U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

Award Number: 2010-DN-BX-K035

Project Title: **Estimation of Age at Death Using Cortical Bone Histomorphometry**

PI: Christian Crowder, PhD, D-ABFA
Office of Chief Medical Examiner
520 First Avenue
New York, NY 10016
Telephone: (212) 447-2761
Fax: (212) 447-4339
Email: ccrowder@ocme.nyc.gov

Research Assistant: Victoria M. Dominguez, MA

Administrative POC:
Samantha Ropiak
Office of Chief Medical Examiner
421 E. 26th Street
New York, NY 10016
Telephone (212) 323-1785
Email: sropiak@ocme.nyc.gov

Financial POC:
Deirdre Snyder
Office of Chief Medical Examiner
421 E. 26th Street
New York, NY 10016
Telephone (212) 323-1737
Email: dsnyder@ocme.nyc.gov

Abstract

Estimating the age at death in the adult skeleton is problematic owing to the biological variability in age indictors and the differential skeletal response to environmental factors over an individual's life. It is particularly difficult to accurately estimate age for individuals over 50 years of age. Thus, it is becoming increasingly important for anthropologists to improve age estimates through the use of multiple age indicators and various modalities of assessment (e.g., macroscopic and microscopic). Previously developed histological methods of age estimation using the femur demonstrate significant methodological issues that affect their reliability and accuracy. This research evaluates histological age estimation using the anterior femur and explores the biological limitations of bone turnover as an age indicator.

The sample includes femur cross-sections from 319 individuals (169 males, 150 females) of known age at death. Prior to this study, research was performed to redefine and validate histological variables. The following variables were collected:

1. Surface Area (Sa.Ar.) per mm^2
2. Intact Secondary Osteons (N.On.):
3. Fragmentary Secondary Osteons (N.Fg.On.)
4. Intact Secondary Osteon Density (OPD(I)) per mm^2
5. Fragmentary Osteon Density (OPD(F)) per mm^2
6. Osteon Population Density (OPD): sum of OPD(I) and OPD(F)
7. Mean Osteonal Cross-Sectional Area (On.Ar) per mm^2
8. Mean Anterior Cortical Width (Ant.Ct.Wi.) per mm^2

The topographic sampling method evaluates ten columns from the periosteal to the endosteal cortex located at the anterior femur midshaft. Using a Merz counting reticule at 200x magnification, 50% of the microscopic fields were evaluated in each column by alternating fields. This sampling strategy accounts for 95% of the remodeling variability

within the anterior cross-section. Osteon areas and cortical widths were calculated using imaging software.

Statistical analyses were performed in SPSS 19 to examine the relationship of the cortical bone histomorphometrics to sex and age. Stepwise linear regression was used to develop the age prediction equation(s). Two variables (OPD(F) and On.Ar.) required log transformation to meet normality requirements. Analysis of observer error was performed using several procedures for evaluating method repeatability.

Pearson correlations show moderate and strong relationships with age for all collected variables except OPD(I). Due to this finding it was determined that the constituent variables for OPD should remain separate in the regression model. One-way ANOVA and ANCOVA analyses indicated that the variables, with the exclusion of OPD(I), demonstrate some significant sex differences at the 0.05 level. Stepwise regression analysis of the male data set produced a model using OPD(F)-log and OPD(I) as predictors, while the female model selected OPD(F)-log, OPD(I), and Ant.Ct.Wi. as predictors. The standard error of the estimate is 11.13 years and 9.77 years, respectively. In the event that sex cannot be determined, a general equation was developed using OPD(F)-log, OPD(I), and On.Ar.-log, providing a standard error of 10.70. Observer error results indicate the method passed repeatability standards set by the authors.

Current histological methods demonstrate significant issues that affect their reliability and accuracy. The method developed from this research demonstrates several advantages over previous methods. The method is based on validated variables, accounts for 90%–95% of the spatial variation in osteons within the anterior cortex, and is not restricted to a specific field size or magnification. The results of the study indicate that histological

analysis of the anterior femur provides reliable age estimates. One of the most prevalent issues regarding adult age estimation is the inability to accurately age older adults. The described regression model is most accurate for individuals over 50 years of age. Bearing in mind that the elderly are a rapidly growing percentage of North American populations and that unidentified adults are a common occurrence in the forensic setting, this research will improve the accuracy of estimating age for older adults.

Table of Contents

EXECUTIVE SUMMARY

Description of the problem

The estimation of age at death is an essential part in the reconstruction of population demographics and the individual analysis of human remains. Estimating the age at death of children and young adults can be performed with greater accuracy owing to methods that are based on the growth and development of the human skeleton; however, the estimation of adult age at death demonstrates a progressive decrease in accuracy as chronological age increases. Most methods for estimating adult age at death are assessments of degenerative changes to gross bone morphology, which demonstrate large variability between individuals. It has been suggested that quantitative cortical bone histology, or histomorphometry, provides a reliable approach to adult age estimation with potential to bridge the 50+ age boundary (Thompson, 1979; Stout and Gehlert, 1982; Ubelaker, 1986). Since the introduction of the first quantitative histological approach for the estimation of age at death by Ellis Kerley in 1965, histological parameters of age-related bone turnover have been well-documented in the anthropological literature. Despite this, histological methods are not widely applied in forensic anthropology owing to the inherent methodological issues identified in the literature (Lynnerup et al., 1998; Villa and Lynnerup, 2010) and the specialized knowledge required in interpreting normal vs. abnormal bone histomorphology.

The premise of histological age estimation is based upon clinical and anthropological research demonstrating that bone turnover occurs in cortical bone at a predictable rate over an individual's lifetime. In theory, the length of time during which

remodeling occurs (chronological age) will be the primary influence on how many secondary osteon creations (intact and fragmentary osteons) accumulate per unit of area (Kerley, 1965; Wu et al., 1967, 1970; Stout and Teitelbaum, 1976). While microscopic age changes are considered to be universal, inconsistencies in the reported accuracy of the methods when they are applied to individuals outside of the reference samples suggest that intrinsic and extrinsic biological factors, such as genetics and a wide range of suggested behaviors, have varying effects on bone microstructure. One such factor is the existence of the osteon asymptote. The asymptote occurs when the bone has become completely remodeled so that new osteon creations simply replace older ones and the proportion of the cortex that is remodeled does not increase (Frost, 1987a, 1987b). According to the literature, ribs reach this asymptote around the 5^{th} or 6^{th} decade of life resulting in the inability to accurately age older individuals. While the histological methods utilizing the ribs (Cho et al., 2002) are currently the most accurate and reliable methods available, they cannot be used with certainty when applied to older individuals because of this biological phenomenon. It is believed that the femur, due to its larger cortical area, does not reach this asymptote until later in life making it a preferred location for histological analysis of older individuals (Stout and Paine, 1992).

As noted by Lazenby (1984), biological issues may all be secondary to concerns relating to the design of histological methods. Methodological concerns include sample size, sample demography, variable definitions, and measurement techniques. Currently there are two preferred methods for estimating adult age at death using the femur: the Kerley method and the Thompson method. Each contains separate, but significant methodological issues affecting their accuracy and reliability. An evaluation of the

Kerley method by Kerley and Ubelaker (1978) noted that the original microscopic field size was incorrectly reported. The authors warned investigators that the variability in field diameters of different microscopes would contribute to "apparent errors" and "unreasonable [age] estimates" (Kerley and Ubelaker, 1978). The use of a smaller field size, as opposed to the original field size, would underestimate age since the sum of recorded structures is always less than that recorded when the regression models were created. Kerley and Ubelaker (1978) suggested applying a correction factor to the analysis if investigators could not reproduce the original field size. Stout and Gehlert (1982) suggest that the use of the correction factor to adjust osteon counts may be of limited value due to the spatial variation of microstructures within the cortex, thus resulting in method inaccuracy if the original Kerley field size is not applied. Because the original field size is still unknown, the conservative approach is that the method should not be used for analysis.

Other issues associated with, but not limited to, the Kerley method include subjective variable definitions and the inability to incorporate remodeling events from the periosteal envelope to the endosteal envelope. The former point is a main factor in controlling the level of observer error associated with how researchers classify osteonal structures or, in other words, how intact and fragmentary osteons are differentiated. For example, Kerley (1965) classified intact osteons as being 80% intact with a complete Haversian canal present, while Stout and Paine (1992) defined an intact osteon as having a Haversian canal that is 90% intact.

In 1979, Thompson published a method designed to minimize the amount of destructive sampling, reduce observer error, and explore the utility of both the lower and

the upper extremities for age estimation. While the Thompson method solved subjectivity issues by developing more objective variable definitions and employing stereological techniques to evaluate histological structures, significant methodological issues are still prevalent. Thompson measured only a small amount of cortex ($4mm^2$), which does not capture enough of the spatial variance known to occur within the cortex (Frost, 1969). The evaluation of the persiosteal aspect limits the utility of the method considering that younger individuals will likely not demonstrate significant remodeling in this area. Applying the principles of stereology in the form of point counting with a 10 x 10 grid reticule was a significant improvement to reduce observer error; however, the method does not provide an accurate assessment of remodeling events within the grid. It is well understood that osteons accumulate over time, thus methods that count the number of osteons per cortical area are likely to provide a more accurate assessment of age. The Thompson method evaluates the percentage of osteonal area, meaning that a field of view containing fewer osteons with larger osteonal areas could provide similar results to a field of view with a higher osteon density but smaller osteonal areas.

It is important to reiterate that the criticisms mentioned above regarding histological methods are not limited to the Kerley and Thompson methods. Furthermore, criticisms of these methods do not lessen the impact these researchers made in the field of anthropology. While recognizing past achievements is important, it should also be noted that there are methodological issues with these pioneering methods that must be contemplated so improvements can be made in future histological techniques.

Purpose, Goals, and Objectives

It is apparent that the accuracy and reliability of histological methods using the femur, in particular, are highly debatable and exhibit significant levels of sampling error and observer error. Methods with vague descriptions of samples, procedures, variables, or potential error rates should not be considered for use in skeletal analysis. Thus, new methods are needed to improve scientific standards within the field. Furthermore, methods that provide accurate age estimates for adult individuals, especially those over 50 years of age, are desperately needed. The primary goal of this research is to develop a new method for the estimation of age using the anterior femur midshaft that reduces the methodological issues previously discussed and improves adult age estimation. Additionally, the goal of this study is to describe the relationships of femur cortical bone histomorphometrics with age and sex for the sample. This research design addresses these issues and strives to produce a new standard for histological age estimation from the femur.

Research Design and Methods

Sample

The research sample includes three histological collections of known age, sex, and ancestry (Tables 1 and 2). Considering that 90.3% of the study sample is composed of individuals of European ancestry, the sample was not subdivided into ancestral groups. The Ericksen sample consists of midshaft femur cross-sections from 286 individuals (144 males, 142 females). Dr. Mary Ericksen developed the original collection with samples removed from George Washington University Medical School cadavers, cemetery

remains from the Dominican Republic, and autopsy specimens from Chilé. The Kerley sample consists of midshaft femur cross-sections from 15 individuals (9 males, 6 females) selected from the original collection consisting of 126 individuals. The Forensic Anthropology Unit (FAU) sample consists of anterior midshaft femur cross-sections from 27 individuals (19 males, 8 females) from forensic cases received by the Office of Chief Medical Examiner-New York City.

In order to perform a validation of the age prediction equation, a developmental set (n = 268) and validation set (n = 60) was extracted from the total femur study sample of 328 individuals. Table 3 provides the adjusted sample numbers for the developmental and validation sets.

Table 1. Descriptive statistics for the femur collections separated by sex.

	Male Sample				Female Sample			
Sample	N	Range	Mean	STDEV	N	Range	Mean	STDEV
Ericksen Collection	144	30-97	67.67	11.93	142	35-96	72.14	12.43
Kerley Collection	9	15-63	31.78	16.71	6	36-76	60.33	17.45
FAU Collection	19	23-87	45.89	16.02	8	19-70	30.00	16.49

Table 2. Ancestry composition of the sample.

	Ericksen		Kerley		FAU		Total	
Ancestry	N	Percentage	N	Percentage	N	Percentage	N	Percentage
White	265	95.0%	9	60%	15	57.7%	289	90.3%
Black	13	4.7%	6	40%	5	19.2%	18	5.6%
Asian	1	0.3%	0	0%	0	0%	1	0.3%
Hispanic	0	0%	0	0%	5	19.2%	5	1.6%
Not reported	0	0%	0	0%	1	3.8%	7	2.2%

Table 3. Final developmental and validation sample separated by sex

	Male Sample				Female Sample			
Sample	*N*	Range	Mean	STDEV	*N*	Range	Mean	STDEV
Total Study Set (*N* = 319)	169	15-97	63.36	15.84	150	19-96	69.53	16.17
Developmental set (*N* = 259)	139	15-97	62.73	16.40	120	19-96	69.88	16.77
Validation set (*N* = 60)	30	30-88	66.27	12.82	30	30-92	68.13	13.65

Histological Methods

Preparation of the femur samples followed standard methods for microscopic analysis (Frost, 1958; Stout and Paine, 1992; Maat et al., 2000). The following histomorphometric variables were collected using an Olympus BX41 transmitted light microscope fitted with a Merz eyepiece reticule:

1. Surface Area (Sa.Ar) in mm^2: Amount of cortical bone evaluated per microscopic field.
2. Intact Secondary Osteons (On.): Number of secondary osteons with an intact Haversian canal bounded by a scalloped reversal line.
3. Fragmentary Secondary Osteons (Fg.On.): Number of secondary osteons with a partially visible Haversian canal that has been breached either by a neighboring osteon or a resorptive bay and secondary osteons with no remnants of a Haversian canal present.
4. Mean Osteonal Cross-Sectional Area (On.Ar) in mm^2: average area of bone contained within the cement lines of structurally complete secondary osteons (reversal lines are intact) calculated as the average cross-sectional area of a minimum of 50 complete osteons per cross-section. Intact osteons with Haversian canals that have maximum diameters more than twice their minimum diameters are excluded.
5. Anterior cortical width (Ant.Ct.Wi.) in mm: A cortical bone thickness measure taken at the anterior aspect of the cross-section from the periosteal to the endosteal surface.

The following variables were calculated using the collected data:

1. Intact Osteon Population Density (OPD(I)) in #/mm^2: number of secondary osteons per unit area divided by the Sa.Ar.
2. Fragmentary Osteon Population Density (OPD(F)) in #/mm^2: number of fragmentary secondary osteons per unit area divided by the Sa.Ar.
3. Osteon Population Density (OPD) in #/mm^2: sum of OPD(I) and OPD(F).

An Olympus DP72 digital camera and associated imaging software were used to capture the digital images from each thin section to facilitate area and width measurements. Measurements for osteon area and anterior cortical thickness were taken from calibrated photomerged images using the image analysis software ImageJ (2009).

The topographic sampling method developed for this research required ten 0.48 mm wide columns from the periosteum to endosteum to be evaluated from the anterior midshaft of the femur (Figure 1). A 5 mm wide field at the anatomical midline (anterior midline) from periosteum to endosteum was marked on each slide. These markings also ensured that observers read the same overall field during the observer error analyses. Alternate fields were examined within each column at 200x magnification with a Merz reticule, forming a checkerboard pattern across the sampling area. This sampling strategy should account for over 90% of the variability of the histological structures within the anterior cortex at midshaft (Iwaniec 1997; Iwaniec et al., 1998).

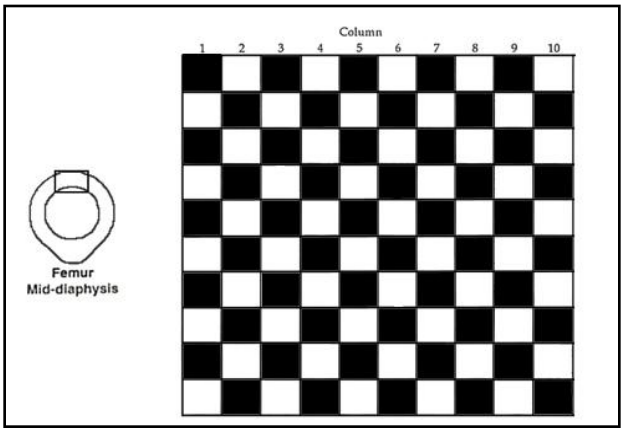
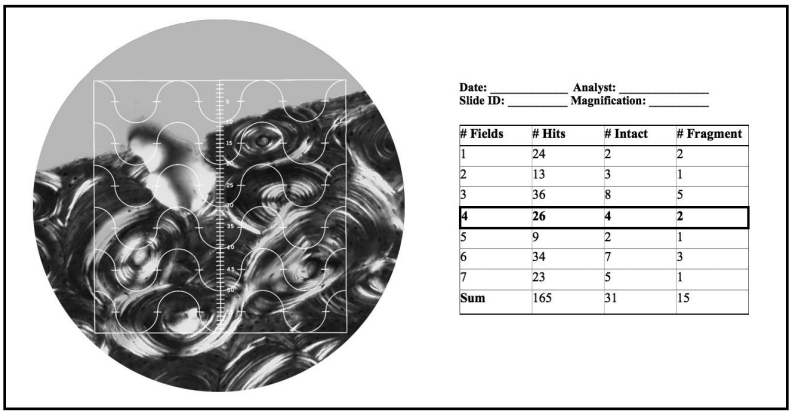


Figure 1. Representation of the sampling method for the proposed research derived from the Iwaniec et al. (1998) study. Each square is equivalent to 0.2304 mm^2 using a 10x ocular and 20x objective.

The Merz reticule contains a square that forms a "region of interest", or ROI, in which histological variables were collected (Figure 2). To calculate the surface area of each "hit" (defined by the tick marks across the line) on the reticule you simply divide the reticule area by the total number of "hits" (36 in this study) or measure the distance between the individual "hits" and calculate the area of each "hit". This point count technique was used to gather the surface area (bone area) data only. Once the surface area is calculated for a field, intact and fragmentary osteons within the ROI of the Merz reticule are counted before moving to the next field. Microstructures, defined previously, that intersect the grid boundary lines will be included within the grid count if more than 50% of the structure is visible within the ROI.



Date: _____ Analyst: _____
Slide ID: _____ Magnification: _____

# Fields	# Hits	# Intact	# Fragment
1	24	2	2
2	13	3	1
3	36	8	5
4	**26**	**4**	**2**
5	9	2	1
6	34	7	3
7	23	5	1
Sum	165	31	15

Figure 2. A superimposed Merz counting reticule containing 36 intersections or "hits". The table represents a partial datasheet for recording the number of hits over cortical bone and the intact and fragmentary osteon counts for each field.

Statistical Methods

Age regression equations have not been established using a sampling protocol that accounts for over 90% of the microstructure variability from the periosteal to the

endosteal surfaces of the femur cortex; therefore, the strength of the relationship between microstructures and chronological age needs to be evaluated in detail. Linear regression and non-linear regression models were examined. Statistical analysis and the generation of age prediction models were performed using Excel and SPSS software.

Quantifying observer error

Two observers (the PI and research assistant) with varying levels of experience performed the observer error analysis. To evaluate inter- and intra-observer error several methods were utilized to provide for a comprehensive analysis. First, the percent mean absolute difference (PMAD) was calculated for OPD, OPD(I), and OPD(F) for both observers (inter-observer) and iterations (intra-observer) using the following equation:

$$PMAD = \sum_{l=1}^{n} \frac{\frac{|x_i - x_j|}{\frac{(x_i + x_j)}{2}} \cdot 100}{n}$$

where *l* indicates the subject, *n* is the total number of samples, *x* is the observation made by a particular observer for OPD, OPD(I), or OPD(F), and *i* and *j* indicate observer 1 and 2 (inter-observer error) or observation 1 and 2 (intra-observer error). A 10% threshold was used to determine acceptable PMAD values in this study.

The second analysis of observer error involved determining the technical error of measurement (TEM) for each variable using:

$$TEM = \sqrt{\frac{\sum_{l=1}^{n} z_l^2}{2n}}$$

where *n* is the number of samples and z_l is the difference between observers/observations 1 and 2 for the *l*th subject (Mueller and Martorell, 1988). The TEM assess repeatability

and provides an approximation of the standard deviation of the differences between paired measurements, thus providing a measurement of the imprecision variance. Additionally, because TEM is in the units of the variable assessed, the TEM of different variables cannot be readily compared. Therefore, the reliability coefficient (R) was calculated to allow for inter-variable comparisons using:

$$R = 1 - \left(\frac{TEM^2}{SD^2}\right)$$

This coefficient indicates the proportion of the measurement variance that is error free. The value ranges from 0 to 1, where *R* values close to 1 indicate a high reliability (Flohr et al., 2010). Because TEM and *R* statistics cannot be used to assess observation bias in the measurements, paired *t*-tests were evaluated and Bland and Altman (1986) plots were examined.

The Bland and Altman (1986) method illustrates the repeatability within and between observers. Repeatability coefficients were calculated by taking the sum of the square differences, dividing by *n*, and then taking the square root to obtain the standard deviation. Because the true measurement value is unknown for the samples, a plot of the difference between values against the mean of the values was performed. Repeatability is achieved if the mean difference is not significantly different from zero and if 95% of the differences are less than two standard deviations. If the mean difference is significantly different from zero, the data cannot be used to assess repeatability.

Variable analysis and the age prediction model

There were several predictor variables of interest: OPD, OPD(I), OPD(F), On.Ar, and Ant.Ct.Wi. Descriptive statistics were calculated and graphed for each variable to

evaluate the normality of the distributions. Next, analysis of covariance (ANCOVA) was used to evaluate the age and sex effect for each variable that did not violate assumptions. With age as the covariate, results will indicate if sex differences are significant. A general linear model accompanied by the Chow test was used to determine if the slopes and intercepts are equal between the sexes.

For comparison with other studies, individuals were grouped by 10 year age cohorts and a two-way analysis of variance (ANOVA), with age, sex, and age-sex interaction terms as prediction variables, was performed. The analysis was run separately for OPD considering the variable is composed of OPD(I) and OPD(F). Next, the variables were examined to determine if a general model could be developed for predicting age for both sexes, or if sex-specific models are required.

Finally, linear regression analysis involved a forward stepwise procedure to develop age prediction models from the developmental set. The equations generated from the developmental set were used to estimate age on the validation set. Inaccuracy and bias values were then calculated for the validation sample. *T*-tests were used to determine if the mean difference in age estimates is significantly different from zero. As a final step, the developmental and validation sets were pooled and used to obtain the final regression model(s).

Findings and conclusions

Observer error

The results for the intra- and inter-observer error analysis are provided in Table 4. Intra-observer error results indicate that repeatability was achieved between OPD counts.

The PMAD values for the OPD component variables (OPD(I) and OPD(F)) are higher (5.6% and 11.8%, respectively), with osteon fragments exceeding the suggested 10% threshold. Based on the remaining criteria, repeatability was achieved within OPD(I) and OPD(F) counts.

Inter-observer error results indicate that repeatability was achieved between OPD counts. The PMAD values for the OPD component variables, OPD(I) and OPD(F), are 5.9% and 14.8%, with osteon fragments again exceeding the suggested 10% threshold. Based on the remaining criteria, repeatability was achieved for OPD(I) and OPD(F) counts.

Table 4. Observer error results for the various tests.

Intra-observer error	PMAD	TEM	R	repeatability coefficient	Mean Difference	sig.
OPD	4.8%	1.01	0.98	±2.9	-0.3	0.33
OPD(I)	5.6%	0.75	0.95	±2.0	-0.3	0.08
OPD(F)	11.8%	0.94	0.96	±2.7	0.08	0.76
Inter-observer error						
OPD	6.4%	1.35	0.97	±3.7	-0.6	0.10
OPD(I)	5.9%	0.74	0.97	±2.1	-0.3	0.17
OPD(F)	14.8%	1.12	0.95	±3.2	0.3	0.29

It should be noted that observer error values will be slightly inflated considering that the exact location of the microscopic region of interests (ROIs) fluctuate between trials. Although the analyst begins the evaluation of the cortex in the same location, it is unlikely that the same ROIs are reproduced. Fragmentary osteon density (OPD(F)) passed one of the two tests for observer agreement. Considering that the exact same fields were not measured, more weight is given to the Bland and Altman test of

repeatability demonstrating that OPD(F) achieved repeatability. Comparing these results to previous analyses of variable error using Stout and Paine's (1992) original variable definitions demonstrates a significant decrease in OPD(I) and OPD(F) error values. Crowder (2005) reported a PMAD value of 11.2% for OPD(I) and 22.8% for OPD(F). The new definitions used in this study cut the counting error approximately in half for each constituent OPD variable. While it is clear that the revised variable definitions significantly reduce observer error in histological analyses, differences in OPD(F) compared to OPD(I) indicates continued problems in differentiating fragmentary osteons.

Variable analysis

Sex, age, and sex-age interaction were examined for the four histological variables: OPD(I), lnOPD(F), lnOn.Ar, and Ant.Ct.Wi. The histological variables all demonstrate relationships with age, similar to previous studies. However, age and sex effects are not consistent between the variables. Surprisingly, the OPD(I) correlation was low for both groups ($r = 0.44$ in males and $r = 0.274$ in females). Equally interesting is the difference in correlation strength between males and females, with females showing weaker correlation to age for OPD(I). These findings suggest that age-related turnover events are best expressed in osteon fragments, especially in females.

Overall, the results of the variable analysis indicate that age and sex differences do exist in the histological variables; however, results are complex and are likely complicated by unequal or small sample sizes within the age cohorts. Regardless, it is reasonable to conclude that separate regression models for males and females are

warranted. Sex differences observed in histological variables are likely related to biological factors involving the endocrine system that affect bone turnover in the female skeleton such as pregnancy, lactation, and menopause. This can be observed in the evaluation of the sex, age cohort, and the sex-age cohort interaction effect for Ant.Ct.Wi, which were significant (p = 0.000, p = 0.000, and p = 0.001, respectively). Males typically have larger cortical thickness values and exhibit less endosteal expansion over time compared to females. A profile plot demonstrates that females exhibit significant cortical bone loss over time (Figure 3).

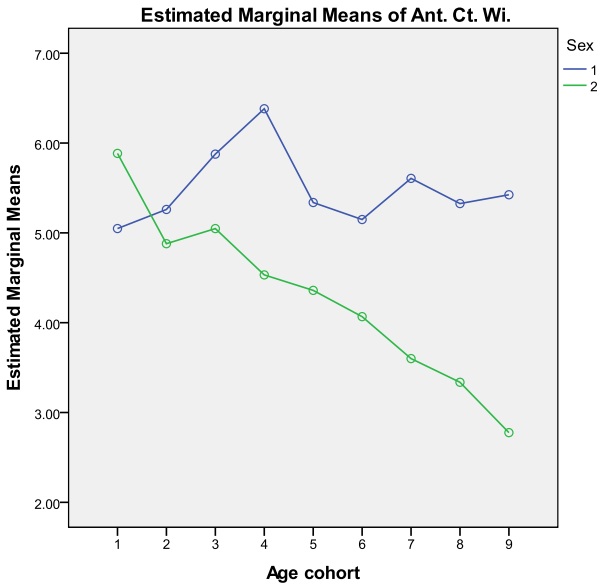


Figure 3. Profile plot for Ant.Ct.Wi. demonstrating cortical bone loss over time (Males = 1, Females = 2). The horizontal axis represents each 10-year age cohort.

Age prediction Model

Stepwise regression analysis for the male sample selected two predictor variables: lnOPD(F) and OPD(I). The regression model produces a standard error of the estimate of

11.24, which is comparable to other histological studies. Stepwise regression analysis for the female sample selected three predictor variables: lnOPD(F), Ant.Ct.Wi., and OPD(I) with a standard error of the estimate of 9.91. Stepwise regression analysis for the pooled sex sample selected three predictor variables: lnOPD(F), OPD(I), and lnOn.Ar. and produced a standard error of the estimate of 10.99.

When the three prediction models were applied to the validation set and the estimated ages were compared to known age at death, the mean differences of the ages do not significantly differ from zero (Tables 5 and 6). Approximately 60% fall within ±10 years of the known age for the male sample and approximately 67% fall within ±10 years of the known age for the female sample. Approximately 66% fall within ±10 years of the known age for the pooled sex model.

The developmental and validation sample were pooled into one reference sample to produce the final age prediction models for males, females, and unknown sex samples. The age prediction equations are provided in Table 7. The standard error of the estimates for males, females, and pooled sex equations are 11.13, 9.77, and 10.70, respectively.

Table 5. Comparison of estimated ages and known ages from the validation set using the sex-specific equations.

	Males (*N*=30)		**Females (*N*=30)**	
	Difference	Absolute Difference	Difference	Absolute Difference
Mean	-1.40	8.60	-0.18	7.67
STDEV	10.79	6.06	7.69	8.09
Standard Error of Mean	1.97	0.87	1.86	1.20
P > |T|	0.484	0.000	0.922	0.000

Table 6. Comparison of estimated ages and known ages from the validation set using the unknown sex equation.

Pooled Sex (*N*=60)	Difference	Absolute Difference
Mean	0.15	7.52
STDEV	9.45	5.64
Standard Error of Mean	1.22	0.73
P > |T|	0.903	0.000

Table 7. Age prediction equations for the three models.

Sex	**Prediction Equation**
Males	Age = 6.638 + 20.355*(lnOPDF)+1.121*(OPDI)
Females	Age = 25.372 + 20.192*(lnOPDF)-3.441*(Ant.Ct.Wi.)+0.714*(OPDI)
Unknown	Age = -11.783 + 20.657*(lnOPDF)+0.617*(OPDI)-7.860*(lnOn.Ar.)

In all models, fragmentary osteon population density was determined to be the best predictor of age at death. The existence of sex and age related differences associated with the histological variables indicate that sex specific age-prediction equations should be applied when estimating age using cortical bone histomorphometry. It should be noted that the equations generated from this study are not accurate with younger individuals. This is likely due to the lack of younger individuals in the reference sample. Furthermore, the majority of the young individuals were obtained from forensic cases, which may represent individuals in poor health due to substance abuse or nutritional issues. Regardless, the focus of this research was to improve age estimation for older individuals and provide a method that may be less affected by the reported asymptotic

value for osteon population density. The results from this study suggest that asymptote in osteon density does not occur in the femur until around 80 to 90 years of age in males, although there is considerable individual variation. This can be observed in further analysis of the validation sample, in which the oldest individuals demonstrate higher differences between estimated and known age. The age estimates tend to significantly underage males 80 to 90 years of age, which supports the hypothesis that the bone has "remodeled out" and reached the asymptote. The female data from this research does not demonstrate the asymptote when evaluating the sample by age cohort, which supports results indicating that the prediction model performs better for females.

Implications for Policy and Practice

Accurate reporting of age at death for the biological profile is imperative in determining the inclusion or exclusion of individuals from a pool of missing persons. It has been suggested that histological age indicators may provide more accurate age estimates considering that they are a product of continuous bone turnover and not the result of degenerative changes in bone morphology; however, the results of this study (and others) suggest that biological variability is significant in bone turnover between and within groups. This suggests that in some aspects, histological methods may perform similar to gross age indicators. Regardless, histological methods appear promising to predict age past the 50+ boundary. While previous histological methods evaluating the femur demonstrate significant methodological issues that affect their reliability, the method developed from this research significantly reduces method error.

This research suggests that the assessment of histological age indicators should be coupled with macroscopic (gross) methods to provide a more comprehensive age estimate. The authors also suggest that the evaluation of gross age indicators should include histological analysis to assist with an assessment of skeletal health, which may indicate why, for some cases, gross indicators do not correlate well with chronological age. Unfortunately, the sample used in this study did not allow for observations of gross structures.

One significant limitation to this study was the skewed age distribution due to the lack of young individuals in the reference sample. While aspects of this were dealt with statistically, further resolution will require additional histological samples. The authors have located an additional sample of younger individuals, which will be evaluated in the near future.

1. INTRODUCTION

The estimation of age at death is an essential part in the reconstruction of population demographics and the individual analysis of human remains. Estimating the age at death of children and young adults can be performed with greater accuracy owing to methods that are based on the growth and development of the human skeleton. Skeletal growth and development are regulated by endocrine and genetic factors producing biological age indicators that have a more predictable relationship with chronological age. The estimation of age at death for adults demonstrates a progressive decrease in accuracy as chronological age increases. Most methods for estimating adult age at death rely on assessments of degenerative changes to gross bone morphology. The traditional gross anatomical age indicators include skeletal features such as the pubic symphyseal face (Brooks and Suchey, 1990), auricular surface (Lovejoy et al., 1985a, 1985b; Buckberry and Chamberlain, 2002), and mid-thoracic sternal rib ends (İşcan et al., 1984; İşcan, 1993). When these surfaces are altered by post-depositional taphonomic factors, anthropologists turn to age indicators that have less defined relationships with chronological age, such as patterns of cortical involution, cranial suture closure, and the presence or absence of degenerative joint disease. It has been suggested that quantitative cortical bone histology, or histomorphometry, provides a reliable approach to adult age estimation (Stout and Gehlert, 1982; Ubelaker, 1986). Furthermore, histological methods show greater potential to bridge the 50+ age boundary that limits the accuracy of the aforementioned methods (Thompson, 1979). It is for these reasons that researchers continuously advocate the use of histological methods in anthropological analyses (Stout, 1986; Stout and Paine, 1992; Cattaneo, 1999; Cho et al., 2002; Crowder, 2005).

Since the introduction of the first quantitative histological approach for the estimation of age at death by Ellis Kerley in 1965, histological parameters of age-related bone turnover have been well-documented in the anthropological literature for various skeletal elements. Crowder (2005) reviewed some of the most frequently employed methods, determining that the assessment of bone histology is a useful method for estimating age, producing accuracy values that are comparable to many gross morphological methods. Considering the myriad of methods available, only a few have received significant attention within the field culminating in multiple validation studies applied to archaeological, cadaveric, and forensic samples. Despite this, histological methods are not widely applied in forensic anthropology owing to the inherent methodological issues identified in the literature (Lynnerup et al., 1998; Villa and Lynnerup, 2010) and the specialized knowledge required in interpreting normal vs. abnormal bone histomorphology. To facilitate the discussion of these issues in applying histological methods, a brief background in bone histology and histological age methods is necessary.

1.1 Cortical Bone Histomorphology

In humans, mature bone consists of organized tissue called lamellar bone. In a transverse cross-section of human cortical bone, five patterns of lamellar bone are recognizable: concentric layers as seen in the secondary osteon, inner and outer circumferential lamellae bordering the periosteum and endosteum, primary interstitial lamellae, and osteonal interstitial lamellae, which are segments of remodeled osteons. Primary lamellar bone consists of inner and outer circumferential lamellae and primary interstitial lamellae. It is laid down *de novo* on a pre-existing bone surface. Secondary lamellar bone results from the resorption of existing bone and consists of concentric lamellae within osteons and interstitial lamellae from pre-existing osteons.

Lamellar bone forms two different types of osteons: primary and secondary. Primary osteons, more correctly referred to as primary vascular canals, are formed by the deposition of fine fibered sheets of concentric lamellae that bend slightly over and under them (Maggiano, 2012). As the bone ages, the primary osteon is resorbed and a new series of concentric lamellae are laid down, forming a secondary osteon (Weiner et al., 1999). The canal that is surrounded by the secondary osteon lamellae is referred to as the Haversian canal. Thus, the structural unit of mature cortical bone is identified as a Haversian system or secondary osteon.

Bone remodeling, or internal bone turnover, is the process of continuous removal of older bone with the replacement of new bone throughout life. Bone remodeling occurs through the localized coupling of osteoclasts and osteoblasts forming an assembly of cells called the Basic Multicellular Units, or BMUs (Frost, 1969; Jaworski, 1984). Bone remodeling is often described as having two functions: microscopic fracture repair and maintaining metabolic homeostatsis of the bone matrix. Microscopic fracture repair is likely the primary function of bone remodeling, allowing the skeleton to adapt to its mechanical environment by reducing the risk of fractures and repairing damage created by repetitive cycles of mechanical loading (Burr 2002). Remodeling can occur within the four bone envelopes: periosteal, Haversian (intracortical), endosteal, and trabecular surfaces (Frost, 1969, 1987a; Parfitt, 2001, 2002). Because each envelope is distinct, remodeling can occur at different times, rates, and magnitudes within the bone.

1.2 Histological Age Estimation

The premise of histological age estimation is based upon clinical and anthropological research demonstrating that bone turnover in cortical bone occurs at a predictable rate over an

individual's lifetime. In theory, the length of time during which remodeling occurs (chronological age) will be the primary influence on how many secondary osteon creations (intact and fragmentary osteons) accumulate per unit of area (Kerley, 1965; Wu et al., 1967, 1970; Stout and Teitelbaum, 1976). Clinical studies evaluating rib tissue indicate this relationship should be evident in a normal adult until remodeling rates begin to fluctuate as homeostasis is compromised by senility (Wu et al., 1970). Eventually, an osteon asymptote may be reached when new osteon creations remove all evidence of older ones (Wu et al., 1970; Frost, 1987; Stout and Stanley, 1991; Stout and Paine, 1994; Robling and Stout, 2000). This phenomenon is discussed in more detail below.

Since the introduction of Kerley's method numerous others have followed; however, the reliability of quantitative bone histology for adult age estimation has yet to be fully demonstrated. The reliability of cortical bone histomorphology in estimating age at death is dependent on measuring the amount of non-stochastic cortical bone remodeling between individuals of the same chronological age within and between populations, the strength and predictability of the relationship between bone remodeling and chronological age, and the levels of accuracy and precision produced by the method. Most research has focused on exploring the first two aspects, while the latter was not systematically tested using a large known population until Crowder's research in 2005. As such, there is a large body of research exploring the changes in human microstructure over chronological age for various skeletal elements (Kerley, 1965; Ahlqvist and Damsten, 1969; Singh and Gunberg, 1970; Thompson, 1979; Thompson and Galvin, 1983; Stout, 1986; Clarke, 1987; Ericksen, 1991; Stout and Paine, 1992; Stout et al., 1994, 1996; Yoshino et al., 1994; Cool et al., 1995; Cho et al., 2002; Curtis, 2004; Kim et al., 2007).

While microscopic age changes are considered to be universal, inconsistencies in the reported accuracy of the methods when they are applied to individuals outside of the reference samples suggest that intrinsic and extrinsic biological factors, such as genetics and a wide range of suggested behaviors, have varying effects on bone microstructure. One such factor is the aforementioned osteon asymptote. The asymptote occurs when the bone has become completely remodeled so that new osteon creations simply replace older ones and the proportion of remodeled cortex does not increase (Frost, 1987a, 1987b). According to the literature, ribs reach this asymptote around the 5^{th} or 6^{th} decade of life resulting in the inability to accurately age older individuals. While histological methods utilizing the ribs (Cho et al., 2002) are currently the most accurate and reliable methods available, they cannot be used with certainty when applied to older individuals because of this biological phenomenon. It is believed that the femur, due to its larger cortical area, does not reach this asymptote until later in life making it a preferred location for histological analysis of older individuals (Stout and Paine, 1992).

As noted by Lazenby (1984), biological issues may all be secondary to concerns relating to the design of methods. Methodological concerns include sample size, sample demography, variable definitions, and measurement techniques. Currently there are two preferred methods for estimating adult age at death using the femur: the Kerley method and the Thompson method. Each contains separate, but significant methodological issues affecting their accuracy and reliability. Kerley's age regression formulas, derived from 126 undecalcified cross-sections taken from the midshaft of the femur, tibia, and fibula of individuals of known age and sex, were based on four predicting variables including intact osteons, osteon fragments, non-Haversian canals, and percentage of circumferential lamellar bone. The variables were observed using four circular fields within the outer third of the cortex adjacent to the periosteal surface of the bone.

The individual variables were counted within each field, including those partly obscured by periphery of the field, and then totaled across all four fields to create a composite value. The percentage of circumferential lamellar bone was averaged for all four fields. These raw counts were then used to develop four different regression models to use when estimating age from a single bone slide.

Kerley and Ubelaker (1978) revised the original Kerley (1965) paper, warning investigators that the variability in field diameters of different microscopes would contribute to "apparent errors" and "unreasonable [age] estimates". Kerley and Ubelaker (1978) realized that using a smaller field size, as opposed to the original field size, would underestimate age since the sum of recorded structures is always less than that recorded when the regression models were created. During this revision, it became apparent that the original microscopes were not available for inspection and a survey of available microscopes suggested that the original field diameter used by Kerley was most likely 1.62 mm at 100x magnification, rather than the previously reported 1.25 mm diameter. A 1.62 mm field diameter results in an area 2.06 mm^2, indicating the method required a field correction factor. Stout and Gehlert (1982) suggest that the use of the correction factor to adjust osteon counts may be of limited value due to the spatial variation of microstructures within the cortex, thus resulting in method inaccuracy if the original Kerley field size is not applied. Because the original field size is still unknown, the conservative approach is that the method should not be used for analysis.

Other issues associated with, but not limited to, the Kerley method include subjective variable definitions and the inability to incorporate remodeling events from the periosteal envelope to the endosteal envelope. The former point is a main factor in controlling the level of observer error associated with how researchers classify osteonal structures or, in other words,

how intact and fragmentary osteons are differentiated. For example, Kerley (1965) classified intact osteons as being 80% intact with a complete Haversian canal present, while Stout and Paine (1992) defined an intact osteon as having a Haversian canal that is 90% intact. Figure 1 demonstrates the issue of selecting observation fields in specific locations rather than sampling from the periosteal envelope to the endosteal envelope. Using a standard field size (indicated by the yellow circle), the amount of remodeling variation captured between the two individuals is not equal. The image on the left represents a young individual with a large cortex and the image on the right represents the cortex of an elderly individual that demonstrates age-related endosteal resorption. In the young individual, histological features associated with the periosteal envelope would be evaluated, while in the older individual the histological variability from the periosteal to the endosteal surfaces is represented by a single field.

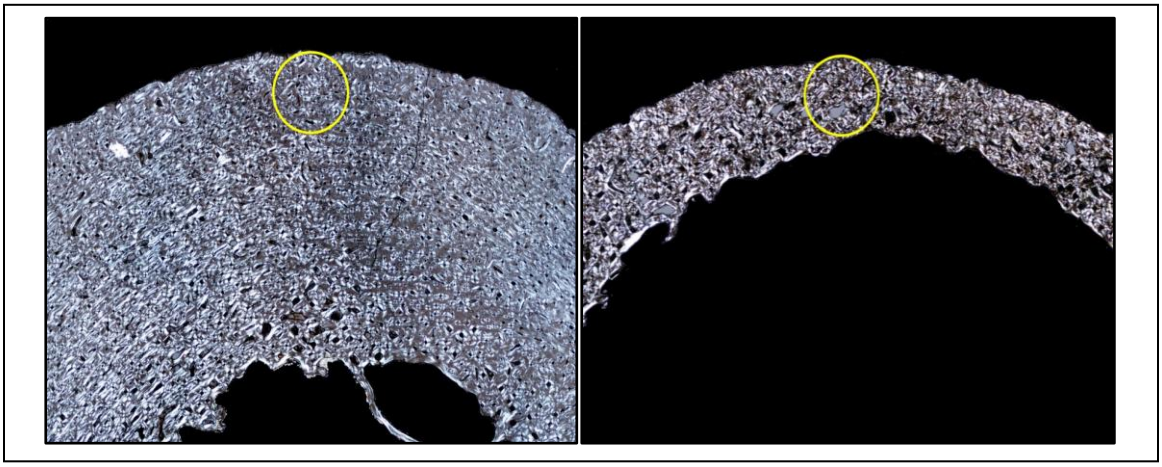


Figure 1. Two individuals demonstrating variation in cortical thickness with a standard field size represented by the yellow circles. The field on the left image evaluates structures located within the periosteal envelope, while the field on the right evalautes the periosteal, intracortical, and endosteal envelopes.

In 1979, Thompson published a method designed to minimize the amount of destructive sampling, reduce observer error, and explore the utility of both the lower and the upper extremities for age estimation. Thompson sampled bone cores (0.4 cm in diameter) from the femora and tibiae of 116 cadavers and the humeri and ulnae of 31 cadavers. Various histomorphometric measurements were recorded through point counting using a 10 x 10 grid eyepiece reticule. Overall, Thompson explored 19 variables, including a number of gross measurements (for example, core weight, cortical thickness, and cortical density). One variable, referred to as the percentage of osteonal area, was determined to be the single best indicator of age. Histological structures were recorded in four contiguous microscopic fields along the core's anterior periosteal surface. Both sex- and side-specific regression equations were developed.

While the Thompson method solved subjectivity issues by developing more objective definitions and employing stereological techniques to evaluate histological structures, significant methodological issues are still prevalent. Thompson measured only a small amount of cortex (4 mm^2), which does not capture enough of the spatial variance known to occur within the cortex (Frost, 1969). The evaluation of the persiosteal aspect limits the utility of the method considering that younger individuals will likely not demonstrate significant remodeling in this area. Applying the principles of stereology in the form of point counting with a 10 x 10 grid reticule was a significant improvement to reduce observer error; however, the method does not provide an accurate assessment of remodeling events within the grid. It is well understood that osteons accumulate over time, thus methods that count the number of osteons per cortical area are likely to provide a more accurate assessment of age. The Thompson method evaluates the percentage of osteonal area, meaning that a field of view containing fewer osteons with larger osteonal areas could provide similar results as a field of view with a higher osteon density but

smaller osteonal areas (Figure 2). This demonstrates that osteon population density would provide a stronger correlation to skeletal age. Lastly, Thompson only records point count "hits" over osteons with an intact or partially visible Haversian canal. This means that osteon fragments lacking a Haversian canal are not included, thus decreasing the amount of observed remodeling events and the correlation strength between bone turnover and skeletal age.

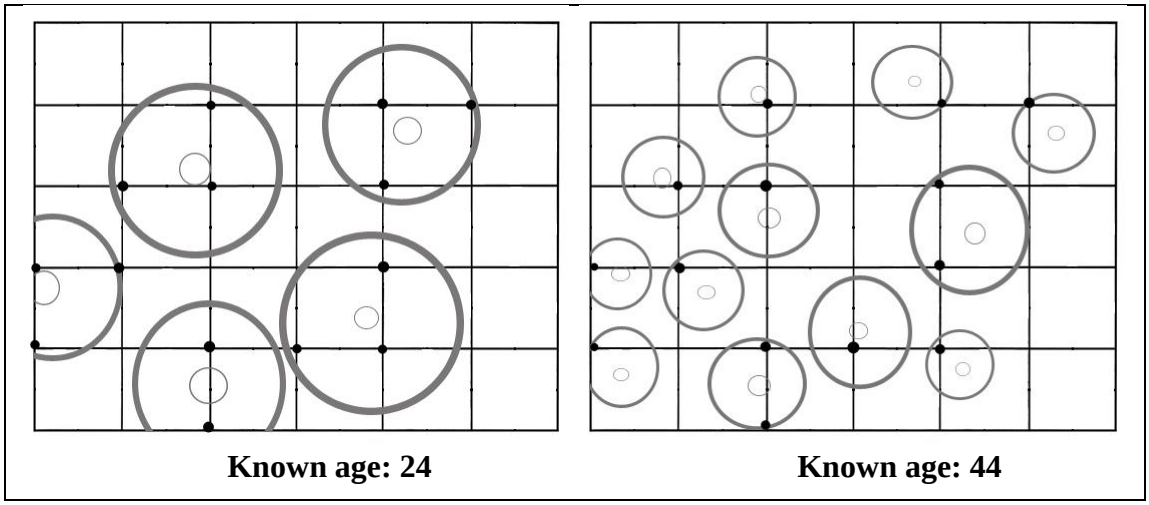


Figure 2. Illustration depicting the point count method to determine percentage of osteonal bone using the same level of magnification. The circles represent osteons with Haversian canals and the small dark circles indicate the number of grid "hits" over osteonal bone. Both fields have a point count of 14 "hits" and would result in identical age estimates with the Thompson (1979) method although there is a 20-year difference between known ages in this hypothetical example.

It is important to reiterate that the criticisms mentioned above regarding histological methods are not limited to the Kerley and Thompson methods. Furthermore, criticisms of these methods do not lessen the impact these researchers made in the field of anthropology. Dr. Kerley provided the first histological method to estimate age which also had the advantage of being based upon multiple skeletal elements. His pioneering work spawned decades of research building upon the original method. Dr. Thompson developed a method designed to minimize the amount of bone sampled from an individual and explored the use of elements from the lower and

upper extremities. While recognizing past achievements is important, it should also be noted that there are methodological issues with these pioneering methods that must be contemplated so improvements can be made in future histological techniques.

It is apparent that the accuracy and reliability of histological methods using the femur, in particular, are highly debatable and exhibit significant levels of sampling error and observer error. Methods with vague descriptions of samples, procedures, variables, or potential error rates should not be considered for use in skeletal analysis. Thus, new methods are needed to improve scientific standards within the field. Furthermore, methods that provide accurate age estimates for adult individuals, especially those over 50 years of age, are desperately needed. The goal of this research is to develop a new method for the estimation of age using the anterior femur midshaft that reduces the methodological issues previously discussed and improves adult age estimation. There is an increasing need for more accurate indicators of adult skeletal age in order to improve the assessment of unidentified human remains from the forensic context. Although various histological methods have been revised in the literature, the fundamental issues concerning the reliability and repeatability of these methods have not been fully addressed. This research design addresses these issues and strives to produce a new standard for histological age estimation from the femur.

2 MATERIALS AND METHODS

2.1 Sample

The research sample includes three histological collections of known age, sex, and ancestry (Tables 1 and 2). Considering that 90.3% of the study sample is composed of individuals of European ancestry, the sample was not subdivided into ancestral groups. Ancestry

was taken from the death records for each sample. The Ericksen sample consists of midshaft femur cross-sections from 286 individuals (144 males, 142 females) selected from the original collection consisting of 328 individuals[1]. Dr. Mary Ericksen developed the original collection with samples removed from George Washington University Medical School cadavers, cemetery remains from the Dominican Republic, and autopsy specimens from Chilé. The Kerley sample consists of midshaft femur cross-sections from 15 individuals (9 males, 6 females) selected from the original collection consisting of 126 individuals[2]. The Forensic Anthropology Unit (FAU) sample consists of anterior midshaft femur cross-sections from 27 individuals (19 males, 8 females) from forensic cases received by the Office of Chief Medical Examiner-New York City.

Table 1. Descriptive statistics for the femur collections separated by sex.

	Male Sample				Female Sample			
Sample	N	Range	Mean	STDEV	N	Range	Mean	STDEV
Ericksen Collection	144	30-97	67.67	11.93	142	35-96	72.14	12.43
Kerley Collection	9	15-63	31.78	16.71	6	36-76	60.33	17.45
FAU Collection	19	23-87	45.89	16.02	8	19-70	30.00	16.49

Table 2. Ancestry composition of the sample.

	Ericksen		**Kerley**		**FAU**		**Total**	
Ancestry	**N**	**Percentage**	**N**	**Percentage**	**N**	**Percentage**	**N**	**Percentage**
White	265	95.0%	9	60%	15	57.7%	289	90.3%
Black	13	4.7%	6	40%	5	19.2%	18	5.6%
Asian	1	0.3%	0	0%	0	0%	1	0.3%
Hispanic	0	0%	0	0%	5	19.2%	5	1.6%
Not reported	0	0%	0	0%	1	3.8%	7	2.2%

[1] While a total of 314 bone samples were received for analysis from the original collection, only 310 could be sectioned or analyzed due to bone fragility. Furthermore, 286 slides were used in the study, owing to the discovery that 28 cross-sections were paired samples from the same individuals (left and right).

[2] The original Kerley collection consisted of 68 femur, 33 tibia, and 25 fibula slides. Due to preservation issues with the mounting medium, only 15 of the 68 femur slides could be evaluated for this study.

In order to perform a validation of the age prediction equation, a developmental set (*n*=268) and validation set (*n*=60) was extracted from the total femur study sample of 328 individuals. During the preparation (see section below) and preliminary evaluation of histological samples (checking for slide clarity) several individuals were identified as exhibiting abnormal histomorphology. Evaluation of the cause of death (COD) indicated diagnoses that could cause the abnormal bone turnover seen in these individuals. Therefore, the individuals were removed from the analysis due to possible bone turnover issues. While other individuals within the sample have documented conditions that could affect bone turnover or show some evidence of senile osteoporosis, the overall histological appearance did not warrant removal. Furthermore, including these individuals will provide for a more robust model. The samples removed exhibit obvious atypical bone histomorphology, which would typically be identified as not applicable to histological age estimation during casework analysis. On a case-by-case basis, practitioners should determine whether remains of unknown individuals are appropriate for age estimation methods. Additionally, several samples did not produce usable slides due to bone integrity issues after slide preparation was complete. Table 3 provides the samples and reasons for removal. Table 4 provides the adjusted sample numbers for the developmental and validation sets.

2.2 Sample Preparation

The Kerley collection slides were prepared previously and did not require any preparation for microscopic analysis. Preparation of the femur samples for the Ericksen and FAU collections followed published methods (Frost, 1958; Stout and Paine, 1992; Maat et al., 2000). Due to the durability of the femur cortical bone, the samples were not embedded in plastic resin prior to

thick-sectioning. Each bone sample was secured in a C-shaped chuck assembly for sectioning in a manner that produced an anatomically transverse section. The 1mm thick-sections were removed using a *Buehler Isomet 1000* saw with a 15 HC (high concentration) diamond-edged blade. Each thick-section wafer was washed in an ultrasonic water bath to remove debris and then allowed to air dry. Once dry, the thick-section was ground to a final thickness of 50-100 µm on a Buehler™ variable-speed grinding unit with a diamond disc. Each thin-section was then covered with a glass cover slip using SECUREMOUNT mounting medium. The following information was recorded on each slide: 1) slide identifier, 2) element name, 3) element side, and 4) anatomical orientation labels (A, P, M, L).

Table 3. Samples removed from the study.

Sample	Reason for removal
M09-1347	Diagenesis
K08-3682	Pathological
E-0020	Pathological
E-0134	Pathological
E-0318	Anterior aspect broken
E-0325	Anterior aspect broken
E-0809	Pathological
E-0820	Pathological
E-1325	Pathological

Table 4. Final developmental and validation sample separated by sex

	Male Sample				Female Sample			
Sample	N	Range	Mean	STDEV	N	Range	Mean	STDEV
Total Study Set (N=319)	169	15-97	63.36	15.84	150	19-96	69.53	16.17
Developmental set (N=259)	139	15-97	62.73	16.40	120	19-96	69.88	16.77
Validation set (N=60)	30	30-88	66.27	12.82	30	30-92	68.13	13.65

2.3 Histological Methods

Histomorphometric data was collected using an Olympus BX41 transmitted light microscope fitted with a Merz eyepiece reticule to record microstructure counts and amount of bone surface area (see below for details). An Olympus DP72 digital camera and associated imaging software were used to capture the digital images from each thin section to facilitate area and width measurements. Measurements for osteon area and anterior cortical thickness were taken from calibrated photomerged images using the image analysis software ImageJ (2009). The histological variables are described in detail below:

1. Surface Area (Sa.Ar): Amount of cortical bone evaluated calculated by the number of reticule "hits" overlaying the cortex for each microscopic field. The sum of the "hits" is multiplied by the area represented by one hit on the reticule, yielding a total area of cortical bone evaluated in mm^2. Thus, the total area is the product of the number of fields observed and the grid factor, which is determined by the size of the eyepiece reticule and combination of oculars and microscope objectives. A 'hit" was recorded only if it overlaid cortical bone or resorption bays of forming osteons. Haversian canals or other areas of porosity and trabecular bone were not included.
2. Intact Secondary Osteons (On.): Number of secondary osteons with an intact Haversian canal bounded by a scalloped reversal line.
    a. If connected to multiple osteons by a clearly defined Volkmann's canal the structures should be counted as separate osteons
    b. If two or more structures appear to share a Haversian canal and/or share a scalloped reversal line due to the plane of sectioning including a branching event, then they are counted as one system

c. If it is difficult to discern whether osteons are branching or connected by a Volkmann's canal, then they are counted as one intact osteon
d. Hemiosteons, i.e., osteons found on the endosteal surface which are formed by osteoclast trenching rather than tunneling, are not counted.
e. Primary vascular canals (primary osteons) are not counted as intact secondary osteons.

3. Fragmentary Secondary Osteons (Fg.On.): Number of secondary osteons with a partially visible Haversian canal that has been breached either by a neighboring osteon or a resorptive bay and secondary osteons with no remnants of a Haversian canal present. Osteon fragments that lack a Haversian canal can be identified by concentric lamellar rings and the presence of a defined reversal line with a scalloped (irregular) margin. It is important to note that fragmentary osteons are distinct from interstitial lamellae, which represent unremodeled primary lamellar bone. It is also useful to observe the orientation of the osteocytic lacunae that are associated with the lamellae of adjacent osteons or osteon fragments. Osteon fragments are best observed using normal transmitted light that allows for better observation of the osteocytic lacunae orientation.
4. Mean Osteonal Cross-Sectional Area (On.Ar): Average area of bone contained within the cement lines of structurally complete secondary osteons (reversal lines are intact) calculated as the average cross-sectional area of a minimum of 50 complete osteons per cross-section. Intact osteons with Haversian canals that have maximum diameters more than twice their minimum diameters are excluded. Drifting osteons are not acceptable for measurement even if Haversian canals meet the required criterion. Drifting osteons are extremely eccentric secondary osteons that appear to

move through bone tissue forming a long tail behind them. Although there is no consensus on how to define or measure them, a reasonable description of these structures is that they have a circular (not oblique) Haversian canal with numerous waves of concentric lamellae forming a tail. Generally, a drifting osteon is considered as one intact osteon and caution must be observed when counting these structures. The numerous waves of concentric lamellae can be confused with osteon fragments, especially under polarized light. Counting these waves will produce exaggerated fragment counts. If there is clear evidence of reversal lines segmenting the tail of the suspected drifting osteon (viewed under nonpolarized light), then it is in reality an intact osteon with closely associated fragments that give the system the appearance of a drifting osteon.

5. Anterior Cortical Width (Ant.Ct.Wi.): A cortical bone thickness measure taken at the anterior aspect of the cross-section from the periosteal to the endosteal surface.

The following variables were calculated using the collected data:

1. Intact Osteon Population Density (OPD(I)): number of intact secondary osteons per unit area divided by the Sa.Ar.
2. Fragmentary Osteon Population Density (OPD(F)): number of fragmentary secondary osteons per unit area divided by the Sa.Ar.
3. Osteon Population Density (OPD): sum of OPD(I) and OPD(F).

These calculated variables were revised from Stout (1986), Stout and Paine (1992), and Cho and colleagues (2002).

The topographic sampling method for the femora was modeled after the research method described by Iwaniec (1997) and Iwaniec and colleagues (1998). The cortical bone from the anterior aspect of the femur was evaluated in columns from the periosteal to the endosteal envelope. According to Iwaniec and colleagues (1998), evaluating 1 mm^2 microscopic fields at 100x magnification within two columns of cortical bone from the periosteum to endosteum accounts for 95% of the variability within the anterior femur. Evaluating alternate columns, or 50% of the sampled area, predicted 98 to 99% of the anterior section total osteon density. The protocol developed for this research required ten 0.48 mm wide columns from the periosteum to endosteum to be evaluated from the anterior midshaft of the femur (Figure 3). A 5 mm wide field from periosteum to endosteum was marked on each slide as a guide. These markings also ensured that observers read the same overall field during the observer error analyses. Alternate fields were examined within each column at 200x magnification with a Merz reticule, forming a checkerboard pattern across the sampling area. This sampling strategy should account for over 90% of the variability of the histological structures within the anterior cortex at midshaft.

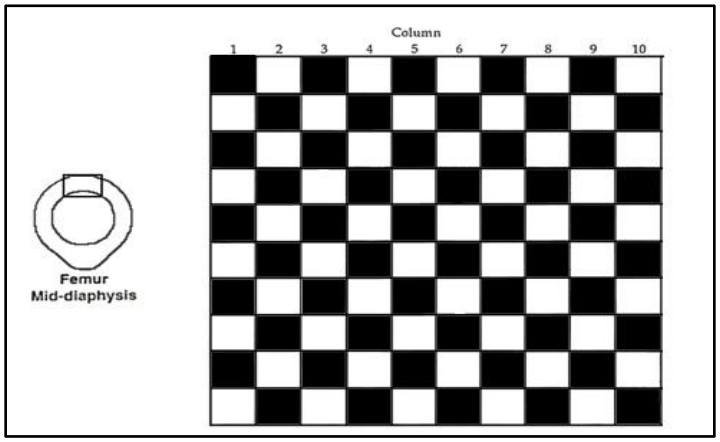


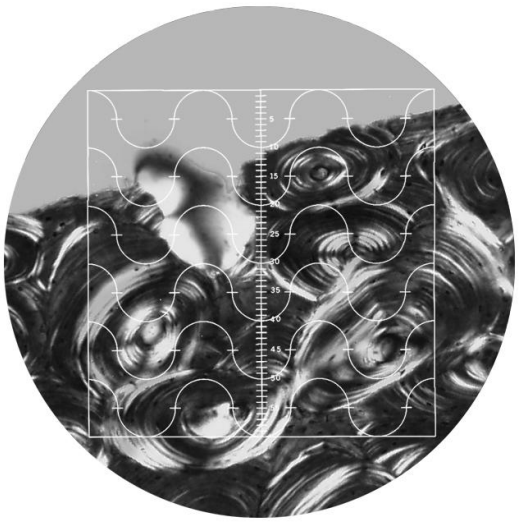
Figure 3. Representation of the sampling method for the research derived from Iwaniec et al. (1998). Each square is equivalent to 0.2304 mm^2 using a 10x ocular and 20x objective.

As mentioned previously, a Merz eyepiece reticule was used for collecting cortical area values and the number of histological structures (intact and fragmentary osteons) within the sampling fields (Figure 4). The Merz reticule contains a square that forms a "region of interest" or ROI in which histological variables will be collected. The square contains six parallel wavy lines with 36 tick marks at specific increments. The tick marks provide "hits" that allow for the counting or measuring of bone area. The average dimension for the 36 "hit" Merz counting reticule used in this study with 10x oculars and a 20x objective is 0.48 mm by 0.48 mm, making the ROI area 0.2304 mm^2. To calculate the area of each "hit" on the reticule you simply divide the reticule area by the total number of "hits" (36) or measure the distance between the individual "hits" and calculate the area of each "hit". At 200x the distance between hits is 0.08 mm. Therefore the area of one hit is 0.0064 mm^2. This point count technique was used to gather the surface area (bone area) data only. The point count technique is based on the stereological principles of morphometry, which allows for the alteration of grid size independent of image magnification. Once the surface area is calculated for a field, intact and fragmentary osteons within the ROI of the Merz reticule are counted before moving to the next field. Microstructures, defined previously, that intersect the grid boundary lines will be included within the grid count if more than 50% of the structure is visible within the ROI.

2.4 Statistical Methods

Observer error for counting and classifying osteons using this sampling method has not been established previously; therefore, observer error analyses were performed to determine the reliability of the method. Previous studies have established error levels for measuring osteon area and cortical thickness and will not be reproduced in this study. Age regression equations

have not been established using a sampling protocol that accounts for over 90% of the microstructure variability from the periosteal to the endosteal surfaces of the femur cortex; therefore, the strength of the relationship between microstructures and chronological age needs to be evaluated in detail. Finally, linear regression models were examined. Statistical analysis and the generation of age prediction models were performed using Excel and SPSS software.



Date: _____ Analyst: _____
Slide ID: _____ Magnification: _____

# Fields	# Hits	# Intact	# Fragment
1	24	2	2
2	13	3	1
3	36	8	5
4	**26**	**4**	**2**
5	9	2	1
6	34	7	3
7	23	5	1
Sum	165	31	15

Figure 4. A superimposed Merz counting reticule containing 36 intersections or "hits". The table represents a partial datasheet for recording the number of hits over cortical bone and the intact and fragmentary osteon counts for each field.

2.4.1 Quantifying observer error

Two observers with varying levels of experience performed the observer error analysis. One observer has 10 years of histological experience, while the other has a few years of experience. Using the definitions outlined in the study, the observers independently read the same slides and separately recorded results. In order to perform a calibration, for the first five slides the observers compared results to rectify any major discrepancies in interpreting the

variable definitions. Such discrepancies would likely result in a clarification of the variable descriptions. This calibration between observers occurred over several weeks, with the remaining observer error analyses performed over several months.

To evaluate inter- and intra-observer error several methods were utilized to provide for a comprehensive analysis. First, the percent mean absolute difference (PMAD) was calculated for OPD, OPD(I), and OPD(F) for both observers (inter-observer) and iterations (intra-observer) using the following equation:

$$PMAD = \sum_{l=1}^{n} \frac{\frac{|x_i - x_j|}{\frac{(x_i + x_j)}{2}} \cdot 100}{n}$$

where *l* indicates the subject, *n* is the total number of samples, *x* is the observation made by a particular observer for OPD, OPD(I), or OPD(F), and *i* and *j* indicate observer 1 and 2 (inter-observer error) or observation 1 and 2 (intra-observer error). It should be noted that guidelines have not been established as to how much observer error is acceptable; therefore, it is usually left to the discretion of the researcher. Commonly, a 10% observer error threshold is adopted in anthropological studies (Nichol and Turner, 1986). Therefore, the 10% threshold will be used to determine acceptable PMAD values in this study.

The second analysis of observer error involved determining the technical error of measurement (TEM) for each variable using:

$$TEM = \sqrt{\frac{\sum_{l=1}^{n} z_l^2}{2n}}$$

where n is the number of samples and z_l is the difference between observers/observations 1 and 2 for the *l*th subject (Mueller and Martorell, 1988). The TEM is a commonly used statistic to assess repeatability and provides an approximation of the standard deviation of the differences

between paired measurements, thus providing a measurement of the imprecision variance. The output from the equation is in the units of the measurement of the variable in question. Both PMAD and TEM describe observer error magnitude, but neither indicates what portion of the variance is error free (Gordon and Bradtmiller, 1992). Additionally, because TEM is in the units of the variable assessed, the TEM of different variables cannot be readily compared. Therefore, the reliability coefficient (R) was calculated to allow for inter-variable comparisons using:

$$R = 1 - \left(\frac{TEM^2}{SD^2}\right)$$

This coefficient indicates the proportion of the measurement variance that is error free. The value ranges from 0 to 1, where *R* values close to 1 indicate a high reliability (Flohr et al., 2010). Because TEM and *R* statistics cannot be used to assess observation bias in the measurements, paired *t*-tests were evaluated and Bland and Altman (1986, 1995) plots were examined.

The Bland and Altman (1986, 1995) method illustrates the repeatability within and between observers. Repeatability coefficients were calculated by taking the sum of the square differences, dividing by *n*, and then taking the square root to obtain the standard deviation. Because the true measurement value is unknown for the samples, a plot of the difference between values against the mean of the values was performed. Repeatability is achieved if the mean difference is not significantly different from zero and if 95% of the differences are less than two standard deviations. If the mean difference is significantly different from zero, the data cannot be used to assess repeatability. Measuring repeatability through evaluating differences against the mean and calculating the repeatability coefficient for observer error quantifies the magnitude of variability that may be masked by correlation, paired *t*-tests and TEM.

2.4.2 Variable analysis and the age prediction model

There were several predictor variables of interest: OPD, OPD(I), OPD(F), On.Ar, and Ant.Ct.Wi. Descriptive statistics were calculated and graphed for each variable to evaluate the normality of the distributions. Previous research performed by the principle investigator suggests that when comparing the linear relationship of OPD, OPD(F), and OPD(I) with age, the predictive power of OPD to estimate age can be largely attributed to changes in OPD(F) with age. The combination of OPD(I) and OPD(F) into OPD does not improve upon the relationship of OPD(F) with age-at-death. Rather, the weaker relationship of OPD with age-at-death suggests that OPD(I) adds noise to this linear relationship. Therefore, as a first step, the OPD and its constituent variables (OPD(I) and OPD(F)) were evaluated in depth to determine if this trend occurs in this sample.

Next, analysis of covariance (ANCOVA) was used to evaluate the age and sex effect for each variable that did not violate assumptions. With age as the covariate, results will indicate if sex differences are significant. A general linear model accompanied with the Chow test was used to determine if the slopes and intercepts are equal between the sexes. This analysis takes the dependent variable Y, the continuous predictor X, and a categorical variable Group to test whether the set of linear regression parameters is equal across the designated groups. The grouping variable was placed in the fixed factor box for the GML analysis menu and the predictor variable x was placed in the covariate box. In order to run the Chow test the design subcommand in the syntax editor window was modified to read: /Design = x Group*x. Including the Group*x interaction causes the GLM to pool the sum of squares and degrees of freedom from the sources Group and Group*x when it reports the F-test for Group*x. The Group term will test differences in intercepts and the Group*x will test differences in slopes.

For comparison with other studies, individuals were grouped by 10 year age cohorts and a two-way analysis of variance (ANOVA), with age, sex, and age-sex interaction terms as prediction variables, was performed for the following variables: OPD, OPD(I), OPD(F), On.Ar, and Ant.Ct.Wi. The analysis was run separately for OPD considering the variable is composed of OPD(I) and OPD(F). Next, the variables were examined to determine if a general model could be developed for predicting age for both sexes, or if sex-specific models are required. Age relationships were further examined using the Pearson correlation matrix.

Finally, linear regression analysis involved a forward stepwise procedure to develop age prediction models from the developmental set. The equations generated from the developmental set were used to estimate age on the validation set. Inaccuracy and bias were calculated for the validation sample. Inaccuracy provides the absolute average error and bias reflects any systematic under or over estimation of age. The *t*-tests were used to determine if the mean difference in age estimates is significantly different from zero. Finally the developmental and validation sets were pooled and used to obtain the final regression model(s).

3 RESULTS

The results are divided into three main sections. Section 3.1 presents the results for the observer error analysis. The level of observer error was established for collecting and classifying intact and fragmentary osteons, as well as the impact to observer error if the variables are combined into one variable (OPD), which has been suggested to decrease classification inconsistencies. Section 3.2 provides the results for the evaluation of the histological variables, specifically examining the relationship of variables with age and sex. Section 3.3 presents the linear regression age prediction models.

3.1 Observer error

It should be noted that observer error values will be slightly inflated considering that the exact location of the microscopic region of interests (ROIs) fluctuate between trials. Although the analyst begins the evaluation of the cortex in the same location, it is unlikely that the same ROIs are reproduced. Furthermore, slight variations to ROI positions will alter repeated counts considering that microstructures along the periphery of the ROI are counted if more than 50% are inside the ROI.

3.1.2 Intra-observer error

The PMAD value for OPD between observations is 4.8%, which is within the 10% acceptance level. The TEM and associate *R* value for OPD is 1.01 #/mm^2 and 0.98, indicating high reliability was achieved with 2% of the variance attributed to measurement error. The Bland and Altman repeatability coefficient is ±2.9 with a mean difference of –0.3 in OPD values (Figure 5). The mean difference in OPD values between the observations is not significantly different from zero ($p = 0.33$). Overall, results indicate that repeatability was achieved between OPD counts.

The PMAD values for the OPD component variables (OPD(I) and OPD(F)) are higher (5.6% and 11.8%, respectively), with osteon fragments exceeding the suggested 10% threshold. The TEM and associate *R* values (0.75 #/mm^2, 0.95 and 0.94 #/mm^2, 0.96; respectively) indicate that 5% of the variance of OPD(I) and 4% of the variance for OPF(F) can be attributed to error. The Bland and Altman repeatability coefficients for OPD(I) and OPD(F) are ±2.0 and ±2.7 with mean differences of -0.3 and 0.08, which are not significantly different from zero ($p = 0.083$ and 0.76, respectively; Figures 6–7). This indicates that repeatability was achieved within OPD(I) counts and OPD(F) counts using this assessment.

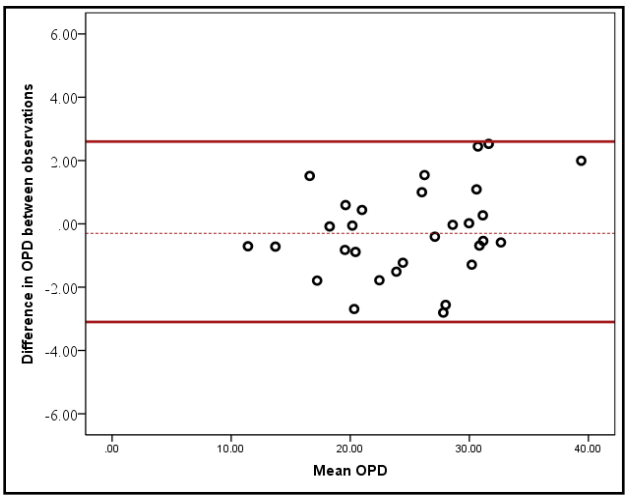


Figure 5. Intra-observer error for OPD values. The graph illustrates the differences between the two trials with the coefficient of repeatability limits represented by the solid lines (±2.9), for which 95% of the variability is expected to fall.

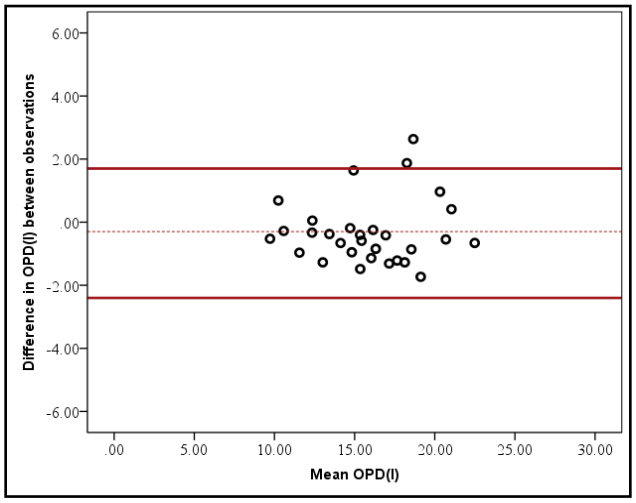


Figure 6. Intra-observer error for OPD(I) values. The graph illustrates the differences between the two trials with the coefficient of repeatability limits represented by the solid lines (±2.0), for which 95% of the variability is expected to fall.

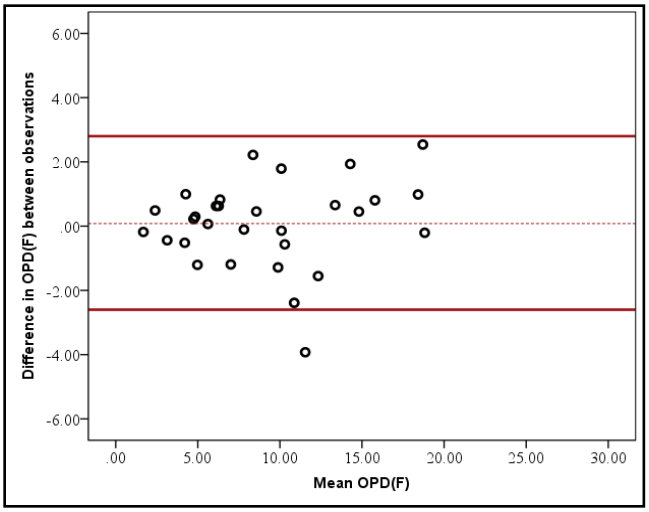


Figure 7. Intra-observer error for OPD(F) values. The graph illustrates the differences between the two trials with the coefficient of repeatability limits represented by the solid lines (±2.7), for which 95% of the variability is expected to fall.

3.1.3 Inter-observer error

The PMAD value for OPD between observers is 6.4%, which is within the 10% acceptance level. The TEM and associate *R* value for OPD is 1.35 #/mm^2 and 0.97, indicating high reliability was achieved with 3% of the variance attributed to measurement error. The Bland and Altman repeatability coefficient is ±3.7 with a mean difference of –0.6 in OPD values (Figure 8). The mean difference in OPD values between the observers is not significantly different from zero (p = 0.098). Overall, results indicate that repeatability was achieved in OPD counts between observers.

The PMAD values for the OPD component variables, OPD(I) and OPD(F), are 5.9% and 14.8%, with osteon fragments again exceeding the suggested 10% threshold. The TEM and associate *R* values (0.74 #/mm^2, 0.97 and 1.12 #/mm^2, 0.95; respectively) indicate that 3% of the variance of OPD(I) and 5% of the variance for OPF(F) can be attributed to measurement error. The Bland and Altman repeatability coefficients for OPD(I) and OPD(F) are ±2.1 and ±3.2 with

mean differences of -0.3 and 0.31, which are not significantly different from zero (p = 0.169 and 0.290, respectively; Figures 9–10). This indicates that repeatability was achieved within OPD(I) counts and OPD(F) counts between observers using this assessment.

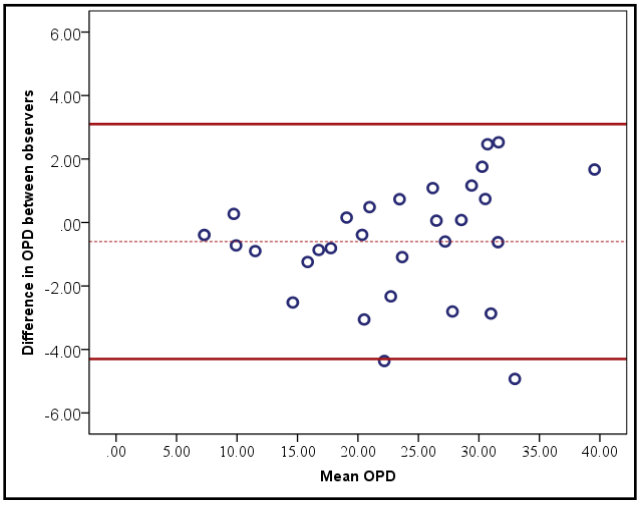


Figure 8. Inter-observer error for OPD values. The graph illustrates the differences between the two observers with the coefficient of repeatability limits represented by the solid lines (±3.7), for which 95% of the variability is expected to fall.

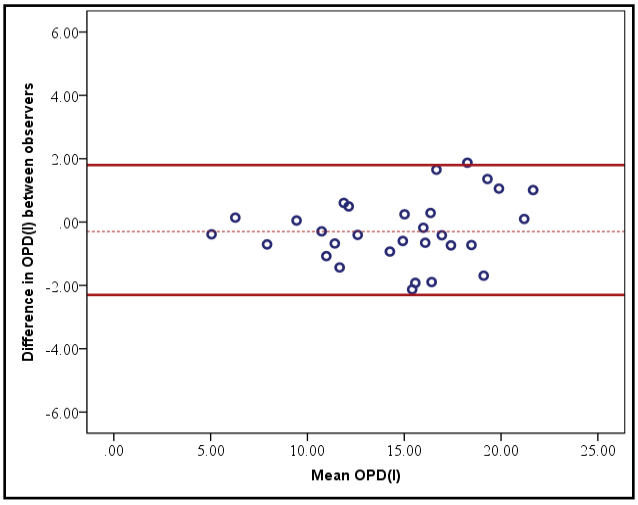


Figure 9. Inter-observer error for OPD(I) values. The graph illustrates the differences between the two observers with the coefficient of repeatability limits represented by the solid lines (±2.1), for which 95% of the variability is expected to fall.

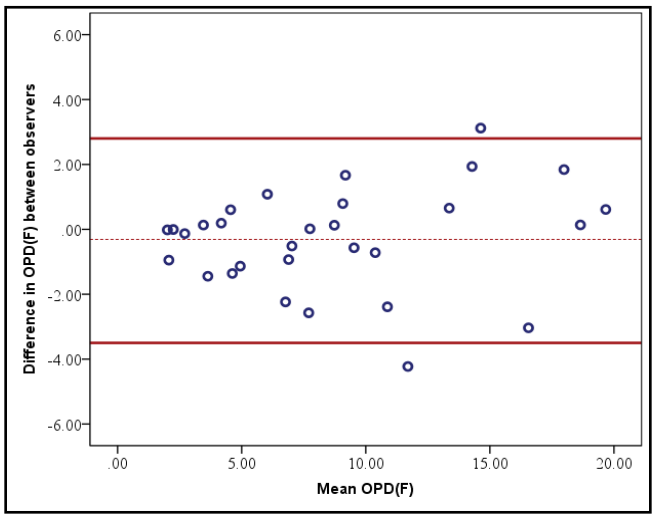


Figure 10. Inter-observer error for OPD(F) values. The graph illustrates the differences between the two observers with the coefficient of repeatability limits represented by the solid lines (±3.2), for which 95% of the variability is expected to fall.

3.2 Variable analysis

Descriptive statistics were analyzed and tests of normality were performed on each histological variable. Variables OPD(F) and On.Ar. did not pass the normality tests and were transformed by the natural log to achieve normalcy. Table 5 presents the descriptive statistics for the histomorphometric variables in the developmental set separated by sex. Variables OPD(F) and On.Ar. are presented in the non-transformed format for this table only to allow for cross-study comparisons.

A one-way ANOVA was run to compare the means of the histological variables between males and females, with the exception of OPD since it is the combination of OPD(I) and lnOPD(F). Results indicate that the mean values for the variables, with exception to OPD(I), are significantly different (Table 6). The Levene statistic indicates that the variables pass the

homogeneity of variances test ($p > 0.05$). Further analysis is warranted to determine the sex and age effects, as well as the sex-age interaction effect.

Table 5. Descriptive statistics of the histological variables for the developmental sample.

	Male Sample (*N*=139)			Female Sample (*N*=120)		
Variable	Mean	STDEV	Std. Error	Mean	STDEV	Std. Error
OPD (#/mm^2)	22.99	5.55	0.471	26.55	6.46	0.590
OPD(I) (#/mm^2)	15.35	3.61	0.306	15.55	3.55	0.324
OPD(F) (#/mm^2)	7.65	3.19	0.270	11.00	5.03	0.459
On.Ar. (mm^2)	0.0443	0.0099	0.0008	0.0399	0.0090	0.0008
Ant.Ct.Wi. (mm)	5.46	1.09	0.093	3.74	1.01	0.093

Table 6. Results for the One-way ANOVA test between sexes.

		Sum of Squares	df	Mean Square	F	Sig.
OPD(I)	Between Groups	1.373	1	1.373	.107	.744
	Within Groups	3300.995	257	12.844		
	Total	3302.368	258			
lnOPD(F)	Between Groups	7.757	1	7.757	32.854	.000
	Within Groups	60.682	257	.236		
	Total	68.439	258			
Ant. Ct. Wi.	Between Groups	194.455	1	194.455	174.315	.000
	Within Groups	286.694	257	1.116		
	Total	481.149	258			
lnOn.Ar.	Between Groups	.709	1	.709	15.318	.000
	Within Groups	11.901	257	.046		
	Total	12.611	258			

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
OPD(I)	.108	1	257	.743
lnOPD(F)	.510	1	257	.476
Ant. Ct. Wi.	.121	1	257	.728
lnOn.Ar.	.015	1	257	.902

An analysis of covariance (ANCOVA) was used to determine if age is a significant covariate of the histological variables and if the group (sex) effect is significant. Before

performing this test the interaction between age and sex in the prediction of the histological variable should be examined to determine the homogeneity of regression slope assumption. A significant interaction between the covariate and the factor suggests that the differences on the dependent variable among groups vary as a function of the covariate. Results indicate that the interaction is significant for Ant.Ct.Wi., which suggest that results from an ANCOVA for this variable would not be meaningful. The interaction for variables OPD(I), lnOPD(F), and lnOn.Ar are not significant and ANCOVA results indicate that age is a significant covariate (Tables 7–9). The sex effect is not significant for OPD(I) ($p = 0.330$), but is significant for lnOPD(F) and lnOn.Ar. ($p = 0.000$ and $p = 0.009$, respectively).

Table 7. ANCOVA results with OPD(I) as the dependent variable

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	438.494[a]	2	219.247	19.598	.000
Intercept	1572.209	1	1572.209	140.539	.000
Sex	10.654	1	10.654	.952	.330
Age	437.121	1	437.121	39.074	.000
Error	2863.874	256	11.187		
Total	64847.908	259			
Corrected Total	3302.368	258			

a. R Squared = .133 (Adjusted R Squared = .126)

Table 8. ANCOVA results with lnOPD(F) as the dependent variable

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	38.848[a]	2	19.424	168.044	.000
Intercept	7.845	1	7.845	67.873	.000
Sex	2.392	1	2.392	20.693	.000
Age	31.091	1	31.091	268.977	.000
Error	29.591	256	.116		
Total	1210.407	259			
Corrected Total	68.439	258			

a. R Squared = .568 (Adjusted R Squared = .564)

Table 9. ANCOVA results with lnOn.Ar. as the dependent variable

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.936[a]	2	1.468	38.844	.000
Intercept	120.191	1	120.191	3180.365	.000
Sex	.259	1	.259	6.844	.009
Age	2.227	1	2.227	58.918	.000
Error	9.675	256	.038		
Total	2646.998	259			
Corrected Total	12.611	258			

a. R Squared = .233 (Adjusted R Squared = .227)

A final analysis was conducted on the variables to further examine sex and age effects with the histological variables. The individuals were grouped into 10-year age cohorts, creating nine age categories (10-19, 20-29, 30-39...), and a two-way ANOVA was performed on each histological variable set as the dependant variable and sex and age cohort set as the independent variables. The results explore if significant mean differences exist between the groups for the two independent variables and for their interaction, Sex*Age Cohort. A sex-age cohort interaction is considered significant at the $p<0.05$ level and age cohort and sex main effects are considered significant at the $p<0.01$ level.

As demonstrated previously, the sex effect and sex-age cohort interaction effect are not significant for OPD(I) (p=0.464, p=0.019, respectively), but the age cohort effect is significant (p=0.000) (Table 10). Mean OPD(I) differs between age groups over time, but with similar values between males and females. The sex and age effect are significant for lnOPD(F) (p=0.007 and p=0.000, respectively), but the sex-age cohort interaction effect is not significant (p=0.354; Table 11). Fragmentary osteon density increases with age and mean values differ between age cohorts. Sex, age cohort, and the sex-age cohort interaction effect are significant for Ant.Ct.Wi (p=0.000, p=0.000, and p=0.001, respectively; Table 12). Males typically have

larger cortical thickness values and exhibit less endosteal expansion over time compared to females. A profile plot demonstrates that females exhibit significant cortical bone loss over time (Figure 11). The sex-age cohort interaction effect and the sex effect are not significant for lnOn.Ar. (p=0.651 and p=0.360, respectively), but the age cohort effect is significant (p=0.000; Table 13). Osteon area decreases in size with age and mean values demonstrate stronger differences when comparing the youngest age cohorts (10–19) to the oldest age cohort (90–99).

A final Two-Way ANOVA was performed using the OPD variable to test if sex and age effects, which were significant for lnOPD(F), would be masked by the OPD variable. Results demonstrate that only the age effect is significant (p=0.000; Table 14), which is predicted considering that OPD increases over time. This indicates that difference between males and females will not be recognized unless the constituent variables are examined. Overall, the results of the two-way ANOVAs suggest that age, sex, and sex-age interaction effects vary between the variables and that the development of male and female regression models will likely improve age estimates. Descriptive statistics for histological variables separated into age cohorts are provided in Table 15. The log transformed variables, OPD(F) and On.Ar., are presented in their non-transformed format for comparison with other studies.

Table 10. Two-Way ANOVA results for OPD(I) as the dependent variable.

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	857.412[a]	17	50.436	4.971	.000
Intercept	17087.828	1	17087.828	1684.352	.000
Age cohort	489.301	8	61.163	6.029	.000
Sex	5.463	1	5.463	.538	.464
Age cohort * Sex	208.384	8	26.048	2.568	.019
Error	2444.956	241	10.145		
Total	64847.908	259			
Corrected Total	3302.368	258			

a. R Squared = .260 (Adjusted R Squared = .207)

Table 11. Two-Way ANOVA results for lnOPD(F) as the dependent variable.

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	39.374[a]	17	2.316	19.205	.000
Intercept	293.305	1	293.305	2432.046	.000
Age cohort	25.464	8	3.183	26.393	.000
Sex	.892	1	.892	7.393	.007
Age cohort * Sex	1.075	8	.134	1.115	.354
Error	29.065	241	.121		
Total	1210.407	259			
Corrected Total	68.439	258			

a. R Squared = .575 (Adjusted R Squared = .545)

Table 12. Two-Way ANOVA results for Ant.Ct.Wi. as the dependent variable.

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	261.814[a]	17	15.401	16.922	.000
Intercept	2042.652	1	2042.652	2244.410	.000
Age cohort	30.735	8	3.842	4.221	.000
Sex	37.335	1	37.335	41.023	.000
Age cohort * Sex	24.966	8	3.121	3.429	.001
Error	219.336	241	.910		
Total	6105.289	259			
Corrected Total	481.149	258			

a. R Squared = .544 (Adjusted R Squared = .512)

Table 13. Two-Way ANOVA results for lnOn.Ar. as the dependent variable.

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3.214[a]	17	.189	4.850	.000
Intercept	844.348	1	844.348	21656.436	.000
Age cohort	1.925	8	.241	6.171	.000
Sex	.033	1	.033	.843	.360
Age cohort * Sex	.233	8	.029	.746	.651
Error	9.396	241	.039		
Total	2646.998	259			
Corrected Total	12.611	258			

a. R Squared = .255 (Adjusted R Squared = .202)

Table 14. Two-Way ANOVA results for OPD as the dependent variable.

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4908.227[a]	17	288.719	13.560	.000
Intercept	40829.926	1	40829.926	1917.629	.000
Age cohort	3091.403	8	386.425	18.149	.000
Sex	56.662	1	56.662	2.661	.104
Age cohort * Sex	224.181	8	28.023	1.316	.236
Error	5131.344	241	21.292		
Total	167296.211	259			
Corrected Total	10039.572	258			

a. R Squared = .489 (Adjusted R Squared = .453)

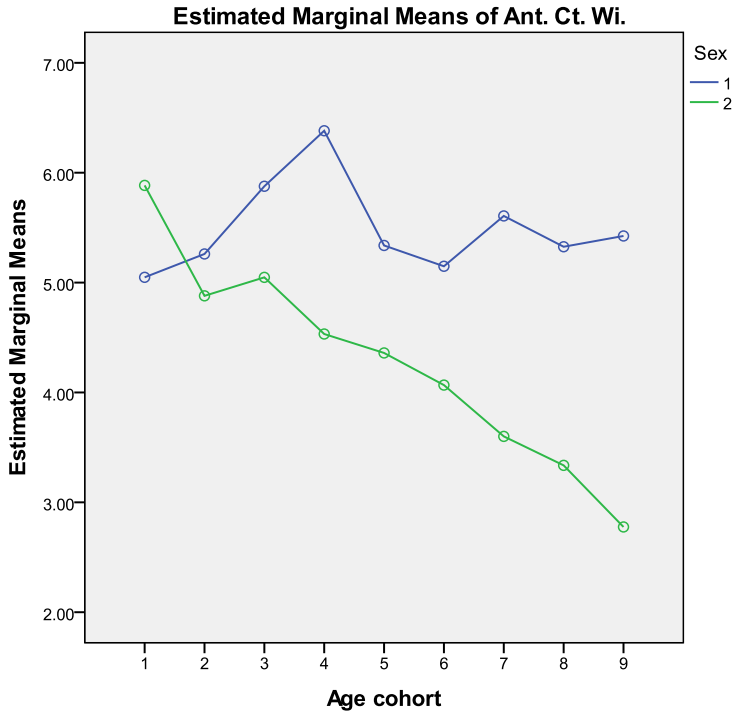


Figure 11. Profile plot for Ant.Ct.Wi. demonstrating cortical bone loss over time (Males = 1, Females = 2). The horizontal axis represents each 10-year age cohort.

Table 15. Descriptive statistics for histological variables separated into age cohorts.

Age Cohort		Males					Females				
		OPD	OPD(I)	OPD(F)	Ant. Ct. Wi.	On.Ar	OPD	OPD(I)	OPD(F)	Ant. Ct. Wi.	On.Ar/
10-19	Mean	11.28	8.12	3.16	5.05	.0561	14.41	10.81	3.60	5.88	.0502
	N	4	4	4	4	4	1	1	1	1	1
	STDEV	3.15	2.77	.82	1.52	.0173	.	.	.	.	.
20-29	Mean	11.13	9.07	2.06	5.31	.0628	12.01	8.40	3.61	4.88	.0563
	N	3	3	3	3	3	5	5	5	5	5
	STDEV	4.58	2.92	1.76	.41	.0172	3.21	1.50	2.89	.73	.0081
30-39	Mean	17.66	13.52	3.69	5.88	.0483	14.54	10.93	3.61	4.63	.0612
	N	7	7	7	7	7	2	2	2	2	2
	STDEV	4.55	3.60	1.08	.52	.0067	9.97	8.04	1.93	1.46	.0267
40-49	Mean	19.35	14.21	5.15	6.64	.0518	22.21	13.76	8.45	4.40	.0438
	N	10	10	10	10	10	4	4	4	4	4
	STDEV	3.55	2.60	2.23	1.24	.0154	8.32	2.30	6.68	.86	.0118
50-59	Mean	21.07	14.86	6.20	5.42	.0463	23.75	17.09	6.65	4.54	.0418
	N	27	27	27	27	27	13	13	13	13	13
	STDEV	4.65	3.96	1.83	1..4	.0113	4.34	2.62	2.25	.81	.0055
60-69	Mean	22.98	14.76	8.21	5.11	.0443	26.67	16.63	10.04	4.02	.0403
	N	34	34	34	34	34	27	27	27	27	27
	STDEV	4.63	3.15	2.91	1.09	.0089	4.58	3.79	3.30	.84	.0063
70-79	Mean	26.41	17.26	9.15	5.61	.0405	27.14	15.53	11.61	3.56	.0387
	N	36	36	36	36	36	30	30	30	30	30
	STDEV	3.77	2.75	2.63	1.04	.0066	4.83	2.95	3.54	.90	.0081
80-89	Mean	27.13	16.80	10.33	5.21	.0395	29.79	15.67	14.11	3.12	.0362
	N	16	16	16	16	16	30	30	30	30	30
	STDEV	3.73	3.06	2.98	.88	.0054	5.21	2.79	4.78	.72	.0068
90-99	Mean	26.00	16.71	9.28	5.42	.0391	32.11	15.29	16.81	2.74	.0349
	N	2	2	2	2	2	8	8	8	8	8
	STDEV	10.09	14.15	5.94	1.56	.0037	4.48	3.42	4.78	.72	.0082
Total	Mean	22.87	15.28	7.60	5.47	.0446	26.55	15.49	11.06	3.72	.0400
	N	139	139	139	139	139	120	120	120	120	120
	STDEV	5.73	3.70	3.24	1.09	.0105	6.46	3.55	5.03	1.01	.0090

Pearson's correlation was used to evaluate the relationship of the predictor variables with age at death for male and female subgroups (Table 16). All of the variables are significantly correlated with age for the females and only Ant.Ct.Wi. is not significantly correlated with age in males. Intact and fragmentary osteon population densities have positive correlations and On.Ar. and Ant.Ct.Wi. have negative correlations. The correlation patterns differ between males and

females. Males show stronger correlation in OPD(I) with age and females show stronger correlations in all other variables.

Table 16. Correlation with age for the histological variables.

	Males			Females		
Variable	*N*	r value	Sig. (2-tailed)	*N*	r value	Sig. (2-tailed)
OPD(I)	139	0.442	0.000	120	0.274	0.002
lnOPD(F)	139	0.675	0.000	120	0.759	0.000
lnOn.Ar.	139	-0.376	0.000	120	-0.495	0.000
Ant.Ct.Wi.	139	-0.104	**0.221**	120	-0.603	0.000

Exploratory regression analysis was performed on each variable for males and females with age set as the dependant variable to examine to what extent the regression parameters (intercepts and slopes) differ between groups. Dummy variables were introduced to indicate the two groups (males = 1, females = 2). Using the General Linear Model function in SPSS, the analysis was built by adjusting the design subcommand for each regression model to include the Sex*variable interaction (/DESIGN = x group*x: see methods section for details). Table 17 provides the results for the significance of the group differences for the single variable regression models. Results demonstrate that the regression slopes for OPD(I), Ant.Ct.Wi., and lnOn.Ar. are significantly different between males and females in relation to age. The regression slope for the lnOPD(F) variable is not significantly different between males and females.

Table 17. Chow test results with age as the dependent variable

Source	Sig.
Sex * OPDI	**.002**
Sex * LnOPDF	.417
Sex * lnOn.Ar	**.050**
Sex * Ant.Ct.Wi.	**.046**

3.3 Age prediction models

Stepwise regression analysis of the developmental sample was performed for the male, female, and pooled sample using four predictor variables: OPD(I), lnOPD(F), lnOn.Ar., and Ant.Ct.Wi.

Stepwise regression analysis for the male sample selected two predictor variables: lnOPD(F) and OPD(I) (Table 18). Two outliers were identified and one was removed after the evaluation of the histomorphology. The individual exhibited large amounts of drifting osteons. Although the individual is young (24 years), the drifting osteons do not appear to be the result of bone modeling nor is femur's cross-sectional geometry abnormal in appearance, suggesting the drift represents a bone turnover issue. The second individual is 92 years old and demonstrates significant age related bone loss, but the cortex does not appear abnormal. The analysis was performed again and the second regression model, which includes both variables, was selected and provides a standard error of the estimate of 11.24. This indicates that approximately 95% of the ages fall within two standard deviations (±22.48 years) of the predicted mean. The confidence intervals will be calculated for the final models and will not be estimated by doubling the SEE.

Stepwise regression analysis for the female sample selected three predictor variables: lnOPD(F), Ant.Ct.Wi., and OPD(I) (Table 19). The third model, which includes all variables, was selected and provides a standard error of the estimate of 9.91. This indicates that approximately 95% of the ages fall within two standard deviations (±19.82 years) of the predicted mean.

Stepwise regression analysis for the pooled sample selected three predictor variables: lnOPD(F), OPD(I), and lnOn.Ar. (Table 20). The third model, which includes all variables, was

selected and provides a standard error of the estimate of 10.99. This indicates that approximately 95% of the ages fall within two standard deviations (±21.98 years) of the predicted mean.

Table 18. Results of the stepwise regression analysis and the age prediction model for the male developmental sample.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
2	.721[a]	.519	.512	11.242	1.113

a. Predictors: (Constant), lnOPD(F), OPD(I)

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
2	Regression	18444.646	2	9222.323	72.977	.000[a]
	Residual	17060.289	135	126.373		
	Total	35504.935	137			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
2	(Constant)	6.559	4.986		1.316	.191
	lnOPD(F)	21.250	2.264	.605	9.385	.000
	OPD(I)	.995	.287	.224	3.469	.001

a. Dependent Variable: Age

Table 19. Results of the stepwise regression analysis and the age prediction model for the female developmental sample.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
3	.812[a]	.659	.650	9.913	1.275

a. Predictors: (Constant), lnOPD(F), Ant. Ct. Wi., OPD(I)

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
3	Regression	22056.991	3	7352.330	74.812	.000[a]
	Residual	11400.134	116	98.277		
	Total	33457.125	119			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
3	(Constant)	36.920	7.809		4.728	.000
	lnOPD(F)	17.827	2.165	.550	8.235	.000
	Ant. Ct. Wi. Mm	-5.371	1.077	-.325	-4.986	.000
	iOPD	.789	.269	.167	2.932	.004

a. Dependent Variable: Age

Table 20. Results of the stepwise regression analysis and the age prediction model for the pooled developmental sample (males and females).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
3	.757[a]	.574	.568	10.993	1.187

a.Predictors: (Constant), lnOPD(F), OPD(I), lnOn.Ar.

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
3	Regression	41279.215	3	13759.738	113.852	.000[a]
	Residual	30697.482	254	120.856		
	Total	71976.698	257			

Coefficients[a,]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
3	(Constant)	-12.371	10.536		-1.174	.241
	lnOPD(F)	21.243	1.492	.655	14.237	.000
	OPD(I)	.523	.232	.112	2.252	.025
	lnOn.Ar.	-8.122	4.025	-.108	-2.018	.045

a. Dependent Variable: Age

3.4 Validation Set Results

When the three prediction models were applied to the validation set and the estimated ages were compared to known age at death, the mean differences of the ages do not significantly differ from zero (see Tables 21 and 22). The male prediction model produced a standard deviation of 10.79, which is slightly lower than the standard error of the estimate produced from the developmental set (11.24). The absolute mean amount that the age estimates vary is 8.6 years. Approximately 37% of the validation age estimates for the males fall within ± 5 years of the known age and approximately 60% fall within ± 10 years of the known age. The female prediction model produced a standard deviation of 7.69, which is slightly lower than the standard error of the estimate produced from the developmental set (9.91). The absolute mean amount that the age estimates vary is 7.67 years. Approximately 47% of the validation set age estimates for the females fall within ± 5 years of the known age and approximately 67% fall within ± 10 years of the known age. The general (pooled) prediction model produced a standard deviation of 10.79. The absolute mean amount that the age estimates vary is 7.52 years. Approximately 66% fall within ± 10 years of the known age.

The developmental and validation sample were pooled into one reference sample to produce the final age prediction models for males, females, and unknown sex samples (Tables 23–25). The age prediction equations are provided in Table 26. The standard error of the estimates for males, females, and pooled sex equations are 11.13, 9.77, and 10.70, respectively.

Table 21. Comparison of estimated ages and known ages from the validation set using the sex-specific equations.

	Males (*N*=30)		**Females (*N*=30)**	
	Difference	Absolute Difference	Difference	Absolute Difference
Mean	-1.40	8.60	-0.18	7.67
STDEV	10.79	6.06	7.69	8.09
Standard Error of Mean	1.97	0.87	1.86	1.20
P > |T|	0.484	0.000	0.922	0.000

Table 22. Comparison of estimated ages and known ages from the validation set using the unknown sex equation.

Pooled Sex (*N*=60)	Difference	Absolute Difference
Mean	0.15	7.52
STDEV	9.45	5.64
Standard Error of Mean	1.22	0.73
P > |T|	0.903	0.000

Table 23. Age prediction model and ANOVA results for males.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
lnOPD(F), OPD(I)	.704	.495	.489	11.133	1.147

ANOVA[c,d]

Model		Sum of Squares	df	Mean Square	F	Sig.
lnOPD(F), OPD(I)	Regression	20079.894	2	10039.947	81.005	.000
	Residual	20450.386	165	123.942		
	Total	40530.280	167			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	6.638	4.717		1.407	.161
	lnOPD(F)	20.355	2.045	.581	9.952	.000
	OPD(I)	1.121	.259	.253	4.325	.000

a. Dependent Variable: Age

Table 24. Age prediction model and ANOVA results for females.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
lnOPD(F), Ant.Ct.Wi., OPD(I)	.801	.642	.635	9.766	1.262

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
lnOPD(F), Ant.Ct.Wi., OPD(I)	Regression	25011.231	3	8337.077	87.417	.000[c]
	Residual	13924.163	146	95.371		
	Total	38935.393	149			

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	25.372	6.675		3.801	.000
lnOPD(F)	20.192	1.888	.623	10.697	.000
Ant. Ct. Wi.	-3.441	.853	-.228	-4.033	.000
OPD(I)	.714	.238	.155	3.003	.003

a. Dependent Variable: Age

Table 25. Age prediction model and ANOVA results for unknown sex.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
lnOPD(F), OPD(I), lnOn.Ar.	.750	.563	.559	10.697	1.234

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
lnOPD(F), OPD(I), lnOn.Ar.	Regression	46321.307	3	15440.436	134.949	.000
	Residual	35926.769	314	114.416		
	Total	82248.075	317			

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-11.783	9.262		-1.272	.204
lnOPD(F)	20.657	1.354	.638	15.258	.000
iOPD	.617	.209	.134	2.959	.003
lnOn.Ar.	-7.860	3.568	-.107	-2.203	.028

a. Dependent Variable: Age

Table 26. Age prediction equations for the three models.

Sex	Prediction Equation
Males	Age = 6.638 + 20.355*(lnOPDF)+1.121*(OPDI)
Females	Age = 25.372 + 20.192*(lnOPDF)-3.441*(Ant.Ct.Wi.)+0.714*(OPDI)
Unknown	Age = -11.783 + 20.657*(lnOPDF)+0.617*(OPDI)-7.860*(lnOn.Ar.)

4 CONCLUSIONS

4.1 Discussion

4.1.1 Observer error

Analysis of the OPD variable indicates that the combination of intact and fragmentary osteon densities does reduce intra- and inter-observer error, compensating for some classification inconsistencies. However, the compensation for classification inconsistencies is misleading and the efficacy of combining the constituent variables in analyses should be carefully considered. Observer error associated with OPD is a combination of the observer agreement associated with the constituent variables OPD(I) and OPD(F). The lack of correlation for observer error values between the constituent variables and OPD indicates that a portion of the observer error for OPD is not explained by either of its constituent variables. The statistically significant relationship between observer error values for OPD(I) and OPD(F) further demonstrates a misclassification of osteon types. The intra-observer results indicate that repeatability was achieved for OPD and OPD(I). Fragmentary osteon density (OPD(F)) passed one of the two tests for observer agreement. Considering that the exact fields were not measured, more weight is given to the

Bland and Altman test of repeatability demonstrating that OPD(F) achieved repeatability. Comparing these results to previous analyses of variable error using Stout and Paine's (1992) original variable definitions demonstrates a significant decrease in OPD(I) and OPD(F) error values. Crowder (2005) reported a PMAD value of 11.2% for OPD(I) and 22.8% for OPD(F). The new definitions used in this study cut the counting error approximately in half for each constituent OPD variable.

The inter-observer results demonstrated slightly higher error values, which are expected considering slight interpretation differences that likely exist between the observers due to experience level. One observer has 10 years of histological experience, while the other has a few years of experience. Similar to the intra-observer results, repeatability was achieved for OPD and OPD(I) using all repeatability methods. Fragmentary osteon density (OPD(F)) passed one of the two tests for observer agreement. It should be noted again that observer error values are likely slightly inflated owing to the fact that field locations varied during analyses. While the observers used the same starting place on each slide, each evaluated field will differ slightly and structures may be determined to be more than 50% outside of the ROI. Therefore, these structures would fall in the adjacent field of view and may not be counted using the checkerboard technique (see Methods section). Considering that the exact same fields were not measured, more weight is given to the Bland and Altman test of repeatability demonstrating that OPD(F) achieved repeatability. Comparing the PMAD values for the original variable definitions reported by Crowder (2005), the new definitions used in this study cut the counting error from 20.6% to 5.9% for OPD(I) and from 20.6% to 14.8% for OPD(F).

While it is clear that the revised variable definitions significantly reduce observer error in histological analyses, differences in OPD(F) compared to OPD(I) indicate continued problems in

differentiating fragmentary osteons. This means that portions of fragmentary osteons are not being recorded and, therefore, observers may not be capturing the full age-related significance of this feature. Analysis of the classification differences between observers and observations identified a systematic bias. The relation of bias and the known age at death were examined as the potential leading factor resulting in the magnitude bias, compounding the effects of variable related error. Results indicate that observer error is affected by age at death, producing an increase in error as age increases. This is not surprising considering that older individuals demonstrate more age-related bone turnover, which results in osteon crowding with higher levels of fragmentary osteons. Identifying intact and fragmentary osteons consistently factors into the success of future histological methods. Thus, difficulties in quantifying bone turnover are amplified by current methods with subjective and less descriptive variable definitions. Therefore, the authors propose that the definitions and selection criteria used in this research will improve future histological methods of age estimation.

4.1.2 Variable analysis

Sex, age, and sex-age interaction were examined for the four histological variables: OPD(I), lnOPD(F), lnOn.Ar, and Ant.Ct.Wi. The histological variables all demonstrate relationships with age, similar to previous studies. However, age and sex effects are not consistent between the variables. In relation to age, lnOPD(F) demonstrates the strongest correlation ($r = 0.68$ in males and $r = 0.76$ in females). Surprisingly, the OPD(I) correlation was low for both groups ($r = 0.44$ in males and $r = 0.274$ in females). Also interesting is the difference in correlation strength between males and females, with females showing weaker

correlation to age for OPD(I). These findings suggest that age-related turnover events are best expressed in osteon fragments, especially in females. Intact osteons, while also related to age, likely represent addition biological relationships (i.e. age, biomechanics, and bone maintenance). Thus, the biological significance of these variables, in relation to age at death, does not appear to be equal.

Results indicate that increasing osteon population density is coupled with a reduction in cortical bone thickness with chronological age in adults, with females demonstrating greater bone loss. This has been identified in other studies evaluating age related bone loss (Carlson et al., 1976; Garn et al., 1992; Bertelsen et al., 1995; Cho and Stout, 2003; to name a few). Both males and females show a negative correlation in anterior cortical thickness with age; however, the relationship was only significant in females ($r = -0.603$, $p = 0.000$). Future analysis of the sample will include a relative measure of cortical thickness to control for size (males tend to have larger cross-sectional areas compared to females), which is likely amplifying the differences.

Osteon area demonstrates a negative relationship with age, with large variances within and between age groups. This age-dependant decrease in osteon area has been reported in many previous studies (Sing and Gunberg, 1970; Ortner, 1975; Stout and Simmons, 1979; Thompson, 1980; Pfeiffer, 1998; Streeter and Stout, 2003; Goliath, 2010). Osteon area is a factor of bone formation rates, which typically decline over age. However, studies suggest that the decrease in osteon area with age may be related to numerous factors, such as biomechanics, cortical area, osteon crowding, or body size. While it appears from the literature that osteon area would be a strong age predictor, the R^2 values for the individual variable regression analysis are 0.159 for males and 0.246 for females. Thus, only a small portion of the variation related to age is

accounted for by osteon area. It also appears that osteon area exhibits different effects between males and females. Examining correlation between osteon area and OPD(I), it is interesting to note that the correlation values are -0.694 for males and -0.461 for females. Further research exploring the relationship between OPD(I) and osteon size is warranted. One suggestion is to evaluate the packing factor or arrangement of osteons, which may be a controlling factor for size and population density (Figure 12). It may be possible to develop a scaling factor that could be used to increase the significance of the relationship of OPD(I) with age. Considering that osteon area demonstrates a relationship negative relationship with OPD(I) and the weak relationship of OPD(I) with age, it may support the hypothesis that smaller osteons allow for a greater number of osteons per unit area.

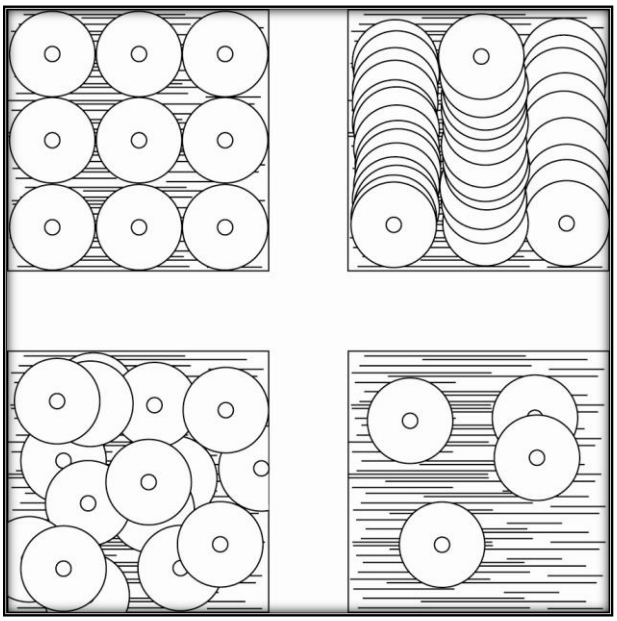


Figure 12. These diagrams, reproduced from Frost 1987a, demonstrate the differences in osteon distribution within a microscopic field of view, which affects the number of osteon contained within the area. Thus the packing factor would have an effect on age estimation models based on osteon population density.

It has been proposed that sex differences in osteon size exist (Burr et al., 1990), but other studies show no indication of sex as a factor of osteon size (Dupras and Pfeiffer, 1996; Pfeiffer, 1998; Cho et al., 2002). This research indicates that osteon area is significantly different between males and females. The two-way ANOVA indicates that the relationship with age and sex is not straight forward. Females tend to have a larger mean On.Ar. in the 20–29 and 30–39 age cohorts when compared to males. While this may be a factor of small sample sizes, it may also be related to age of parity. For the other age-cohorts, a qualitative examination of the means indicates that females demonstrate smaller osteon area compared to males.

The results indicate that age and sex differences do exist in the histological variables; however, the effects are somewhat complex and are likely complicated by unequal or small sample sizes within the age cohorts. Regardless, it is reasonable to conclude that separate regression models for males and females are warranted. While it has been documented that bone density and the rate of bone remodeling differ between males and females, there is no agreement among studies that these differences exist with measurable consistency. This is apparent within the literature in that many histological age estimation methods provide sex-specific equations, while others indicate no significant differences between groups. More specifically, research models based on the 6^{th} rib typically do not produce sex-specific equations (Stout, 1986, Stout and Paine 1992, Cho et al 2002); however, age estimation models based on the femur do provide sex-specific models (Ericksen, 1991; Thompson, 1979). This may be the result of sampling error, skeletal element evaluated, or the selection of histological variables. As stated previously, sex differences observed in histological variables are likely related to biological factors involving the endocrine system that affect bone turnover in the female skeleton such as pregnancy, lactation, and menopause. Females experience bone loss associated with a drop in estrogen

levels following menopause, which manifests in the loss of trabecular connectivity and increased porosity within the Haversian envelope of cortical bone. Biocultural factors (such as fecundity, breastfeeding practices, types of food consumed, or activity levels) may produce or extenuate histological differences between males and females that significantly impact bone biology.

4.1.3 Regression model

Sex specific age prediction equations were generated, as well as a general model to be used when sex is unknown or if the analyst prefers to use the pooled sample model. In both sex specific models, lnOPD(F) was determined to be the best predictor of age at death. The significance of this observation was described in the previous section. The standard errors of the estimates for the equations are large, but consistent with previous histological studies with large sample sizes (Table 27).

Table 27. Reported Standard Errors of the Estimates for Various Histological Methods.

Methods	Elements	Sample Parameters: N, Age Range, Sex [M,F]	Mean Age in Years	SEE
Kerley (1965)	Femur	67, 0-95, [41, 6]	41.6	9.39-13.85*
Ahlqvist & Damsten (1969)	Femur	20, 4-89, [?]	55.4	6.71
Singh & Gunberg (1970)	Femur	33, 39-87, [33, 0]	62.3	3.24-3.82*
Thompson (1979)	Femur	116, 30-97, [64, 52]	69.5	6.41-9.69*
Samson & Branigan (1987)	Femur	58, 16-91, [31, 27]	NA	6-16
Ericksen (1991)	Femur	328, 14-97, [174, 154]	62.8	9.96-12.21*
Stout & Paine (1992)	6th Rib	40, 13-62, [32, 7]	28.6	NA
Stout et al. (1994)	4th sternal Rib	59, 11-88, [?, ?]	39.2	10.43
Watanabe et al. (1998)	Femur	98, 0-92, [72, 26]	50.4 (M) 48.8 (F)	3.16-11.50
Cho et al. (2002)	6th rib	154, 17-95, [?, ?]	50.4	12.22
This Study	Femur	168, 15-97, [M] 150, 19-96,[F]	66 68	11.13 9.77

*Contains multiple regressions, for more detail see original publications.

The existence of sex and age related differences associated with the histological variables indicate that sex specific age-prediction equations should be applied when estimating age using cortical bone histomorphometry. Though the standard error of the estimates for the generated sex specific models and the pooled sex model are similar, the sex specific models selected different variables for males and females which support existing research indicating that bone histomorphometry differs between males and females. This, coupled with the fact that the validation consists of a small sample, suggests it is appropriate to maintain sex specific equations. It should be noted that the equations generated from this study are not accurate with younger individuals. This is likely due to the lack of younger individuals in the reference sample. Furthermore, the majority of the young individuals were obtained from forensic cases, which may represent individuals in poor health due to substance abuse or nutritional issues. Regardless, the focus of this research was to improve age estimation for older individuals and provide a method that may be less affected by the reported asymptotic value for osteon population density.

As chronological age increases, the cortex becomes crowded with complete and fragmentary secondary osteons (Robling and Stout, 2000). In theory, the length of time during which remodeling occurred (chronological age) will be a major influence on how many secondary osteon creations (complete and fragmentary osteons) accumulate per unit of area. This linear relationship should be evident in a normal individual until remodeling rates begin to fluctuate after the sixth decade of life, as homeostasis is compromised by senility (Wu et al., 1970). It has been suggested that the ribs are an ideal sampling location for histological studies, in part because of the minimal biomechanical variation of the mid-thoracic region compared to the variation seen in the appendicular skeleton. One drawback to using the rib for histological

analysis is possible remodeling rate differences compared to that of the femur. A higher remodeling rate, coupled with the smaller cortical area of the ribs produces an earlier asymptote. The results from this study suggest that this phenomenon does not occur in the femur until around 80 to 90 years of age in males (Figure 13), although there is considerable individual variation. This can be observed in further analysis of the validation sample, in which the oldest individuals demonstrate higher differences between estimated and known age. The age estimates tend to significantly underage males 80 to 90 years of age, which supports the hypothesis that the bone has "remodeled out" and reached the asymptote. The female data from this research does not demonstrate the asymptote when evaluating the sample by age cohort, which supports results indicating that the prediction model performs better for females. Figure 13 provides the plot for total OPD from the two-way ANOVA, demonstrating the difference in males and females. This is interesting considering that females have significantly thinner cortical area. The male samples in these age cohorts are smaller, thus a larger sample may provide different results.

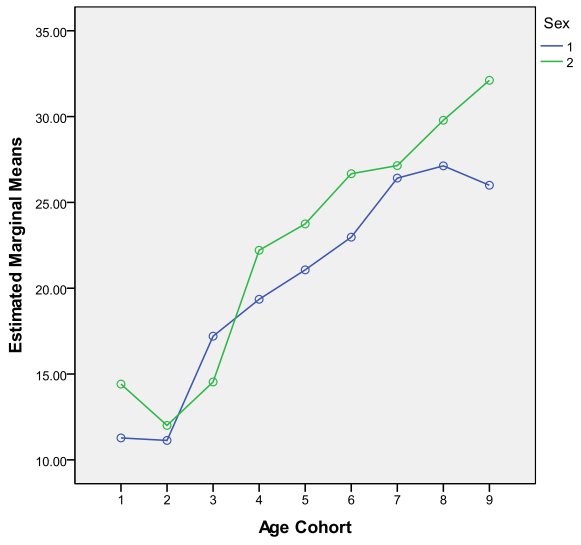


Figure 13. Profile plot for total OPD demonstrating difference in males (1) and females (2) over time. The horizontal axis represents each 10-year age cohort.

4.1.4 Final Comments

Current histological methods demonstrate significant issues that affect their reliability and accuracy. The method developed from this research demonstrates several advantages over previous methods. The method is based on validated variables, accounts for 90–95% of the spatial variation in osteons within the anterior cortex, and is not restricted to a specific field size or magnification. Therefore, this method provides more objective histological variable definitions and reduces the error associated with histological analysis.

One issue which may have adverse effects on accurate evaluation of the cortex is diagenesis. Diagenetic agents can structurally alter bone micro-morphology, affecting the reliability of histological age estimates. Although diagenesis, when present, may not affect all areas within the cortex a bone sample, one must adhere as closely as possible to the sampling protocol outlined by the method being employed. The sampling protocol as defined for this method, while requiring a larger area of bone, accounts for significant variation in the cortex while allowing flexibility in which fields are evaluated. In the presence of extensive diagenesis, however, histological age methods should not be applied.

Another issue in the evaluation of the cortical bone was identifying the boundary between trabecular and cortical bone at the endosteum. This was particularly difficult as trabecularization and endosteal expansion increased with age. The developed method requires the presence of cortical bone remodeling and did not include voids indicative of bone porosity (with the exception of resorptive bays) when calculating cortical bone surface area.

Analysis of the histological variables indicates that they demonstrate complex interactions with age, sex, health, and biomechanics. This should serve as a caution to researchers in producing “simple” models of histological age estimation and explore biological

reasons for histological variation. It may be possible to elucidate more information by evaluating the spatial relationship of osteons within the cortex, for example evaluating differences between OPD in the periosteal and endosteal envelopes.

One of the most prevalent issues regarding adult age estimation is the inability to accurately age older adults. The results of this study indicate that histological analysis of the anterior femur provides reliable age estimates for older individuals. The described regression model is most accurate for individuals over 50 years of age and it is currently not accurate for use with young adults (< 30 years). This is likely due to a combination of factors. First, the age distribution is skewed toward older individuals and second, the strength of intact osteons with age is considerably less than fragmentary osteons. Despite this, the standard error in this study is similar to that of previous histological studies with large sample sizes and to methods that use gross bone age indicators. Bearing in mind that the elderly are a rapidly growing percentage of North American populations and that unidentified adults are a common occurrence in the forensic setting, this research will improve the accuracy of estimating age for older adults.

4.2 Implications for Policy and Practice

The development of new or improved standards for adult age estimation is greatly needed for the anthropological assessment of unidentified remains. The accurate reporting of age for the biological profile is imperative in determining the inclusion or exclusion of individuals from a pool of missing persons. It has been suggested that histological age indicators may provide more accurate age estimates considering that they are a product of continuous bone turnover and not the result of degenerative changes in bone morphology; however, the results of this study (and others) suggest that biological variability is significant in bone turnover between and within

groups. Regardless of this and despite previous methods evaluating the femur demonstrating significant methodological issues, histological methods appear promising to predict age past the 50+ boundary. The method developed from this research significantly reduces error in histological analysis.

This research suggests that the assessment of histological age indicators should be coupled with macroscopic (gross) methods to provide a more comprehensive age estimate. The Forensic Anthropology Unit within the OCME-NYC receives 30–40 skeletal cases per year in which histological age estimation is performed in conjunction with various gross age indicators. The authors suggest that evaluation of gross indicators should include histological analysis to assist with an assessment of skeletal health, which may indicate why, for some cases, gross indicators do not correlate well with chronological age. Unfortunately, the sample used in this study did not allow for observations of gross structures.

One significant limitation to this study was the skewed age distribution due to the lack of young individuals in the reference sample. While aspects of this were dealt with statistically, further resolution will require additional histological samples.

4.3 Implications for further research

This research produces several recommendations for further research. First, the relationship of intact osteons with age and other biological factors needs to be explored. This research demonstrated a strong correlation between Intact Osteon Population Density (OPD(I)) and osteon area (On.Ar.). As OPD(I) increased, On.Ar. decreased. While decreasing On.Ar. over time has been observed in numerous studies, the relationship with OPD(I) has not been thoroughly investigated. The authors recommend the development of a scaling factor for On.Ar.

to investigate if the age relationship between intact osteons and chronological age could be improved. Associated with this issue of OPD(I) and On.Ar. is the presence of drifting osteons (see section 2.3). Currently, drifting osteons are not included in osteon area data collection. This may be an oversight considering that drifting osteons can produce large areas of bone remodeling and decrease the number of osteon counts within a microscopic field. Future research should investigate including drifting in measurements or as a factor of osteon packing (see discussion in section 4.1.2).

Second, evaluation of sex differences in bone turnover should be carefully performed in histological studies. Considering that bone turnover is affected by changes in the endocrine system, sex differences should be prevalent in other skeletal elements. Methods that use the 6th rib midshaft do not demonstrate sex-specific responses in intracortical bone turnover based on osteon counts. This may be due to the variables selected or the use of OPD in the regression equations rather than evaluating intact and fragmentary osteons separately.

Third, it was mentioned that a significant limitation of this research was due to the lack of young individuals in the reference sample. Recently the authors have located a younger sample to add to this database.

Finally, it is apparent that the linear regression model approach should be reconsidered for future methods. A non-linear approach was explored and the results did not indicate that this approach provides a more useful model for histological age estimation. At the practitioner level, a non-linear model is more difficult to apply and the results would not significantly improve the accuracy of the age estimates. Additional exploration may prove useful, but the authors believe that a more appropriate approach will be the use of a Bayesian model.

5 REFERENCES

Ahlqvist J, Damsten O. 1969. A modification of Kerley's method for microscopic determination of age in human bone. J Forensic Sci 14:205-212.

Bertelsen PK, Clement JG, Thomas CDL. 1995. A morphometric study of the cortex of the human femur from early childhood to advanced old age. Forensic Sci Intl 74:63-77.

Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 8476: 307–310.

Bland J, Martin, Altman DG. 1995. Comparing methods of measurement: Why plotting difference against standard method is misleading. Lancet 346(8982):1085-1089.

Brooks ST, Suchey JM. 1990. Skeletal age determination based on the os pubis: a comparison of the Ascadi-Nemeskeri and Suchey-Brooks methods. Hum Evol 5:227–238.

Buckberry JL, Chamberlain AT. 2002. Age estimation from the auricular surface of the ilium: A revised method. Am J Phys Anthropol 119:231–239.

Burr DB. 2002. Targeted and nontargeted remodeling. Bone 30(1):2-4.

Burr DB, Ruff CB, Thompson DD. 1990. Patterns of skeletal histologic change through time: Comparison of an Archaic Native American population with modern populations. Anat Rec 226:307–313

Carlson DS, Armelagos GJ, Van Gerven DP. 1976. Patterns of age-related cortical bone loss (osteoporosis) within the femoral diaphysis. Human Biol 48(2):295-314.

Cattaneo C, DiMartino S, Scali S, Craig OE, Grandi M and Sokol RJ. 1999. Determining the Human Origin of Fragments of Burnt Bone: A Comparative Study of Histological, Immunological and DNA Techniques. Forensic Sci Int. 102(2-3):181-191.

Chan A, Crowder C, Rogers T. 2007. Variation in Cortical Bone Histology within the Human Femur and its Impact on Estimating Age at Death. Am J Phys Anthropol 132(1):80-8.

Cho H, Stout SD. 2003. Bone remodeling and age-associated bone loss in the past: A histomorphometric analysis of the Imperial Roman skeleton population of Isola Sacra. In Agarwal S. and Stout S.D. (editors) Bone Loss and Osteoporosis: An Anthropological Perspective. New York: Kluwer Academic/Plenum pp. 207-229.

Cho H, Stout SD, Madsen RW, Streeter MA. 2002. Population-specific histological age-estimating method: A model for known African-American and European-American skeletal remains. J Forensic Sci 47(1):12-18.

Clarke DF. 1987. Histological and radiographic variation in the parietal bone in a cadaveric population. Thesis, Anatomy Department, The University of Queensland.

Cool SM, Hendrikz JK, Wood WB. 1995. Microscopic age changes in the human occipital bone. J Forensic Sci 40(5):789-796.

Crowder C. 2005. Evaluating the use of quantitative bone histology to estimate adult age at death. Ph.D. dissertation, University of Toronto, Department of Anthropology.

Curtis J. 2004. Estimation of age at death from the microscopic appearance of the frontal bone. Master's Thesis, University of Indianapolis, Indiana.

Dupras TL, Pfeiffer SK. 1996. Determination of sex from adult human ribs. Can Soc Forens Sci 29(4):221-231.

Ericksen MF. 1991. Histological estimation of age at death using the anterior cortex of the femur. Am J Phys Anthropol 84:171-179.

Flohr S, Leckelt J, Kierdorf U, and Kierdorf H. 2010. How Reproducibly Can Human Ear Ossicles Be Measured? A Study of Inter-Observer Error. Anatomical record (Hoboken, NJ : 2007) 293(12):2094-2106.

Frost, H.M., 1958. Preparation of thin undecalcified bone sections by rapid manual method. Stain Technol 33(6):273–7.

Frost HM. 1969. Tetracycline based histological analysis of bone remodeling. Calcif Tissue Res 3:211-237.

Frost HM. 1987a. Secondary osteon populations: an algorithm for determining mean bone tissue age. Yrbk Phys Anthropol 30:221-238.

Frost HM. 1987b. Secondary osteon population densities: an algorithm for estimating missing osteons. Yrbk Phys Anthropol 30:239-254.

Garn SM. 1970. The earlier gain and the later loss of cortical bone. Charles C. Thomas: IL.

Garn SM, Sullivan TV, Decker SA, Larkin FA, Hawthorne VM. 1992. Continuing bone expansion and increasing bone loss over a two-decade period in men and women from a total community sample. Am J Hum Biol 4:57-67.

Gordon C, Bradtmiller B. 1992Interobserver Error in a Large-Scale Anthropometric Survey". Am J of Human Biology V4(2): 253-263.

Goliath, J.R., 2010. Variation in osteon circularity and its impact on estimating age at death. MA Thesis. The Ohio State University.

ImageJ, 2009. U.S. National Institutes of Health, Bethesda, Maryland, USA, http://rsbweb.nih.gov/ij/.

İşcan MY. 1993. *Casts of age phases from the sternal end of the rib for white males and females.* France Casting, Fort Collins, Colorado.

İşcan MY, SR Loth and R K Wright 1984. Age estimation from the rib by phase analysis: white males. J Forensic Sci 29(4):1094-104.

Iwaniec UT, Crenshaw TD, Scheninger MJ, Stout SD, Ericksen MF. 1998. Methods for improving the efficiency of estimating total osteon density in the human anterior mid-diaphyseal femur. Am J Phys Anthropol 107:13–24.

Iwaniec, UT, 1997. Effects of Dietary Acidity on Cortical Bone Remodeling: A Histomorphometric Assessment. Ph.D. Dissertation, University of Wisconsin-Madison.

Jaworski ZFG. 1984. Coupling of bone formation to bone resorption: A broader view. Calcif Tissue Int 36:531–535.

Kerley ER. 1965. The microscopic determination of age in human bone. Am J Phys Anthropol 23:149-164.

Kerley ER, Ubelaker DH. 1978. Revisions in the microscopic method of estimating age at death in human cortical bone. Am J Phys Anthropol 49:545-546.

Kim Y, Kim D, Park D, Lee J, Chung N, Lee W and S Han. 2007. Assessment of histomorphological features of the sternal end of the fourth rib for age estimation in Koreans. J Forensic Sci 52(6):1237-1241.

Lazenby RA. 1984. Inherent deficiencies in cortical bone microstructural age estimation techniques. OSSA 9-11: 95-103.

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. 1985a. Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of age at death. Am J Phys Anthropol 68:15–28.

Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ. 1985b. Multifactorial determination of skeletal age at death: a method and blind tests of its accuracy. Am J Phys Anthropol 68:1–14.

Lynnerup N, Thomsen JL, Frohlich B. 1998. Intra- and Inter-observer variation in histological criteria used in age at death determination based on femoral cortical bone. Forens Sci Intl 91:219-230.

Maggiano, C.M., 2012. Making the Mold: A Microstructural Perspective on Bone Modeling During Growth and Mechanical Adaptation. In: Crowder, C., Stout, S. (eds.) Bone Histology: An anthropological approach. CRC Press, Boca Raton: pp. 45–90.

Maat, G.J.R., Van den Bos, R.P.M., Aarents, M.J., 2000. Manual for the preparation of ground sections for the microscopy of bone tissue. Barge's Anthropologica, Nr. 7., Leiden.

Mueller W, Martorell R. 1988. Reliability and accuracy of measurement. Anthropometric Standardisation Reference Manual:83-86.

Nichol CR, Turner CG. 1986. Intra- and Interobserver concordance in classifying dental morphology. Am J Phys Anthropol 69:299-315.

Ortner DJ. 1975. Aging effects on osteon remodeling. Calcif Tissue Res 18:27–36.

Parfitt AM. 2001. Skeletal heterogeneity and the purpose of bone remodeling: implications for the understanding of osteoporosis. In Marcus R., Feldman D., and Kelsey J., (Eds): Osteoporosis. 2nd Edition. Academic: San Diego, CA; 433-447.

Parfitt AM. 2002. Misconceptions (2): Turnover is always higher in cancellous than in cortical bone. Bone 30(6):807-809.

Pfeiffer S. 1992. Cortical bone age estimates from historically known adults. Z Morph Anthrop 79(1):1-10.

Pfeiffer S. 1998. Variability in osteon size in recent human populations. Am J Phys Anthropol 106:219-227.

Robling AG, Stout SD. 2000. Histomorphometry of human cortical bone: Applications to age estimation. In Katzenberg S, Saunders S. (Eds): Biological anthropology of the human skeleton. Wiley-Liss: NY.

Samson C, Branigan K. 1987. A new method of estimating age at death from fragmentary and weathered bone. In Boddington A, Garland AN, Janaway RC. (Eds): Death, decay, and reconstruction: Approaches to archaeology and forensic science. Manchester University Press.

Singh IJ, Gunberg DL. 1970. Estimation of age at death in human males from quantitative histology of bone fragments. Am J Phys Anthropol 33:373-382.

Stout SD, Teitelbaum SL. 1976. Histological analysis of undecalcified thin sections of archaeological bone. Am J Phys Anthropol 44:263–270.

Stout SD, Gehlert SJ. 1982. Effects of field size when using Kerley's histological method for determination of age at death. Am J Phys Anthropol 58:123-125.

Stout SD, Simmons DJ. 1979. Use of histology in ancient bone research. Yrbk Phys Anthropol 44:263-270.

Streeter M, Stout D. 2003. The histomorphometry of the subadult rib: age associated changes in bone mass and the creation of peak bone mass. In Agarwal S. and Stout S.D. (editors) Bone Loss and Osteoporosis: An Anthropological Perspective. New York: Kluwer Academic/Plenum pp. 91-101

Stout SD. 1986. The use of bone histomorphometry in skeletal identification: the case of Francisco Pizarro. J Forensic Sci 31(1):296–300.

Stout SD, Stanley SC. 1991. Percent osteonal bone verses osteon counts: The variable of choice for estimating age at death. Am J Phys Antrhopol 86:515–519.

Stout SD, Paine RR. 1992. Brief communication: Histological age estimation using rib and clavicle. Am J Phys Anthropol 87:111-115.

Stout SD, Paine RR. 1994. Brief communication: Bone remodeling rates: A test of an algorithm for estimating missing osteons. Am J Phys Anthropol 93: 123–129

Stout SD, Dietz WH, Işcan MY, Loth SR. 1994. Estimation of age at death using cortical histomorphometry of the sternal end of the fourth rib. J Forensic Sci 39(3):778-784.

Stout SD, Marcello AP, Perotti B. 1996. Brief communication: A test and correction of the clavicle method of Stout and Paine for histological age estimation of skeletal remains. Am J Phys Antrhopol 100:139-142.

Thompson DD. 1979. The core technique in the determination of age at death in skeletons. J Forensic Sci 24(4):902-915.

Thompson DD. 1980. Age changes in bone mineralization, cortical thickness, and Haversian canal area. Calcif Tissue Int 31:5-11.

Thompson DD, Galvin CA. 1983. Estimation of age at death by tibial osteon remodeling in an autopsy series. Forensic Sci Int 22:203-211.

Ubelaker D. 1986. Estimating Age at Death from Immature Human Bone. In: Age Markers in the Human Skeleton. Ed. Işcan MY. Springfield, Illinois: Charles C. Thomas.

Villa C., Lynnerup N. 2010. Technical note: A stereological analysis of the cross-sectional variability of the femoral osteon population. Am J Phys Antrhopol 142:491-496.

Watanabe Y, Konishi M, Shimada M, Ohara H, Iwamoto S. 1998. Estimation of age from the femur of Japanese cadavers. Forensic Sci Int 98:55-65.

Weiner S, Traub Wolfie, Wagner HD. 1999. Lamellar bone: Structure-Function relations. J Struct Biol 126:241-255.

Wu K, Jett S, Frost H. 1967. Bone resorption rates in physiological, senile and postmenopausal osteoporoses. J Lab Clin Med 69:810-818.

Wu K, Schubeck H, Frost M, Villanueva A. 1970. Haversian bone formation rates determined by a new method in a mastodon, and in human diabetes mellitus and osteoporosis. Calc Tissue Res 6:204-219.

Yoshino M, Imaizumi K, Miyasaka S, Seta Sueshige. 1994. Histological estimation of age at death using microradiographs of humeral compact bone. Forensic Sci Int 64:191-198.

6 DISSEMINATION OF RESEARCH FINDINGS

Results were presented to the American Academy of Forensic Sciences (AAFS):

Crowder, C.M. and V.M. Dominguez
2012 A New Method for Histological Age Estimation of the Femur. Presented at the American Academy of Forensic Sciences Annual Meetings, Atlanta, GA, February 20-24.

This research received the Kerley Award, which is given by the Kerley Forensic Sciences Foundation for the paper best demonstrating originality, creativity, depth of research, innovation, new methodologies, research design, significance to the field, and/or potential impact on the practices of forensic anthropology.

Following the AAFS presentation, the research will be submitted for publication in the *Journal of Forensic Sciences*. Additional publications are being considered for the *American Journal of Physical Anthropology* and the *Journal of Bone and Tissue Research*. The results will also be presented to the Scientific Working Group for Forensic Anthropology or SWGANTH. This working group consists of professionals from the forensic anthropology community with the goal to identify and recommend "best practice" within the forensic anthropology discipline. The SWGANTH has created Committees, which are populated by U.S. and international forensic anthropologists, to examine targeted issues for the purpose of identifying what is best practice for the profession to follow. A committee is currently being considered to develop guidelines for histological analysis.