Exploring Statistical Control of Specimen Size: Examples using Portuguese Water Dogs from "The Georgie Project"

J. Hunter Marshall [1], John G. Skedros [1], Kevin Chase [2]
[1]University of Utah Dept. of Orthopaedic Surgery; [2] Dept. of Biology, Salt Lake City, Utah, USA
Email of corresponding author: jskedrosmd@uosmd.com

Disclosures: The authors have nothing to disclose.

INTRODUCTION: A common confounding variable when comparing morphometric parameters is variation of specimen body size. There is a long history of using various ratios to 'control' for size, but ratio adjustments in morphologic studies have since been criticized heavily because they often lead to erroneous/illusionary correlations and may not actually control for specimen size [1-3]. Ratio adjustments of data create a 'new' measurement that represents a proportion or relative size measurement as in the case of the body mass index (BMI), which can make interpreting the 'new' variable difficult. The statistical sense of 'control' involves partitioning the variance of the response variable, which is best accomplished by multivariable methods such as multiple regression. Multiple regression analyses are often used to control for specimen size and have been reported to be superior to ratio adjustments in the context of QTL genetic analyses [3]. In this study of age- and sex-related changes of the mid-femoral canine diaphysis, we evaluate various methods employed by researchers to control for specimen size, and illustrate the differing conclusions that could be interpreted based on the particular statistical model chosen to control for body size – emphasizing the importance of selecting an appropriate statistical model. In this study, we used the femora from a large sample of Portuguese Water Dogs (PWDs) to determine the most appropriate statistical model to control for specimen size when the variable of interest is correlated with size. This is important if researchers hope to elucidate the effects of age and/or sex of a given variable while also controlling for the size variations that are common between males and females.

MATERIALS AND METHODS: PWD carcasses were obtained from animals from "The Georgie Project" [4-6]. We used the right femur from 415 mature PWDs; age range: 2-16 years old. Transverse cross-sections of the diaphysis at 50% of total bone length (Ltot) were scanned and cortical (CA), medullary (MD.Ar), and total cross-sectional areas (T.Ar) were measured digitally using imageJ. The second polar moment of area (Jpol) was calculated using these measurments. Statistical methods of controlling for specimen size as a confounding variable included: allometric ratios, multiple regression, and size stratification, and were analyzed using the statistical software NCSS 2020. The model used in the multiple regression analysis is represented by the equation, , where sex was defined as a dichotomous variable where males were assigned a value of zero and females were assigned a value of one. No interaction terms were significant, and therefore were excluded from the model. All independent variables were centered around the mean (i.e.,) to control for multicolinearity as determined by the calculated variance inflation factor.

RESULTS: Table 1 illustrates correlations and p-values between the parameter and age, for the various methods.

DISCUSSION: The results of this study provide the important ground work to more conclusively determine the best statistical model for controlling specimen size when the variable of interest is correlated with size. This is an essential step in our ongoing age- and osteoporosis-related research on femora in the PWD model. Three of eight analyses showed that age was statistically significant but the fits of the models were poor, and thus the significance of age is not very useful in explaining the variation of the parameters. As **Table 1** illustrates the various statistical approaches can influence an observer's interpretation and conclusion. Ratio adjustments can be misleading because either the numerator or denominator could be influencing the outcome but one would have no way of discerning which variable had the larger effect. Regression models can have a poor fit, which reduces the ability to make conclusions based on the model given. This was the case in a few of our models shown in **Table 1**. We then turned to stratifying a subset of our sample according to size [7]. By selecting specimens of similar size to compare age– and sex-related differences we could 'control' for specimen size in the data selection process. However, the statistical power is reduced because the sample size is reduced in order to stratify, and the variation of the parameter of interest is artificially constricted. The non-adjusted category shows that females had a negative percent change in both the median and mean values with total area. This observation is most likely attributed to the fact that the older females in our sample happen to be smaller than the younger specimens, and thus illustrates the importance in controlling for body size during the analysis process. Monte Carlo simulations are needed to unveil which method would most accurately control for specimen size when size is correlated with the parameter of interest. At this time stratification seems most promising.

SIGNIFICANCE: Properly controlling for specimen size to asses for age, sex, or other related effects when the variable of interest is correlated with size can improve the accuracy of conclusions in many fields of study. Currently we are performing Monte Carlo simulations to conclusively recommend a method for controlling for specimen size when the variable measured is correlated with size.

REFERENCES: [1] Albrecht et al.(1993) Am J Phys Anrthop 91:441-; [2] Jasienski and Bazzaz (1999) Oikos 84:321-; [3] Smith (2005) Curr Anthrop 46:249-; [4] Chase et al. (2002) PNAS USA 99: 9930-; [5] Chase et al. (2005) Am J Med Genet A. 135: 334-; [6] Chase et al. (2011) AGE 33:461–; [7] Looker et al. (2001) JBMR 16: 1291-.

Table 1

	Allometric Ratio				Multiple Regression					Stratified				Non-Adjusted			
	Pearson r-value	Pearson p-value	Median % change	Mean % change	Adj. R^2	Pearson Ltot	Pearson Weight	Pearson Sex	Pearson Age	Pearson r-value	Pearson p-value	Median % change	Mean % change	Pearson r-value	Pearson p-value	Median % change	Mean % change
CA (mm^2)	$(Ltot)^{2.99}$																
male	-0.16	0.06	0.0%	0.0%	0.34	0.24***	0.19***	-7.15***	-0.5**	-0.2	0.06	0.8%	-0.9%	-0.2	0.06	1.0%	-0.7%
female	-0.7	0.3	4.4%	4.3%						-0.09	0.5	-4.5%	-10%	-0.3	0.001***	-10.0%	-11.0%
ANOVA†	0.2						N/A			<0.001***				<0.001***			
MD.Ar (mm^2)	$(Ltot)^{4.18}$																
male	0.1	0.2	6.5%	8.1%	0.12	0.52***	0.18	0.95***	0.97	0.2	0.2	5%	8%	0.15	0.07	7.0%	7.0%
female	0.2	0.01**	15.5%	14.2%						0.2	0.06	13%	13%	0.1	0.1	9.0%	7.0%
ANOVA†	<0.001***						N/A			0.05*				<0.001***			
T.Ar (mm^2)	$(Ltot)^{2.76}$																
male	0.03	0.8	3.5%	3.7%	0.3	0.78***	0.36***	-6.39**	0.42	-0.01	0.9	0.3%	4%	0.05	0.6	2.0%	3.0%
female	0.1	0.2	5.7%	3.9%						0.14	0.3	1%	2%	-0.06	0.4	-0.4%	-1.4%
ANOVA†	0.03*						N/A			<0.001***				<0.001***			
Jpol (mm^4)	$(Ltot)^{3.17}$																
male	-0.05	0.6	1.2%	3.9%	0.31	21.8***	12.45***	-294.9***	-1.67	-0.06	0.6	-1%	4%	-0.01	0.9	1.3%	3.0%
female	0.03	0.4	0.3%	1.7%						0.04	0.7	-1.3%	-5%	-0.2	0.03*	-6.2%	-10.0%
ANOVA†	0.5						N/A			<0.001***				<0.001***			

† ANOVA was calculated with age as a covariate, meaning the differences between males and females was calculated at the average age between the two sexes. This was done to control for confounding effects of age when calculating means that otherwise could skew the mean to be either larger or smaller.
* p< 0.05, ** p< 0.01, *** p< 0.001