Evaluation of Regression Equations to Estimate Age at Death Using Cranial suture Closure

Carlos J Zambrano BA*, Archeology and Forensics Laboratory, University of Indianapolis

University of Indianapolis
Archeology and Forensics Laboratory
http://archlab.uindy.edu

Abstract

Learning Objective: To present to the forensic anthropological community the utility of regression formulae to estimate age at death using cranial suture closure.

During a skeletal analysis, age at death is one of the primary components of the biological profile that the forensic anthropologist constructs. There are several indicators used to estimate age at death for adults, however, some are employed more frequently than others. 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 an age estimator has been limited. This study examines 6 of the regression equations (Equations 1, 2, 3, 4, 7, and 8) created by Nawrocki using recently deceased individuals. The test sample contains 366 individuals (111 females, 255 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 27 landmarks were scored on each specimen: 16 ectocranial (external) surface, 7 endocranial (internal) surface, 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. The endocranial and palatal sutures were scored following Nawrocki's (1998) guidelines. Age was estimated for each individual using up to 4 different equations: 2 general equations (EQ 1 and 2) and 2 group-specific equations (e.g., all females, European females). Inaccuracy and bias were calculated for each equation to assess its performance. The percentage of individuals whose actual age falls within each equation's 2SE interval was also calculated. An analysis of covariance (ANCOVA) was used to determine if suture closure is influenced by an individual's sex.

Inaccuracy is the average absolute error of the estimate. For Equation 1, inaccuracy was 12.81 yrs (M + F). Inaccuracy for Equation 2 was 13.44 yrs (M + F). The 2 male-specific equations (EQ 4 and EQ 8) had inaccuracies of 12.58 and 15.64 years respectively. The inaccuracy for the 2 female-specific equations (EQ 3 and EQ 7) was 22.18 and 18.84 years respectively. Bias is a measure of the overall under- or over-estimation. Bias for EQ 1 was -5.61 yrs (M + F). Bias for EQ 2 was -7.61 yrs (M + F). Bias for the 2 male-specific equations (EQ 4 and EQ 8) was -2.21 and -8.25 years respectively. For the 2 female-specific equations (EQ 3 and EQ 7), bias was -19.13 and -4.21 years respectively. The percentage of individuals falling within 2SE for the general equations ranged from 73.6 to 93.5 percent. For the male-specific equations (EQ 4 and EQ 8), 84.4% and 74.7% respectively fell within the 2SE. The percentage for the female-specific equations (EQ 3 and EQ 7) was 51.6% and 49.5% respectively. The ANCOVA results suggest that summed suture score is influenced by sex. This study found that the general equations performed well and in general had better results than the group-specific equations. The male-specific equations performed better than the female-specific equations. The poorer performance for the females may be due to a larger number of older individuals in that subgroup compared to the males. Cranial suture closure does correlate with age, however, sex influences that relationship and needs to be accounted for when using sutures to estimate age. In conclusion, cranial suture closure can be a useful tool to estimate age and the long-lasting skepticism should be reconsidered.

Introduction

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 to employ. 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. Most 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.

When estimating age, it is necessary to account for factors that could influence the prediction, such as sex, ancestry, and secular changes. 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 calottes from males and females of European ancestry ranging in age from 58 to 102 years old.

Nawrocki's study is only one of a handful that has found and accounted for the variation in aging due to sex and ancestry. Although carefully constructed, his 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 amount of population variation in suture closure. However, the equations created with the smaller sample sizes have larger adj. r^2 values. This suggests that they should perform better.

Since the caseloads of forensic anthropologists include many young adults and middle-aged individuals, the older test sample used by Nawrocki may have little bearing on the application of his equations to many modern forensic cases. This poster examines the performance of 6 of Nawrocki's equations (Table 1) using recently deceased individuals of European ancestry. The null hypotheses for this study include:

Hn1: The Nawrocki equation with the largest adj. r^2 value will perform the best;

Hn2: Sex will not have a significant effect on age estimation.

TABLE 1: 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; inaccuracy = 10.6 yrs; se = 12.9 yrs)
Equation 2	All Groups	AGE = 5.86(PLQ) + 6.42(BRZ) + 4.91(TRP) + 24.3 (adj. r^2 = 0.56; inaccuracy = 9.6 yrs; se = 12.1 yrs)
Equation 3	All Females	AGE = 5.29(CRQ) + 7.38(PRQ) + 8.84(TRP) + 26.8 (adj. r^2 = 0.65; inaccuracy = 8.6 yrs; se = 10.9 yrs)
Equation 4	All Males	AGE = 7.0(PLQ) – 6.08(ASQ) + 6.83(TRQ) + 9.12(BRZ) + 28.3 (adj. r^2 = 0.61; inaccuracy = 8.6 yrs; se = 11.5 yrs)
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; inaccuracy = 5.6 yrs; se = 8.2 yrs)
Equation 8	Euro-American Males	AGE = 15.01(PLQ) – 6.76(ASQ) + 37.9 (adj. r^2 = 0.61; inaccuracy = 8.2 yrs; se = 11.0 yrs)
SUMALL is the summed suture score. Table 3 has the landmark abbreviations.		

Materials and Methods

Sampling. The specimens used for this study are curated at three institutions: University of Indianapolis (UIndy), University of Tennessee (Bass Donated Collection), and the University of New Mexico (UNM Documented and Forensic Collections). All of the specimens were scored by the author, except for some of the specimens from the University of Indianapolis. The suture closure data for the UIndy sample were retrieved from the case files of Dr. Nawrocki. Only specimens free of significant cranial trauma with known age, sex and ancestry were examined. This study only examines individuals of European ancestry. Table 2 displays the distribution of individuals in the study sample by sex, decade and collection.

TABLE 2: Distribution of Individuals by Decade, Sex, 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
<20	1	0	0	0	0	0	1	0	2
20-29	2	1	2	4	4	1	3	2	19
30-39	9	4	1	1	13	2	1	1	32
40-49	5	1	2	0	26	3	1	0	38
50-59	20	4	1	1	45	11	0	0	82
60-69	16	9	0	0	36	14	0	1	76
70-79	17	13	0	0	23	10	0	1	64
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	6	5	366

Scoring suture fusion. Before data collection the author was trained in the proper scoring techniques for cranial suture closure under the guidance of Dr. Stephen Nawrocki at the University of Indianapolis. Up to 27 landmarks were scored on each specimen (Table 3): 16 ectocranial (external) surface, 7 endocranial (internal) surface, and 4 from the palate. Following Meindl and Lovejoy (1985), one-centimeter segments along the cranial sutures are scored from 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 suture 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 suture extending onto the alveolar surface is ignored (Nawrocki, 1998). Note that the intermaxillary suture has a tendency to have bony mounding occurring parallel to it on either side of the suture. This mounding is ignored and the suture was scored below the level of the mounding. 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, sex, and ancestry was concealed until after the specimen had been scored.

TABLE 3: Cranial Suture Landmarks and Abbreviations.

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	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
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

Statistical analysis. Age was estimated for each specimen using up to four equations. Inaccuracy and bias statistics were calculated for each equation for the sample as a whole and by sex. 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. The percentage of individuals in the study sample whose actual age fell within the $\pm$ 2 standard error interval (2SE) was calculated for each equation. An analysis of covariance (ANCOVA) was used to determine if suture closure is influenced by an individual's sex using the following model:

SUMALL = SEX + AGE

Where the dependent variable SUMALL is the summed score for all landmarks SEX is a categorical independent variable, and AGE is the continuous covariate.

Results

Table 4: Inaccuracy and Bias Statistics in Years for Nawrocki (1998) Equations 1 through 4, 7 & 8 Using the Current Study Sample.

Equation	n	Inaccuracy	Bias
EQ1 (M+F)	302	12.81	- 5.61
EQ1 (M)	217	11.45	- 2.90
EQ1 (F)	85	16.26	- 12.51
EQ2 (M+F)	332	13.44	- 7.61
EQ2 (M)	241	11.70	- 4.69
EQ2 (F)	91	18.04	- 15.34
EQ4 (M)	243	12.58	- 2.21
EQ8 (M)	249	15.64	- 8.25
EQ3 (F)	93	22.18	- 19.13
EQ7 (F)	105	18.84	- 4.21

Table 5: Percentage of Individuals Whose Actual Age Falls Within the 2SE Interval of Each Equation.

Equation	n	Percentage
EQ1 (M+F)	302	89.1
EQ1 (M)	217	93.5
EQ1 (F)	85	77.6
EQ2 (M+F)	332	83.7
EQ2 (M)	241	87.5
EQ2 (F)	91	73.6
EQ4 (M)	243	84.4
EQ8 (M)	249	74.7
EQ3 (F)	93	51.6
EQ7 (F)	105	49.5

Table 6: ANCOVA Results With SUMALL as the Dependant Variable.

	df	F	Sig.
Age	1	82.78	.000
Sex	1	25.9	.000
R squared = .267 (Adj. R squared = .264) n = 302			

Discussion

Performance of Equation 1. In general EQ1 worked best for the male sample with an inaccuracy of 11.45 yrs and bias of -2.90 yrs. The male sample had 93.5% of the individuals with their actual age falling within the 2SE interval. With males and females combined EQ1 inaccuracy and bias rose to 12.81 yrs and -5.61 yrs respectively. Eighty-nine percent of these individuals had their actual age falling within the 2SE interval. Equation 1 performed worse on the female sample with an inaccuracy of 16.26 yrs and bias of -12.51 yrs. Only 77.6% of the female sample fell within the 2SE interval.

Performance of Equation 2. Equation 2 performed best on the male sample with an inaccuracy of 11.7 yrs and bias of -4.69 yrs. For the male sample 87.5% of the individuals had their actual age falling within the 2SE interval. With the males and females combined EQ2 inaccuracy and bias rose to 13.44 yrs and -7.61 yrs. For males and females combined, 83.7% of them had their actual age falling within the 2SE interval. EQ2 performed worse on the female sample with an inaccuracy of 13.44 yrs and bias of -7.61 yrs. Only 73.6% of the female sample fell within the 2SE interval.

Performance of male specific equations 4 & 8. Equation 4 (all males equation) performed better than Equation 8 (Euro-American male equation). Equation 4 had slightly better inaccuracy and bias results (12.58 yrs and -2.21 yrs respectively) than either of the general all groups equations (EQ1 & EQ2) with the males and females combined. However, when only males were considered EQ4 performed slightly worse than the two general equations. The percent of individuals with their actual age falling within the 2SE interval was in general lower than EQ1 and EQ2 with only 84.4% falling within the interval. Equation 8 had larger inaccuracy (15.64 yrs) and bias (-8.25 yrs) statistics than Equations 1, 2 & 4. Also only 74.7% of individuals had their actual age falling within the 2SE interval.

Performance of female specific equations 3 & 7. Overall EQ3 & 7 performed the worst. Equation 3 had an inaccuracy of 22.18 yrs and bias of -19.13 yrs. Equation 7 had and inaccuracy of 18.84 yrs and bias of -4.21 yrs. For either equation approximately only 50% of the individuals had their actual age falling within the 2SE interval.

Several trends are present:

1) **Overall the Nawrocki (1998) equations tested performed better when applied on males**. This is evident due to lower values for inaccuracy and bias as well as the higher percentages of individuals whose actual age falls within the 2SE interval. This suggests that females are more variable in their suture closure than males, or that the female equations are less accurate.

2) **The tested equations on average tend to underestimate age**. This is evident by the negative bias values.

3) **As the number of landmarks used in the regression decreases, inaccuracy and bias tend to increase. Also, as the number of landmarks decreases so does the percentage of individuals whose actual age falls within the 2SE interval**. A greater number of landmarks appears to increase the robusticity of the equations.

4) **The general all groups equations (EQ1 & 2) perform better than the group-specific equations (e.g., all male or Euro-American males).** This is counterintuitive because the adjusted r^2 values for the group-specific equations are larger than the all group equations, which would suggest that the group-specific equations should perform better.

5) **Sex appears to significantly affect the pattern or rate of suture closure as evidenced by the ANCOVA results.**

The above trends suggest that both null hypotheses should be rejected.

In theory, the equation with the largest adjusted r^2 value should perform the best. However, as the number of landmarks is decreased the equation becomes less robust to variation in suture closure and/or observer error. This is most likely why EQ1, which uses the largest number of landmarks (27), but has the lowest adjusted r^2 value, performed better than the other equations. Equation 2 requires only 3 landmarks. However, one must also consider the fact that the general equations were created using both sexes and individuals of Euro- and African-American ancestry. The variation within and between these groups most likely makes the general equations more resilient to error.

There are two reasons that could account for the poorer performance of the equations on the new female sample and their average underestimation of age: 1) the mean age of Nawrocki's Terry Collection sample may differ compared to the sample used in this study; 2) there could be a difference in the rate of suture closure between the Terry Collection and recently deceased individuals (i.e., a secular trend). Table 7 displays the summary statistics for Nawrocki's Terry Collection sample and the current study sample whose ages were estimated using EQ1. Table 8 displays inaccuracy and bias for EQ1, separated by sex and decade.

Table 7: Summary Statistics for Nawrocki's Terry Collection Sample and the Current Study Whose Ages Were Estimated using EQ1.

	n	Min	Max	Mean	Std. Dev.
Terry M	50	21	85	53.76	18.50
Terry F	50	22	84	53.66	18.30
Terry M + F	100	21	85	53.71	18.30
M + F	302	18	101	61.55	14.01
M	217	18	101	59.83	16.68
F	85	21	98	65.94	17.16
Note: The Terry sample contains individuals of Euro- & African-American ancestry.					

Table 8: Inaccuracy and Bias for EQ1, Separated by Sex and Decade. (n = 302)

	Inaccuracy		Bias		n	
	M	F	M	F	M	F
18-29	17.34	15.86	17.34	15.86	8	4
30-39	12.54	5.65	12.17	3.08	17	4
40-49	8.50	7.53	8.12	7.53	31	3
50-59	6.29	5.52	3.32	0.93	54	14
60-69	6.78	12.79	-4.14	-11.81	46	22
70-79	14.23	18.35	-14.16	-18.35	36	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

From Table 7 it is clear that Nawrocki's sample has a mean age that is younger than the current study sample. The larger inaccuracy and bias values for females seen in Table 4 is partially a result of their larger mean age when compared to the Terry sample (Table 7). New trends are visible when inaccuracy and bias are broken down by decade. On average EQ 1 performed slightly better on females under the age of 60, having slightly smaller inaccuracy and bias values. After age 60, inaccuracy and bias appears only slightly larger for the females. This is evidence suggesting that the inaccuracy and bias values in Table 4 are inflated due to the difference in mean age between Nawrocki's Terry sample and the current study sample. The mean age for the males in this study is close to Nawrocki's Terry male mean age (Table 7). This would explain why the equations appear to work better for males. Also, visible in Table 8 is the tendency of regression to overestimate younger individuals and to underestimate older individuals. The bias values show this phenomenon quite well. Bias is large and positive for the earlier decades, approaches zero close to the equation's mean (53 yrs), and decreases into negative values for the later decades. Comparing the new inaccuracy values (Table 4) to Nawrocki's values (Table 1) it is clear that for many, the new inaccuracy values are only slightly larger. The differences between sample means may account for some of the differences between these inaccuracy values. Considering these factors, overall the equations work relatively well.

Conclusions

Of the 6 equations tested, the 2 general all group equations (EQ1 & 2) performed the best. Overall the equations do work relatively well and can play a role in age estimation. In conclusion, cranial suture closure does correlate with age, however, sex influences that relationship and needs to be accounted for when using sutures to estimate age. Also, it is necessary to account for the errors caused by a difference in the mean age between samples. Some of the current results are reassuring that suture closure can play a role in age estimation, it is also evident that more study is needed to clarify the differences between the sexes.

Works Cited

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.

Acknowledgments

I would like to thank: Dr. Nawrocki & Dr Schmidt from the University of Indianapolis for their advice during this project; Dr. Edger & Dr Komar of the University of New Mexico for permitting me access to the Collections; Chris Grivas for helping me track down individuals; Dr. LM Jantz of the University of Tennessee for permitting me access to the Bass Collection. Heather Walsh-Haney for keeping me sane during long days of data collection. I would also like to thank my fellow Graduate students at the University of Indianapolis who offered editing advice and provided me with distractions.