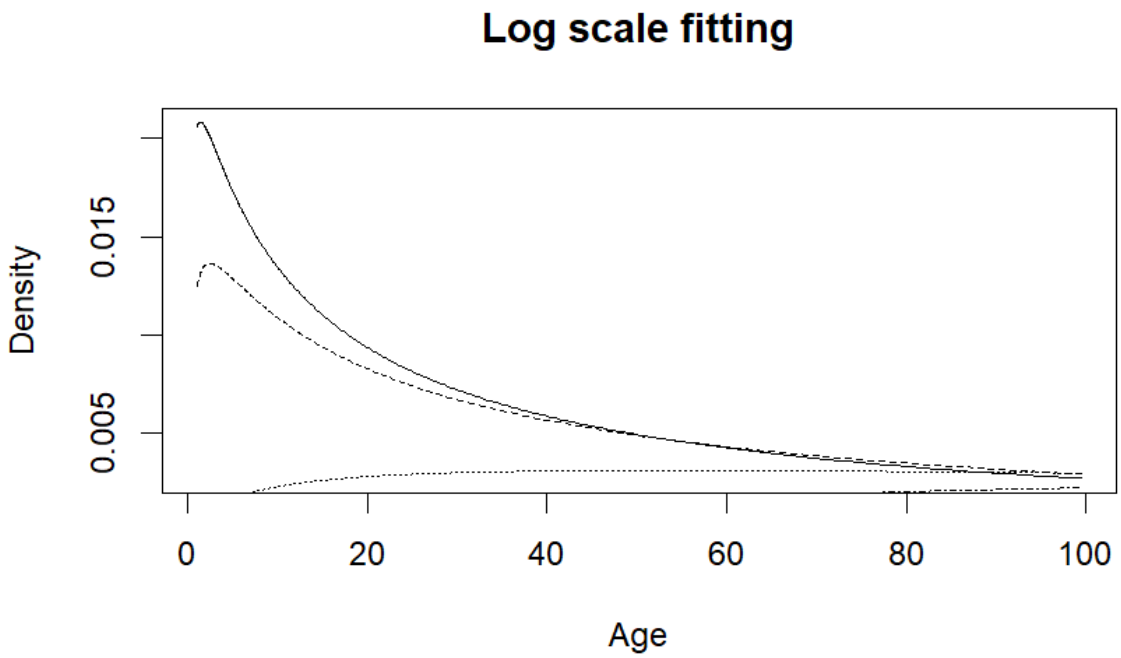
Transition curve 5: Male cranial suture.



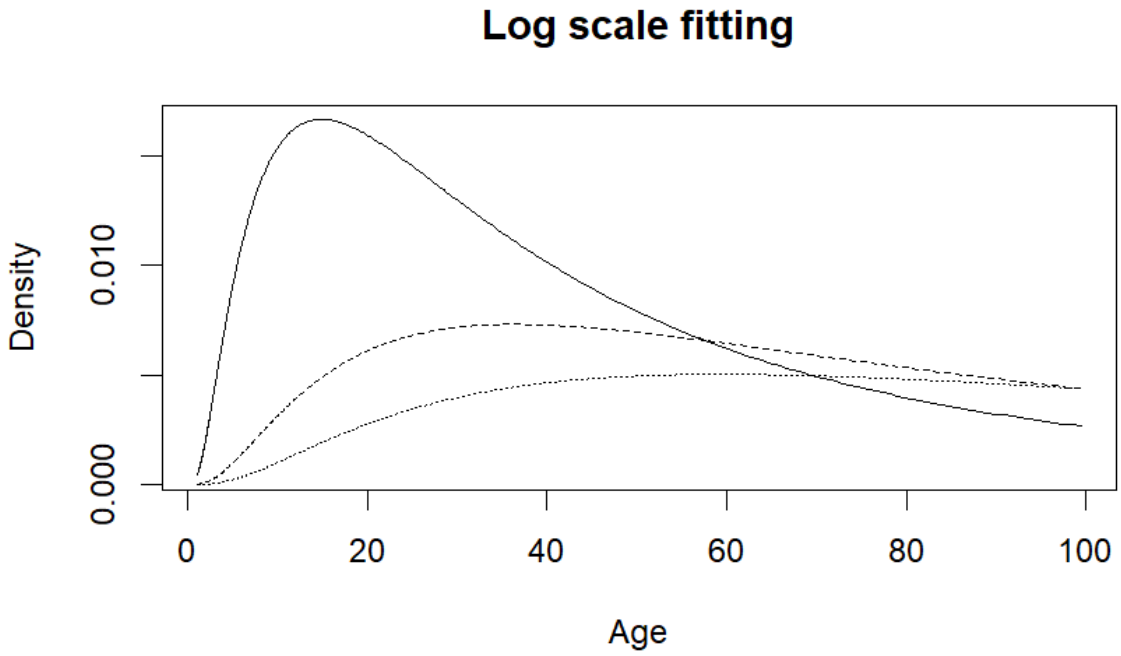
Transition curve 6: Male pubic symphysis.



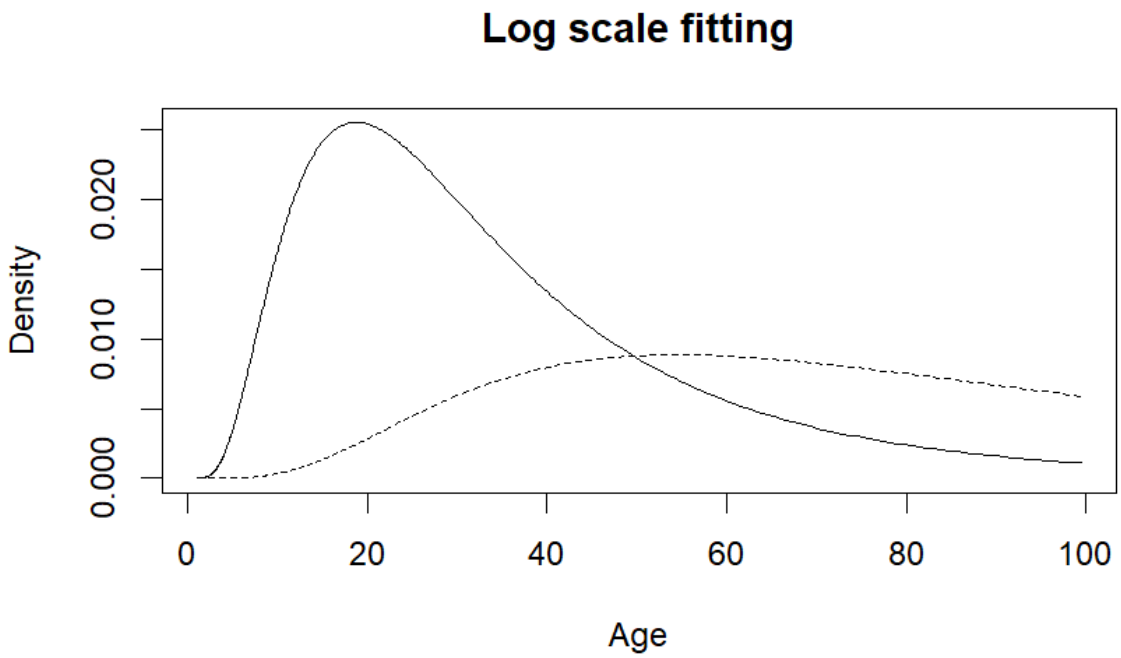
Transition curve 7: Male proximal humerus.



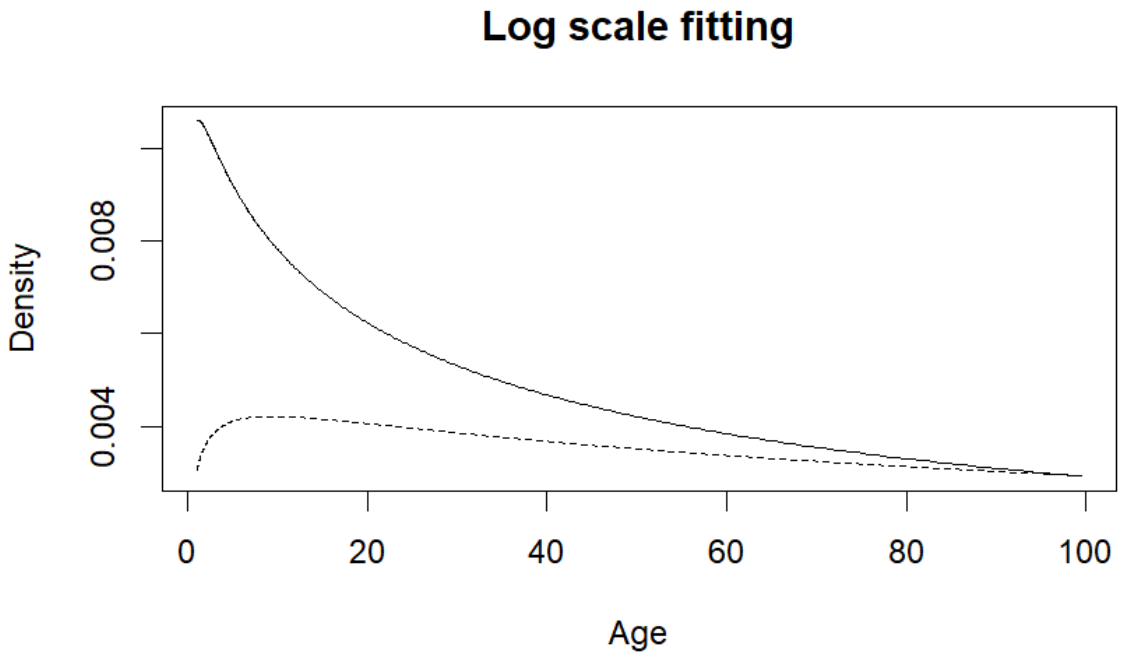
Transition curve 8: Male proximal femur.



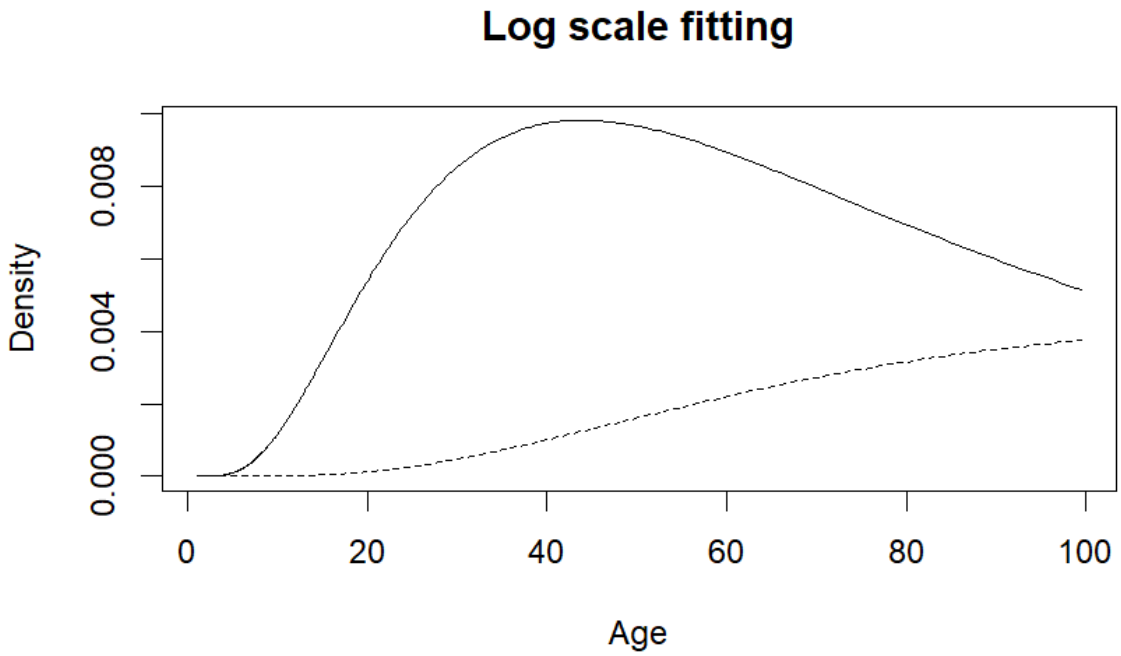
Transition curve 9: Collapsed Bayesian Method female cranial sutures.



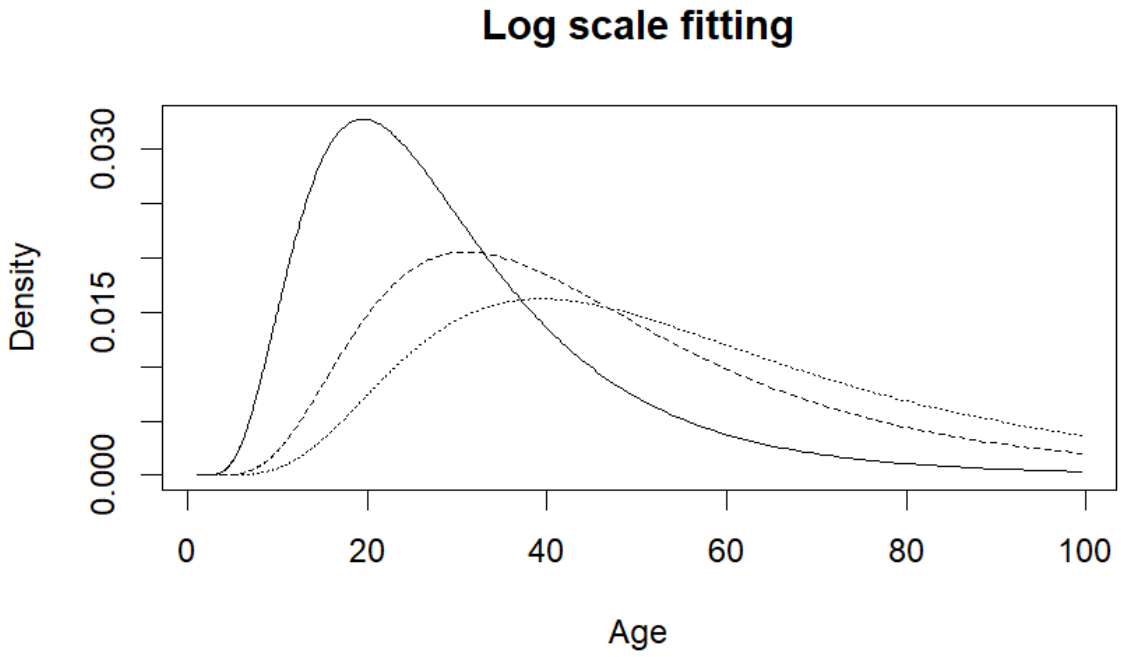
Transition curve 10: Collapsed Bayesian Method female pubic symphysis.



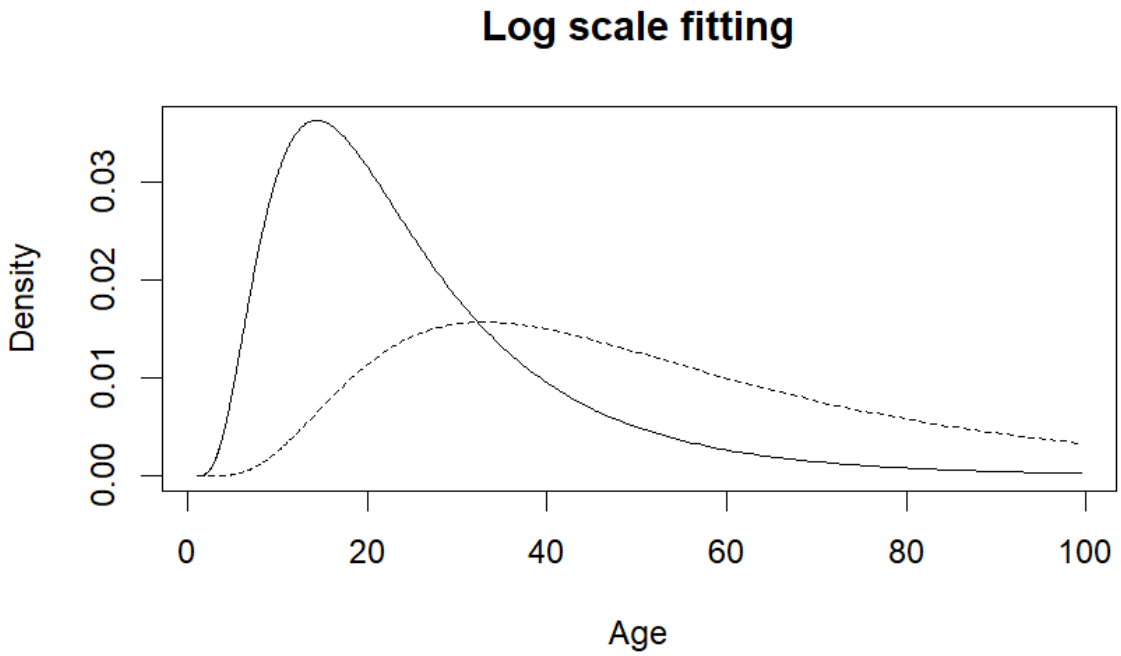
Transition curve 11: Collapsed Bayesian Method female proximal humerus.



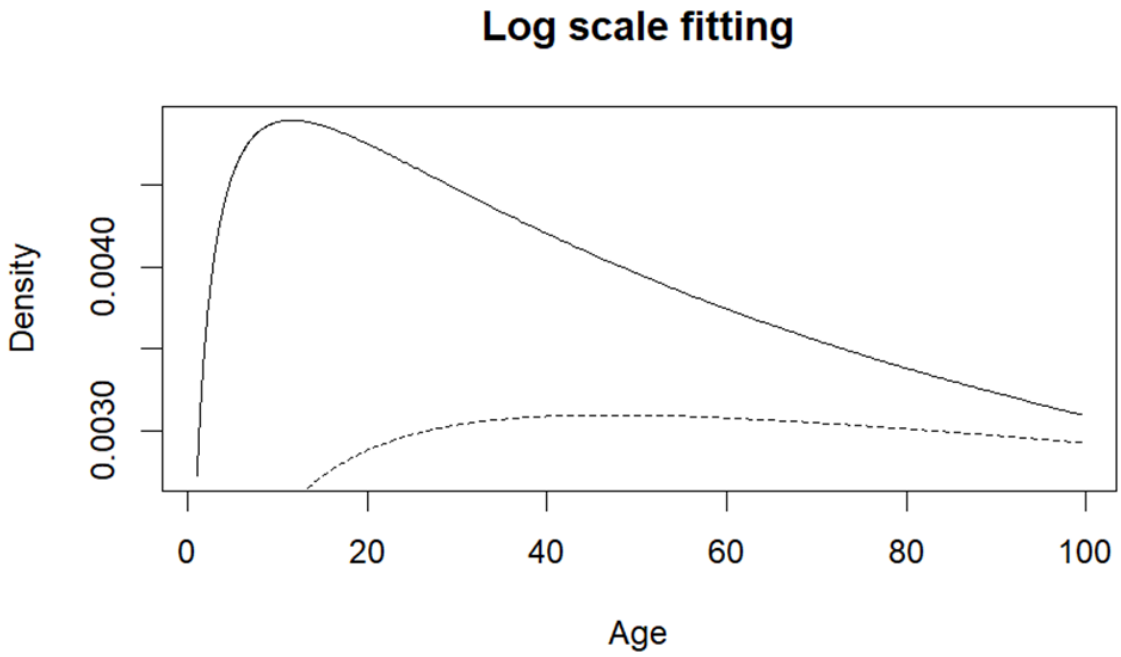
Transition curve 12: Collapsed Bayesian Method female proximal femur.



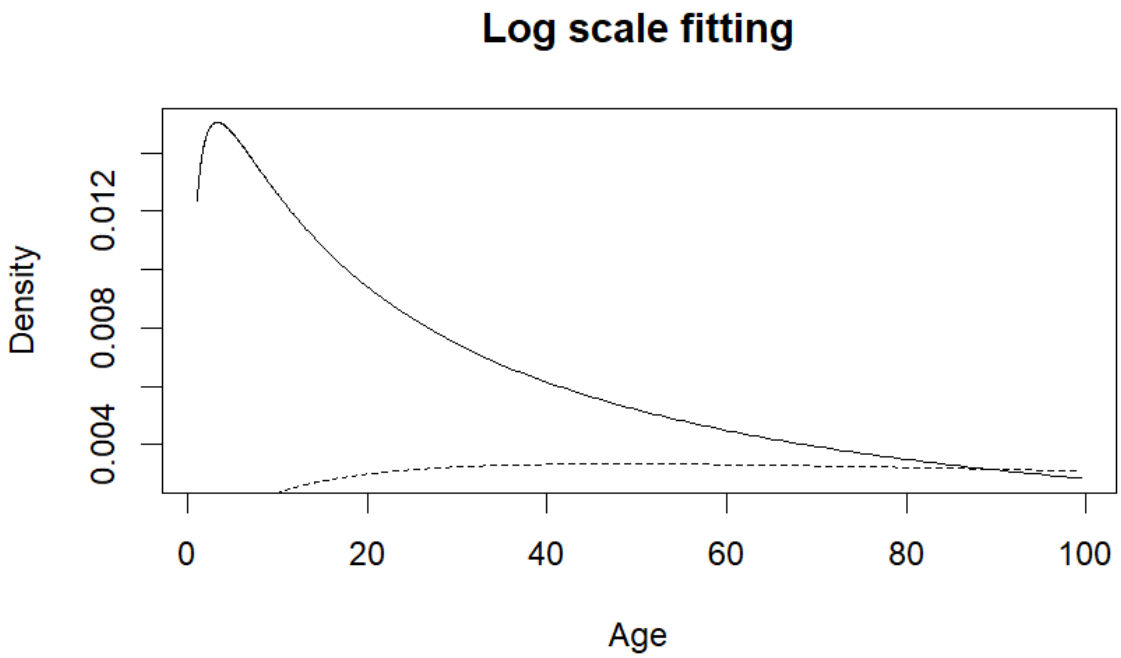
Transition curve 13: Collapsed Bayesian Method male cranial sutures.



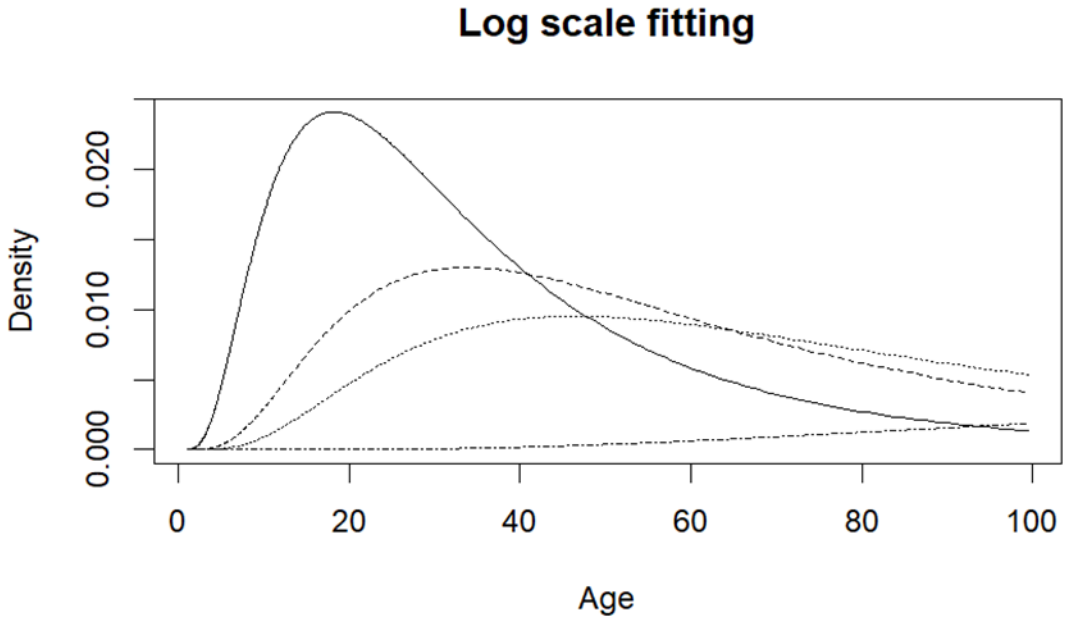
Transition curve 14: Collapsed Bayesian Method male pubic symphysis.



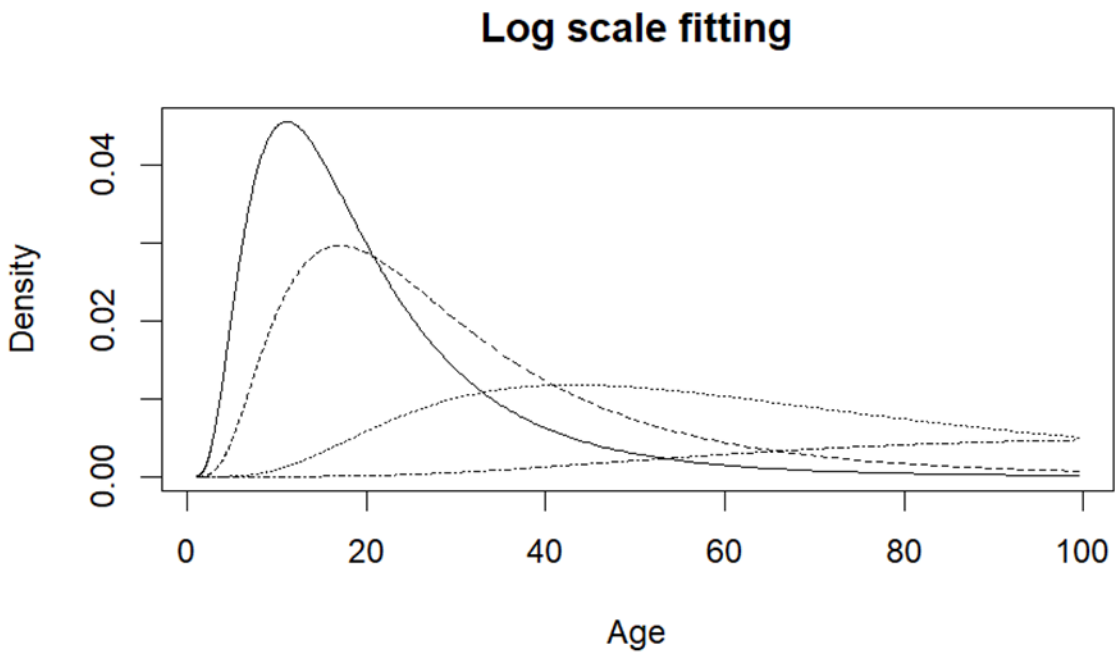
Transition curve 15: Collapsed Bayesian Method male proximal humerus.



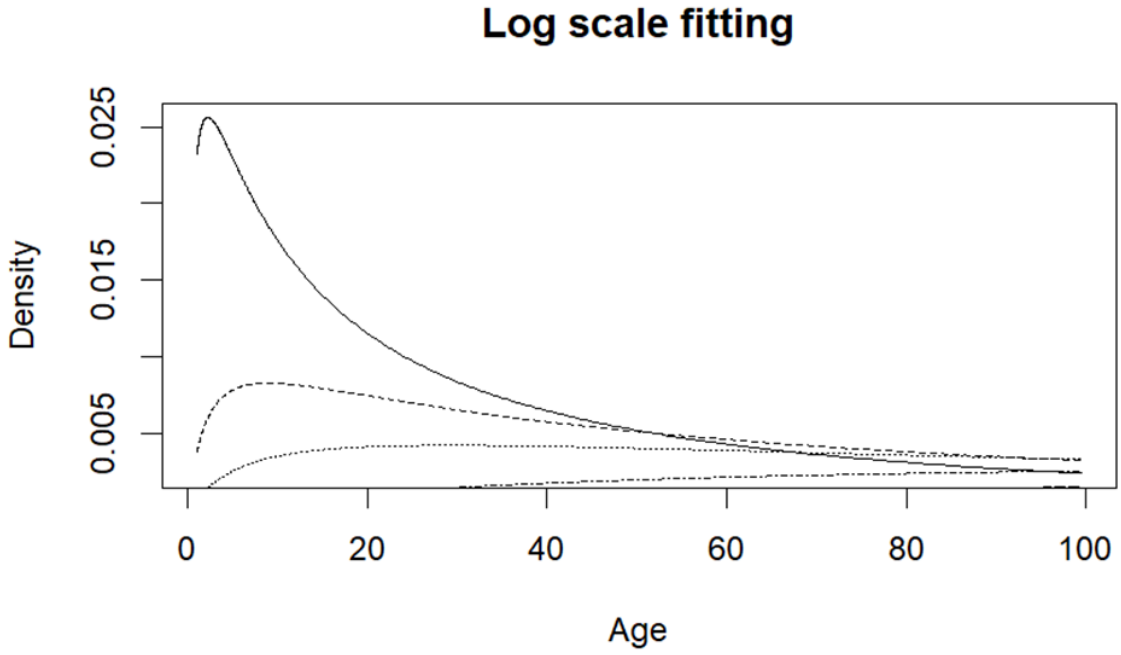
Transition curve 16: Collapsed Bayesian Method male proximal femur.



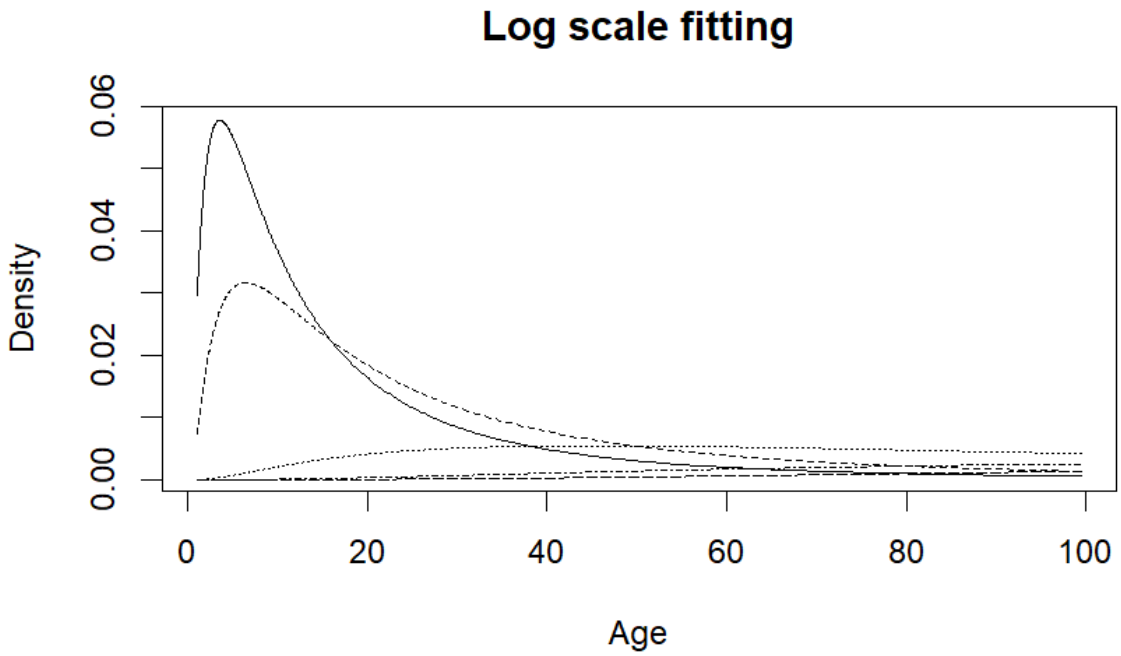
Transition curve 17: Pooled Bayesian Method pooled sexes' cranial sutures.



Transition curve 18: Pooled Bayesian Method pooled sexes' pubic symphysis.



Transition curve 19: Pooled Bayesian Method pooled sexes' proximal humerus.



Transition curve 20: Pooled Bayesian Method pooled sexes' proximal femur.

APPENDIX B

Gompertz parameters

```
GompertzFit = function(method="nlminb", maxit=1000, lower=0.001, upper=0.1, pa3=0.01,
pb3=0.05){
x= read.table(choose.files(caption = 'Select the file containing the ages of the informative
prior'), head=1, sep="\t")
ages=as.matrix(x); shift=min(ages)
library(survival); library(optimx)

Gf = function(x){
   a3= x[1]; b3= x[2]
   t =ages; nr=nrow(ages)
   s = 0
   for(i in 1:nr){
      G  = a3*exp(b3*(t[i]-shift))*exp(a3/b3*(1-exp(b3*(t[i]-shift))))
      s=s+log(G)}
   return(s)}

r=optimx(c(pa3,pb3), Gf,method=method, lower=lower,upper=upper, hessian=T,
control=list(maxit=maxit, maximize = TRUE))
a3= r[,1]; b3= r[,2]

plot(survfit(Surv(ages)~1), xlim=c(1,100), main='Gompertz plot', xlab='Age',
ylab='Survivorship')
hi=matrix(0, (100-shift+1))
for(i in 1:(100-shift+1)){hi[i] = exp(a3/b3*(1-exp(b3*(i-1))))}
lines(shift:100,hi,lwd=2)
cat("\n"); cat("\n")
cat(noquote("RESULTS:")); cat("\n"); cat("\n")
cat(noquote("shift = ")); cat(shift); cat("\n")
cat(noquote("a3 = ")); cat(a3); cat("\n")
cat(noquote("b3 = ")); cat(b3); cat("\n"); cat("\n")
```

```
cat("\n"); cat("\n"); cat("Results for direct copy-paste in the HPDR/MHPDR/MHPDRC functions:")
cat("\n"); cat("\n"); cat("Gom=c("); cat(shift); cat(","); cat(a3); cat(","); cat(b3); cat(")")
cat("\n"); cat("\n")
}
```

GompertzFit()

Maximizing -- use negfn and neggr

RESULTS:

shift = 20

a3 = 0.01302971

b3 = 0.04473718

Results for direct copy-paste in the HPDR/MHPDR/MHPDRC functions:

Gom=c(20,0.01302971,0.04473718)

Transition Analysis

```
TA=function(){

x= read.table(choose.files(caption = 'Select the file containing age vs. age indicator data'),
head=1, sep="\t")

nf=max(x[,2])


library(VGAM)


v=vglm(formula = x[,2] ~ log(x[,1]), family = cumulative(link = "probit", parallel = TRUE,
reverse = TRUE))

y=coef(v, matrix = TRUE); yy=coef(summary(v), matrix = TRUE)


cat("\n"); cat("\n")

cat(noquote("RESULTS using Log scale and a common slope")); cat("\n")

cat(noquote(" (key: AAT = Age-at-transition)")); cat("\n"); cat("\n")

cat(noquote("Phases")); cat("\t");
cat(noquote("Log(AAT)"));cat("\t");cat(noquote("SE"));cat("\t"); cat(noquote("AAT"));
cat("\t"); cat(noquote("SE(AAT)")); cat("\n")

for(j in 1:(nf-1)){

   mu= abs(y[1,j]/y[2,1])

   se=mu*(yy[j,2]/abs(y[1,j])+ yy[nf,2]/abs(y[2,1]))

   cat(j,(j+1)); cat("\t"); cat(mu);cat("\t");cat(se);cat("\t"); cat(exp(mu))

   cat("\t"); cat(exp(mu)*se); cat("\n")}

cat(noquote("SD"));cat("\t");cat(1/y[2,1]); cat("\t")

yy=coef(summary(v), matrix = TRUE)
```

```
cat(noquote("SE(SD)"));cat("\t");cat(yy[nf,2]/y[2,1]^2); cat("\n")
yf= dnorm(seq(0,6,.01), abs(y[1,1]/y[2,1]),1/y[2,1])
c= exp(seq(0.01,6,.01))- exp(seq(0,6-0.01,.01))
s=0
for(j in 1:599) {s=s+c[j]*(yf[j+1]+yf[j])/2}
plot( exp(seq(0,4.6,.01)), dnorm((seq(0,4.6,.01)), abs(y[1,1]/y[2,1]),1/y[2,1])/s, type='l',
xlab='Age', ylab='Density', main="Log scale fitting")
for(j in 2:(nf-1)){
   yf= dnorm(seq(0,6,.01), abs(y[1,j]/y[2,1]),1/y[2,1])
   s=0
   for(jj in 1:599){ s=s+c[jj]*(yf[jj+1]+yf[jj])/2}
   lines( exp(seq(0,4.6,.01)), dnorm(seq(0,4.6,.01), abs(y[1,j]/y[2,1]),1/y[2,1])/s, lty=j)
}
pv=c(); m=0
for(j in 1:nf) {
   pv[j]=yy[j,4]
   if(pv[j] > 0.05) m=1}
cat("\n"); cat(noquote("p-values of fitting coefficients:"));cat("\n");cat(pv)
cat("\n"); cat("\n")
if(m==1) {
  cat(noquote("BE CAREFULL: p-value(s) statistical non-significant"));cat("\n")}
cat("\n"); cat("\n"); cat("Results for direct copy-paste in the HPDR function:")
cat("\n"); cat("(They can be also used in the MHPDR/MHPDRC functions)")
cat("\n"); cat("\n"); cat("mu=c(");
for(j in 1:(nf-2)){
```

```
    mu= abs(y[1,j]/y[2,1])
    cat(mu);cat(",");}
mu= abs(y[1,nf-1]/y[2,1])
cat(mu);cat(")"); cat("\n")
cat("sdev=");cat(1/y[2,1]); cat("\n"); cat("\n")
}
```

Male_cranial_sutures

```
> TA()
```

RESULTS using Log scale and a common slope
(key: AAT = Age-at-transition)

Phases	Log(AAT)	SE	AAT	SE(AAT)
1 2	3.01008	1.751628	20.2890	35.53883
2 3	3.453715	1.896562	31.6176	59.96477
3 4	3.6804	1.971079	39.6622	578.17743
4 5	5.06421	2.448013	158.2554	387.4113
SD	0.5098705	SE(SD)	0.132519	

p-values of fitting coefficients:
0.001899851 0.0005452926 0.0002859166 7.658291e-06 0.0001193154

Results for direct copy-paste in the HPDR function:
(They can be also used in the MHPDR/MHPDRC functions)

```
mu=c(3.01008,3.453715,3.6804,5.06421)
sdev=0.5098705
```

Highest Posterior Densities and Highest Posterior Density Regions

```
HPDR=function (Gom, mu, sdev, a=0.95, top=120){
library(nleqslv)
shift=Gom[1]; a3=Gom[2]; b3=Gom[3]


int=function(x, a, b){
   step=x[4,1]-x[3,1]
   n=length(x[,1])
   j1=a;j2=b
   for(j in 1:(n-1)){
      if((a>=x[j,1]) && (a<x[j+1,1])){
        j1=j
        break}}
   j2=j1
   for(j in 1:(n-1)){
      if((b> x[j,1]) && (b<= x[j+1,1])){
        j2=j
        break}}
   if(j2==j1){
     I=0
   } else{
   I=0
   for(j in j1:j2){
       I=I+(x[j,2]+x[j+1,2])*step/2}
   }
```

```
   return(I)}

cat('\n');cat('INPUT DATA:'); cat('\n'); cat('\n')
cat('Gompertz model parameters:'); cat('\n'); cat(paste('h = ',shift));cat('\t');
cat(paste('a3 = ',a3));cat('\t'); cat(paste('b3 = ',b3)); cat('\n')
cat('mu vector:'); cat('\t'); cat(mu); cat('\n')
cat('SD:'); cat('\t'); cat(sdev); cat('\n')

imax=length(mu)+1
for(ip in 1:imax) {
   probit=function (t){
      if(ip==1) return(1-pnorm(log(t),mu[1],sdev))
      if(ip==imax) return(pnorm(log(t),mu[imax-1],sdev))
      return(pnorm(log(t),mu[ip-1],sdev)-pnorm(log(t),mu[ip],sdev)) }

   f1 = function(t){
      p = probit(t)
      Gf = a3*exp(b3*(t-shift))*exp(a3/b3*(1-exp(b3*(t-shift))))
      return(Gf*p) }

   hpd = optimize(f1,c(shift,top),maximum=T)$maximum

   f2 = function(t){
      p = probit(t)
```

```
    Gf = a3*exp(b3*(t-shift))*exp(a3/b3*(1-exp(b3*(t-shift))))
    return(Gf*p/integrate(f1,shift,top)$value) }

x= c(shift, shift:top); y=c(0,f2(shift:top))
x0= c(shift, seq(from=shift, to=top,by=0.1)); y0=c(0,f2(seq(from=shift, to=top,by=0.1)))
xy=cbind(x0,y0)
windows(record=T)
plot(x,y, type='l',xlab='Age',ylab='Density',xlim=c(shift,top))
title(paste('ip = ',ip, '  ', a*100,'% HPDR'))

cutf2= function(tt) {return(integrate(f2,top,tt)$value+(1-a)/2)}
cutf = function(tt) {return(integrate(f2,top,tt)$value+(1-a))}
m=-0.5
for(j in 1:20){
    m=m+0.5
    if(cutf(hpd-m)<0) break }

hi = uniroot(cutf,c(hpd-m,top))$root

fs = f2(shift)
if(f2(hi) <= fs) {
   low = shift
   cutf2= function(tt){
   return(integrate(f2,top,tt)$value+(1-a))}
```

```
  hi = uniroot(cutf2,c(hpd,top))$root
} else {
  f = function(x) {
  y = numeric(2)
  if(x[1]<shift) x[1]=shift
  y[1] = f2(x[2])-f2(x[1])
  y[2] = integrate(f2,x[1],x[2])$value - a
  y}
ol=nleqslv(c(hpd-1,hpd+1), f)$x
low=ol[1]; hi=ol[2]
}

if(ip==1) {
    cat('\n'); cat('\n'); cat('RESULTS:'); cat('\n')
     cat('\n');cat('Phase');cat('\t'); cat('HPD');cat('\t'); cat(a*100,'%Low \t');
     cat(a*100,'%Hi \t'); cat('ABCP');cat('\t'); cat('AL');cat('\t'); cat('AH');cat('\n') }

cat(ip); cat('\t'); cat(round(hpd,digits=2)); cat('\t');
cat(round(low,digits=2)); cat('\t'); cat(round(hi,digits=2)); cat('\t');
cat(round(int(xy,low,hi), digits=2)); cat('\t'); cat(round(int(xy,shift,low) , digits=4));
cat('\t'); cat(round(int(xy,hi,top-1), digits=4));cat('\n')
lines(c(0,top),c(0,0))
lines(c(hi,hi), c(0,f2(hi)), lwd=1, lty=2)
lines(c(low,low), c(0,f2(low)), lwd=1, lty=2)
```

```
}
cat('\n')
}
> Gom=c(20,0.01302971,0.04473718)
> mu=c(3.01008,3.453715,3.6804,5.06421)
> sdev=0.5098705
> HPDR(Gom=Gom,mu=mu,sdev=sdev)
```

INPUT DATA:

Gompertz model parameters:

h = 20 a3 = 0.01302971 b3 = 0.04473718

mu vector: 3.01008 3.453715 3.6804 5.06421

SD: 0.5098705

RESULTS:

Phase	HPD	95 %Low		95 %Hi	ABCP	AL	AH
1	20	20	53.55	0.95	0	0.0503	
2	30.97	20	63.31	0.95	0	0.05	
3	40.38	20	68.3	0.95	0	0.0507	
4	53.97	25.84	79.66	0.95	0.0214	0.0293	
5	66.92	41.88	88.24	0.95	0.0286	0.0217	

ETHICS CLEARANCE CERTIFICATE

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG | FACULTY OF HEALTH SCIENCES | Anatomical Sciences

SCHOOL OF ANATOMICAL SCIENCES ETHICS WAIVER CLEARANCE LETTER

Re: In terms of Chapter 8, sections 62-64 of the National Health Act No 61 of 2003 donated bodies and their tissues may be used for, among other purposes, health and research. Use of such Material is subject only to permission from the responsible person in the School of Anatomical Sciences – the Head or person designated by the Head.

Human Research Ethics Committee (Medical) Clearance Certificate: **W-CBP-220504-01**

This letter serves to confirm that the Head of School, or the person designated by the head based in the School of Anatomical Sciences, Faculty of Health Sciences, has reviewed the research proposal entitled:

An evaluation of the Acsádi-Nemeskéri Complex Method for adult age estimation on a black South African sample by Amanda Mokoena an MSc student supervised by Dr Anja Meyer.

and has granted clearance to access the blanket ethics waiver to conduct the abovementioned research study.

Professor Amadi O. Ihunwo
Head of School: Anatomical Sciences

20/06/2022
Dated

PLAGIARISM DECLARATION

UNIVERSITY OF THE WITWATERSRAND
FACULTY OF HEALTH SCIENCES
SCHOOL OF ANATOMICAL SCIENCES

DECLARATION ON PLAGIARISM AND RESEARCH MISCONDUCT

Research misconduct is defined as the fabrication, falsification, or plagiarism in the proposing, performing, reviewing or reporting of research, where:

Fabrication is making up data or results and recording or reporting them;
Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented; and
Plagiarism is the appropriation of another person's ideas, processes, results, or words without giving appropriate credit.

I, the undersigned, declare that this work (1) is my own. It is being submitted for, or in contribution to, the fulfillment of the requirements of the Bachelor of Science degree in the Faculty of Science at the University of the Witwatersrand, Johannesburg. The research presented here is an accurate account and there has been no attempt to fabricate, falsify or misrepresent data or results. In addition, this work has not been submitted before for, or as a part of, any degree or examination in any other course in this or another academic institution.

I have read and understood the School's/University's definition of plagiarism and research misconduct. I understand that plagiarism or research misconduct in any form is a serious offence that among other actions may lead to a mark of zero and/or expulsion from the University.

I certify that the studies contained herein were approved by the Animal Ethics Committee and /or the Human Research Ethics Committee (Medical) of the University of the Witwatersrand, as appropriate.
Tick where applicable: Yes ..✓..... ; No .....; Not applicable.....

1. Nature of the work:

a. Essay .....
b. Project .....
c. Dissertation ..✓.....
d. Thesis .....
e. Other written work .....

Name Amanda Valeria Mokoena..... Date 22/02/2024.....

Class/Degree Program Master of Science in Medicine.....

Student Registration Number 1624784 .....
Signature: .....

A copy of this form MUST be completed and included with all work handed in for assessment. No work will be marked without it.

TURNITIN REPORT

