Sex estimation of infants between birth and one year through discriminant analysis of the humerus and femur

Authors: Kyra E. Stull[1], MS, Kanya Godde, PhD [2,3]

[1]Department of Anatomy

University of Pretoria

Private Bag x323

Arcadia, South Africa 0007

[2]University of California, Merced

School of Social Sciences, Humanities, and Arts

School of Natural Sciences

5200 N. Lake Rd

Merced, CA 95340

[3]Department of Anthropology

University of Tennessee, Knoxville

250 South Stadium Hall

Knoxville, TN 37996

Corresponding author:

Kyra E. Stull

Department of Anatomy

University of Pretoria

Private Bag x323

Arcadia, 0007

Email: kstullster@gmail.com

Telephone: +27 012 319 2438

Mobile Number (USA): (864) 230-2301

Fax: +27 012 319 2240

ABSTRACT

Estimation of sex when investigating subadult skeletal remains is largely problematic due to unreliable and inaccurate results. Despite the limitations encountered with skeletal material, the medical literature clearly demonstrates differences between males and females in utero that persist through life. The current study investigates sexual dimorphism in the long bones of the

humerus and femur for individuals between birth and one year of age. A radiographic sample amassed from Erie County Medical Examiner's office includes 85 femoral and 45 humeral images for analysis in relation to sex. Measurements for lengths and breadths were collected through morphometric software. Discriminant analysis proved to be the most successful method, with error rates of 3% when utilizing maximum breadth at midshaft of the femur and 11% with humerus maximum distal breadth. This research demonstrates that it is possible to correctly classify sex of unknown subadult remains when comparing them to a known sample.

Keywords: forensic science, forensic anthropology, sex estimation, subadults, discriminant analysis, radiographs, femur, humerus

When faced with unidentified skeletal remains, the forensic anthropologist and bioarchaeologist assist with identification through the development of a biological profile. Techniques applied to mature remains can detect many attributes that assist in positive identification, such as age, sex, and ancestry. Overall, these techniques are considered more accurate and are widely accepted within the forensic anthropological community. However, when working with subadult remains, age is the only parameter routinely estimated. Methods applied to subadult remains rarely receive the same amount of scrutiny and are applied with varying degrees of accuracy. In part, this is due to the idea that diagnostic criteria indicating sex and ancestral differences are absent in the developing skeletal system. It is thought that these types of traits are not fully established until the completion of the adolescent pubertal hormonal stage. The medical literature reveals clear differences between sexes beginning very early in utero.

Sex differences begin around the eighth week after conception, when the embryo experiences a rise in hormonal levels (1–3). By week twelve, the process of sexual differentiation is largely complete; although the critically sensitive period, (the time in which tissue development can be modified by environmental influences) (3), lasts up to twenty-four weeks in utero (4). The next period of increased hormone levels, with median levels equivalent to the pubertal hormonal surge, is referred to as the neonatal hormonal surge (1, 3). During this period, levels of estrogen and testosterone, specific to chromosomally determined sex, are triggered by gonadotropin levels (3). The neonatal surge is considered to be as potent as the pubertal hormonal surge and lasts through the first year of life (1, 3). The hormone levels peak between the third and fourth month after birth for males while females have increased estradiol production from birth through early childhood (1, 3). Hormonal surges that create size differences between males and females and their occurrences, shortly after conception through the first year of life, are well documented in the medical literature. Thomas et al. (5) showed that female neonates had lower average birth weights, lengths, and head circumferences than male neonates. Additional studies corroborate the finding that males are consistently born heavier, longer, and with bigger head circumferences than females (6–10).

Many researchers have demonstrated the direct connection between the extent of fetal development and the pace of development during the first year of life (2, 11). Thomas et al. (5) also noted that sexual differences in growth velocity were apparent after the 36th week in utero, at which point skeletal development was already approximately three weeks more advanced in females than males. By birth, maturational variance between sexes increases to approximately a four to six week difference (2). Essentially, quantitative anthropometric differences between

males and females originate during the prenatal period and continue to advance throughout development.

Anthropological research on the developing skeleton has traditionally focused on the exploration of sex differences and sex estimation through shape analyses of the innominate and the mandible (12–20). While there have been some successful techniques which discern differences between the sexes, they may not be applicable to dry bone (18, 20, 21). Other studies have been published with forensically significant results, however validation studies of the original articles have rendered much lower accuracy rates (21–23). The skeletal elements continually investigated are areas known to change due to the pubertal hormonal surge, which assumes the same differences would be recognizable in the subadult skeleton prior to puberty. Factors that affect the pelvis are derived mainly from females preparing for reproductive maturation during the pubertal surge. Conclusions of results by Cardoso and Saunders (24) note that since the pelvis has late developmental pattern and adult size it may be that the sexually dimorphic features are an expression of development and not distinguishable prior to puberty. This observation may explain why anthropologists have been unsuccessful in quantifying morphological sex differences in subadult pelves.

The metric analyses of long bones to determine sex differences in subadult remains have been uncommon in the literature. One exception is Choi and Trotter's (25) investigation of the weight and length of fetal long bones in both blacks and whites. There was no significant difference found between ancestries in the fetal skeletons; however, Choi and Trotter (25) correctly classified the sex of 72% of the fetal remains when using discriminant analysis. Another area of metric research involves both the permanent and deciduous dentition by

examining the sexually dimorphic features of crown dimensions (26, 27). Although the recently revisited morphological studies of the mandible and pelvis fail to achieve success rates similar to those produced in the original articles (21–23), the publications analyzing metrics of long bones and dental dimensions consistently yield a high percent correct for sex estimation.

There are multiple issues that may cause discrepancies in the previously noted results, mainly data sources and age distribution. Substantial collections are rare which include modern subadult skeletons with known demographics (28, 29). Most large museums with research collections do not actively collect subadult skeletal material and even in collections that are actively growing, donations of deceased children are exceedingly rare. Most available material is either of archaeological origin or is extremely limited in terms of demographic variation. Forensic anthropologists can pursue radiographic data sources in the absence of modern subadult skeletal collections.

An issue that is rarely addressed in subadult research is the over-use of large age ranges. During growth there are recognized differences between infant, childhood, and juvenile ages following velocities and specific focuses (30). It is certainly possible that observation of large chronological age ranges may cause non-metric and metric results to fluctuate. For example, Franklin et al. (31) investigated sex differences of the mandible by application of geometric morphometrics in individuals of 1 to 17 years of age. Results of this study concluded the mandible is unfeasible for sex estimation (31). However, the enormous amount of growth and development that occurs during this period may drastically blur the morphological structures that are being analyzed. Research investigating changes within more narrow age ranges may elucidate sex differences that are being confounded by noise in the data. For this reason the

authors focused on a limited age range and separated the data set into smaller cohorts in order to determine how narrow age intervals should be to detect sexual dimorphism in subadults.

Results of medical, clinical, and physiological research demonstrate unequivocal differences between males and females in size and proportions beginning early in life (2, 30). Similar results should therefore be detected metrically in the subadult skeletal system. Soft tissue and skeletal differences between the sexes within the first year of life were first examined by Stull in 2008 (32). The current study builds upon Stull (32) by focusing solely on skeletal structures. The current study uses a radiographic sample of modern subadults to investigate sexual dimorphic differences in infant long bones. The elements evaluated are the left femur and humerus. These bones were chosen for numerous reasons, including their role in overall stature, for being sexually dimorphic throughout life, and their prevalence in being routinely imaged and easy to measure in radiographs (2, 33). The aim of the present study is to provide insight into sexually dimorphic differences of long bone lengths and breadths, between sexes, within the first year of life. As the above literature review points out, there have been significant differences between the sexes documented in the medical literature. With demonstrable differences in soft tissue structures, such as lengths, we expect skeletal tissue to exhibit a similar pattern.

Materials and Methods

Radiographs of 85 individuals (n= 49 males, n= 36 females) were used for the femoral analyses and 45 individuals (n=24 males, n= 21 females) were used for the humeral analyses. The radiographs were acquired from the Erie County Medical Examiner's Office in Buffalo, New York. The sample includes individuals between birth and one year of age (Figure 1). The

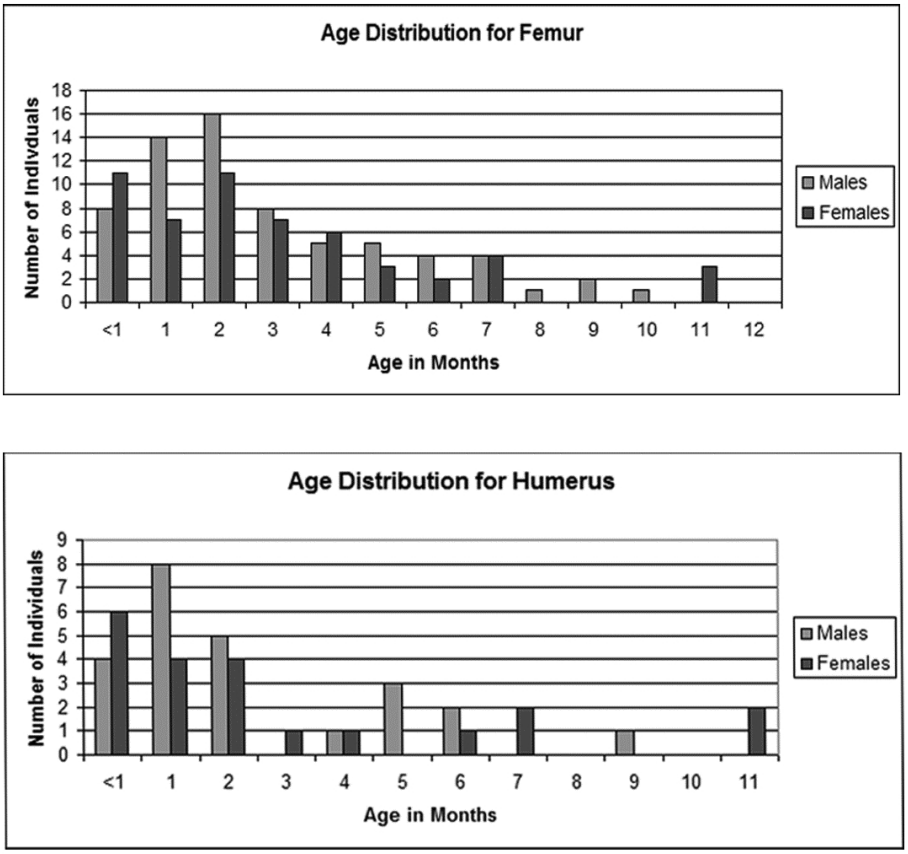

Figure 1. Age distribution tables separated by femur (above) and humerus (below).

range was limited in order to contend with timing of the neonatal hormonal surge and the numerous changes that occur in this developmental period. Due to the data source and the high rate of infant mortality for the first six months of age the data is somewhat skewed towards younger individuals. Sex, ancestry, and other biological parameters, such as height, weight, and cause and manner of death were known. The individuals were born between the years of 1998 and 2007. Digital images of the radiographs were collected and utilized. Considering the results

of Choi and Trotter (25) and the small dataset, ancestry groups were pooled to allow for the largest possible sample.

Two measurements were obtained from the femur for analysis: (1) maximum length (FMXL) and (2) maximum breadth at midshaft (FMSBRD). On the humerus, four measurements were examined: (1) maximum proximal breadth (HPBRD), (2) maximum length (HMXL), (3) maximum breadth at midshaft (HMSBRD), and (4) maximum distal breadth (HDBRD). Maximum length, breadth, and distal femoral and humeral measurements were assessed following Fazekas and Kósa (12) definitions. Maximum proximal humeral breadth was created specifically for the humerus in this study. The measurement is taken at the maximum projections on the proximal diaphysis while the bone is in anatomical position. The landmarks were demarcated in Photoshop CS3 and the document was then uploaded into tpsDig (34) and digitized. After the files were digitized, coordinate data was amassed using tpsUtil (35). The recorded measurements have been analyzed in pixels, rather than being converted to a metric scale.

The amount of distortion that is acquired during the radiographic process was tested by using a feature displayed both on the cartridge and within the radiographic image. This landmark was measured on both the cartridge and the film, in a random sample of thirty images. Measurements derived from the radiographs were scaled to real dimensions with an error of 0.2 mm. The error measurement reflects the distortion and magnification inherently produced during the process of radiography. This was only possible due to standardized operating procedures within the Erie County Medical Examiner's Office. If radiographs were introduced from another office, the images would all have to be appropriately scaled.

Statistical analyses were performed separately on the humeral and femoral data. Pearson product of moment correlations were calculated to determine if there was a statistical correlation between age, sex, and the measurements. A Kruskal-Wallis test was utilized as a means to determine if there were significant differences in the variables (measurements), among the sexes. The Kruskal-Wallis test is a nonparametric ANOVA executed on the ranks of the data and tests the null hypothesis that means are equal across samples. The data here are of unequal size and not necessarily normally distributed (depending on how the data is partitioned temporally), which requires this nonparametric approach. Only two measurements were calculated from femoral landmarks, and thus variable selection was not necessary for examining regression and discriminant analysis (DA) models, as there are only 3 possible models to explore. Conversely, four measurements were available to model sex from the humerus. As this will produce numerous models to inspect, McHenry's algorithm for variable selection was utilized to find the best performing models [35]. Small changes in Wilks' Lambda indicate the most appropriate models for testing.

As sex is a categorical trait, logistic regression (36, 37) was utilized as a means to model the measurements as predictors of sex. Logistic regression is a regression alternative that allows for assessment of categorical data or combined categorical and continuous data. Logistic regression can produce results that measure the significance of the model through the analysis of deviance. Moreover, it also generates the percentage of observations correctly classified, based on the variables in the model. The preceding analyses were completed in NCSS (38), as were plots to examine patterns in the data. A fifth analysis was also undertaken to understand the distribution of measurements across the sexes: discriminant analysis. Discriminant analysis is a

model free clustering method that classifies individual observations into a preselected number of groups, based on the variables in the model. It produces an error rate, which when validated, estimates the number of individuals that were misclassified in relation to the known group assignment.

Discriminant analysis has several options for proceeding, based on data structure. In order to determine which DA is the best fit, several screening procedures should be undertaken. First, multivariate normality (MVN) must be determined. Both linear discriminant analysis (LDA) (39) and quadratic discriminant analysis (QDA) (40) require MVN for reliable results. If the data is not MVN, a nonparamentric approach, for example nearest neighbor discriminant analysis (NNDA) (41), is selected to generate reliable results. Linear discriminant analysis also requires that the variance-covariance matrices (s matrices) of the different groups (sex) in the sample are not significantly different from one another. Thus, a chi-square test was applied to the variance-covariance matrices of the sexes to determine if LDA was an appropriate analysis. If MVN is achieved, but the variance-covariance matrices are not equal, QDA is the best analysis for the data, with the caveat that QDA does not perform well on small sample sizes. Thus, LDA should be substituted for QDA when dealing with small samples, even if the s matrices are significantly different. Both QDA and LDA produce Wilks' Lambda estimates, whose p-values indicate whether or not there are significant differences among the groups (sex) in the samples under study.

The nonparametric approach selected here for dealing with data not MVN was NNDA (41). Nearest neighbor discriminant analysis is a kernal method that classifies observations based on the value of its nearest neighboring observation(s). Due to the small sample sizes in

this paper, k=3 was selected (42) as the number of neighboring observations by which each observation should be compared to. Validation of DA results should be conducted to reduce overfitting the model to the training sample. As the original data sets were too small to create holdout samples, validation was completed using the jackknife method (Tukey's method; see Miller 1974 (43) for a review of the methodology). The jackknife method of validation is more conservative than resubstitution, and thus gives a more realistic estimation of DA performance via the error rate. Discriminant analyses, and its subsequent validation, were performed in SAS 9.1.2 (44).

Results

Plots of age in relation to size for the measurements most highly correlated to sex (FXLN and HDBRD), reveal that males and females do not cluster separately (Figures 2 and 3, respectively). Rather, observations from males and females are interspersed among one another. The femoral measurements were weakly negatively correlated (Pearson product of moments correlations) to sex (Table 1), while weak positive correlations existed between humeral variables and sex, with the exception of HPBRD, which was negatively correlated with age (Table 2). Temporal trends in the data were suspected as the cause for the low correlations. Thus, new Pearson product of moment correlations were generated with the addition of an age variable (in weeks). These new tests revealed a strong correlation between age and the femoral

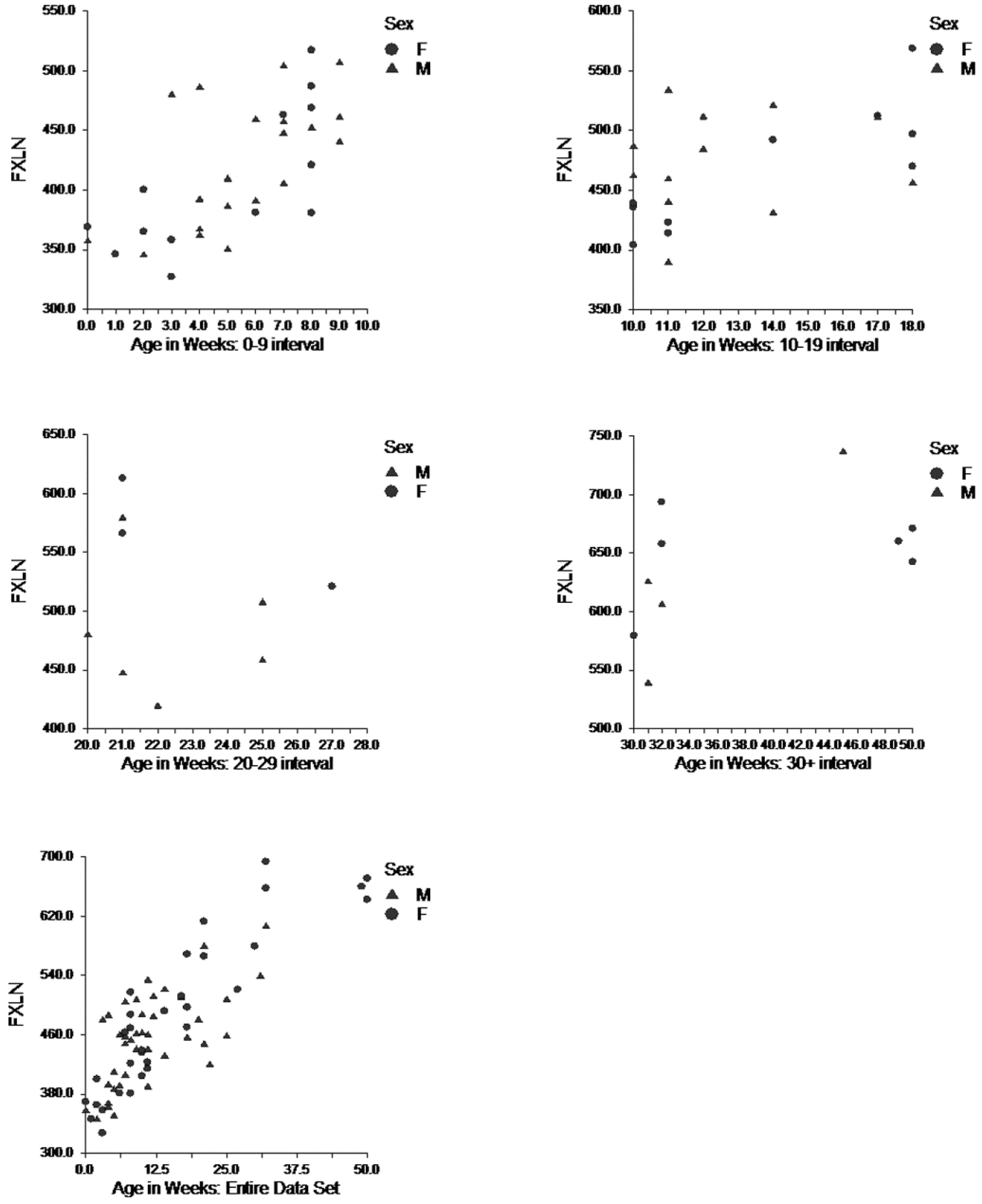

Figure 2. Plots of FMXL by age, by category, and entire data set.

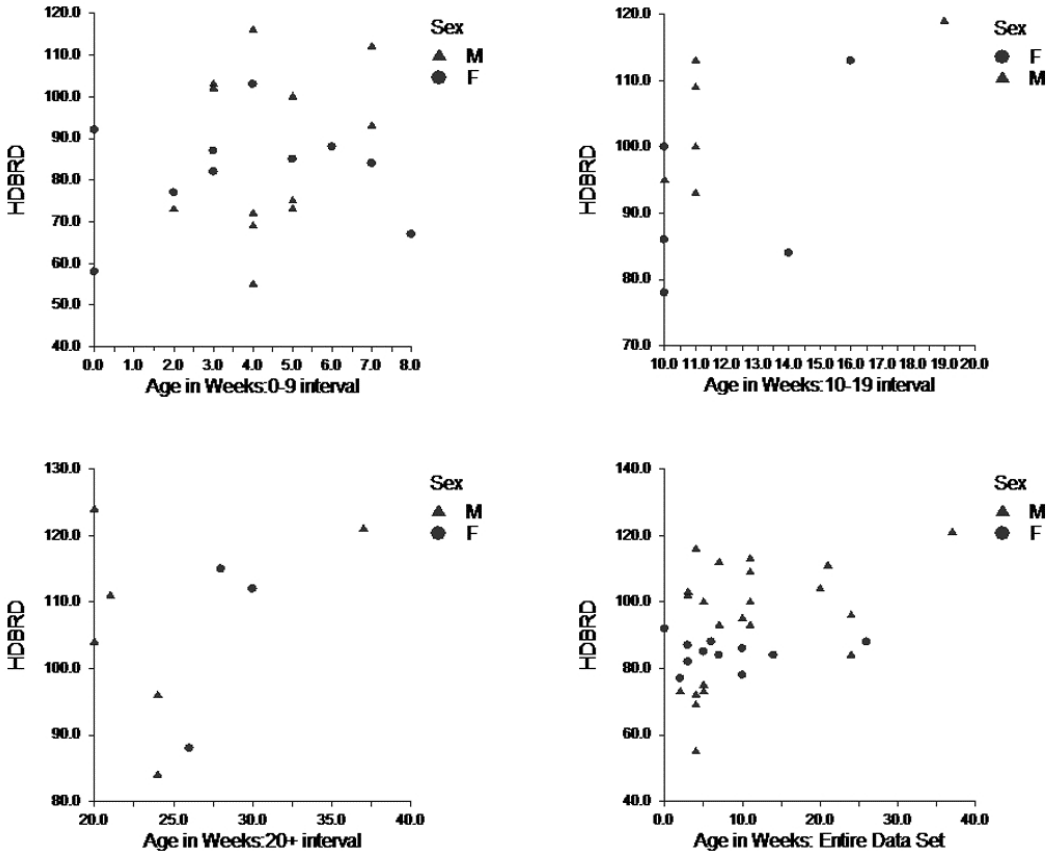

Figure 3. Plots of HDBRD by age, by category, and entire data set.

measurements (Table 1), and a lesser correlation among the humeral measurements and age (Table 2).

Table 1—*Pearson product of moment correlations of entire femur data set.*

	Sex	Age (in weeks)	FXLN	FMSBRD
Sex	1			
Age (in weeks)	-0.202793	1		
FXLN	-0.184317	0.822406	1	
FMSBRD	-0.148145	0.63188	0.849724	1

Table 2—*Pearson product of moment correlations of entire humerus data set.*

	Sex	Age (in weeks)	HPBRD	HXLN	HMSBRD	HDBRD
Sex	1					
Age (in weeks)	0.15571	1				
HPBRD	-0.01112	0.450851	1			
HXLN	0.128586	0.603943	0.75602	1		
HMSBRD	0.155099	0.595816	0.706582	0.716005	1	
HDBRD	0.293188	0.408755	0.446893	0.574872	0.663273	1

With confirmation of a temporal trend, the data were separated into 10-week cohorts (unless sample size per cohort numbered less than 5) to dissect out any sexual dimorphism in the data. A Kruskal-Wallis one-way ANOVA on the age cohorts confirmed the partitioning of the cohorts was appropriate through significant p-values ($p=0.000004$ for femoral age cohorts, $p=0.004393$ for humeral age cohorts). The age cohorts were also subject to Kruskal-Wallis one-way ANOVA analyses in relation to sex. No significant values were produced among the femoral or humeral data (Tables 3 and 4).

Table 3—*Kruskal-Wallis one-way ANOVA on femur data.*

Data set	N	Variable	Chi-Square	p-value
Entire	72	FXLN	0.713909	0.398149
Entire	72	FMBRD	0.6215424	0.430475
0-9 weeks	33	FXLN	0.4900452	0.483907
0-9 weeks	33	FMBRD	0.4178888	0.517992
10-19 weeks	22	FXLN	0.3521739	0.552885
10-19 weeks	22	FMBRD	2.73E-02	0.868735
20-29 weeks	9	FXLN	3.266667	0.070701
20-29 weeks	9	FMBRD	1.103448	0.293511
30+ weeks	10	FXLN	0.7272727	0.393769
30+ weeks	10	FMBRD	0.1829268	0.66887

Table 4—Kruskal-Wallis one-way ANOVA on humerus data.

Data set	*N*	*Variable*	*Chi-Square*	*p-value*
Entire	*33*	*HPBRD*	*1.32E-02*	*0.908681*
Entire	*33*	*HXLN*	*1.025385*	*0.311245*
Entire	*33*	*HMSBRD*	*0.5601866*	*0.454185*
Entire	*33*	*HDBRD*	*3.364732*	*0.066606*
0-9 weeks	*22*	*HPBRD*	*0.039375*	*0.842708*
0-9 weeks	*22*	*HXLN*	*0.736446*	*0.390802*
0-9 weeks	*22*	*HMSBRD*	*7.06E-02*	*0.790463*
0-9 weeks	*22*	*HDBRD*	*0.1839033*	*0.66804*
10-19 weeks	*11*	*HPBRD*	*3.33E-02*	*0.855132*
10-19 weeks	*11*	*HXLN*	*0.3*	*0.583882*
10-19 weeks	*11*	*HMSBRD*	*3.38E-02*	*0.854145*
10-19 weeks	*11*	*HDBRD*	*2.152905*	*0.1423*
20+ weeks	*9*	*HPBRD*	*3.266667*	*0.070701*
20+ weeks	*9*	*HXLN*	*1.361345*	*0.243305*
20+ weeks	*9*	*HMSBRD*	*2.440678*	*0.118225*
20+ weeks	*9*	*HDBRD*	*0*	*1*

The best logistic regression models from each data set (entire and age cohorts) are reported in Tables 5 and 6 for femoral and humeral measurements, respectively. The results from logistic regression on the femoral measurements only yielded one model whose variables were significant: FMXL in the 20-29 week cohort. This model classified the observations into the correct sex 78% of the time. The other models and cohorts ranged from 50-90% classification rate. In the humerus, two models were significant: HMXL in 0-9 week age cohort and HPBRD in the 20+ week cohort. The classification rates for these models were 63.636% and 66.667%, respectively.

Table 5—*Results from logistic regression on femur data.*

Dataset	Model	% Correctly Classified	Analysis of Deviance	Overall p-value	p-value of Terms
Entire	FXLN	56.944	96.92	0.12	
0-9 weeks	FMSBRD	63.636	44.25	0.61	
10-19 weeks	FXLN	50	30.32	0.67	
20-29 weeks	FXLN	77.778	11.46	**0.04**	
30+ weeks	FXLN, FMSBRD	90	13.46	0.18	FXLN = 0.12, FMSBRD = 0.08

Bold indicates significant at the 0.05 level

Table 6—*Results from logistic regression on humerus data.*

Dataset	Model	% Correctly Classified	Analysis of Deviance	Overall p-value
Entire	HDBRD	75.758	42.01	0.09
0-9 weeks	HXLN	63.636	30.32	**0.4**
10-19 weeks	HDBRD	63.636	15.16	0.09
20+ weeks	HPBRD	66.667	11.46	**0.03**

Bold indicates model is significant at the 0.05 level

The appropriate DA was selected for each femoral and humeral data set, based on the selection criteria discussed above. The results of the DA analysis selection criteria for the best models of each dataset are listed in Tables 7 and 8. In the femur, error rates ranged from 0-14%. The humeral DA were not as accurate; error rates ranged from 8-38%.

Table 7—*Best performing DA per femur data set. Decision criteria, error rates, and misclassifications are listed.*

Data set	Variable	Analysis	Error Rate	# of F misclassified	# of M misclassified	Chi-square for s matrices	Wilks' Lambda	MVN
Entire	FMSBRD	NNDA	0.0313	0	2	na	na	Absent
0-9 weeks	FMSBRD	NNDA	0.1429	0	4	NA	NA	Absent
10-19 weeks	FMSBRD	QDA	0.3	0	6	<.0001	0.6187	Present
20-29 weeks	FMSBRD, FXLN	NNDA	0	0	0	na	na	Absent
30+ weeks	FMSBRD	LDA	0.1	0	1	0.0009	0.0911	Present

Table 8—*Best performing DA per humerus set. Decision criteria, error rates, and misclassifications are listed.*

Data set	Variable	Analysis	Error Rate	# of F misclassified	# of M misclassified	Chi-square for s matrices	Wilks' Lambda	MVN
Entire	HDBRD	QDA	0.1136	1	2	<.0001	0.0977	Present
0-9 weeks	HPBRD, HXLN, HDBRD	NNDA	0.375	3	5	NA	NA	Absent
10-19 weeks	HXLN, HDBRD	LDA	0.2833	1	2	0.416	0.0356	Present
10-19 weeks	HDBRD	LDA	0.3667	2	2	0.5445	0.1221	Present
20+	HMSBRD, HDBRD	NNDA	0.0833	1	0	NA	NA	Absent

Discussion

Initiation of this research was due to the overwhelming medical and clinical data illustrating sexual differences within the fetal period as well as the first year of life. It is logical to assume there would be manifestations of these differences in long bones, which is the primary reason they were chosen. The results of this study demonstrate that there are sexually dimorphic differences in the humerus and femur within the first year of life; however this was only successfully observed through discriminant analysis. Within the Pearson product of moment correlations, the strongest result was between age and long bone lengths of the humerus and femur. This is an expected result as long bone length is a variable forensic anthropologists rely on for estimation of age-at-death in subadults when soft tissue and/or dentition is not present (45). Narrowing age to less than one year is not necessary and impractical when estimating sex from long bone measurements (see below). Despite the fact that a logistic regression model yielded 78% correct classification, it is not the most appropriate approach, as the variables were not significant, suggesting a poor fit of the long bone measurements to an equation that estimates sex.

Discriminant analysis's performance is better than regression, indicating that measurements on an unknown infant can be used to estimate sex when compared to a sample of known sex infants. After running all possible models for the femoral data on the entire dataset and the individual age cohorts, femur midshaft breadth (FMSBRD) on the entire data set was the best model (using NNDA), yielding a 3% error rate. Although two age categories (20-29 and 30+ weeks) also have low error rates, one then introduces dependence upon age-at-death estimation and higher error rates among other age cohorts. Averaging the error rate across all age cohorts, to take into account the possible misclassification of the age-at-death technique, the average error rate is much higher, at 14%. Therefore, classification by age into smaller categories is not as accurate as using a 0-1 year age cohort. The same is true for the humeral DA analysis.

The suggested McHenry model for the humerus is the three-variable model that generated an error rate of 11%. Similar to the femoral data, the humeral 20+ age category has a lower error rate (~1%) than the entire data set; however the same issues arise when including the potential problems associated with age-at-death estimation techniques. Averaging across the age cohorts, the error rate increases to 20%. Therefore, the error rate for analyzing juvenile remains independent of age cohort is superior to breaking the age category down smaller than 0-1 year old.

The age-at-death techniques currently applied within the field may not be able to accurately place unknown remains into age categories more specific than 0-1 year, and thus not separating by smaller age cohorts is more appropriate in anthropological contexts. Although the current study did attempt to separate the data sample into 10-week categories (based on the results from the correlation matrix), the statistics were unable to discern the slightest expression

of sexual dimorphism within the sample. The peak of the neonatal hormone surge (3-4 months), if detectable in long bone measurements, should have resulted in a sexual dimorphic difference in the corresponding age cohort (12-19 weeks). Even though this hormonal difference is not evident in the long bone measurements input into DA, it is not to say that this event does not have an effect on long bone dimensions, or would be prevalent in different samples. Larger sample sizes may allow for a clearer insight into true growth differences between males and females. Additionally, investigating fetuses may elucidate more nebulous differences, considering the neonatal hormonal surge begins early in utero.

Plots reveal a potential bias in this study when depicting the relationship of size, sex, and age within the data sample (Figure 2). There is no evident trend between the sexes in regards to size as it relates to age, which may be an artifact of small sample size. Small sample size may also have influenced the outcome of this research in regards to the excellent error rates reported within. Thus, caution should be exercised when applying the methodology to other samples, until these results can be verified on a much larger, more diverse sample. Additional error was inherently introduced due to the nature of radiography. This should be considered whenever doing metric analysis on radiographic images. However, the measurements obtained on each element are repeatable on dry bone and the results of this study demonstrate that applying DA statistics with a known sample should result in high classification rates.

The greatest strength of this research lies in the promising results from analyzing the data with DA. This paper demonstrates there is no need to separate by age within the first year of life, as the strongest results overall were obtained using the entire sample. Therefore, this eliminates error introduced by estimating age-at-death into highly specific age categories. DA

using the aforementioned models for both the femur and the humerus illustrates that an unknown subadult skeleton can be assigned a sex with high degree of certainty, when compared to a sample of known sex infants. This result corroborates other literature, such as Choi and Trotter (25), which also demonstrates a high correct classification rate of fetal sex when utilizing DA of long bone measurements.

Conclusions

Past medical and clinical studies have demonstrated there are sexual differences very early during growth and development. The anthropological literature has been less persuasive in the matter. Although numerous studies have investigated sex differences within the subadult skeleton, most have focused on morphological variation and less on metric analysis. Moreover, most studies include large age intervals, rather than focusing on more specific cohorts that are partitioned based on known events during growth and development (30). This approach has led to results where the accuracy of the technique have either 1) not been replicable, or 2) are not acceptable within forensic anthropological standards.

The current study demonstrates that it is possible to correctly sex unknown skeletal remains less than one year of age through metric analyses of the humerus and femur. The DA results boast low error rates with the elimination of introducing bias from age-at-death estimation techniques. This has major implications in the forensic anthropological field by providing a possibility to expand the biological profile to incorporate sex when assessing unknown subadult skeletal remains. Future studies should be directed towards extending the presented research

model to all individuals, prior to the onset of epiphyseal fusion and expression of secondary sexual characteristics.

The two weaknesses of the current study are the small data sample and the use of radiographic images. As for the small sample size, it may have had implications on the logistic regression results, as well as possibly concealing any trends that may be associated with the neonatal hormonal surge. As for the radiographic analysis, although there is always an issue with distortion and applying the technique on dry bone, it seems to be the only option for research involving subadult skeletal remains.

The research model presented in this study can be used by forensic anthropologists in conjunction with a dataset of known individuals to correctly determine sex of unknown subadult skeletal remains. In an effort to augment existing subadult skeletal data sources, the National Institute of Justice awarded funding (2008-DN-BX-K152) for the development of a national radiographic database of known modern subadults between birth and twenty years of age. The data were collected from geographically diverse medical examiners offices across the United States. The radiographic images were scanned and all demographic and biological information available was collected. This large collection of subadult radiographs will be available online and is recommended as a data source to all researchers interested in trends in modern growth and development. This reference collection, along with the individuals utilized in this research project, will be amassed to compose a known data set for comparison. Future research also promises the reexamination of the soft tissue data as well as age-at-death estimations through discriminant analysis in hopes to further reveal differences in the subadult skeletal systems.

Acknowledgements

Many thanks go to the Erie County Medical Examiner's Office, specifically Dr. Dianne Vertes, for providing access to a radiographic data source. This research would never have been possible without the guidance, education, and support from the faculty members at Mercyhurst College which include Drs. Dennis Dirkmaat, Steve Symes, Steve Ousley, and Mr. Luis Cabo-Perez. Additional thanks go to Kathryn Frazee and Stephanie Super.

References

1. Fechner P. The biology of puberty: New developments in sex differences. In: Hayward C, editor. Gender differences at puberty. Cambridge: Cambridge University Press; 2003:17–28.

2. Tanner J. Foetus into Man: Physical Growth from Conception to Maturity. Cambridge: Howard University Press; 1989.

3. Knickmeyer R, Baron-Cohen S. Fetal testosterone and sex differences. Early Hum Dev 2006;82:755–60.

4. Wilhelm D, Palmer S, Koopman P. Sex determination and gonadal development in mammals. Physiol Rev 2007;87:1–28.

5. Thomas P, Peabody J, Tunier V, Clark R. A new look at intrauterine growth and the impact of race, altitude, and gender. Pediatrics 2000;106(2)

6. Largo R, Walli R, Duc G, Fanconi A, Prader A. Evaluation of perinatal growth. Presentation of combined intra- and extrauterine growth standards for weight, length and head circumference. Helvetica Paediatrica Acta 1980;35(5):419–36.

7. Garn S, Keating M. Effect of various prenatal determinants on size and growth through seven years. Ecol Food Nutr 1980;9:109–12.

8. Lubchenco L, Hansman C, Dressler M, Boyd E. Intrauterine growth as estimated from liveborn birth-weight data at 24 to 42 weeks of gestation. Pediatrics 1963;32:793–800.

9. Lubchenco L, Hansman C, Boyd E. Intrauterine growth in length and head circumference as estimated from live births at gestational ages from 26 to 42 weeks. Pediatrics 1966;37(3):403–8.

10. Olsen I, Groveman S, Lawson M, Clark R, Zemel B. New Intrauterine Growth Curves Based on United States Data. Pediatrics 2010;125:214–24.

11. Roche A, Sun S. Human Growth: Assessment and Interpretation. Cambridge: Cambridge University Press; 2003.

12. Fazekas S, Kósa F. Forensic Fetal Osteology. Budapest: Akademiai Kiado; 1978.

13. Schutkowski H. Sex determination of fetal and neonatal skeletons by means of discriminant analysis. Int J Anthropol 1987;2:347–52.

14. Schutkowski H. Sex determination of infant and juvenile skeletons: I. morphognostic features. Am J Phys Anthropol 1993;90:199–205.

15. Molleson T, Cruse K, Mays S. Some sexually dimorphic features of the human juvenile skull and their values in sex determination in immature remains. J Archaeol Sci 1998;25:719–28.

16. Holcomb S, Konigsberg L. Statistical study of sexual dimorphism in the human fetal sciatic notch. Am J Phys Anthropol 1995;97:113–25.

17. Sutter R. Nonmetric Subadult Skeletal Sexing Traits: l. A Blind Test of the Accuracy of Eight Previously Proposed Methods Using Prehistoric Known-Sex Mummies from Northern Chile. J Forensic Sci 2003;48(5):927–35.

18. Reynolds E. The Bony Pelvic Girdle in Early Infancy. Am J Phys Anthropol 1945;3:321–54.

19. Sundick R. Age and sex determination of subadult skeletons. J Forensic Sci 1977;22:141–4.

20. Boucher B. Sex differences in the foetal pelvis. Am J Phys Anthropol 1957;15:581–600.

21. Scheuer L. A Blind Test of Mandibular Morphology for Sexing Mandibles in the First Few Years of Life. Am J Phys Anthropol 2002;119:189–91.

22. Loth S, Henneberg M. Sexually dimorphic mandibular morphology in the first few years of life. Am J Phys Anthropol 2001;115:179–86.

23. Vlak D, Roksandic M, Schillaci M. Greater sciatic notch as a sex indicator in juveniles. Am J Phys Anthropol 2008;137:309–15.

24. Cardoso H, Saunders S. Two arch criteria of the ilium for sex determination of immature skeletal remains: A test of their accuracy and an assessmend of intra- and inter- observer error. Forensic Sci Int 2008;178:24–9.

25. Choi S, Trotter M. Statistical study of the multivariate structure and race-sex differences of American White and Negro fetal skeletons. Am J Phys Anthropol 1970;33:307–12.

26. Black T. Sexual dimorphism in the tooth-crown diameters of the deciduous teeth. Am J Phys Anthropol 1978;48:77–82.

27. Cardoso HF. Sample-Specific (Universal) Metric Approaches for Determining the Sex of Immature Remains Using Permanent Tooth Dimensions. J Archaeol Sci 2008;35(1):158–68.

28. Feldesman M. Femur/stature ratio and estimates of stature in children. Am J Phys Anthropol 1992;87(4):447–59.

29. Shapiro D, Richtsmeier J. Brief communication: A sample of pediatric skulls available for study. Am J Phys Anthropol 1997;103:415–6.

30. Bogin B. Patterns of Human Growth. Cambridge: Cambridge University Press; 1999.

31. Franklin D, Oxnard C, O'Higgins P, Dadour I. Sexual dimorphism in the subadult mandible: Quantification using geometric morphometrics. J Forensic Sci 2007;52(1):6–10.

32. Stull K. Sexual dimorphism of the infant skeleton. In: 60th Annual Meeting of the American Academy of Forensic Sciences. Washington, DC.: 2008

33. France DL. Observation and Metric Analysis of Sex in the Skeleton. In: Reichs KJ, editor. Forensic Osteology: Advances in the indentification of human remains. Springfield: Charles C. Thomas; 1998

34. Rohlf F. tpsDig. State University at Stony Brook, Stony Brook, New York: 2010.

35. Rohlf F. tpsUtil. State University at Stony Brook, Stony Brook, New York: 2010.

36. Agresti A. A Survey of Strategies for Modeling Cross-Classifications Having Ordinal Variables. J Am Stat Assoc 1983;78:184–98.

37. Agresti A. Analysis of Ordinal Categorical Data. New York: John Wiley & Sons; 1984.

38. Hintze J. NCSS, PASS, and GESS. Kaysville: 2006.

39. Fischer R. The Use of Multiple Measurements in Taxonomic Problems. Ann Eugen 1936;7:179–88.

40. Smith C. Some Examples of Discrimination. Ann Eugen 1947;13:272–82.

41. Fix E, Hodges J. Discriminatory Analysis - Nonparametric Discrimination: Consistency Properties. 1951;

42. Rencher A. Methods of Multivariate Analysis. New York: John Wiley; 1995.

43. Miller R. The Jackknife -- A Review. Biometrika 1974;61:1–15.

44. SAS. SAS 9.1 for Windows SAS Institute, Inc. Cary, North Carolina: 2002.

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

Additional Information and Reprint Requests:

Kyra E. Stull, M.S.

Department of Anatomy

University of Pretoria

Private Bag x323

Arcadia, South Africa 0007

Email: kstullster@gmail.com