Evaluation of Regression Equations Used to Estimate Age at Death from Cranial Suture Closure

Carlos J. Zambrano

B.A., University of Florida at Gainesville, 2000

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Human Biology in the Graduate School of the University of Indianapolis

August 2005

FORM B

Accepted by the faculty of the Graduate School, University of Indianapolis, in the partial fulfillment of the requirements for the Master of Science degree in

HUMAN BIOLOGY

Stephen P. Nawrocki

Thesis Advisor – Dr. Stephen P. Nawrocki, D.A.B.F.A

Christopher W. Schmidt

Reader – Dr. Christopher W. Schmidt

8-15-05

Date

ABSTRACT

During a skeletal analysis, age at death is one of the primary components of the biological profile that the forensic anthropologist constructs. The study of cranial suture closure and its relationship with age dates back to the 16th century. However, since that time and continuing into the present, there have always been doubts about the applicability of suture closure to age estimation. Even with this skepticism, researchers continue to examine suture closure as an indicator of age at death. Most recently, Nawrocki (1998) introduced 14 regression equations to estimate age using cranial suture closure. Testing of the performance of these equations as well as their applicability as age estimators has been limited. This study examines 6 of the regression equations (Equations 1, 2, 3, 4, 7, and 8) created by Nawrocki using a new sample of recently-deceased individuals. This test sample contains 388 individuals (121 females, 267 males) of European ancestry. The majority of the test sample is derived from documented skeletal collections curated by the University of Tennessee and the Maxwell Museum at the University of New Mexico. A small percentage of individuals were forensic cases processed by the University of Indianapolis Archeology and Forensics Laboratory.

Up to 31 landmarks were scored on each specimen in the test sample: 18 ectocranial (external) surface, 7 endocranial (internal) surface, 2 facial, and 4 from the palate. Following Meindl and Lovejoy (1985), one-centimeter segments along the cranial sutures were scored from 0 to 3, where 0 is no closure, 1 is 1-50% closure, 2 is 51-99% closure, and 3 is complete obliteration of the suture. The palatal sutures were scored following Nawrocki's guidelines. Kendall's and Pearson's correlation matrices were calculated and compared to values obtained from Nawrocki's Terry Collection sample. Age was estimated for each individual using up to 4 different equations: 2 general equations and 2 group-specific equations (e.g., all females, European females, etc.). Inaccuracy and bias statistics were calculated for each equation to assess its performance. The percentage of individuals whose estimated ages fell within each equation's ± 1 and ± 2 standard error intervals was also calculated. An analysis of covariance (ANCOVA) was used to determine if suture closure is influenced by an individual's sex. Then the sample was culled to create an even distribution by age and the statistical analyses were repeated. This procedure was conducted in order to correct for any biases caused by a different mean age at death for the test sample when compared to Nawrocki's original Terry sample.

The correlation matrices show that correlation strength is dependent on sex and suture location. Inaccuracy was greater for both the original and culled test samples when compared to Nawrocki's published values. Bias ranged from 1.31 to –17.07 years and tended to be negative, indicating systematic underestimation of age. However, these error rates decreased when culling the sample to more closely match the mean age of Nawrocki's Terry sample. The ANCOVA results suggest that summed suture score is influenced by sex. Of the 6 equations tested, the 2 general all group equations (Equations 1 & 2) performed best, followed by Equation 4 (all males). However, overall, the original equations work relatively well and can play a role in age estimation in modern American populations. In conclusion, cranial suture closure does correlate with age and sex influences the pattern and/or rate of suture closure, although the effects of sex are not so great that they preclude one from combining the sexes into a single equation. Also, while the modern sample is more variable (making age estimation less accurate), there seem to be no systematic secular trends that would prevent the use of Nawrocki's (1998) equations on modern individuals from forensic contexts. This study reaffirms the need to carefully control the reference samples used to test age estimation methods.

TABLE OF CONTENTS

Signature Page i

Abstract ii

Chapter 1: Introduction 1

Chapter 2: Literature Review 7

Chapter 3: Materials and Methods 25

Chapter 4: Results 31

Chapter 5: Discussion and Conclusions 65

Literature Cited 73

CHAPTER 1: INTRODUCTION

Age at death is one of the primary components of the biological profile that an osteologist attempts to determine during a skeletal analysis. The aging process is continuous and can be divided into the three broad phases: growth (or development), maintenance, and degeneration (or atrophy) (Shipman et al., 1985; Angel et al., 1986). Age estimation of individuals in their growth phase is based upon developmental phenomena such as the appearance of ossification centers, tooth development, and epiphysial union. Age estimation for juveniles has reasonable accuracy due to the consistent and rapid developmental changes that occur early in life. Age estimation for adults is typically based on the morphological appearance of joint surfaces that undergo daily wear and tear. However, the rate of degeneration is slower than the rate of growth and varies more between individuals, which reduces the accuracy of adult age estimates compared to those made for juveniles.

Estimating age at death from skeletal remains is not a simple task. Regardless of what phase of life an individual is in, when assigning an age at death for a specimen, the osteologist is attempting to predict chronological age based on the organism's biological age. Chronological age is the number of years lived from birth (e.g., Elvis was 42 when he died). Biological age (also known as skeletal or physiological age) cannot be measured in years but is instead measured as the degree of senescence of the organism (Acsadi and Nemeskeri, 1970). Biological age is an assessment of how old the organism looks. Unfortunately, biological age does not perfectly correlate with chronological age. Moreover, the degree of correlation varies in different areas of the body and across the lifespan (Acsadi and Nemeskeri, 1970). This inconsistent relationship between biological and chronological age explains in part why it is much more difficult to assign an age estimate once an individual has surpassed the growth phase.

As age increases, the correlation between the biological and chronological age decreases (Nawrocki, 1998). Therefore, due to the variable rate of biological aging between individuals, it is difficult to extrapolate a chronological age from the appearance of the skeleton.

There are several skeletal indicators that are used to estimate age at death for adults. Cranial suture closure is the oldest and the most controversial age indicator. Cranial suture closure (or suture synostosis) has been used since the 16th century, and since that time there has been considerable debate about its applicability and reliability in age estimation (Todd and Lyon, 1924; Montague, 1934; Masset, 1989). Cranial sutures are classified as fibrous joints, meaning that they lack a synovial cavity and the bones are held together firmly by fibrous connective tissue. Sutures develop when adjacent bones of the skull come into close proximity. During infancy, the suture is straight and the gap between the bones is filled with fibrous connective tissue, which gradually reduces in thickness, bringing the bones closer together. Once tissue reduction has occurred, the suture becomes more complex with bony interdigitations. The functions of sutures are (1) to prevent separation of the bones when external forces are applied (e.g., muscle function or trauma), and (2) to allow some movement to occur between bones during rapid growth of the cephalic viscera (Bennett, 1967). Suture fusion (“closure,” or synostosis) frequently begins as ossification along the transverse fibers uniting the bones, however, these fibers are usually only present after adulthood is attained (Herring, 1993). Suture closure has been attributed to vascular, hormonal, genetic, mechanical, and local factors, however, to date there are no clear explanations of why suture fusion occurs (Cohen, 1993).

Recently, Nawrocki (1998) introduced 14 regression equations for determining age at death using cranial suture closure. These equations were created using specimens from the Terry Collection, which is a cadaver population collected from dissecting rooms in the St. Louis area during the early part of the 20th century. Although carefully constructed, Nawrocki's sample is relatively small (n = 100) for all groups combined, and once the groups are divided by sex and ancestry, sample sizes drop to about 25 for each of the four subgroups examined. With subgroup sample sizes being so small, it is reasonable to argue that Nawrocki's study sample only minimally represents the range of population variation in suture closure. The most appropriate way to evaluate an age estimation method is to test it on a sample not used in the construction of the method. Testing the performance of Nawrocki's equations has been limited. Of the 14 equations created, Nawrocki was able to test only five equations using a modern cadaver sample from dissection rooms in Syracuse, New York and Indianapolis, Indiana. The test sample consisted of 61 calvaria from males and females of European ancestry ranging in age from 58 to 102 years old. However, since the caseloads of forensic anthropologists include many young adults and middle-aged individuals, the older sample used by Nawrocki may not represent an adequate test of his equations.

Estimating Chronological Age

The biological age of an individual is influenced by intrinsic and extrinsic factors. Intrinsic factors are innate characteristics, such as sex or ancestry. Extrinsic factors are external influences, such as the environment. In order to increase the level of accuracy and precision for an age estimation method, it is necessary to account for extrinsic and intrinsic factors that can influence age predictions. In addition, one must make sure that the experimental design and sampling procedures account for factors that may introduce bias.

Sex. A review of the literature finds that sex differences in suture obliteration have been both demonstrated and emphatically denied (Masset, 1989). A difference between the sexes was suggested as early as 1866 by Welcker (Masset, 1989). More recent studies have suggested that there is no difference between the sexes (Nermeskeri et al., 1960; Meindl and Lovejoy, 1985). Meindl and Lovejoy (1985) examined almost 200 individuals and their results indicate that there is no measurable bias due to sex. Key and colleagues (1994) examined three suture closure methods applicable to males and females, including the Meindl and Lovejoy method. Their results show that endocranial suture closure patterns are not affected by sex, however, ectocranial suture closure is. They found that the differences involve both the rate of suture closure as well as the pattern of closure. Masset (1989) suggests that the more recent studies did not discover any sex differences because of a lack of female specimens in the 30-49 age group, which is where the differences between males and females is the greatest. Nawrocki (1998) found sex differences in the summed scores and suggests that using 27 landmarks instead of the 10 used by Meindl and Lovejoy allows for more subtle relationships to be tested.

Ancestry. The influences that ancestry could have on age estimation has been ignored or not detected in most age estimation studies, even when a distinction has been made for sex. The Todd and Lyon studies in the 1920's examined ancestral differences, however, in the end they concluded that there was no significant difference between Africans and Europeans. Singer (1953) also suggested that no "racial" differences existed in suture closure patterns. Meindl and Lovejoy (1985) found no bias in aging due to ancestry in their study sample. Mann (1987) examined the palatal sutures and concluded that the difference between ancestral groups is negligible. However, in 1984, Baker examined cranial suture closure in 144 males and did find differences between individuals of European and African descent. Nawrocki's (1998) data suggested that ancestry alone does not have an effect, however, the interaction between ancestry and sex does in the summed scores.

Environmental effects. Secular trends are non-genetic biological (phenotypic) changes that occur in a population over short periods of time, usually attributed to environmental and nutritional factors. A number of authors have discussed the occurrence of various types of secular trends in skeletal samples (Trotter and Glesser, 1951; Jantz and Moore-Jansen, 1988; Masset, 1989; Meadows and Jantz, 1995; Nawrocki, 1998; Ousley, 1995). Masset (1989) has suggested the possibility of a secular trend in Europe that has increased the rate of synostosis. Nawrocki tested for differences between his reference population (Terry Collection) and his test population (modern cadavers). Finding none, he concluded that no secular trends had occurred in the past 100 years. However, his examination was limited to individuals of European ancestry, using the sutures present on the upper portion of the cranial vault only, and using individuals older than 57 years of age.

Suture closure has progressed substantially by age 60 and it is in the early decades that there appears to be more activity in closure. Therefore, Nawrocki may have failed to find a secular trend because the primarily closed sutures of most older individuals may mask differences present earlier in life. In addition, the older test sample used by Nawrocki may have been more similar to the Terry Collection than a younger, more modern population. A more appropriate test would have compared recent young adults with Terry young adults from the turn of the century. If differences were found, Nawrocki's original formulae based on the Terry sample may not be appropriate for use on modern specimens in the earlier decades of life.

Reference population size and composition. The size and distribution of the reference population has a significant effect on age prediction. It is necessary for the reference population used to construct an aging method to adequately represent the range of variability of the age indicator. Because the relationship between age and suture closure changes with increasing age, an even distribution of the reference population is necessary to control for error that could arise from having a very high or very low mean age in the sample (Masset, 1989).

Hypotheses

Since their seems to be little agreement between researchers on what factors influence suture closure, it is necessary to continue examining how sutures are influenced by age. Due to the limited testing and the small sample sizes used to construct Nawrocki's equations, it is unknown whether they are appropriate for use in modern forensic cases. The current study follows the data collection guidelines of Nawrocki (1998) for cranial suture closure and examines the performance of some of his equations. Suture closure data are collected on individuals of European ancestry of recent origin and with a broader age range than the original test sample. The null hypotheses for the study include:

H_n1: Errors in age estimation using a large modern sample will not differ from those obtained by Nawrocki using the Terry Collection;

H_n2: Sex will not have a significant effect on suture closure in a large modern sample;

H_n3: The correlation between suture closure and age will not differ between the Terry Collection and a large modern sample when controlling for sex.

Chapter 2 is a literature review of cranial suture closure studies. Chapter 3 describes the material and methods used in this study. Chapter 4 presents the results of using Nawrocki's (1998) equations on a modern sample. Chapter 5 discusses the results, presents conclusion, and gives suggestions for future study.

CHAPTER 2: LITERATURE REVIEW

Cranial suture closure is probably the oldest age at death estimator in existence. The observation of cranial sutures dates back to the time Hippocrates and Aristotle, however, it was not until 1542 that Vesale (Vesalius) suggested that there is a relationship between age and cranial suture synostosis (Todd and Lyon, 1924; Montagu, 1938; Masset, 1989). In the mid- to late 1800's there was an explosion of research in Europe regarding suture closure and age. Unfortunately, language barriers prevented the exposure of the more extensive studies to the scientific community at large (Masset, 1989). It was not until 1890 that Dwight published the first work in the United States examining suture closure. For a review of European authors before and after this date one should refer to Montagu (1938) or Masset (1989).

Early studies. Dwight's (1890) paper entitled "The closure of the cranial sutures as a sign of age" examined 100 crania from paupers, and he admits that some of the known ages are suspect. Dwight examined the coronal, sagittal, and lambdoidal sutures and gave a brief summary of his observations for each individual. Unfortunately, Dwight did not apply his terminology or make his observations in a consistent manner, nor did he always examine the endocranial surface. Dwight states that the progress of ossification that he observed is, "if anything understated" because of these inconsistencies. Even so, Dwight comes to 3 conclusions: (1) sutures begin to close much earlier than had been previously stated, with fusion occurring before age 30; (2) suture closure typically begins on the endocranial surface; and (3) the time of closure of any particular part of a suture, and the order in which the process of fusion advances, is very uncertain.

In 1905, Parsons and Box published a paper that examined the vault sutures (endo- and ectocranial) of 82 lower and middle class individuals from England. In their conclusion they agreed with Dwight that suture closure does occur before the age of 30, adding that over 30 years of age there is some closure on the endocranial surface and that at 50 and usually always over 60 all the endocranial sutures are obliterated. Parson and Box suggest that no estimate of age should be made from the ectocranial sutures because they are too variable, and when possible, age estimates should be made using the endocranial sutures. In addition, they state that a difference between the sexes exists in that male sutures become obliterated earlier than females'.

The first standard. In 1924, Todd and Lyon began to publish a series of papers that examined cranial suture closure using a large skeletal sample of 514 crania collected from the dissection rooms at Case Western University in Cleveland, Ohio. The Todd and Lyon series of papers was a hallmark at the time because it examined endo- and ectocranial suture closure as well as the influence that ancestry has on suture closure. The first part of the series, entitled "Endocranial suture closure: its progress and age relationship, Part I - adult males of white stock," set the guidelines for their 4-part series and established a general method for age estimation using suture closure. Todd and Lyon divided the sutures into segments and scored each segment using a five-point scale ranging from 0 to 4, where 0 indicates no union, 4 indicates complete closure, and 1, 2, and 3 refer to the amount of union in quarters. Todd and Lyon's preliminary survey of 307 Euro-American males consisted of plotting the degree of union for each suture segment for each specimen. They then averaged the closure for each suture and plotted new graphs for right and left sides. In this manner, Todd and Lyon were able to visualize a trend in the progress of suture closure with relation to age. The graphs also gave them the opportunity to see "anomalous" specimens and to remove them before the final graphs were created. Using the graphic data, Todd and Lyon made a rough draft of what they termed the "modal type" of suture union. They continued with this methodology and used their Euro-

American male sample as a standard for comparison. In this manner they eliminated specimens with abmodal progress in suture union in the other subgroups they examined. Their final graphs were smoothed using a three-year moving average.

After the removal of what they felt were abmodal specimens, their Euro-American male sample decreased from 307 to 267 individuals. They determined that: (1) the vault sutures close in the order of sagittal, coronal, and lambdoid; (2) the closure pattern between the vault and the circum-meatal sutures differs (union in the circum-meatal group commences later in life); (3) suture closure exhibits a periodicity, with the most extreme activity occurring between 26 and 30 years of age and subsidiary periods of activity in the 50’s and 70’s; and (4) the variability in suture union is too great to be used alone as an age marker.

In 1925, Todd and Lyon continued their series with Part II, which examines ectocranial suture closure of males of European ancestry, following the methodologies from Part I. They (1925a) conclude that: (1) suture closure in general occurs on both the ecto- and endocranial surfaces at the same time; (2) ectocranial closure progresses at a slower rate with more variation than endocranial closure and thus is not as reliable as endocranial closure; (3) the years between 26 and 30 is the period when most union takes place, with secondary periods of activity in the late 30’s and at 60; and (4) the closure sequence of the ectocranial sutures is generally the same as the endocranial suture closure pattern.

Part III (Todd and Lyon, 1925b) examines endocranial suture closure in African American males following the methods described in their 1924 publication. Their initial sample size was 120, however, after discarding abmodal individuals their sample size was reduced to 79 individuals. Their results suggest that: (1) there is no fundamental difference in the graph patterns or in age relationships between individuals of European and African ancestry;

(2) individual variability is greater in African Americans; and (3) there is a modal type of suture closure unaffected by ancestry. Part IV (Todd and Lyon 1925c) examines ectocranial suture closure in African American males using the 79 individuals from Part III. Their conclusions are in essence identical to the ectocranial study performed on individuals of European ancestry in Part II.

Several problems have been raised with respect to the Todd and Lyon series. Their removal of specimens that were atypical compared to their preconceived ideal of suture fusion is suspect (Singer, 1953; Baker, 1984). By removing specimens that did not conform to their modal progression of suture closure, they basically made the African American sample appear similar to the European American sample and masked any inter-group variation. This procedure most likely led to their conclusion that there is no difference between these two ancestral groups, even though they do state that individual variability is greater in African Americans. Also, the manner in which they used a three-year age interval instead of reported age to smooth out the graphs has been criticized (Singer, 1953; Baker, 1984). Using the three-year interval further obscures differences between samples. The Todd and Lyon series of studies only examined males because they decided to discard their small sample of females; however, they somehow conclude that no difference exist between the sexes. Also, there are problems in the certainty of the age information of the individuals used in their study. Having questionable data on actual age compromises the strength of their study.

New criticism ensues. Up until and subsequent to the Todd and Lyon studies, the use of sutures for age estimation was suggested to be unsuitable due to the extreme variability in fusion. Beginning in the 1950’s, a new round of this criticism ensued. In 1953, Singer published the first paper specifically concerned with the unreliability of cranial suture closure and age estimation. His paper, entitled “Estimation of age from cranial suture closure: a report on its unreliability,” concerns itself with the application of sutures to estimate age at death for medico-legal purposes. Singer describes that various textbooks concerning forensic medicine contained charts and tables for estimating age at death from suture closure. He states that the reference sources for these charts are not always clear. However, upon investigation, those sources can be traced back to Todd and Lyon and earlier investigators. Singer criticizes these texts and suggests that the information presented within them is misleading, and he worries that the information may be readily quoted as fact in a court of law.

Singer’s concerns revolved around the fact that the variability in suture closure has been understated since the earliest of studies. He also suggests that the work done by Todd and Lyon is flawed due to their methodology (see above) and that no attention is given to these problems in the reprinting of their work in textbooks on forensic medicine. Singer observed 430 individuals of varying geographic origins in order to demonstrate the unreliability in suture closure, although he only presents suture closure data on 11 individuals of known age. These individuals deviated from the expected suture closure pattern presented in textbooks at the time by displaying too little or too much closure for their known ages. Singer (1953) concludes that “assessment regarding the precise age at death of any individual, gauged only on the degree of closure of the vault sutures of the skull, is a hazardous and unreliable procedure” (p. 59).

Brooks (1955) also attacked suture closure as an age estimator in her study comparing suture closure to pubic symphysis morphology. She compares the Todd 10 phase pubic symphysis method to the Todd and Lyon cranial suture closure method using individuals from an archeological skeletal population as well as the individuals used to construct the original Todd methods. Her study suggests that pubic symphysis morphology has a higher correlation with age than does cranial suture closure. It also is clear in her study that age related sexual dimorphism exists. Brooks states that cranial sutures worked better to estimate age for her male sample than for her female sample. This finding is not surprising since the correlation coefficients for suture closure and the pubic symphysis were always higher for her male sample. Brooks suggests that the sutures in females tend to remain open throughout life or close much later than the Todd and Lyon studies indicate. In addition, Brooks points out the common occurrence of individuals displaying the suture closure pattern of someone less than 25 years of age, while other indicators suggested an age greater than 35 years.

In 1957, Mckern and Stewart presented a study examining suture closure using 369 soldiers killed during the Korean War. Their sample is predominantly composed of individuals of European ancestry with ages ranging between 17 and 50 years, the majority being individuals under the age of 30. Mckern and Stewart scored suture closure following the guidelines set by Todd and Lyon (1924), with a modification following Singer (1953). In addition to the calvarial sutures, Mckern and Stewart also examined the sutures of the face and the basilar suture. They initially present their closure data in tabular form, which clearly demonstrates the amount of variability present in calvarial suture closure. They point out that with advancing age, the percentage of individuals with open sutures does decrease. However, they suggest that their tables show that there is no clear pattern of when suture closure begins or ends, therefore limiting the applicability of sutures for age estimation. While the Mckern and Stewart study sample is large, it lacks older individuals, thus making it difficult to see long-term trends.

Mckern and Stewart's (1957) examination of the facial sutures suggests that progression of closure is more regular in the face than in the cranium. However, a reliable trend that could be translated into a workable age estimation method is absent. They suggest that the only reliable suture for age estimation is the basilar suture, with complete fusion occurring at age 20. To further examine the total pattern of suture closure, Mckern and Stewart employed regression analysis. Their initial equation used the main vault sutures (i.e., coronal, sagittal, and lambdoid). However, the correlation between the vault sutures and age (r = 0.39) was less than the correlation between all the sutures and age (r = 0.49). Mckern and Stewart derived three conclusions from their regression analysis: (1) suture closure seems to progress in a fairly uniform manner; (2) sutures and their closure patterns are related to each other, however, they are not reliable enough to be applied in cases of individual age determination; and (3) age estimates using overall suture closure are more accurate than using the vault sutures alone.

Genoves and Messmancher (1959) examined 101 known adult males from Mexico to determine if this population followed the suture patterns described by Todd and Lyon and to determine the validity and errors associated with the use of suture closure in age estimation. They found that the errors between estimated and actual age range between 9 and 13 years and suggest that the process of suture obliteration does not follow the pattern traditionally defined for age estimation. They concluded that the estimation of age using the cranium lacks validity in at least 50% of the cases.

Powers (1962) examined 271 individuals of know age, sex, and ancestry from the British Museum of Natural History. All but 19 of these individuals were young adult males from several European countries, and most died in the middle of the nineteenth century. Powers primarily examined suture closure following the guidelines prescribed by the Todd and Lyon studies. The histograms comparing the estimated ages to the actual ages were noticeably different. The greatest errors of overestimation occurred between 25 and 35 years, and the greatest errors of underestimation occurred for individuals over 60 years. Powers suggests that the most reliable estimates are made between the ages of 20 and 28 years and that individuals over 68 years will give completely unreliable results. While the samples examined by Genoves and Messmancher and Powers are large, there is no certainty that the rate or pattern in suture closure would be uniform through time or between the different geographic regions that their specimens are derived from. Also, no controls for differences in mean age at death between samples were employed (refer to Masset below).

Later studies. In 1970, Ascadi and Nemeskeri reported on a long-term study that they had previously published in Hungarian only. They scored the vault sutures of 402 modern cadavers from Budapest using a five-point scale (0 to 4) and calculated the average score for each of the sutures. Their initial examination of the calottes suggested that no sexual dimorphism existed. They disregarded any "pathological" cases and reduced their sample to 352 individuals. Suture closure commences endocranially following the sequence of coronal, sagittal, and lambdoid sutures. Ectocranial closure commences with the sagittal suture and finishes with the coronal suture. For the coronal suture they determined that 80 percent of cases were completely closed by age 30 on the endocranial surface, while ectocranially they found that stage 2 occurred at every age and that stages 3 and 4 occurred in only about 1/3 of the cases. The sagittal suture does not completely close ectocranially, displaying closure stages between 2 to 3.9 over the age of 60; endocranially the sagittal suture is closed in individuals over 70 years for 67 percent of their sample. The lambdoid suture closes slowly and displays a great deal of variation, with stage 1 occurring up to the 70's. Completely closed sutures occurred in only 6 individuals older than 55 years.

In an attempt to create an age estimation method using endocranial suture closure, Acsadi and Nemeskeri (1970) examined 285 symmetrically closing crania. Their crosstabulations suggest that as age increases so does mean suture closure. Also, the dispersion around the means is very large. The regression curve based on five-year intervals suggests that closure is uniform, occurs rapidly at first, then slows down. However, individual data is widely dispersed about the curve. Acsadi and Nemeskeri give 5 mean closure ranges with corresponding ages. They concluded that age determined on the basis of cranial sutures is possible using wide age limits and that suture closure is an important factor when used with other aging indicators.

Johnson (1976) examined 213 individuals that from two nineteenth century burial sites in Lancashire, which were exhumed due to road widening. Johnson's study attempted the use of discriminant function analysis to improve age estimation using suture closure and dental attrition. He used published ages for the commencement and cessation of closure and made modifications so that 50% closure is the midpoint of this range. Johnson calculated correlation coefficients relating actual to estimated age, which ranged from 0.47 to 0.77. The higher correlations were obtained when multiple indicators were used (i.e., all sutures and dental attrition). Johnson concludes that discriminant function analysis using tooth attrition and suture closure give marginally more accurate age estimates than the use of either indicator individually. Johnson's assumption that the midpoint of an age rage corresponds to 50% closure has no bearing on the actual phenomena of suture closure and may have compromised his study. He should have attempted to relate age to his new scoring strategy using the known ages of his sample.

Research continues. Perizonius (1983) examined 256 cadavers dissected in Amsterdam between 1883 and 1909. Following Ascadi and Nemenskeri (1970), Perizonius scored the vault sutures (ecto- and endocranial surfaces) and calculated mean closure values. Student t-tests suggest that no significant differences in mean suture closure exists between males and females.

Mean suture closure appears to increase up to age 69 after which the degree of closure diminishes (Perizonius, 1983). Because of this phenomenon Perizonius divided his sample into two age groups, less than 50 (“selec young”) and greater than 49 years (“selec old”). He calculated Spearman rank correlations between mean suture closure stage and age for the entire sample and for the 2 age subgroups. For the entire sample Perizonius found that only the coronal suture had a significant correlation with age (r = 0.23); for the subsamples the strongest correlations were in the selec young (20 to 49 yrs) age group (n = 40), with correlation coefficients ranging from 0.50 to 0.71 for the endocranial sutures. Spearman rank correlation coefficients were calculated for each suture segment of the selec young age group and they ranged between 0.27 to 0.66. The selec old age group (50 to 99 yrs; n = 216) had correlation coefficients ranging from – 0.15 to 0.13, with most of the ectocranial segments having insignificant correlations. Based on these results, Perizonius selected the best suture segments for the two age groups: 10 for the selec young system, and 5 for the selec old system. The selec old system had a low correlation with age (0.26), while the selec young system had a value of 0.70. By dividing his sample into old and young groups, Perizonius has artificially inflated the correlation values for his young group and has removed valuable information from the old group, making the selec old system of segments practically useless for age estimation.

The majority of the previous studies have not determined if suture closure differs due to sex or ancestry. There has been no clear determination of the existence of these differences because several of the studies contradict one another. Also, most of the studies have had small sample sizes representing the different sex and ancestry groups. Baker (1984) attempted to shed some light on these problems. He examined ecto- and endocranial suture closure on a modern sample of cranial vaults sectioned during autopsy at the Department of the Chief Medical

Examiner-Coroner, Los Angeles, California. This sample consisted of 195 individuals (144 males and 51 females), the majority of which were of European (which included Mexican Americans) or African ancestry. Baker examined the three main vault sutures (lambdoidal, sagittal, and coronal). These sutures were scored along their entirety using the following simplified scale: O = non-union (open suture); P = partial union; and C = complete union (obliteration).

Baker used crosstabulations to create tables that summarize the pattern of fusion that his sample displayed. While Baker's scoring scheme and statistical analysis are simple, he observed that: (1) narrower age ranges may be obtained when controlling for sex and ancestry; and (2) fusion tends to occur first on the endocranial surface, so endocranial suture fusion will usually be more advanced than ectocranial fusion. For general age determinations regardless of sex, Baker gives 2 rules: (1) if the endocranial sutures are open, the individual is 27 years or younger; and (2) if complete endocranial closure has occurred, the individual is 26 years or older. Baker finds evidence to suggest that sex and ancestry does influence age estimates, but his sample of females is small, especially for African Americans (n = 12). Also, since there were no attempts to control the mean age at death between his subgroups, it is difficult to determine if his test for sex and ancestry differences is statistically valid.

Mann (1987) examined maxillary (palatine) suture obliteration as a method for estimating age at death. He collected data on 186 individuals with known age, sex, and ancestry, the majority of which are from the Terry Collection. The length of obliteration was measured and divided by the total length of the suture to calculate the percentage of closure. This procedure was done for each of the four maxillary sutures and then they were combined to give a total score. Mann's correlation results suggest that ancestry has a minor influence on suture closure and that sex significantly affects suture closure. Analysis of variance confirmed that sexual dimorphism does exist in the Terry sample. Regression analysis suggested that a great deal of variation exists in suture obliteration and that males display a greater degree of obliteration than females for any given age. Age predictions were more accurate for females, and males tended to have a greater difference between actual and predicted age. However, neither sex was greatly over- or underestimated (Mann, 1987).

Masset (1989) built on his previous works (see Bocquet-Appel and Masset, 1982; Masset, 1982), which point out the methodological problems in age estimation that researchers had not realized or had just decided to ignore. He points out that since the middle of the 19th century, researchers have described sex differences in the degree of suture obliteration, and more recent researchers have not resolved the issue. He notes that most researchers use collections that tend to have older individuals and that these samples lack individuals in the critical 30 to 49 age group, which is where the largest difference exists between the sexes. Authors that did not find sexual dimorphism (e.g., Acsadi and Nemeskeri, 1970; Perizonius, 1983) used samples with a small number of individuals in the critical age group.

Masset also demonstrates that the reference population used to construct the age estimation method introduces systematic biases that must be controlled for. He states that the age composition of samples has been habitually overlooked. Having a reference population whose mean age is very young or very old will affect the age estimation results. When researchers compare two populations and report differences, these differences can usually be attributed to the difference in the mean age between the populations. Masset suggests that the solution to this problem is to have an evenly distributed reference population, with each age category having the exact same number of individuals. By doing this, most of the systematic error can be eliminated and will not affect the age estimate. However, "attraction to the middle" still occurs. The "attraction to the middle" (i.e., movement of the estimated age towards the reference population mean) is the reason for the tendency of overestimation of age for young individuals and the underestimation of age for older individuals. This phenomenon is a natural result of regression (Nawrocki 1996). In order to overcome the attraction to the middle, Masset suggests the use of a probability matrix relating the specimen to the different age categories. Masset also describes three examples of a possible secular trend in the rate of synostosis in Europe.

For age estimation Masset (1989) scored the vault sutures using a five-point scale identical to the Todd and Lyon method. The scores for bilateral landmarks were averaged. All of the scores were summed and divided by 10 (the total number of landmarks) to create an "obliteration coefficient" or the **S** value. Masset obtained correlation coefficients with age of 0.63 (males) and 0.51 (females) for the endocranial sutures and 0.49 (males) and 0.46 (females) for the ectocranial sutures. He created a total of 20 sex specific regression equations for use on the endo- and ectocranial sutures. These equations were designed for use on populations with different life expectancies. Masset states that the correlations between suture closure and age are weak, however, there is useful information obtained from them.

The current standard. It was not until 1985 that a new method was developed that replaced the Todd and Lyon method as the standard for estimating age at death from cranial suture closure. The Meindl and Lovejoy method is the primary manner to estimate age at death using cranial suture closure in North America and is reproduced in most manuals for this purpose (Byers, 2002; White, 2000; Ubelaker, 1999; Buikstra and Ubelaker, 1994). Meindl and Lovejoy (1985) examined 236 crania from the Hamann-Todd Collection and selected 10 ectocranial landmarks that they determined to be the most useful for age estimation. They scored 1 cm segments on each landmark from 0 to 3, where 0 = no closure, 1 = 1 to 50% closure, 2 = 51 to 99% closure, and 3 = complete obliteration. The landmarks were grouped into two systems of sutures, the lateral-anterior and vault systems, and composite scores (summed scores) were calculated for each system and are used for age estimation. The age estimates consist of the average age and ranges for the composite scores (6 for the vault system, 7 for the lateral-anterior system).

Meindl and Lovejoy suggest that the lateral-anterior system of sutures gives the best estimates and is more useful in the upper age ranges. The measures of dispersion around mean age within each stage of the composite scores have a great deal of overlap and variability (Meindl and Lovejoy, 1985). Meindl and Lovejoy also examined the effects of ancestry and sex, and their ANOVA results suggest that no measurable bias exists in age predictions using either the lateral-anterior or vault systems. However, they used decades as the age variable instead of actual ages, thus reducing the sensitivity of their ANOVAs. Another problem is their omission of non-active crania from the ANOVAs. This omission would hinder the detection of rate or pattern differences in synostosis between their subgroups, especially in the termination phase of suture fusion. In addition, Meindl and Lovejoy do not give any indication of their sample composition or age distribution, which is most likely skewed in the direction of elderly individuals of European ancestry, and the majority of them are probably males.

More recent studies. Key and colleagues (1994) tested three cranial suture age estimation methods: Acsadi and Nermeskeri (1970), Meindl and Lovejoy (1985), and Perizonius (1983). Part of their study was to examine the relationship between suture closure and age, taking into consideration the variation in the pattern of suture closure due to sexual dimorphism and differences within and between populations. Key and colleagues examined 183 crania with known age at death from the Spitalfields collection, which is derived from a London cemetery dating to the 18th and early 19th centuries. Their examination of the Acsadi and Nermeskeri (AN) method showed that only 3 out of the 5 stages are statistically different from each other, and therefore the AN method can only roughly sort the Spitalfields crania into general groups of young, middle, and old age. Using Wilcoxon signed rank tests they determined that no sexual dimorphism is present when using the AN method. Also, since the AN method only uses endocranial fusion, Key et al. (1994) suggest that this method is only appropriate for younger individuals because most endocranial fusion is complete by 50 years.

Key and colleagues' (1994) analysis of the Meindl and Lovejoy (ML) method for ectocranial cranial sutures suggests that interpopulational differences do exist. The mean age at death is significantly different between the Hamann-Todd Collection sample and the Spitalfields sample. The Spitalfields sample shows delayed suture closure compared to the Hamann-Todd sample. Also, Key and colleagues suggest that the ML lateral-anterior system shows no trend of increasing age for stages 1 to 7, and the vault system only shows a significant difference between stages 1 and 6 (i.e., the extremes). This investigation suggests that differences exist between the sexes in the rate and pattern of suture closure in the Spitalfields sample. Females tended to have more suture locations that had significant correlations with age, and males tended to display more advanced closure than females.

The Spitalfields sample only had 53 individuals that were young enough (less than 50) to apply Perizonius' selec young system. Perizonius system failed to place over half of these individuals into the less than 50 group. The selec old system was applied to the remaining 130 individuals, however, age did not systematically increase with increasing suture closure stages

(Key et al., 1994). The lack of increasing age with suture closure for the selec old system should have been expected since Perizonius (1983) had previously demonstrated that his selec old system has a poor correlation with age. Key and colleagues (1994) point out the fact that the application of an age estimation method developed on one sample does not always result in accurate ages when applied to another sample. They highlight that the differences between samples are demonstrated by their observation of sexual dimorphism in ectocranial suture closure, whereas Perizonius (1983) and Meindl and Lovejoy (1985) found none. While most studies caution against the application of suture closure for age estimation due to the large amount of variation in closure, Key and colleagues (1994) believe that the primary problem with existing techniques is their reliance on mean closure scores instead of assessments of individual suture closure sites. Therefore, Key and colleagues proposed a new method.

Using a simplified three-point scale, the Spitalfields sample was scored and divided into two age groups, less than 50 and over 50 years. Only the ectocranial sutures that had the best correlations with age were used. Key and colleagues produced a table for the two age groups, divided by sex to displaying the landmark, score, and corresponding age threshold (older than 50 or less than 50) and gave guidelines on how to classify an individual into the age group. Testing the method on a South African sample, age predictions were accurate for 70% of the males and 65% of the females. Unfortunately, Key and colleagues performed the same mistake of Perizonius by dividing their sample in two. Important information that defines the relationship between age and suture closure is lost when the sample is divided in this manner, which ultimately reduces the performance of their method, especially for the older than 50 group. Perizonius and Key and colleagues should have realized this when the correlation strength between age and suture closure was weakened in their above 50 age groups

Kemkes-Grottenthaler (1996) examined several age indicators using two archeological samples dating to the middle ages. While the samples did not allow for the comparison of estimated to actual age, Kemkes-Grottenthaler does conclude that sex-specific variation is present using the Acsadi and Nermeskeri method.

Hershkovitz and colleagues (1997), in an attempt to determine if suture closure is a pathologic condition or an age related change, examined 3,636 crania from the Hamann-Todd and Terry Collections. They measured the length of the sagittal suture and recorded the percent closure along its length. The suture was placed into one of the following categories: (1) totally closed, (2) partially closed, (3) totally open, and (4) partially open. Using tables and figures to display the frequency and distribution of the four conditions, Hershkovitz et al (1997) suggest that beyond 35 years of age, the distribution of the categories were virtually similar in all age groups. Therefore, they argue that suture closure is not an ongoing, continuous transition from open to closed states. However, the authors do state that the frequency of open sutures steadily decreased with age, while the relative frequency of totally closed sutures increased with age, and females displayed a different pattern of closure. Both Perizonius (1983) and Ascadi and Nemeskeri (1970) suggested that the sagittal suture frequently remains open, but they still found it useful for age estimation. Hershkovitz and colleagues conclude that suture closure is an independent permanent phenomenon and that it is neither a pathologic condition nor the result of the aging process. For the first time a metric approach was used on a cranial suture; unfortunately, they simplified their metric data into four categories and lost valuable information (which is what they were trying to avoid in the first place). Calculating correlation coefficients for their data would have given them a better understating of suture closure and aging instead of interpreting complicated graphs.

Galera and colleagues (1998) examined 963 individuals from the Terry collection and applied the methods of Acsadi and Nemeskeri (AN), Masset, Baker, and Meindl and Lovejoy (ML). Each method was applied and the correlation coefficient between age at death was calculated, controlling for sex and ancestry. Tests of significance were conducted using the Z-transformation of the correlations. Inaccuracy and bias statistics were calculated, controlling for sex and ancestry. All four methods displayed a significant correlation with age, with the endocranial data having the greatest correlations. Differences due to ancestry were found in all four methods, African Americans displaying stronger correlations than European Americans. Sexual dimorphism appeared to be ancestry dependent for the AN and Masset methods, but no sexual dimorphism occurred for the ML or Baker methods. Overall, methods that are based on endocranial suture closure performed the best. The AN and Masset methods that necessitate the calculation of a closure index or an obliteration coefficient performed better (i.e., lower inaccuracy and bias values) than the simplified scoring methodologies used for the ML and Baker methods.

Summary

Extensive research has been conducted on cranial suture closure and its relationship with age. The greatest problems with the body of work as a whole are the lack of control over the mean age at death for each sample and the unbalanced composition of the samples by sex and ancestry. While some studies contradict one another regarding the existence of differences between subgroups (i.e., by ancestry and sex), more recent studies are suggesting that these factors do influence age predictions. However, it is still not clear if overall age estimates are hindered if these group differences are not accounted for.

CHAPTER 3: MATERIALS AND METHODS

Sampling. The specimens used in this study are curated at three institutions: the University of Indianapolis, the University of Tennessee at Knoxville, and the University of New Mexico at Albuquerque. The University of Indianapolis (UIndy) sample comes from 33 recent forensic cases processed from 1991 to the present at the Archeology and Forensics Laboratory (18 Euro-American males and 15 Euro-American females).

The Maxwell Museum Laboratory of Human Osteology at the University of New Mexico (UNM) curates a Documented and a Forensic Collection. The Documented Collection contains over 235 individuals from the Southwestern U.S. whom have died in the last 15 years. During the week of April 5th 2004, the author scored 126 individuals from the Documented Collection (78 Euro-American males and 48 Euro-American females) and 12 individuals from the Forensic Collection (6 Euro-American males and 6 Euro-American females). The University of Tennessee is the home of the Bass Donated Skeletal Collection. It contains over 400 known individuals mostly of European and African ancestry, with ages raging from fetal to 101 years. These individuals came primarily from Tennessee and surrounding areas. During June 2nd through 11th of 2004, the author scored 217 individuals from the Bass Collection (165 Euro-American males and 52 Euro-American females). Table 3.1 presents the distribution of individuals by sex, decade, and collection. Table 3.2 gives summary statistics for the entire study sample.

Scoring suture fusion. Before data collection, the author was trained in the proper scoring techniques for the cranial sutures under the guidance of Dr. Stephen Nawrocki at the University of Indianapolis. All crania used for this study were scored by CJZ except for

TABLE 3.1. Distribution of Individuals by Sex, Decade, and Collection.

Age Group	**UNM Documented Collection**		**UNM Forensic Collection**		**Bass Donated Collection**		**UIndy Forensic Cases**		
	M	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**Total**
18-19	1	0	0	0	0	0	2	0	**3**
20-29	2	1	2	4	4	1	8	7	**29**
30-39	9	4	1	1	13	2	1	3	**34**
40-49	5	1	2	0	26	3	3	3	**43**
50-59	20	4	1	1	45	11	1	0	**83**
60-69	16	9	0	0	36	14	1	1	**77**
70-79	17	13	0	0	23	10	2	1	**66**
80-89	6	11	0	0	14	10	0	0	**41**
90-99	2	4	0	0	3	1	0	0	**10**
100-109	0	1	0	0	1	0	0	0	**2**
Total	**78**	**48**	**6**	**6**	**165**	**52**	**18**	**15**	**388**

TABLE 3.2. Summary Statistics for the Study Sample. Columns 2 through 4 are in years.

Sample	**n**	**Min**	**Max**	**Mean**	**St dev.**
All Males and Females	388	18	101	58.97	18.617
All Males	267	18	101	57.59	17.489
UNM Documented Males	78	19	96	59.81	16.987
UNM Forensic Males	6	24	51	37.83	10.907
Bass Donated Males	165	25	101	59.61	15.762
UIndy Forensic Males	18	18	77	36.00	18.815
All Females	121	20	101	62.03	20.641
UNM Documented Females	48	22	101	70.35	18.109
UNM Forensic Females	6	20	59	31.00	14.710
Bass Donated Females	52	20	94	65.33	14.935
UIndy Forensic Females	15	21	74	36.40	16.326

some of the specimens from the University of Indianapolis, the scores for which were retrieved from the case files of Dr. Nawrocki. Only specimens free of significant cranial trauma with known sex, ancestry, and age were examined. Up to 31 landmarks were scored on each specimen (Table 3.3): 18 on the ectocranial surface, 7 on the endocranial surface, 2 facial, and 4 palatal. Following Meindl and Lovejoy (1985), one-centimeter segments at each landmark along the cranial sutures are scored on a scale of 0 to 3 using the following criteria:

0 = no observable closure
1 = 1 to 50% closure
2 = 51 to 99% closure
3 = 100% closure in the observed segment

The suture closure scores are based entirely on the degree of bony bridging across the suture's surface. The sutural gap could be in close proximity or fusion may have already occurred deep within the suture, however, if there is no bridging across the surface of the suture, it is scored as being open (i.e., scored as a zero). The endocranial sutures were scored following Nawrocki (1998). The four palatal sutures were assessed in their entirety using the same scale; however, any extent of the palatine sutures extending onto the alveolar surface was ignored (Nawrocki, 1998). Note that the intermaxillary suture has a tendency to have bony mounding occurring parallel with it on either side of the suture. This mounding was ignored and the suture was scored below the level of the mounding.

In addition to Nawrocki's (1998) 27 landmarks, the inferior coronal and the zygomaxillary sutures were scored. The inferior coronal suture consists of the most inferior 1 cm segment of the coronal suture. The zygomaxillary suture's most central two quarters were scored. Both the inferior coronal and zygomaxillary sutures were scored bilaterally following the criteria for the ectocranial sutures.

TABLE 3.3. Cranial Suture Landmarks and Abbreviations (after Nawrocki 1998)**.**

Ectocranial	LLQ mid-lambdoid left LRQ mid-lambdoid right LAQ lambda OBQ obelion ASQ anterior sagittal BRQ bregma CLQ mid-coronal left CRQ mid-coronal right CLI inferior coronal left	CRI inferior coronal right PLQ pterion left PRQ pterion right SLQ sphenofrontal left SRQ sphenofrontal right ILQ inferior sphenotemporal left IRQ inferior sphenotemporal right TLQ superior sphenotemporal left TRQ superior sphenotemporal right
Facial	ZLM zygomaxillary left	ZRM zygomaxillary right
Endocranial	LLZ mid-lambdoid left LRZ mid-lambdoid right LAZ lambda SAZ mid-sagittal	BRZ bregma CLZ mid-coronal left CRZ mid-coronal right
Palatine	ICP incisive AMP intermaxillary	TRP palatomaxillary PMP interpalatine

All sutures except for the endocranial sutures were scored using a 10x hand lens. The endocranial sutures were viewed through the foramen magnum using an array of flashlights and fiber optic light benders for illumination. To maintain consistency in the endocranial suture scores, all endocranial sutures were scored through the foramen magnum even if the skullcap had been sectioned. During the scoring process, the known age and sex was concealed until after the specimen had been scored. The data were entered into *SPSS 12.0* for statistical analysis and was examined as a whole and separated by sex. The performance of Nawrocki's (1998) equations 1 through 4, 7, and 8 (Table 3.4) were assessed.

Age for each of these specimens was estimated using up to 4 equations: 2 equations for all groups and 2 group-specific equations (e.g., all males, Euro-American males). The study sample was then carefully culled to create an even age distribution of individuals to prevent any skewness and to facilitate comparison to Nawrocki's Terry Collection sample. Approximately 6

TABLE 3.4. Nawrocki (1998) Equations 1 through 4, 7, and 8.

Equation 1	Summed sutures All Groups	AGE = 0.71(SUMALL) + 25.3 (adj. r^2 = 0.51; se = 12.9 years)
Equation 2	All Groups	AGE = 5.86(PLQ) + 6.42(BRZ) + 4.91(TRP) + 24.3 (adj. r^2 = 0.56; se = 12.1 years)
Equation 3	All Females	AGE = 5.29(CRQ) + 7.38(PRQ) + 8.84(TRP) + 26.8 (adj. r^2 = 0.65; se = 10.9 years)
Equation 4	All Males	AGE = 7.0(PLQ) – 6.08(ASQ) + 6.83(TRQ) + 9.12(BRZ) + 28.3 (adj. r^2 = 0.61; se = 11.5 years)
Equation 7	Euro-American Females	AGE = 9.78(LRQ) + 12.27(OBQ) + 9.93(SLQ) – 12.94(SAZ) + 40.0 (adj. r^2 = 0.80; se = 8.2 years)
Equation 8	Euro-American Males	AGE = 15.01(PLQ) – 6.76(ASQ) + 37.9 (adj. r^2 = 0.61; se = 11.0 years)

females were chosen randomly from each decade of life, beginning in the third decade and ending in the seventh decade (20's through 80's). Approximately 20 males were chosen from each decade in the same fashion. The resulting culled samples numbered 39 females and 128 males (Table 3.5).

Statistical Analysis

Known age and summed suture closure were plotted and examined for trends. Kendall's Tau-b correlation matrices were calculated between the summed suture scores and actual age and between suture landmark and age to examine the relationships between these variables. For each of Nawrocki's equations, the percentage of individuals whose actual age falls within the ± 1 and ± 2 standard error intervals was calculated for the sample as a whole and by sex. Inaccuracy and bias statistics are calculated for each sample as a whole, by sex, and per decade. Inaccuracy measures the average difference between estimated and actual values and is calculated as:

$$\sum \left| estimated \quad age - actual \quad age \right| / n$$

Bias determines the average over- or under- estimation of an equation and is calculated as inaccuracy but without absolute value bars. To determine if sex does indeed influence suture closure in Euro-Americans, an analysis of covariance (ANCOVA) was performed using the following model:

SUMALL = SEX + AGE

The dependent variable for this model is SUMALL, which is the summed suture closure score for the 27 landmarks used by Nawrocki (1998). The main effect is SEX and age is the continuous covariate. To test for population differences, Nawrocki's original suture closure data for Europeans in the Terry Collection was compared to the new culled study sample using the above ANCOVA model, modified to include COLLECTION as an additional main effect. In addition, a new SUMALL equation was generated using the culled study sample and compared to Nawrocki's SUMALL equation to determine if better age estimates could be achieved.

TABLE 3.5. Summary Statistics for the Culled Study Sample.
Columns 2 through 4 are in years.

	n	Min	Max	Mean	St dev.
Males and Females	167	18	89	55.80	19.602
Males	128	18	89	56.60	19.506
Females	39	21	85	53.15	19.936

CHAPTER 4: RESULTS

Results for the Entire Sample

Plots. The summed suture scores (SUMALL) for Nawrocki's (1998) 27 landmarks were plotted against age for males and females combined and separated by sex. In general, as age increases, SUMALL tends to increase as well (Figures 4.1, 4.2, & 4.3), and the ranges for the SUMALL values appear to be relatively large for any given decade. This funnel-shaped plot is typical for most age indicators and is due to the differential rate of aging between individuals, producing greater errors as the years pass. Males (Figure 4.2) seem to be clustered more closely together than females (Figure 4.3).

Correlations. Kendall's Tau-b (rank order) correlations were calculated to examine the relationships between age and suture closure. Correlations were compared directly to those obtained by Nawrocki (1998). Table 4.1 compares Nawrocki's (1998) Euro- and African-American individuals to the current study sample of Euro-Americans. The gray cells indicate the highest between-sample correlations for a given landmark. Bolded values within the table represent the highest within-sample correlation for a given landmark. Overall, the Terry sample tends to have slightly to moderately higher correlation coefficients than the current study sample, as indicated by the greater number of gray cells in the three right columns. The few instances when the correlation values between age and a landmark were larger for the current study sample nearly always occurred when the Terry sample did not have a significant (ns) correlation for that landmark. Females tend to have higher correlation coefficients than males in both the current study and Terry samples, but this difference is more marked in the current sample. Females in both samples have higher correlation coefficients for the vault series of sutures (LLQ through

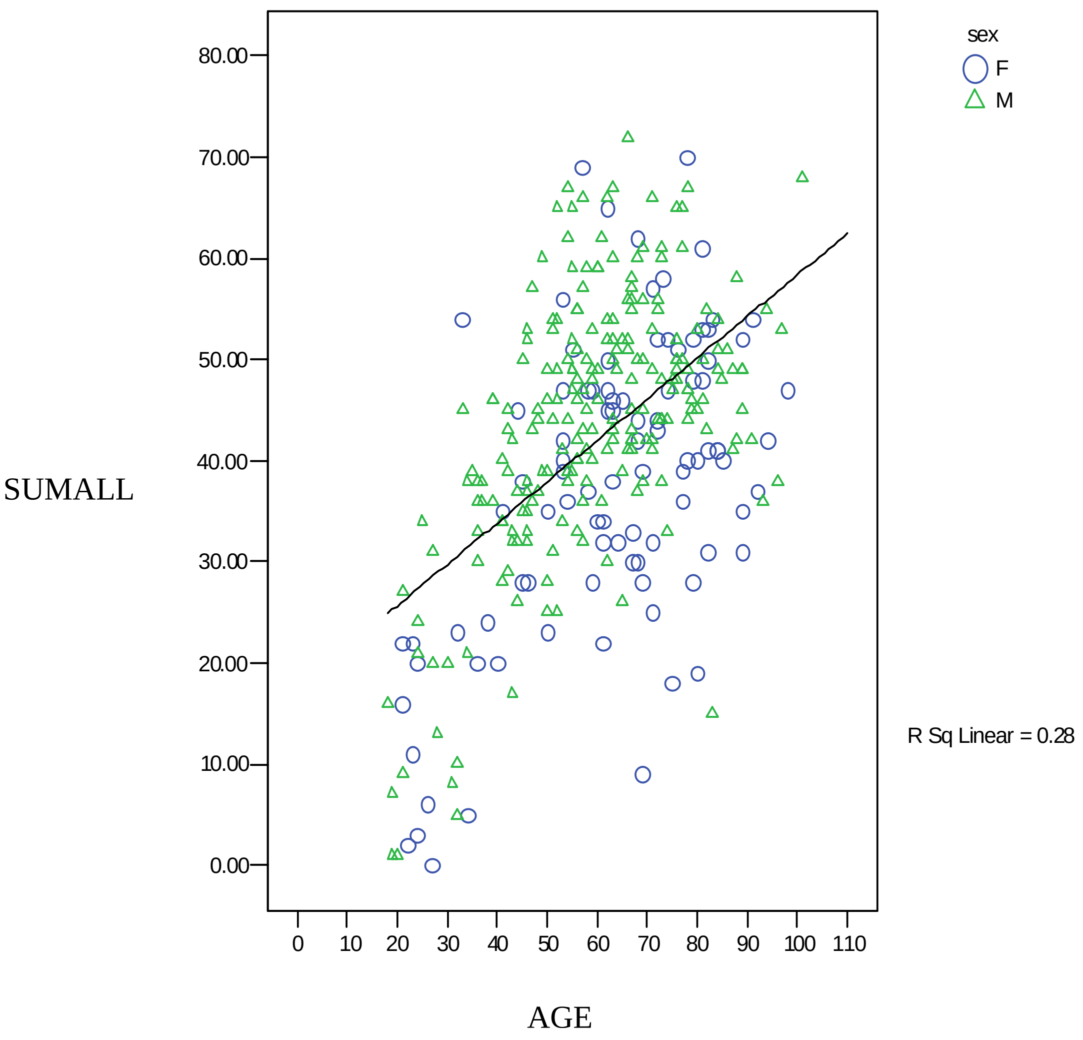

FIGURE 4.1. Plot of SUMALL vs. Age for Males and Females Combined (n = 319). The regression line ($y = 0.685x + 31.35$) is superimposed on the scatterplot.

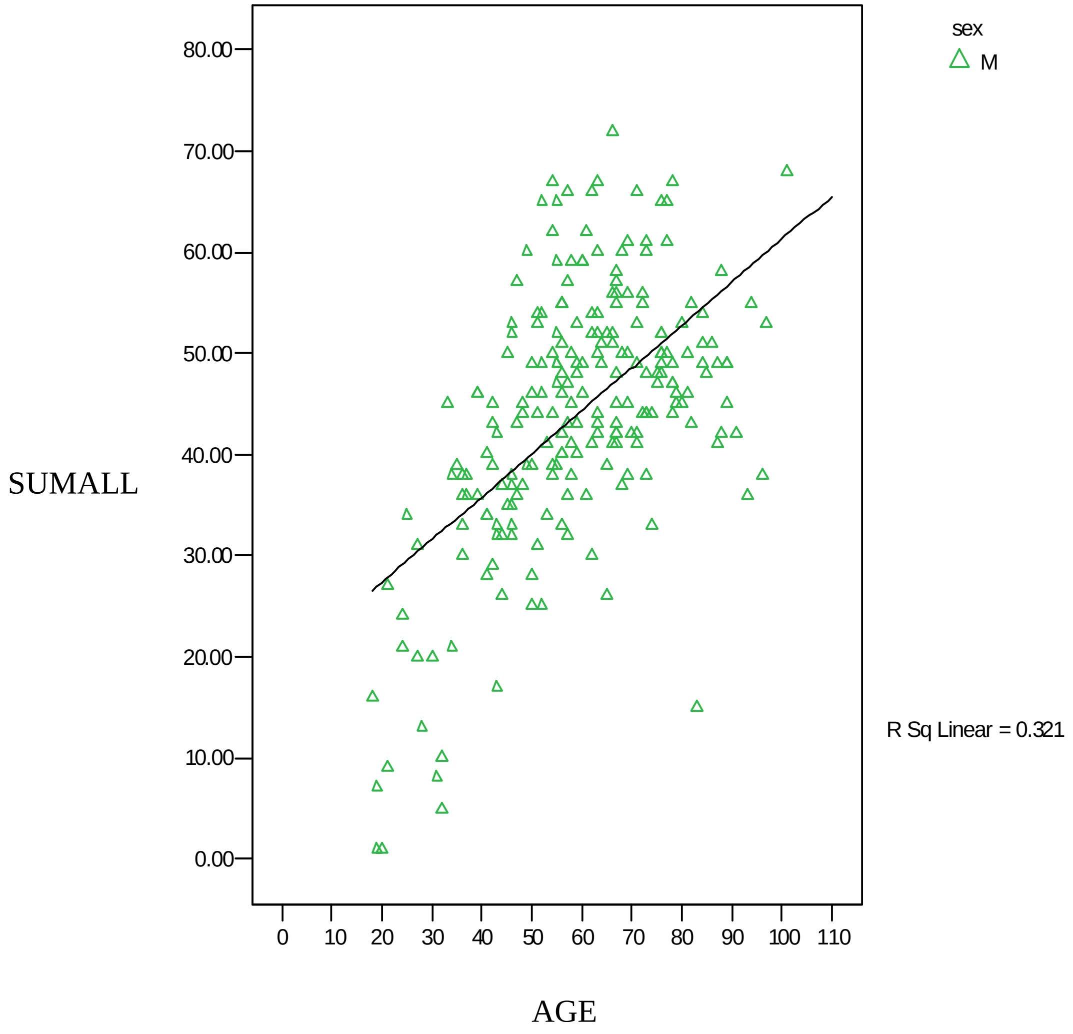

FIGURE 4.2. Plot of SUMALL vs. Age for Males (n = 225). The regression line (y = 0.756x + 25.92) is superimposed on the scatterplot.

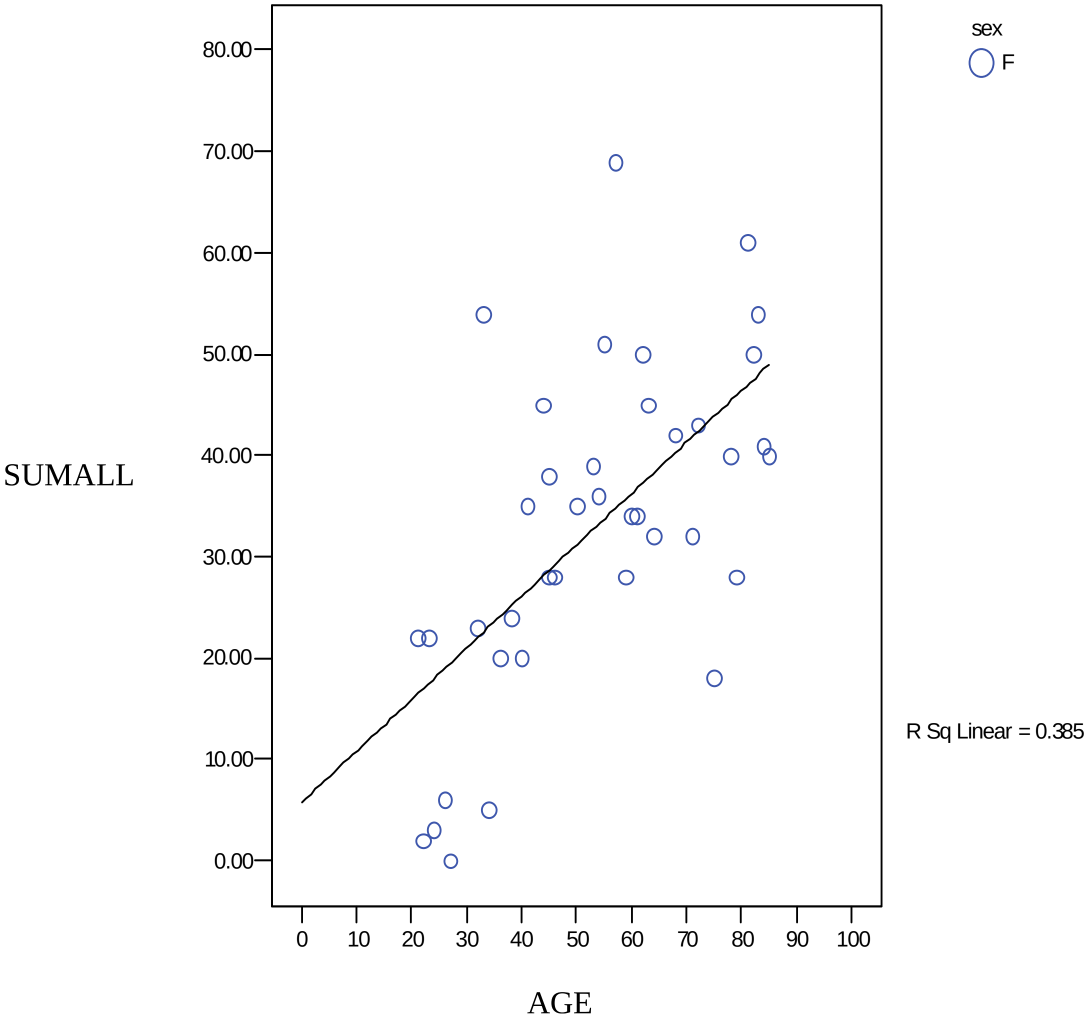

FIGURE 4.3. Plot of SUMALL vs. Age for Females (n = 94). The regression line (y = 0.718x + 35.64) is superimposed on the scatterplot.

TABLE 4.1. Comparison of Rank-Order Correlations (Kendall's Tau-b) Between Each Suture Landmark with Age. Bold values are the largest within-sample correlations; gray cells are the largest between-sample correlations; ns = not a significant correlation. (**Note**: the Terry sample contains individuals of European and African ancestry).

Landmark	**Current study M +F (n)**	**Current study M only (n)**	**Current study F only (n)**	**Terry sample M+F (n)**	**Terry sample M only (n)**	**Terry sample F only (n)**
LLQ	0.239 (385)	0.160 (264)	**0.330** (121)	0.379 (100)	0.368 (50)	**0.389** (50)
LRQ	0.276 (386)	0.200 (265)	**0.387** (121)	0.443 (100)	**0.452** (50)	0.432 (50)
LAQ	0.191 (384)	0.157 (263)	**0.270** (121)	0.423 (100)	0.415 (50)	**0.460** (50)
OBQ	ns (386)	ns (266)	**0.181** (120)	0.361 (100)	0.278 (50)	**0.447** (50)
ASQ	0.177 (387)	0.134 (267)	**0.292** (120)	0.308 (100)	0.219 (50)	**0.416** (50)
BRQ	0.236 (388)	0.202 (267)	**0.298** (121)	0.374 (100)	0.284 (50)	**0.451** (50)
CLQ	0.162 (385)	ns (266)	**0.320** (119)	0.413 (100)	0.293 (50)	**0.541** (50)
CRQ	0.194 (386)	0.104 (266)	**0.315** (120)	0.431 (100)	0.330 (50)	**0.539** (50)
CLI	0.218 (346)	0.211 (243)	**0.223** (120)	n/a	n/a	n/a
CRI	0.271 (343)	**0.285** (241)	0.226 (102)	n/a	n/a	n/a
PLQ	0.250 (376)	**0.326** (260)	0.239 (116)	0.535 (100)	**0.539** (50)	0.513 (50)
PRQ	0.215 (373)	**0.313** (257)	0.145 (116)	0.496 (100)	0.426 (50)	**0.555** (50)
SLQ	0.260 (380)	**0.303** (262)	0.289 (118)	0.466 (100)	**0.468** (50)	0.458 (50)
SRQ	0.256 (372)	**0.318** (256)	0.249 (116)	0.449 (100)	0.397 (50)	**0.481** (50)
ILQ	0.246 (382)	**0.344** (263)	0.167 (119)	0.202 (100)	**0.331** (50)	ns (50)
IRQ	0.267 (379)	**0.386** (256)	0.160 (120)	0.263 (100)	**0.402** (50)	ns (50)
TLQ	0.158 (372)	**0.276** (258)	ns (114)	0.174 (100)	**0.346** (50)	ns (50)
TRQ	0.189 (374)	**0.330** (258)	ns (116)	0.272 (100)	**0.460** (50)	ns (50)
LLZ	0.283 (382)	0.266 (265)	**0.336** (117)	0.495 (100)	**0.516** (50)	0.475 (50)
LRZ	0.276 (384)	0.210 (266)	**0.414** (118**)**	0.456 (100)	**0.503** (50)	0.430 (50)
LAZ	0.356 (382)	0.339 (264)	**0.376** (118)	0.401 (100)	0.368 (50)	**0.428** (50)
SAZ	0.292 (382)	0.257 (264)	**0.365** (118)	0.346 (100)	**0.367** (50)	0.336 (50)
BRZ	0.359 (385)	0.321 (267)	**0.419** (118)	0.516 (100)	0.461 (50)	**0.557** (50)
CLZ	**0.325** (384)	0.318 (266)	**0.325** (118)	0.476 (100)	0.439 (50)	**0.509** (50)
CRZ	0.305 (383)	0.288 (266)	**0.326** (117)	0.467 (100)	0.460 (50)	**0.471** (50)
ICP	0.255 (370)	0.265 (259)	**0.287** (111)	ns (100)	ns (50)	ns (50)
AMP	0.244 (369)	**0.347** (258)	0.209 (111)	0.381 (100)	**0.457** (50)	0.312 (50)
TRP	0.246 (361)	**0.298** (256)	0.246 (105)	0.408 (100)	0.400 (50)	**0.459** (50)
PMP	0.226 (368)	0.270 (257)	**0.329** (111)	0.365 (100)	**0.451** (50)	0.280 (50)
ZLM	ns (338)	**0.169** (237)	ns (101)	n/a	n/a	n/a (50)
ZRM	ns (341)	**0.121** (241)	ns (100)	n/a	n/a	n/a (50)
SUMALL	0.312 (319)	**0.356** (225)	0.318 (94)	0.516 (100)	0.506 (50)	**0.542** (50)

CRQ), whereas males in both samples have higher correlation coefficients for the lateral-anterior series of sutures (CLI through TRQ). Compared to males, females in the current study sample have higher correlation coefficients for all of the endocranial suture landmarks (LLZ through CRZ). No trend is present in the Terry endocranial landmarks or the palatine landmarks (ICP through PMP) in either sample. The correlation coefficients for males and females combined for both samples are always lower than the coefficients for either males or females alone, with the exception of CLZ, for which there is a tie. The Terry sample has higher correlation values for SUMALL, with females having the highest within-sample value. The males in the current study sample have the highest within-sample value for SUMALL but still lower than those obtained for the Terry sample.

Table 4.2 compares Terry sample Euro-Americans to the current study sample. As before, the Terry sample tends to have slightly to moderately higher correlation coefficients than the current study sample, but there is less disparity than was seen when including African-Americans. Females tend to display stronger correlations than males in both samples. Terry sample Euro-American females have higher correlations than Terry sample Euro-American males for the vault series (LLQ through CRQ) of sutures (most of these correlations are insignificant in males). The pattern evident for the combined Terry Euro- and African-American males in the lateral-anterior series (CLI through TRQ) is not as clear in the Terry Euro-American male subgroup. There is no clear pattern of sexual dimorphism in endocranial or palatine suture closure between Terry Euro-American males and females. Once again, the Terry sample has higher correlation values for SUMALL, with females having the highest within-sample value.

Pearson’s correlations were calculated for SUMALL for both samples (Table 4.3). The Terry sample has higher correlation coefficients than the current study sample. Males in the

TABLE 4.2. Comparison of Rank-Order Correlations (Kendall's Tau-b) Between Each Suture Landmark with Age. Bold values are the largest within-sample correlations; gray cells are the largest between-sample correlations; ns = not a significant correlation. (**Note**: only Euro-Americans are included in the Terry sample).

Landmark	Current study M +F (n)	Current study M only (n)	Current study F only (n)	Terry sample M+F (n)	Terry sample M only (n)	Terry sample F only (n)
LLQ	0.239 (385)	0.160 (264)	**0.330** (121)	0.385 (49)	ns (24)	**0.479** (25)
LRQ	0.276 (386)	0.200 (265)	**0.387** (121)	0.434 (49)	0.365 (24)	**0.458** (25)
LAQ	0.191 (384)	0.157 (263)	**0.270** (121)	0.384 (49)	0.398 (24)	**0.430** (25)
OBQ	ns (386)	ns (266)	**0.181** (120)	0.345 (49)	ns (24)	**0.526** (25)
ASQ	0.177 (387)	0.134 (267)	**0.292** (120)	0.240 (49)	ns (24)	**0.390** (25)
BRQ	0.236 (388)	0.202 (267)	**0.298** (121)	0.248 (49)	ns (24)	**0.371** (25)
CLQ	0.162 (385)	ns (266)	**0.320** (119)	0.334 (49)	ns (24)	**0.566** (25)
CRQ	0.194 (386)	0.104 (266)	**0.315** (120)	0.369 (49)	ns (24)	**0.508** (25)
CLI	0.218 (346)	0.211 (243)	**0.223** (120)	n/a	n/a	n/a
CRI	0.271 (343)	**0.285** (241)	0.226 (102)	n/a	n/a	n/a
PLQ	0.250 (376)	**0.326** (260)	0.239 (116)	0.484 (49)	**0.619** (24)	0.393 (25)
PRQ	0.215 (373)	**0.313** (257)	0.145 (116)	0.446 (49)	**0.470** (24)	0.440 (25)
SLQ	0.260 (380)	**0.303** (262)	0.289 (118)	0.463 (49)	0.461 (24)	**0.501** (25)
SRQ	0.256 (372)	**0.318** (256)	0.249 (116)	0.403 (49)	0.362 (24)	**0.460** (25)
ILQ	0.246 (382)	**0.344** (263)	0.167 (119)	ns (49)	ns (24)	ns (25)
IRQ	0.267 (379)	**0.386** (256)	0.160 (120)	0.249 (49)	**0.411** (24)	ns (25)
TLQ	0.158 (372)	**0.276** (258)	ns (114)	ns (49)	ns (24)	ns (25)
TRQ	0.189 (374)	**0.330** (258)	ns (116)	ns (49)	**0.377** (24)	ns (25)
LLZ	0.283 (382)	0.266 (265)	**0.336** (117)	0.526 (49)	**0.581** (24)	0.518 (25)
LRZ	0.276 (384)	0.210 (266)	**0.414** (118)	0.447 (49)	**0.516** (24)	0.451 (25)
LAZ	0.356 (382)	0.339 (264)	**0.376** (118)	0.359 (49)	ns (24)	**0.431** (25)
SAZ	0.292 (382)	0.257 (264)	**0.365** (118)	0.279 (49)	**0.369** (24)	ns (25)
BRZ	0.359 (385)	0.321 (267)	**0.419** (118)	**0.480** (49)	0.473 (24)	0.472 (25)
CLZ	**0.325** (384)	0.318 (266)	**0.325** (118)	0.426 (49)	ns (24)	**0.543** (25)
CRZ	0.305 (383)	0.288 (266)	**0.326** (117)	0.382 (49)	0.342 (24)	**0.442** (25)
ICP	0.255 (370)	0.265 (259)	**0.287** (111)	**0.300** (49)	ns (24)	ns (25)
AMP	0.244 (369)	**0.347** (258)	0.209 (111)	0.465 (49)	**0.509** (24)	0.436 (25)
TRP	0.246 (361)	**0.298** (256)	0.246 (105)	0.365 (49)	ns (24)	**0.469** (25)
PMP	0.226 (368)	0.270 (257)	**0.329** (111)	0.323 (49)	0.362 (24)	**0.383** (25)
ZLM	ns (338)	**0.169** (237)	ns (101)	n/a	n/a	n/a
ZRM	ns (341)	**0.121** (241)	ns (100)	n/a	n/a	n/a
SUMALL	0.312 (319)	**0.356** (225)	0.318 (94)	0.439 (49)	0.369 (24)	**0.550** (25)

TABLE 4.3. Comparison of Pearson's Correlations Between SUMALL and Age. Bold values signify the largest within-sample correlation values.

Sample	**SUMALL**	**n**
Current Study Sample (M + F)	0.529	319
Current Study Sample (M only)	**0.566**	225
Current Study Sample (F only)	0.556	94
Terry Sample (M + F)	0.715	100
Terry Sample (M only)	0.702	50
Terry Sample (F only)	**0.751**	50
Terry Sample Euro-Americans (M+ F)	0.633	49
Terry Sample Euro-American (M only)	0.609	24
Terry Sample Euro-American (F only)	**0.738**	25

current study sample have a higher correlation than females, although the difference is probably not significantly different. Terry sample females (Euro- & African-Americans combined) have the highest within-sample correlation value and the largest overall value.

In summary, it appears that correlations between cranial suture closure and age at death are higher in the Terry sample than the current (modern) study sample, and that the inclusion of African-Americans in the Terry sample increases the correlations. In addition, correlations are higher for females in the vault system while males are higher for the lateral-anterior system.

Performance of the Equations

Equation 1. Equation 1 (Nawrocki, 1998) worked best for males in the current study sample (Tables 4.4 & 4.5), with an inaccuracy of 11.34 years and a bias of -2.62 years. For the male sample, 65% of the individuals have their actual age falling within the ± 1 se interval as given for Nawrocki's equation, and 94% fall within the ± 2 se interval. With the males and females combined, inaccuracy and bias increased slightly to 12.54 and -4.99 years, respectively.

TABLE 4.4. Inaccuracy and Bias Statistics in Years for Nawrocki's (1998) Equations 1 through 4, 7, & 8 for the Current Study Sample.

Equation	n	Inaccuracy	Bias
Equation 1 (M + F)	319	12.54	-4.99
Equation 1 (M only)	225	11.34	-2.62
Equation 1 (F only)	94	15.43	-10.65
Equation 2 (M + F)	351	13.10	-6.98
Equation 2 (M only)	251	11.52	-4.42
Equation 2 (F only)	100	17.07	-13.41
Equation 4 (all M)	252	12.38	-1.92
Equation 8 (Euro-American M)	260	15.54	-7.81
Equation 3 (all F)	102	20.80	-17.01
Equation 7 (Euro-American F)	114	19.48	-2.10

TABLE 4.5. Percentage of Individuals Whose Actual Age Falls Within the ± 1 se & ± 2 se Interval for Each Equation.

Equation	% in 1 se	% in 2 se	n
Equation 1 (M+F)	60	89	319
Equation 1 (M)	65	94	225
Equation 1 (F)	48	79	94
Equation 2 (M+F)	58	85	351
Equation 2 (M)	64	89	257
Equation 2 (F)	42	76	100
Equation 4 (M)	62	88	252
Equation 8 (M)	44	75	260
Equation 3 (F)	31	55	102
Equation 7 (F)	26	49	114

For males and females combined, 60% of these individuals have their actual ages falling within the ± 1 se interval and 89% fall within the ± 2 se interval. Equation 1 performs worse on the female sample, with an inaccuracy of 15.43 years and a bias of -10.65 years. Only 48% of females have their actual ages falling within the ± 1 se interval, and only 79% fall within the ± 2 se interval. Although it appears that Equation 1 works best on the male sample, when inaccuracy and bias are separated by sex and calculated by decade (Table 4.6), different trends become apparent. The female sample actually has lower inaccuracy and bias values for ages between 18 and 59 years. For ages above 60 years, females consistently have higher inaccuracy and bias values. Overall, the values for inaccuracy and bias are high in the earlier decades and tend to decrease as the 6th decade of life is approached. After the 6th decade, the error values begin to increase again. Bias values are positive in the earlier decades (indicating systematic overestimation of age) and approach zero in the 6th decade. After the 6th decade, bias becomes negative (indicating systematic underestimation of age) and steadily increases. Errors are particularly high after 80 years.

Equation 2. Equation 2 also performs better on the males in the current study sample (Tables 4.4 & 4.5) with an inaccuracy of 11.52 years and a bias of -4.42 years. For the male sample, 64% have their actual age falling within the ± 1 se interval, and 89% fall within the ± 2 se interval. With males and females combined, inaccuracy and bias rises slightly to 13.10 and -6.98 years, respectively. For males and females combined, only 58% of these individuals have their actual ages falling within the ± 1 se interval, and 85% fall within the ± 2 se interval. Equation 2 performs worse on the female sample, with inaccuracy and bias values increasing to 17.07 and -13.41 years, respectively. Only 42% of the females have their actual age falling within the ± 1 se interval, and only 76% fall within the ± 2 se interval. When inaccuracy and

TABLE 4.6. Inaccuracy and Bias Statistics for Equation 1 for the Current Study Sample, Separated by Sex and Decade.

Decade	Inaccuracy		Bias		n (n = 319)	
	M	F	M	F	M	F
18-29	14.62	10.28	14.62	9.90	12	9
30-39	12.54	10.65	12.16	8.59	17	5
40-49	8.39	5.20	8.03	4.76	32	6
50-59	6.36	5.52	3.44	0.93	55	14
60-69	6.79	12.78	-4.21	-11.81	47	22
70-79	13.96	18.34	-13.89	-18.34	37	18
80-89	26.01	27.46	-26.01	-27.46	19	16
90-99	37.09	36.50	-37.09	-36.50	5	4
100-109	27.42	--	-27.42	--	1	0

TABLE 4.7. Inaccuracy and Bias Statistics for Equation 2 for the Current Study Sample, Separated by Sex and Decade.

Decade	Inaccuracy		Bias		n (n = 351)	
	M	F	M	F	M	F
18-29	7.57	7.52	7.49	6.72	16	11
30-39	9.73	10.53	7.84	6.65	22	5
40-49	7.72	7.22	4.85	3.14	35	6
50-59	7.42	5.22	1.31	5.22	62	14
60-69	8.88	14.64	-5.56	-14.28	52	23
70-79	14.16	21.34	-14.16	-21.34	39	19
80-89	29.99	30.32	-29.99	-30.32	19	18
90-99	40.35	41.84	-40.35	-41.84	5	4
100-109	40.75	--	-40.75	--	1	0

bias were calculated by decade (Table 4.7), a trend similar to that described for Equation 1 becomes evident. The female sample tends to have slightly lower inaccuracy and bias values for ages between 18 and 49 years. For ages above 60 years, females tend to have higher inaccuracy and bias values. The values for inaccuracy and bias tend to decrease and approach zero as the 6thth decade of life is approached. These values increase after the 6th decade. After the 6th decade, bias becomes negative and steadily increases. Errors are particularly high after 80 years.

<u>Equation 4</u>. This males-only equation has an inaccuracy of 12.38 years and a bias of -1.92 years (Tables 4.4 & 4.5). Only 62% have their actual ages falling within the ± 1 se interval, and 88% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, the trends seen in the previous equations are apparent for bias (Table 4.8). Inaccuracy values between the ages of 18 and 79 years are similar in magnitude. Errors are particularly high after 80 years.

<u>Equation 8</u>. This Euro-American male equation has the highest inaccuracy and bias values for any of the equations when applied to the male sample (Tables 4.4 & 4.5). Inaccuracy and bias values are 15.54 and -7.81 years, respectively. Only 44% of the males have their actual ages falling within the ± 1 se interval, and only 75% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, trends similar to that seen for Equation 4 are present (Table 4.9). However, bias becomes negative in the 5th decade, earlier than for the other equations. Errors are particularly high after 80 years.

<u>Equation 3</u>. Equation 3 for all females is the equation with the highest overall error values, with an inaccuracy of 20.80 years and a bias of -17.01 years (Tables 4.4 & 4.5). Only 31% of the females have their actual ages falling within the ± 1 se interval, and only 55% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, trends similar to the ones seen for Equations 1 and 2 are seen (Table 4.11). Bias approaches zero in the 5th decade and becomes negative in the 6th decade. The error values are particularly high after 70 years.

Equation 7. Equation 7 for Euro-American females (Tables 4.4 & 4.5) has the second highest inaccuracy value at 19.48 years. However, it has the second lowest bias value at -2.10 years. Only 26% of females have their actual ages falling within the ± 1 se interval, and only 49% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, the usual trends occur, however, error values are particularly high between ages 20-39 years and over 80 years (Table 4.11).

ANCOVA results. As expected, age has a significant effect on overall suture closure, as does sex (Tables 4.12, 4.13, & 4.14). There is not a significant difference between the Terry sample and the current study sample, however, the moderately high F value is suggestive in both tests (Tables 4.13 & 4.14). The interaction between sex and sample (Sex*Sample) is not significant in either test (Tables 4.13 & 4.14). These ANCOVA results reaffirm the sexual dimorphism in the sample as reflected by the Kendall and Pearson correlations discussed earlier.

TABLE 4.8. Inaccuracy and Bias Statistics by Decade for Equation 4, All Males.

Decade	Inaccuracy	Bias	n (n =252)
18-29	9.57	9.35	16
30-39	10.62	6.60	23
40-49	9.26	4.20	34
50-59	12.14	4.70	61
60-69	10.25	-3.25	51
70-79	9.56	-7.27	41
80-89	27.22	-27.22	20
90-99	34.68	-34.68	5
100-109	35.92	-35.92	1

TABLE 4.9. Inaccuracy and Bias Statistics by Decade for Equation 8, Euro-American Males Only.

Decade	Inaccuracy	Bias	n (n = 260)
18-29	12.91	12.79	17
30-39	13.43	6.69	23
40-49	13.50	-1.01	36
50-59	11.73	-3.19	64
60-69	12.27	-10.24	52
70-79	16.43	-16.18	42
80-89	33.96	-33.96	20
90-99	43.40	-43.40	5
100-109	53.36	-53.36	1

TABLE 4.10. Inaccuracy and Bias Statistics by Decade for Equation 3, All Females.

Decade	Inaccuracy	Bias	n (n = 102)
20-29	9.50	9.45	10
30-39	7.78	5.45	5
40-49	4.12	0.39	6
50-59	10.32	-5.66	14
60-69	18.81	-17.19	24
70-79	25.44	-25.44	20
80-89	35.58	-35.58	18
90-99	44.05	-44.05	4
100-109	42.11	-42.11	1

TABLE 4.11. Inaccuracy and Bias Statistics by Decade for Equation 7, Euro-American Females Only.

Decade	Inaccuracy	Bias	n (n = 114)
20-29	28.27	28.27	12
30-39	22.37	22.37	8
40-49	14.92	5.46	7
50-59	16.11	8.16	16
60-69	13.14	13.14	24
70-79	19.72	-14.40	21
80-89	22.97	-18.05	21
90-99	25.59	-25.59	5

TABLE 4.12. ANCOVA Results for the Current Study Sample (n = 319) with SUMALL as the Dependent Variable.

Source	df	F	Sig.
Age	1	146.238	0.000
Sex	1	32.396	0.000
$r^2 = 0.347$ (Adj. $r^2 = 0.342$)			

TABLE 4.13. ANCOVA Results Comparing the Current Study Sample (n = 319) to the Terry Sample (n = 100) with SUMALL as the Dependent Variable. (**Note:** the Terry sample contains individuals of Euro- and African-American ancestry).

Source	df	F	Sig
Age	1	240.832	0.000
Sex	1	26.186	0.000
Sample	1	3.812	0.052
Sex*Sample	1	0.580	ns
$r^2 = 0.394$ (Adj. $r^2 = 0.388$)			

TABLE 4.14. ANCOVA Results Comparing the Current Study Sample (n = 319) to the Terry Euro-American Sample (n = 49) with SUMALL as the Dependent Variable.

Source	df	F	Sig
Age	1	182.444	0.000
Sex	1	34.496	0.000
Sample	1	2.992	ns
Sex*Sample	1	1.987	ns
$r^2 = 0.377$ (Adj. $r^2 = 0.370$)			

Results for the Culled Study Sample

The preceding statistical analyses suggest that differences exist in the rate or pattern of suture closure due to sex and possibly between the current study sample and Nawrocki's (1998) Terry sample. In order to verify that the differences are real and not due to sampling biases, it is necessary to cull the current study sample. By culling the sample, errors caused by differences between the samples in their mean ages at death should be reduced. Samples with higher mean age at death would be expected to perform more poorly regardless of the applicability of the equations.

Plots. The summed suture scores (SUMALL) for Nawrocki's (1998) 27 landmarks were plotted against age for males and females combined and separated by sex. As before, as age increases, SUMALL tends to increase as well (Figures 4.4, 4.5, & 4.6), and the ranges for the SUMALL values appear to be relatively large for any given decade. Males (Figure 4.5) appear to be clustered more closely together than females (Figure 4.6). The plots for the culled sample are very similar in appearance to the plots for the un-culled sample.

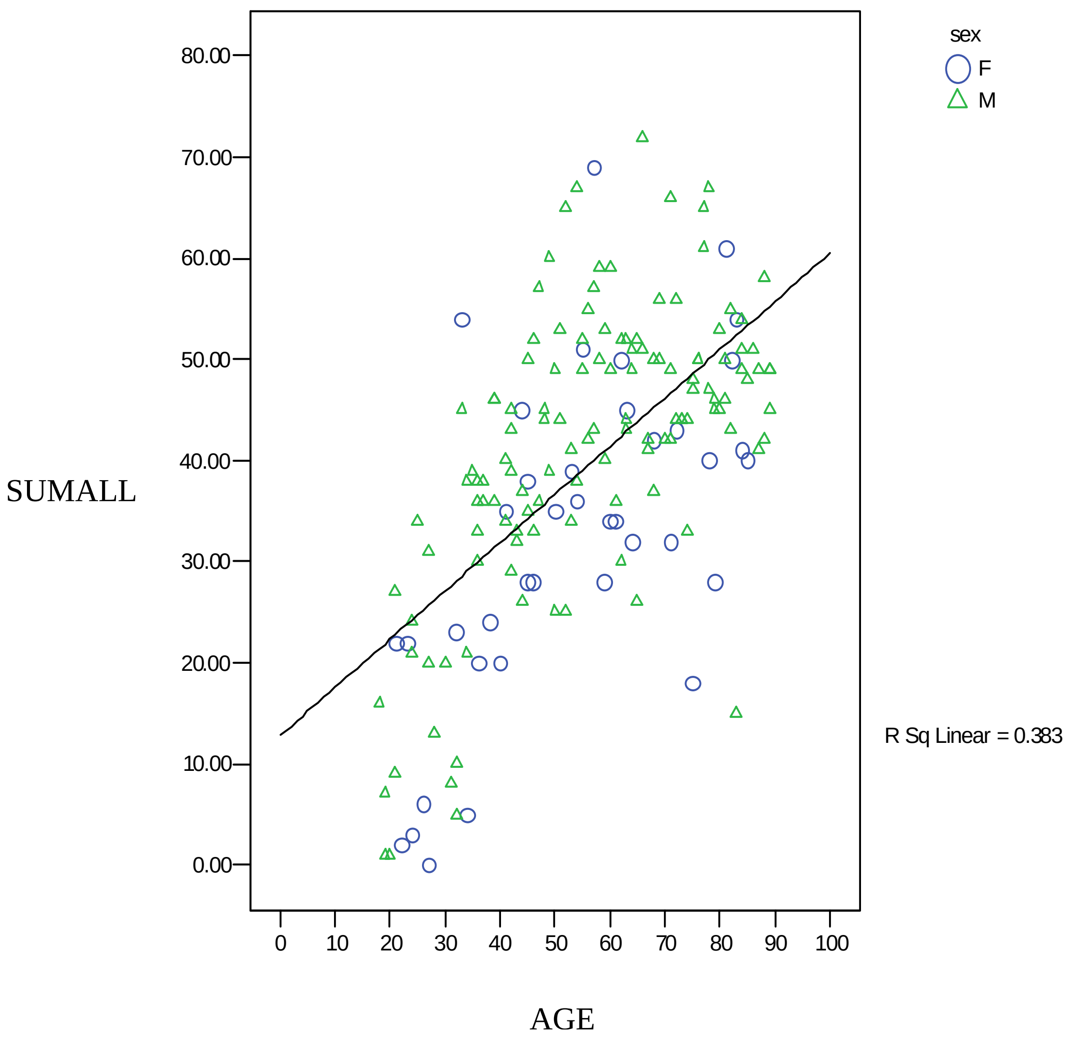

FIGURE 4.4. Plot of SUMALL vs. Age for Males and Females of the Culled Study Sample (n = 167). The regression line (y = 0.803x + 24.14) is superimposed on the scatterplot.

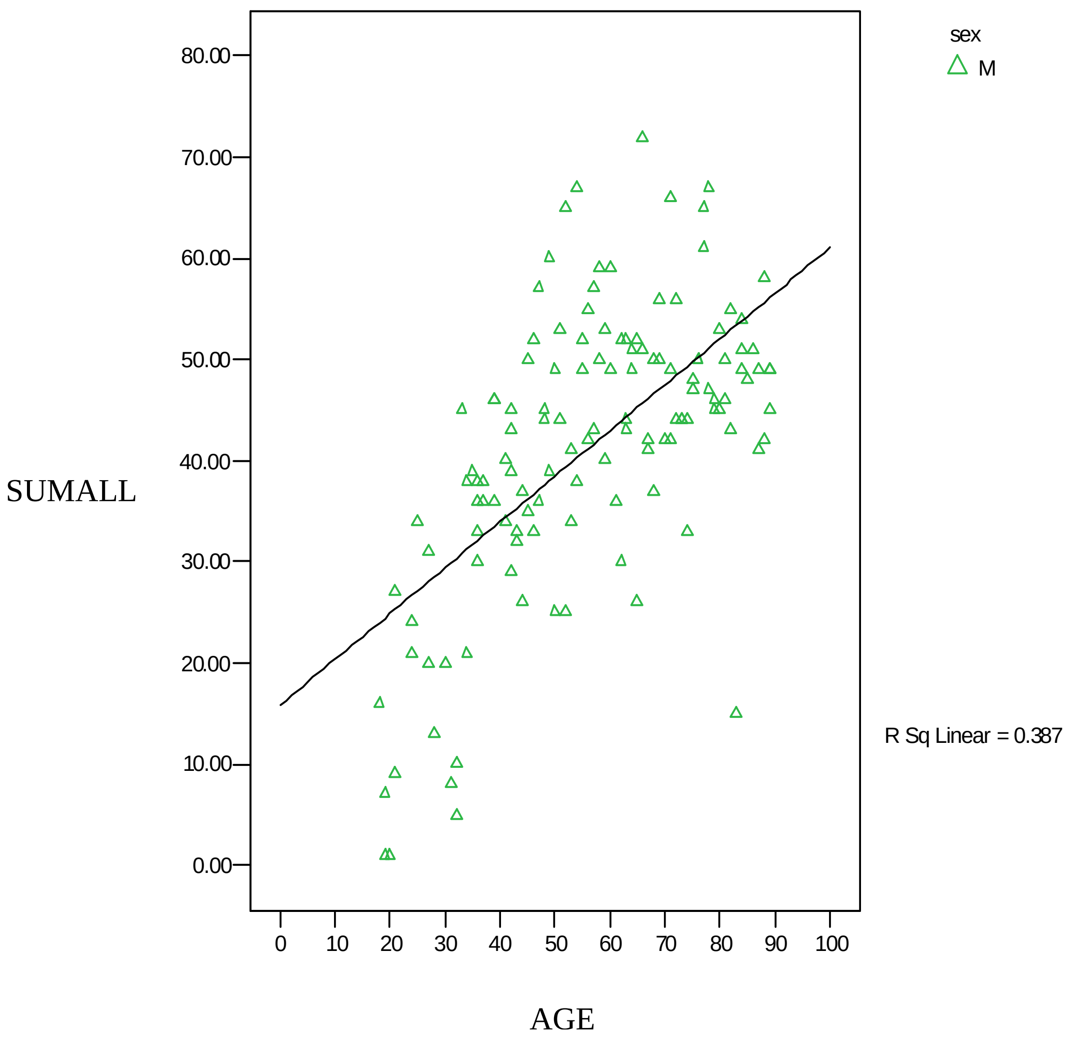

FIGURE 4.5. Plot of SUMALL vs. Age for Males of the Culled Study Sample (n = 128). The regression line (y = 0.887x + 21.07) is superimposed on the scatterplot.

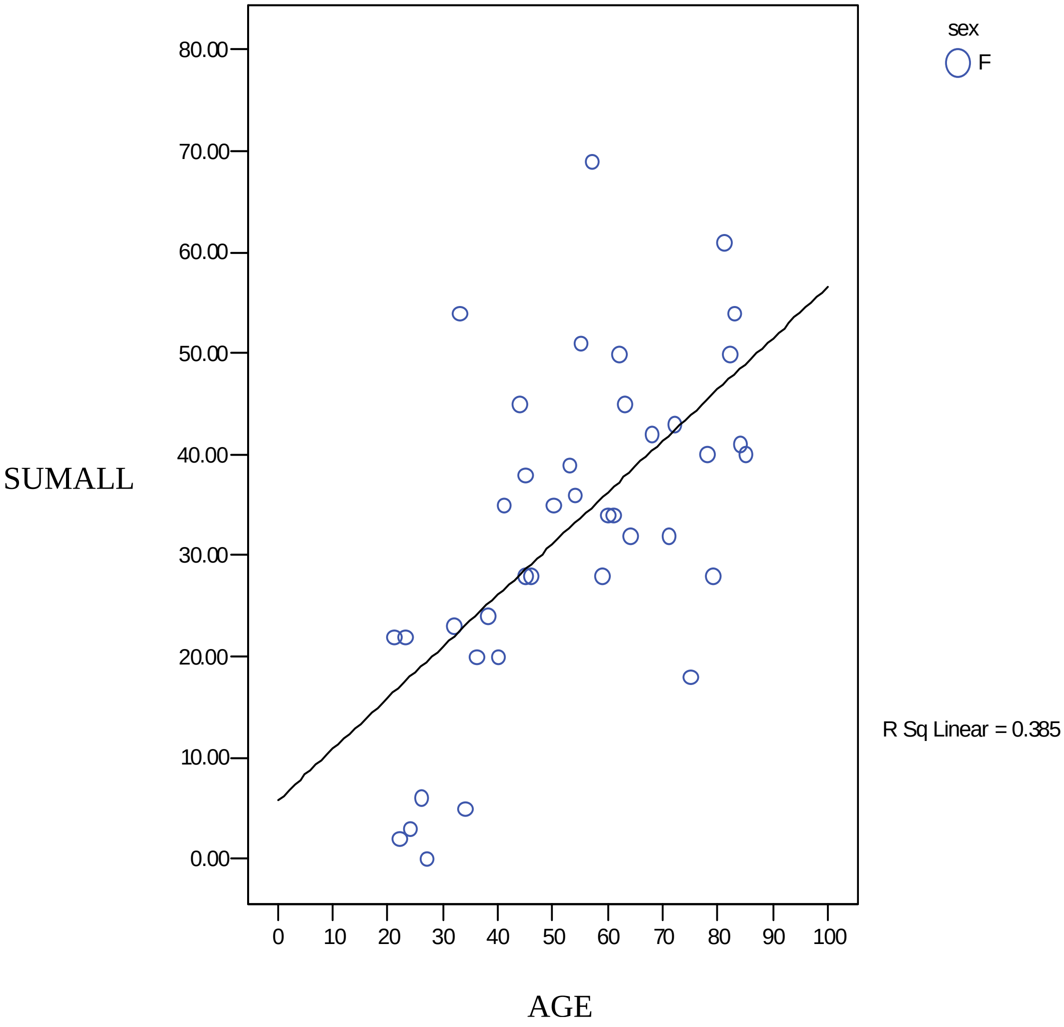

FIGURE 4.6. Plot of SUMALL vs. Age for Females of the Culled Study Sample (n = 39). The regression line (y = 0.758x + 28.33) is superimposed on the scatterplot.

Correlations. Kendall's Tau-b (rank-order) correlations were calculated to examine the relationship between age and suture closure. Correlations were compared directly to Nawrocki's (1998) Terry Collection sample (Tables 4.14 & 4.15). Table 4.15 compares Nawrocki's (1998) Euro- and African-American individuals to the culled study sample of Euro-Americans. The general trends seen in Table 4.1 are also apparent in Table 4.15. The gray cells indicate the highest between-sample correlations for a given landmark. Bolded values within the table represent the highest within-sample correlation for a given landmark. Overall, the Terry sample still has slightly to moderately higher correlation coefficients than the culled study sample, as indicated by the greater number of gray cells in the 3 right columns. Females of the culled study sample tend to display higher correlation coefficients than the males of the culled sample. Females continue to display higher correlation coefficients for the vault suture series (LLQ through CRQ), whereas males have higher values for the lateral-anterior suture series (CLI through TRQ). Females also have higher correlation coefficients than males for all of the endocranial landmarks (LLQ through CRZ).

The number of significant correlations has been reduced in the culled study sample, particularly for the females. The trends in the palatine sutures are similar to the un-culled sample. Culled females have the highest within-sample correlation coefficient for SUMALL, whereas males have the highest value for the un-culled sample. The culled study sample has higher between-sample correlation coefficients for endocranial lambda (LAZ). The correlation coefficients for the culled males and females combined are lower than the coefficient for either individual subgroup (except for CLI, CRI, & TRP). In general, when compared to the un-culled sample (Table 4.1), the correlation coefficients have increased for the culled study sample (Table 4.15), however, these values are still consistently lower than for the Terry sample.

TABLE 4.15. Comparison of Rank-Order Correlations (Kendall's Tau-b) Between Each Suture Landmark with Age. Bold values are the largest within-sample correlation; gray cells are the largest between-sample correlations; ns = not a significant correlation. (**Note**: the Terry sample contains individuals of European and African ancestry).

Landmark	Current study M +F (n)	Current study M only (n)	Current study F only (n)	Terry sample M+F (n)	Terry sample M only (n)	Terry sample F only (n)
LLQ	0.194 (167)	ns	**0.397** (39)	0.379 (100)	0.368 (50)	**0.389** (50)
LRQ	0.259 (167)	0.199 (128)	**0.430** (39)	0.443 (100)	**0.452** (50)	0.432 (50)
LAQ	0.200 (167)	ns	**0.411** (39)	0.423 (100)	0.415 (50)	**0.460** (50)
OBQ	ns	ns	Ns	0.361 (100)	0.278 (50)	**0.447** (50)
ASQ	0.204 (167)	0.167 (128)	**0.318** (39)	0.308 (100)	0.219 (50)	**0.416** (50)
BRQ	0.237 (167)	0.206 (128)	**0.340** (39)	0.374 (100)	0.284 (50)	**0.451** (50)
CLQ	0.156 (167)	ns	**0.326** (39)	0.413 (100)	0.293 (50)	**0.541** (50)
CRQ	0.182 (167)	0.149 (128)	**0.315** (39)	0.431 (100)	0.330 (50)	**0.539** (50)
CLI	**0.220** (167)	0.217 (128)	Ns	n/a	n/a	n/a
CRI	**0.382** (167)	0.369 (128)	0.336 (39)	n/a	n/a	n/a
PLQ	0.396 (167)	**0.427** (128)	0.282 (39)	0.535 (100)	**0.539** (50)	0.513 (50)
PRQ	0.352 (167)	**0.417** (128)	Ns	0.496 (100)	0.426 (50)	**0.555** (50)
SLQ	0.355 (167)	**0.393** (128)	Ns	0.466 (100)	**0.468** (50)	0.458 (50)
SRQ	0.359 (167)	**0.407** (128)	Ns	0.449 (100)	0.397 (50)	**0.481** (50)
ILQ	0.304 (167)	**0.332** (128)	Ns	0.202 (100)	**0.331** (50)	ns (50)
IRQ	0.319 (167)	**0.369** (128)	Ns	0.263 (100)	**0.402** (50)	ns (50)
TLQ	0.285 (167)	**0.324** (128)	Ns	0.174 (100)	**0.346** (50)	ns (50)
TRQ	0.354 (167)	**0.399** (128)	Ns	0.272 (100)	**0.460** (50)	ns (50)
LLZ	0.349 (167)	0.320 (128)	**0.433** (39)	0.495 (100)	**0.516** (50)	0.475 (50)
LRZ	0.325 (167)	0.255 (128)	**0.531** (39)	0.456 (100)	**0.503** (50)	0.430 (50)
LAZ	0.402 (167)	0.381 (128)	**0.460** (39)	0.401 (100)	0.368 (50)	**0.428** (50)
SAZ	0.310 (167)	0.270 (128)	**0.429** (39)	0.346 (100)	**0.367** (50)	0.336 (50)
BRZ	0.374 (167)	0.342 (128)	**0.477** (39)	0.516 (100)	0.461 (50)	**0.557** (50)
CLZ	0.340 (167)	0.339 (128)	**0.363** (39)	0.476 (100)	0.439 (50)	**0.509** (50)
CRZ	0.310 (167)	0.294 (128)	**0.378** (39)	0.467 (100)	0.460 (50)	**0.471** (50)
ICP	0.318 (167)	0.289 (128)	**0.396** (39)	ns (100)	ns (50)	ns (50)
AMP	**0.332** (167)	**0.332** (128)	0.312 (39)	0.381 (100)	**0.457** (50)	0.312 (50)
TRP	**0.323** (167)	0.320 (128)	0.314 (39)	0.408 (100)	0.400 (50)	**0.459** (50)
PMP	0.359 (167)	0.363 (128)	**0.390** (39)	0.365 (100)	**0.451** (50)	0.280 (50)
ZLM	ns	ns	ns	n/a	n/a	n/a
ZRM	ns	ns	ns	n/a	n/a	n/a
SUMALL	0.420 (167)	0.421 (128)	**0.441** (39)	0.516 (100)	0.506 (50)	**0.542** (50)

TABLE 4.16. Comparison of Rank-Order Correlations (Kendall's Tau-b) Between Each Suture Landmark with Age. Bold values are the largest within-sample correlation; gray cells are the largest between sample correlations; ns = not a significant correlation. (**Note**: only Euro-Americans are included in the Terry sample).

Landmark	Current study M +F (n)	Current study M only (n)	Current study F only (n)	Terry sample M+F (n)	Terry sample M only (n)	Terry sample F only (n)
LLQ	0.194 (167)	ns	**0.397** (39)	0.385 (49)	ns (24)	**0.479** (25)
LRQ	0.259 (167)	0.199 (128)	**0.430** (39)	0.434 (49)	0.365 (24)	**0.458** (25)
LAQ	0.200 (167)	ns	**0.411** (39)	0.384 (49)	0.398 (24)	**0.430** (25)
OBQ	ns	ns	ns	0.345 (49)	ns (24)	**0.526** (25)
ASQ	0.204 (167)	0.167 (128)	**0.318** (39)	0.240 (49)	ns (24)	**0.390** (25)
BRQ	0.237 (167)	0.206 (128)	**0.340** (39)	0.248 (49)	ns (24)	**0.371** (25)
CLQ	0.156 (167)	ns	**0.326** (39)	0.334 (49)	ns (24)	**0.566** (25)
CRQ	0.182 (167)	0.149 (128)	**0.315** (39)	0.369 (49)	ns (24)	**0.508** (25)
CLI	**0.220** (167)	0.217 (128)	ns	n/a	n/a	n/a
CRI	**0.382** (167)	0.369 (128)	0.336 (39)	n/a	n/a	n/a
PLQ	0.396 (167)	**0.427** (128)	0.282 (39)	0.484 (49)	**0.619** (24)	0.393 (25)
PRQ	0.352 (167)	**0.417** (128)	ns	0.446 (49)	**0.470** (24)	0.440 (25)
SLQ	0.355 (167)	**0.393** (128)	ns	0.463 (49)	0.461 (24)	**0.501** (25)
SRQ	0.359 (167)	**0.407** (128)	ns	0.403 (49)	0.362 (24)	**0.460** (25)
ILQ	0.304 (167)	**0.332** (128)	ns	ns (49)	ns (24)	ns (25)
IRQ	0.319 (167)	0.369 (128)	ns	0.249 (49)	**0.411** (24)	ns (25)
TLQ	0.285 (167)	**0.324** (128)	ns	ns (49)	ns (24)	ns (25)
TRQ	0.354 (167)	**0.399** (128)	ns	ns (49)	**0.377** (24)	ns (25)
LLZ	0.349 (167)	0.320 (128)	**0.433** (39)	0.526 (49)	**0.581** (24)	0.518 (25)
LRZ	0.325 (167)	0.255 (128)	**0.531** (39)	0.447 (49)	**0.516** (24)	0.451 (25)
LAZ	0.402 (167)	0.381 (128)	**0.460** (39)	0.359 (49)	ns (24)	**0.431** (25)
SAZ	0.310 (167)	0.270 (128)	**0.429** (39)	0.279 (49)	**0.369** (24)	ns (25)
BRZ	0.374 (167)	0.342 (128)	**0.477** (39)	**0.480** (49)	0.473 (24)	0.472 (25)
CLZ	0.340 (167)	0.339 (128)	**0.363** (39)	0.426 (49)	ns (24)	**0.543 (25)**
CRZ	0.310 (167)	0.294 (128)	**0.378** (39)	0.382 (49)	0.342 (24)	**0.442** (25)
ICP	0.318 (167)	0.289 (128)	**0.396** (39)	**0.300** (49)	ns (24)	ns (25)
AMP	**0.332** (167)	**0.332** (128)	0.312 (39)	0.465 (49)	**0.509** (24)	0.436 (25)
TRP	**0.323** (167)	0.320 (128)	0.314 (39)	0.365 (49)	ns (24)	**0.469** (25)
PMP	0.359 (167)	0.363 (128)	**0.390** (39)	0.323 (49)	0.362 (24)	**0.383** (25)
ZLM	ns	ns	ns	n/a	n/a	n/a
ZRM	ns	ns	ns	n/a	n/a	n/a
SUMALL	0.420 (167)	0.421 (128)	**0.441** (39)	0.439 (49)	0.369 (24)	**0.550** (25)

Table 4.16 compares the Terry sample's Euro-Americans to the culled study sample. Overall, the Terry sample still has slightly to moderately higher correlation coefficients than the culled study sample. The trends in Table 4.16 are virtually identical to the trends in Table 4.2 and to Table 4.15. The only mentionable differences occur when the culled study sample has a non-significant correlation (ns) for a landmark that was significant in the un-culled sample (LLQ & LAQ for males; OBQ, CLI, PRQ, SLQ, SRQ, ILQ, & IRQ for females), or when the Terry sample has a non-significant correlation for a landmark and the culled study sample has a significant correlation at the corresponding landmark, thus giving the culled study sample the higher between-sample correlation. Incidences of the culled study sample having the higher between-sample correlations are similar to the ones in Table 4.2, except for the interpalatine (PMP) and incisive (ICP) sutures, which now have slightly higher correlation values in the culled study sample.

Pearson's correlations were calculated for SUMALL using the culled study sample (Table 4.17). As before, the Terry sample has higher correlation coefficients than the culled study sample (Table 4.17). The culled males have the highest within-sample correlation value, however, this value is probably not significantly higher than those obtained for the culled female subgroup or the combined culled male and female groups. The culled study sample has stronger correlation values than the un-culled sample and has correlation values similar to (but still lower than) the Terry Euro-American sample. Terry sample females still have the highest overall correlation coefficients.

In summary, the pattern of correlations obtained for the culled study sample is almost identical to that obtained for the un-culled sample, although the culled correlations tend to be higher than for the un-culled sample. These results suggest that the higher mean age at death for

TABLE 4.17. Comparison of Pearson's Correlations Between SUMALL and Age. Bold values signify the largest within-sample correlation value.

Sample	**SUMALL**	**n**
Culled Study Sample (M + F)	0.619	167
Culled Study Sample (M only)	**0.622**	128
Culled Study Sample (F only)	0.621	39
Un-culled Study Sample (M + F)	0.529	319
Un-culled Study Sample (M only)	**0.566**	225
Un-culled Study Sample (F only)	0.556	94
Terry Sample (M + F)	0.715	100
Terry Sample (M only)	0.702	50
Terry Sample (F only)	**0.751**	50
Terry Sample Euro-Americans only (M+ F)	0.633	49
Terry Sample Euro-American (M only)	0.609	24
Terry Sample Euro-American (F only)	**0.738**	25

TABLE 4.18. Inaccuracy and Bias Statistics in Years for Nawrocki (1998) Equations 1 through 4, 7, & 8 for the Culled Study Sample.

Equation	**Inaccuracy**	**Bias**	**n**
Equation 1 (M & F)	12.76	-2.53	167
Equation 1 (M only)	12.88	-1.88	128
Equation 1 (F only)	12.37	-4.60	39
Equation 2 (M & F)	12.62	-5.82	167
Equation 2 (M only)	12.47	-5.47	128
Equation 2 (F only)	13.10	-6.94	39
Equation 4 (all M)	12.29	-3.60	128
Equation 8 (Euro-American M)	16.30	-9.46	128
Equation 3 (all F)	15.95	-9.46	39
Equation 7 (Euro-American F)	20.26	1.31	39

the un-culled sample are affecting the overall results, but the Terry sample still displays a tighter relationship between age and suture closure.

Performance of the Equations Using the Culled Data Set

Equation 1. Equation 1 has the lowest inaccuracy value of 12.37 years using the culled female sample (Tables 4.18 & 4.19), however, females have the highest bias at -4.60 years. Only 59% of females have their actual ages falling within the ± 1 se interval, and 87% fall within the ± 2 se interval. The male culled sample has the highest inaccuracy at 12.88 years and the lowest bias at -1.88 years. Only 56% of the males have their actual ages falling within the ± 1 se interval, and 94% have their actual ages falling within the ± 2 se interval. The combined culled male and female sample have an inaccuracy of 12.76 years and a bias of -2.53 years. Only 56% percent of the combined sample has their actual ages falling within the ± 1 se interval, and 92% fall within the ± 2 se interval. Compared to the un-culled sample, inaccuracy has increased slightly for the male subgroup and for the combined male and female group. Inaccuracy has decreased for female subgroup. Inaccuracy values for the subgroups of the culled data set are very similar in magnitude. When compared to the un-culled sample, bias values for the culled study sample have been reduced.

When inaccuracy and bias are calculated by decade, the same trends seen in the un-culled sample are apparent (Table 4.20). Females tend to have lower inaccuracy and bias values for ages between 18 and 59 years. For ages above 60 years, females tend to have higher inaccuracy and bias values. The values for inaccuracy and bias are larger in the earlier decades, and these values decrease as the 6th decade of life is approached. After the 6th decade, the error values begin to increase again. Bias values are positive in the earlier decades (indicating systematic

TABLE 4.19. Percentage of individuals Whose Actual Age Falls Within the ±1 se & ±2 se Intervals for the Culled Study Sample.

Equation	% in 1 se	% in 2 se	n
Equation 1 (M+F)	56	92	167
Equation 1 (M)	56	94	128
Equation 1 (F)	59	87	39
Equation 2 (M+F)	56	92	167
Equation 2 (M)	56	94	128
Equation 2 (F)	59	87	39
Equation 4 (M)	61	84	128
Equation 8 (M)	43	70	128
Equation 3 (F)	46	67	39
Equation 7 (F)	23	49	39

TABLE 4.20. Inaccuracy and Bias Statistics for Equation 1, Separated by Sex and Decade for the Culled Study Sample.

Decade	Inaccuracy		Bias		n (n = 167)	
	M	F	M	F	M	F
18-29	14.62	8.54	14.62	7.97	12	6
30-39	12.54	10.65	12.17	8.59	17	5
40-49	9.34	5.20	9.32	4.76	20	6
50-59	7.06	6.82	4.20	1.16	20	6
60-69	7.65	9.65	-5.86	-9.65	20	6
70-79	14.22	26.84	-14.10	-26.84	20	5
80-89	26.01	22.77	-26.14	-22.77	19	5

overestimation of age) and approach zero in the 6th decade. After the 6th decade, bias becomes negative (indicating systematic underestimation of age) and tends to increase. Errors per decade are similar to the un-culled sample.

Equation 2. Equation 2 performs the best on the culled male sample (Tables 4.18 & 4.19), with inaccuracy and bias values of 12.47 and -5.47 years, respectively. For males, only 56% have their actual ages falling within the ± 1 se interval, and 94% fall within the ± 2 se interval. The combined culled male and female sample has an inaccuracy of 12.62 years and a bias of -5.82 years. Only 56% of the combined sample has their actual ages falling within the ± 1 se interval, and 92% fall within the ± 2 se interval. The culled female sample has an inaccuracy of 13.10 years and a bias of -6.94 years. Only 59% percent of the culled female sample has their actual ages falling within the ± 1 se interval, and 87% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, trends similar to the ones seen in the un-culled sample for Equation 2 are apparent (Table 4.21). Errors per decade are similar to the un-culled sample and become particularly large after 70 years.

Equation 4. This all males equation has an inaccuracy value of 12.29 years and a bias value of -3.60 years (Tables 4.18 & 4.19). Only 61% of males have their actual ages falling within the ± 1 se interval, and 84% fall within the ± 2 se interval. No clear trend is present when inaccuracy and bias are calculated by decade (Table 4.22). Inaccuracy values are similar in magnitude between ages 30 and 79. Bias values are similar between 18 and 59 years. Compared to the un-culled sample, inaccuracy has slightly decreased while bias has slightly increased. Errors are particularly high after 80 years. Errors per decade are similar to the un-culled sample.

TABLE 4.21. Inaccuracy and Bias Statistics for Equation 2, Separated by Sex and Decade for the Culled Study Sample.

Decade	Inaccuracy		Bias		n (n = 167)	
	M	**F**	**M**	**F**	**M**	**F**
18-29	7.14	5.96	7.14	4.49	12	6
30-39	8.71	10.50	6.41	6.24	17	5
40-49	6.32	7.19	5.24	3.12	20	6
50-59	8.00	4.04	2.00	0.96	20	6
60-69	10.31	13.17	-8.79	-13.17	20	6
70-79	14.95	28.94	-14.95	-28.94	20	5
80-89	30.04	26.33	-30.04	-26.33	19	5

TABLE 4.22. Inaccuracy and Bias Statistics for Equation 4, All Males.

Decade	Inaccuracy	Bias	n (n = 128)
18-29	6.94	6.64	12
30-39	9.02	4.43	17
40-49	9.23	5.36	20
50-59	11.67	4.63	20
60-69	10.83	-7.34	20
70-79	10.10	-8.48	20
80-89	26.29	-26.29	19

Equation 8. This Euro-American male equation has inaccuracy and bias values of 16.30 and -9.46 years, respectively (Tables 4.18 & 4.19). Only 43% have their actual ages falling within the ± 1 se interval, and 70% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, no clear trends are present for inaccuracy (Table 4.23). Bias becomes negative in the 5th decade and increases afterwards. Errors are particularly high after age 70. Compared to the un-culled sample, overall errors have increased, but errors per decade have slightly decreased.

Equation 3. This all females equation has an inaccuracy of 15.95 years and a bias of -9.46 years (Tables 4.18 & 4.19). Only 46% have their actual ages falling within the ± 1 se interval, and only 67% fall within the ± 2 se interval. When inaccuracy and bias are calculated by decade, trends similar to the ones seen for the un-culled sample are evident (Table 4.24). Errors are particularly high after age 70. Compared to the un-culled sample, inaccuracy and bias have decreased.

Equation 7. This Euro-American female equation has the highest inaccuracy value at 20.26 years, however, it also has the smallest bias value at 1.31 years (Tables 4.18 & 4.19). Only 23% of the females have their actual ages falling within the ± 1 se interval, and only 49% have their actual age falling within the ± 2 se interval. When inaccuracy and bias are calculated by decade, trends similar to the ones seen in the previous equation are evident (Table 4.25). Inaccuracy and bias have slightly decreased for the culled females. Errors are particularly high between ages 20-39 and over age 70.

ANCOVA results. As expected, age has a significant effect on overall suture closure, as does sex (Tables 4.26, 4.27, & 4.28). This time, however, the differences between the Terry sample and the current (culled) sample have reached significance in both tests (Tables 4.27 &

TABLE 4.23. Inaccuracy and Bias Statistics for Equation 8, Euro-American Males Only.

Decade	Inaccuracy	Bias	n (n = 128)
18-29	10.37	10.20	12
30-39	11.73	3.41	17
40-49	13.22	-1.03	20
50-59	10.70	-5.09	20
60-69	13.77	-13.01	20
70-79	55.34	-19.21	20
80-89	32.89	-32.89	19

TABLE 4.24. Inaccuracy and Bias Statistics for Equation 3, All Females.

Decade	Inaccuracy	Bias	n (n = 39)
20-29	9.96	9.89	6
30-39	7.78	5.45	5
40-49	4.12	0.39	6
50-59	11.53	-5.08	6
60-69	19.81	-19.46	6
70-79	34.24	-34.24	5
80-89	27.92	-27.92	5

TABLE 4.25. Inaccuracy and Bias Statistics for Equation 7, Euro-American Females Only.

Decade	Inaccuracy	Bias	n (n = 39)
20-29	29.24	29.24	6
30-39	24.77	24.77	5
40-49	13.53	6.99	6
50-59	15.92	0.24	6
60-69	8.99	-7.27	6
70-79	28.96	-28.96	5
80-89	23.10	-20.63	5

TABLE 4.26. ANCOVA Results for the Culled Study Sample (n = 167) with SUMALL as the Dependent Variable.

Source	df	F	Sig.
Age	1	103.037	0.000
Sex	1	11.209	0.001
$r^2 = 0.423$ (Adj. $r^2 = 0.415$)			

TABLE 4.27. ANCOVA Results Comparing the Culled Study Sample (n = 167) to the Terry Sample (n = 100) with SUMALL as the Dependent Variable.
(**Note:** the Terry sample contains individuals of Euro- and African-American ancestry).

Source	df	F	Sig
Age	1	199.182	0.000
Sex	1	15.089	0.000
Sample	1	4.476	0.035
Sex*Sample	1	0.057	ns
$r^2 = 0.457$ (Adj. $r^2 = 0.449$)			

TABLE 4.28. ANCOVA Results Comparing the Culled Study Sample (n = 167) to Terry Euro-Americans (n = 49) with SUMALL as the Dependent Variable.

Source	df	F	Sig
Age	1	139.262	0.000
Sex	1	24.522	0.000
Sample	1	4.057	0.045
Sex*Sample	1	2.218	ns
$r^2 = 0.448$ (Adj. $r^2 = 0.437$)			

4.28). The interaction between sex and sample (Sex*Sample) is not significant in either test (Tables 4.27 & 4.28). These ANCOVAs basically reaffirm the results for the un-culled sample. Also, it is now clear that a significant difference exits between the current sample and the Terry sample.

New Regression Equations

Since differences exist between the Terry sample and the more recently-deceased individuals in the current study sample, it may be appropriate to construct new regression equations for use on modern individuals. To facilitate comparison to Nawrocki's published equations, a new equation was first generated for SUMALL with males and females combined (Table 4.29). The culled study sample was used to minimize errors caused by unbalanced samples (i.e., high mean age at death). Bias for the new equation nearly equals zero as a forced result of the regression and thus is not a reflection of the applicability of the equation. However, inaccuracy, standard error, and r^2 values should give indications of the utility of the equation. As can be seen, this new equation is not appreciably better compared to Nawrocki's original Equation 1 (Table 4.18) with respect to inaccuracy, and the standard error and adjusted r^2 values are no better either (Table 3.4). Splitting the culled sample by decade and recalculating inaccuracy and bias allows for further comparisons. Since the SUMALL equation is not sex specific it is most appropriate to compare inaccuracy and bias without regards to sex. Table 4.30 gives inaccuracy and bias by decade for Nawrocki's equation 1 and the new equation and Table 4.31 displays inaccuracy and bias separated by sex. Comparing the inaccuracy and bias results

TABLE 4.29. New Regression Equation Using the Culled Study Sample.

New Equation 1	Age = 0.80(SUMALL) + 24.14 (Adj. r^2 = 0.379; se = 15.44 yrs; inaccuracy = 12.63 yrs); bias = -0.13

TABLE 4.30. Comparison of Inaccuracy and Bias Statistics for Nawrocki's Equation 1 and the New Equation Separated by Decade for the Culled Study Sample (n = 167).

	Nawrocki Equation 1		New Equation		
Decade	**Inaccuracy**	**Bias**	**Inaccuracy**	**Bias**	**n**
18-29	12.59	12.40	12.86	12.54	18
30-39	12.11	11.35	13.78	12.86	22
40-49	8.39	8.27	10.57	10.58	26
50-59	7.00	3.50	8.94	6.49	26
60-69	8.11	-6.73	6.95	-3.81	26
70-79	16.74	-16.65	14.46	-13.67	25
80-89	25.34	-25.34	22.30	-22.23	24
Overall	12.76	-2.53	12.63	-0.13	167

TABLE 4.31. Inaccuracy and Bias Statistics for the New Equation, Separated by Sex and Decade for the Culled Study Sample (n = 167).

	Inaccuracy		**Bias**		**n**	
Decade	**M**	**F**	**M**	**F**	**M**	**F**
18-29	14.99	8.59	14.99	7.94	12	6
30-39	14.99	12.04	13.79	9.70	17	5
40-49	11.80	6.51	11.80	6.51	20	6
50-59	9.10	8.38	7.28	3.87	20	6
60-69	6.64	7.97	-2.78	-7.26	20	6
70-79	11.40	25.10	-10.81	-25.10	20	5
80-89	22.94	19.50	-22.94	-19.50	19	5

for Nawrocki's Equation 1 (Table 4.30) to the results for the new equation, it appears that the new equation performs slightly better for individuals older than 60 years. When controlling for sex the usual trends occur in that females have lower inaccuracy and bias values under 60 years and errors are particularly high after 70 years. While differences do exist between the new study sample and Nawrocki's Terry Collection sample, the generation of new equations is not justifiable, since equations based on the current study sample would have lower r^2 values and similar error rates.

CHAPTER 5: DISCUSSION AND CONCLUSIONS

The purpose of this study was to test the following three null hypotheses:

H_n1: Errors in age estimation using a large modern sample will not differ from those obtained by Nawrocki using the Terry Collection.

H_n2: Sex will not have a significant effect on suture closure in a large modern sample.

H_n3: The correlation between suture closure and age will not differ between the Terry Collection and a large modern sample when controlling for sex.

Performance of the Equations

The performance of the equations was examined by calculating their error values (inaccuracy and bias) as well as the percentage of individuals whose actual ages fell with the ± 1 se and ± 2 se intervals. Table 5.1 compares the inaccuracy and bias of the un-culled and culled study samples to Nawrocki's published inaccuracy values. For Equations 1, 2, and 4, the differences in inaccuracy values tend to be relatively small, under 4 years. The differences in inaccuracy for Equations 3, 7, and 8 are higher and range from 7 to 14 years. The general all groups equations (Equations 1 & 2) performed the best overall, having greater percentages of individuals falling within the ± 2 se interval (Tables 4.5 & 4.19) as well as relatively low inaccuracy and bias values. While the all males Equation 4 has slightly lower inaccuracy and bias values, it also has lower percentages of individuals falling within the ± 2 se interval. Equations 3, 7, and 8 performed poorly, having relatively higher inaccuracy and bias values as well as having low percentages of individuals falling within the ± 1 se and ± 2 se intervals. While the error values have increased slightly when using the current study sample with Equations 1, 2, and 4, these values are probably not significantly higher than Nawrocki's given that the current study sample has a mean age at death that is slightly higher than the Terry

TABLE 5.1. Comparison of Inaccuracy and Bias Statistics (in years) for Nawrocki's (1998) Equations 1 through 4, 7, & 8.

Equation	Un-culled Study Sample			Culled Study Sample			Nawrocki's Published Values	
	n	Inaccuracy	Bias	n	Inaccuracy	Bias	n	Inaccuracy
Equation 1 (M & F)	319	12.54	-4.99	**167**	12.76	-2.53	**100**	**10.60**
Equation 1 (M only)	225	11.34	-2.62	128	12.88	-1.88	--	--
Equation 1 (F only)	94	15.43	-10.65	39	12.37	-4.60	--	--
Equation 2 (M & F)	351	13.10	-6.98	167	12.62	-5.82	100	9.60
Equation 2 (M only)	251	11.52	-4.42	128	12.47	-5.47	--	--
Equation 2 (F only)	100	17.07	-13.41	39	13.10	-6.94	--	--
Equation 4 (all males)	252	12.38	-1.92	128	12.29	-3.60	50	8.60
Equation 8 (Euro-American males)	260	15.54	-7.81	128	16.30	-9.46	24	8.20
Equation 3 (all females)	102	20.80	-17.01	39	15.95	-9.46	50	8.6
Equation 7 (Euro-American females)	114	19.48	-2.10	39	20.26	1.31	25	5.9

Collection sample. Therefore, **Hn1** is not rejected for Equations 1, 2, and 4. With regards to Equations 3, 7, and 8, **Hn1** must be rejected because the errors are too great to be the result of the slight difference in mean ages between the samples. Equations 3, 7, and 8 have inaccuracy values that appear to be substantially higher than those published by Nawrocki (1998) and are slightly higher than those obtained for Equations 1, 2, and 4. Equations 3, 7, and 8 may still be of some value, however especially when the condition of the specimen limits the number of available landmarks for analysis.

Inaccuracy and bias for Nawrocki's Equations 1, 2, and 4 are comparable to published inaccuracy and bias values for the Acsadi and Nemeskeri (1970) and Masset (1989) methods and are smaller than those for the Meindl and Lovejoy (1985) method (see Key et al., 1994). The performance of the equations is somewhat counterintuitive because the equation with the highest r^2 value (Equation 7, $r^2 = 0.80$) should have performed the best, however, the equations with lower r^2 values (Equations 1 & 2, $r^2 = 0.51$ & 0.56, respectively) performed the best. It appears that the equation that incorporates a greater number of landmarks performs better than equations with fewer landmarks. Equation 1 uses the sum of 27 landmarks and most likely simulates a continuous variable, thus the regression performs better in the "real world." The other equations use semicontinuous variables and do not perform as well. Equations 2 and 4, which use 3 and 4 landmarks respectively, perform well probably because the reference population of these equations contains individuals of both African and European ancestry. The combination of sex and ancestry groups probably makes these equations more robust to error than the ancestry-specific equations (e.g., Equations 7 & 8). Another consideration when using all of the equations is that at 70 years of age and above (sometimes at 60) all of the equations have high error values. Error values tend to be high in the older decades for all adult age indicators.

Sexual Dimorphism

Throughout the years, researchers have published contradicting accounts of the influence of sex on suture closure (see Chapter 2). The current study clearly demonstrates that sex does influence suture closure. The differences that exist between the sexes are minimally due to the pattern of suture closure. This effect is seen in the rank-order correlation tables (Tables 4.1, 4.2, 4.15, & 4.16) that show how correlation strength is clustered by suture series and sex. Females have stronger correlations for the vault and endocranial series of sutures. Males have the strongest correlations for the lateral-anterior series of sutures. The ANCOVA results reinforce the fact that differences due to sex exist in the rate or pattern of suture closure (Tables 4.12 - 4.14 & 4.26 - 4.28). Therefore, **Hn2** is rejected because the effect of sex is significant.

Why sexual dimorphism has not been clearly demonstrated in earlier studies is not clear. Perhaps Masset (1989) is correct that most other studies failed to find a difference due to their lack individuals in the 30 to 40 year age range. While several authors (Baker, 1984; Masset, 1989; Key et al., 1994, Nawrocki, 1998) have demonstrated and warned that sex needs to be taken into account when estimating age at death using suture closure, no attempts to enhance age estimates have been made with this knowledge. Having methods to estimate age that combine samples, such and the Menidl and Lovejoy method and Nawrocki's (1998) Equations 1 and 2, may not reduce the overall accuracy of the age estimates. Because of the great deal of overlap in variation between the sexes, the combined sex equations performed better overall compared to sex specific equations. While sexual dimorphism does exist, it may not practically affect age estimates using cranial suture closure and taking a broader approach may actually help the performance of the age estimation method.

Correlation Differences Between the Terry and Current Study Samples

The Terry Collection sample tends to have higher correlation coefficients than the current study samples (culled and un-culled) (Tables 4.1, 4.2, & 4.15 – 4.17). However, when controlling for the sample distribution, i.e., the culled study sample, the differences in the correlation values do not appear to be that large (Tables 4.15 – 4.17). This fact is especially evident with the Pearson's correlation values when comparing the culled study sample to Terry Euro-Americans (Table 4.17). To determine if the difference between the correlations between SUMALL and age are significant, we can compare them by using Fisher's transformation:

$$z = \frac{r_{F1} - r_{F2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

where r_F is the transformed r-value, n is the sample size, and the subscripts refer to the sample (Mattson, 1993). As can be seen in Table 5.2, there are no significant differences between Nawrocki's Terry Collection sample correlations and the current culled study sample correlations. Although not displayed here, when controlling for sex, the un-culled study sample correlations are not significantly different from the Terry correlations either. Therefore, **Hn3** is not rejected and there does not appear to be a major difference between the Terry Collection and a large modern sample with regards to the correlation between cranial suture closure and age. This finding is very reassuring given the concern of possible secular trends in the rate of suture closure (see Masset, 1989). The higher correlation values that the Terry sample displays may be caused by one of the following: (1) the Terry Collection is much more homogenous than the current study sample, which is derived from 3 geographic regions as well as more diverse socioeconomic backgrounds; (2) there is more random variation displayed in the current study sample due to the larger sample size; and/or (3) the inclusion of African Americans may improve

TABLE 5.2. Comparison of Z-Transformed r-Values Between the Terry Collection and the Culled Study Sample for SUMALL vs. Age. (**Note**: the Terry Sample contains African Americans; EA = Euro-American only; ns = no significant difference).

Samples Compared	Terry sample r	Culled sample r	z	Difference
Terry M + F vs. Culled Study Sample M + F	0.715	0.619	1.06	ns
Terry EA M + F vs. Culled Study Sample M + F	0.633	0.619	0.14	ns
Terry M vs. Culled Study Sample M	0.702	0.622	0.84	ns
Terry EA M vs. Culled Study Sample M	0.609	0.622	0.09	ns
Terry F vs. Culled Study Sample F	0.751	0.621	1.13	ns
Terry EA F vs. Culled Study Sample F	0.738	0.621	0.81	ns

correlation strength because Terry African Americans may have higher or tighter correlations with age than Terry European Americans (see Galera et al., 1998).

New Landmarks

This study examined two new landmarks, the inferior coronal and zygomaxillary sutures. The strength of the rank-order correlations for the inferior coronal suture (left and right) were comparable to surrounding landmarks of the vault for both the un-culled and culled study samples. The zygomaxillary suture (left and right) displayed a weak correlation for the un-culled male sample. No other significant correlations were found for the zygomaxillary suture for either the un-culled or culled study samples. Experimenting with stepwise regression analysis (results not presented here) on the culled study sample, the right inferior coronal suture was selected for most of the prediction models produced, suggesting that it is particularly useful for age prediction. The left zygomaxillary suture was selected three times (once for combined sexes and twice for males only) during the stepwise regression. The selection of the zygomaxillary suture is interesting because it did not have any significant rank-order correlations. Pearson's correlation was run on this variable and no significant relationships were found with age. The regression probably finds that the zygomaxillary suture is useful to tease out age for a certain range on the regression line and is therefore included in some of the prediction models.

Conclusions

Of the 6 equations tested, the 2 general all group equations (Equations 1 & 2) performed the best, followed by Equation 4 (all males equation). However, overall the equations do work relatively well and can play a role in age estimation. In conclusion, cranial suture closure does correlate with age and sex influences the pattern and/or rate of suture closure, although the effects of sex do not necessarily preclude one from combining the sexes into single equations. Also, there seems to be no secular trends that would prevent the use of Nawrocki's (1998) equations on modern individuals from forensic contexts. This study reaffirms the need to control the reference populations used to generate age estimation methods as well as the samples that are used to test and compare methods.

Future Studies

This study only examined European Americans and therefore more study is needed on an African American sample to determine if and how the patterns and rates of suture closure may differ from European Americans. Nawrocki's equations should be tested on an independent African American sample. While this study has confirmed that sexual dimorphism does exist, it has done little to help us understand suture closure as a biological phenomenon. Most studies that use suture closure to estimate age at death score suture closure using a broad scale based on a percentage range of closure (e.g., 1 to 50% closure). These broad scales are used to facilitate replication and reduce inter- and intraobserver error. However, by using these broad scales we are losing information on the nature of suture closure. Using a finer scale (more stages), calculating obliteration indices, quantifying the amount of closure, or examining the sutures using technology (e.g., CT scans or digital imaging) may provide better insight into the relationship between suture closure and age.

LITERATURE CITED

Acsadi GY, Nemeskeri J. 1970. Human lifespan and mortality. Akademiai Kiada, Budapest.

Angel JA. 1984. Variation in estimating age at death of skeletons. Coll Anthropol 8:163-168.

Angel JA, Suchey JM, Iscan MY, Zimmerman MR. 1986. Age at death estimated from the skeleton and viscera. In Zimmerman MR, Angel JL editors. Dating and Age Determination of Biological Materials. London, Croom Helm p179-220.

Ashley-Montagu MF. 1938. Aging of the skull. Am J Phys Anthropol 23:355-375.

Aykroyd RG, Lucy D, Pollard AM, Roberts CA. 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. Am Antiquity 64: 55-70.

Baker RK. 1984. The relationship of cranial suture closure and age analyzed in a modern multi-racial sample of males and females. Masters Thesis, California State University, Fullerton.

Bennett KA. 1967. Craniostenosis: a review of the etiology and a report of new cases. Am J Phys Anthropol 21:1-9

Bocquet-Appel JP, Masset C. 1982. Farewell to paleodemography. J Hum Evol 11:321-333.

Brooks ST. 1955. Skeletal age at death: the reliability of cranial and pubic age indicators. Am J Phys Anthropol 13:567-597.

Buikstra JE, Ubelaker DH. 1994. Standards for data collection from human skeletal remains. Arkansas archeological survey research series No. 44

Byers SN. 2002. Introduction to forensic anthropology. Boston, Allyn and Bacon.

Cohen MM. 1993. Sutural Biology and the correlates of craniosynostosis. Am J Med Gen 47:581-616

Dwight T.1890. The closure of the cranial sutures as a sign of age. Boston Medical and Surgical Journal 122:389-392.

Galera V, Ubelaker DH, Hayek LC. 1998. Comparison of macroscopic cranial methods of age estimation applied to skeletons from the terry collection. J Forensic Sci 43:933-939.

Genoves ST, Messmacher M. 1959. Valor de los patrones tradicionales para la determinacion de la edad por medio de las suturas en craneos mexicanos (indigenas y mestizos). Cuadernos del Instiiuto de Histora, Serie Antropologica No 7, Universidad Nacional Autonoma de Mexico

Herring SW. 1993. Epigenetic and functional influences on skull growth. In Hanken J and Hall BK editors. The skull Volume 1: Development. Chicago The University of Chicago Press p153-206.

Hershkovitz I, Latimer B, Dutour O, Jellema LM, Wish-Bartz S, Rothschild C, Rothschild B. 1997. Why do we fail in aging the skull from the sagittal suture? Am J Phys Anthropol 103:393-399.

Jantz R, Moore-Jansen, P. 1998. A database for forensic anthropology. Knoxville, University of Tennessee department of anthropology Report of Investigations No. 47

Johnson JS. 1976. A comparison of age estimation using discriminant function analysis with some other age estimations of unknown skulls. J Anat 121:475-484.

Kemkes-Grottenthaler A. 1996. Critical evaluation of osteomorphognostic methods to estimate adult age at death: a test of the complex method. Homo 46:280-292.

Key CA, Aiello LC, Molleson T. 1994. Cranial suture closure and its implications for age estimation. Int J Osteoarch 4:193-207.

Mann RW. 1987. Maxillary suture obliteration: a method for estimating skeletal age. Masters Thesis, University of Tennessee, Knoxville.

Masset C. 1982. Estimation de l'age au deces par demographique observees dans quelques sepultures neolitheques. Ixe Congres de l'Union Int Sci Prehistoriques et Protohistoriques.

Masset C. 1989. Age estimation on the basis of cranial sutures. In Iscan MY editor. Age markers in the human skeleton. Springfield CC Thomas Publisher LTD p71-103.

Mattson DE. 1993. Statistics: Difficult concepts, understandable explanations. Wauconda, Bolchazy-Corducci Publisheres.

Meadows L, Jantz RL. 1995. Allometric Secular Change in the Long Bones from the 1800's to the Present. J Forensic Sci 40:762-767.

McKern TW, Stewart TD. 1957. Skeletal age changes in young American males. Environmental Protection Res Div, Quartermaster Res & Dev Center U.S. Army, Nantick Mass. Tech Rep EP-45.

Meindl RS, Lovejoy CO. 1985. Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures. Am J Phys Anthropol 68:57-66.

Nawrocki SP. 1998. Regression formulae for estimating age at death from cranial suture closure. In: Reichs KJ editor. Forensic Osteology 2nd edition. Springfield CC Thomas Publisher LTD p276-292.

Nemeskeri J, Harsanyi L, Ascadi G. 1960. Methoden zur diagnose des lebensalters von skelettfuden. Anthoropol Anzeiger 24:70-90.

Ousley S. 1995. Should we estimate biological or forensic stature? J Foren Sci 40:768-773

Parson FG, Box CR. 1905. The relation of the cranial sutures to Age. J Anthropol Inst Great Britain and Ireland 35:30-38.

Perizonius WRK. 1984. Closing and non-closing sutures in 256 crania of known age and sex from Amsterdam (A.D. 1883-1909). J Hum Evol 13:201-216.

Powers SR. 1962. The disparity between known age and age as estimated by cranial suture closure. Man 83:52-55.

Shipman P, Walker A, Bichell D. 1985. The human skeleton. Cambridge, Harvard University Press.

Singer R. 1953. Estimation of age from cranial suture closure. J Foren Med 1:52-59.

Todd TW, Lyon DW. 1924. Endocranial suture closure, its progress and age relationship: part I adult males of the white stock. Am J Phys Anthropol 7:325-384.

Todd TW, Lyon DW. 1925a. Cranial suture closure, its progress and age relationship: part II ectocranial closure in adult males of the white stock. Am J Phys Anthropol 8:23-45.

Todd TW, Lyon DW. 1925b. Cranial suture closure, its progress and age relationship: part III endocranial closure in adult males of negro stock. Am J Phys Anthropol 8:47-71.

Todd TW, Lyon DW. 1925c. Cranial suture closure, its progress and age relationship: part IV ectocranial closure in adult males of the negro stock. Am J Phys Anthropol 8:149-168.

Ubelaker DH. 1999. Human skeletal remains: Excavation, analysis, interpretation. Washington, Taraxacum.

WhiteTD. 2000. Human Osteology. San Diego, Academic Press.