Missing data estimation in tyrannosaurid dinosaurs: Can diameter take the place of circumference?

Gavin J. Bradley[*], Michael E. Burns, Philip J. Currie

Department of Biological Sciences, University of Alberta, 11455 Saskatchewan Drive, Edmonton, Alberta T6G 2E9, Canada

A B S T R A C T

Missing data, due to taphonomic deformation, inaccessibility of specimens, or human error in collecting, cataloguing and measuring features, is a formidable problem in current palaeontological studies. Missing values within a data set can undermine confidence in analyses, skew results in promoting analyses of small portions of a population, not necessarily representative of the entire data set, and drastically decrease sample sizes. Missing data estimation methods, however, may reduce the effects of these missing values and potentially boost sample sizes for palaeontological studies. Here, six missing data estimation models for the prediction of femoral circumferences in tyrannosaurids were statistically tested and their predictive success measured against true circumferences, and other models. The statistical analyses suggest that estimation models based on anteroposterior diameter values of tyrannosaurid femora were poor predictors of circumference, whereas those based on mediolateral diameters were much more successful. Three out of the six models, were presented as viable alternatives to missing measured circumferences and may be used to boost tyrannosaurid samples with significant levels of missing data.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Missing data are a major limitation in palaeontological studies, decreasing confidence in analyses and drastically reducing sample sizes (Kearney & Clark, 2003). The problem of missing data is far reaching within many subsets of the discipline, and affects a host of palaeontological methodologies. In fossil reconstructions, missing data can limit the accuracy and extent to which an extinct organism can be recreated (Grillo and Azevedo, 2011). It can also impact the rate at, and specificity with, which new specimens may be described and assigned a phylogenetic position; missing data have prevented more exclusive phylogenetic resolution for the potential Chinese tyrannosauroid *Chingkangosaurus fragilis*, known only from fragmentary scapula remains (Brusatte et al., 2013), and delayed a comprehensive phylogenetic analysis of *Deinocheirus mirificus* for fifty years (Lee et al., 2014).

The potential impact of missing data in morphometric analyses is well known, and alternative missing data estimation techniques, such as 'Bayesian Principal Component Analysis' and 'mean substitution' have been proposed and compared (Campione and Evans, 2011; Brown et al., 2012). Phylogenetic analyses are also complicated by the effect of missing data (Wilkinson, 2003; Wiens, 2003; Wiens, 2006; Wiens and Morrill, 2011; Lemmon et al., 2009), and in this subfield of palaeontology, missing data estimation models have also been devised to reduce the effect of spaces in the sample (Norell & Wheeler, 2003).

The most common cause of missing data in a palaeontological sample is taphonomic deformation (Chapman, 1990; Dilkes, 2001; Miyashita et al., 2011; Arbour & Currie, 2012; Tschopp et al., 2013; Hedrick & Dodson, 2013). The geological pressures that distort and destroy fossils as they form are beyond the control of palaeontologists; however, paucity in sample size can also be compounded by anthropological factors. Human error might lead to the destruction of fossils due to accidental damage or vandalism (Lipps, 2009), the misplacement of specimens or accompanying information between excavation and preparation, or even inappropriate measurement of specimens. Missing data can also be generated by lack of access to the complete specimen due to backlogs in fossil preparation laboratories, its inclusion in a mounted museum or university display, or even poor relations between palaeontological institutions.

This investigation looks at femoral circumference in tyrannosaurid dinosaurs as a case study in the examination of missing data

* Corresponding author. Tel.: +1 7807072077.
  E-mail address: gbradley@ualberta.ca (G.J. Bradley).

sets within palaeontology. Load bearing bones like femora form an important component of the mammalian-like erect posture that allowed dinosaurs to attain great speeds and vast body sizes (Persons & Currie, 2011a; Persons & Currie, 2011b; Benton et al., 2014; Seymour et al., 2011; O'Gornan & Hone, 2012). More specifically, femoral circumferences are important in several subfields of palaeontology. Their proportions compared to other limb bones have been used in descriptive studies, such as that of *Gigantoraptor erlianensis* (Xu et al., 2007), and allometric scaling studies have also made use of femoral dimensions (Christiansen, 1998, 1999; Christiansen and Fariña, 2004; Carrano, 2001; Bybee et al., 2006; Kilbourne & Makovicky, 2010; Funston et al., 2015). Lee et al. (2008) also used femoral circumferences along with histological growth lines to create a life history model for *Hypacrosaurus*.

Missing femoral data are a problem with tyrannosaurids particularly because they are bipedal and thus have a reduced number of bones which can be used for bone loading (e.g. Farlow et al., 1995), locomotion (e.g. Heinrich et al., 1993; Persons & Currie, 2011a; Persons & Currie, 2011b) and body mass studies (e.g. Erickson et al., 2004; Campione & Evans, 2012; Campione et al., 2014; Benson et al., 2014). Tyrannosaurids face an additional problem in that they are also popular mounted exhibits in museums, and thus access to measure femoral circumference will often be impossible.

Herein, we used anteroposterior and mediolateral diameter measurements to predict femoral circumference in tyrannosaurids through six different missing data estimation models, which may be divided crudely into two main groups: generalised geometric equation models and regression based equation models. By comparing the predictive success of the model generated circumferences to a set of true circumferences, and the relative performance of each model by juxtaposing error values, it was hoped to assess whether or not tyrannosaurid femoral diameters may be used to predict circumferences in data sets with missing measurements. If so, this could increase the sample sizes for studies involving femoral circumference in tyrannosaurids and minimise the effects of missing data.

2. Materials and methods

Six models were tested in this study: (1) **CAP**, the circumference of a circle using the diameter of the anteroposterior shaft width (x); (2) **CML**, circumference of a circle using the diameter of the mediolateral shaft width (y); (3) **ELL**, circumference of an ellipse using both anteroposterior and mediolateral diameters; (4) **APR**, linear regression equation for the anteroposterior diameter of the shaft and true circumference; (5) **MLR**, linear regression equation for the mediolateral diameter of the shaft and true circumference; (6) **ELLR**, multiple linear regression equations for both the true anteroposterior and mediolateral diameters of the shaft and true circumference. The equation for the circular models, **CAP and CML** was:

$$C = \pi D$$

Estimating the femoral circumference using **ELL**, the anteroposterior and mediolateral diameters (x and y) were expressed as a radical fraction in the standard equation:

$$C \approx 2\pi \frac{\sqrt{x^2 + y^2}}{2}$$

To create predictive equations for **APR** and **MLR**, true diameters (x axis) were plotted against corresponding true circumferences (y axis) in bivariate linear regression graphs using Microsoft Excel 2007. The linear relationship between the true circumferences and

their corresponding diameters was expressed in the standard regression format: $y = mx + b$, where b is the 'y axis' intercept, m is the slope of the line, and 'y' and 'x' represent the circumference and diameter values respectively (mm). These equations may be used by substituting a known diameter in place of 'x' and calculating a predicted circumference at 'y'. Because regressions of raw data introduced violations of the underlying assumptions of linear regressions (Seber and Lee, 2003), all data were log-transformed before regression. Log-transformation of data has been demonstrated to reduce the influence of outliers in previous palaeontological morphometric regression analyses (Cawley and Janacek, 2010; Campione and Evans, 2012). Because the experimental design was unbalanced, due to the range in sample sizes of each genus (McDonald, 2014), each raw, true diameter and circumference was appropriately weighted using the following formula in order to fulfil the statistical assumption of independence:

$$WD = R \cdot \frac{1}{n}$$

where: WD = Weighted Datum, R = Raw Datum and n = sample size of genus.

Paired two-tailed Student's t-tests were carried out between each model's set of predicted and true circumferences using Microsoft Excel 2007 and repeated using Graph Pad 'Quick Calcs'. Prior to carrying out these t-tests, data from the regression based formulae were required to be back-transformed and weighted back to the scale of the original raw values. The latter was performed using the inverse of the previous equation:

$$R = \frac{WD}{\left(\frac{1}{n}\right)}$$

Percent Prediction Error (PPE) and Standard Error of the Estimate (SEE) values were calculated for each set of predictions using the 'ppe' and 'see' functions in 'MASSTIMATE' (Campione, 2013; Campione et al., 2014): a package within 'R' (R Development Core Team, 2014). Unpaired two-tailed t-tests were carried out between the PPE values of all models, as well as between SEE values. For the latter set of t-tests, five SEE values, one from each genus of tyrannosaurid, were used for each model.

Due to the high number of t-tests carried out on PPE and SEE values, it was necessary to take into account the increased potential for false discoveries, or, the 'False Discovery Rate' (Benjamini and Hochberg, 1995). For this, adjusted p-values, or 'q-values' were returned using the 'p.adjust' function within the 'stats' package in 'R' (R Development Core Team, 2014). These values were then compared to a statistical significance cut-off value (c), calculated at $\alpha = 0.05$, using the formula suggested by Benjamini and Hochberg (1995), to identify false positives from the multiple t-tests.

Measurements were taken with a tape measure or digital callipers. The anteroposterior and mediolateral diameters were measured at the minimum shaft width of the femur, which in tyrannosaurids is distal to the mid-length of the femur. The original sample set consisted of 71 tyrannosaurid femora; however, only 51 had measured circumferences and therefore could be used as comparisons for predictive models (Appendix 1). The taphonomic alteration of each specimen dictated which diameters could be measured and hence, which models could be tested for which specimen. Thus, 45 anteroposterior diameters were available to test the **CAP** and **APR** models, 41 could be tested using **CML** and **MLR**, and only 35 were complete enough to be tested using **ELL** and **ELLR**. The femora represent five tyrannosaurid genera: *Albertosaurus, Daspletosaurus, Gorgosaurus, Tarbosaurus* and *Tyrannosaurus*. One specimen, BMNH 2002.004.001, previously designated

to *Nanoytrannus* was incorporated into the *Tyrannosaurus* data set on the grounds that an extra sample for this genus would be of greater statistical use than an extra sample set, with only one specimen. Carr (1999) has suggested that *Nanotyrannus* specimens represent juvenile forms of *Tyrannosaurus*.

3. Institutional abbreviations

AMNH, American Museum of Natural History, New York, New York, U.S.A.; **BMNH**, Burpee Museum of Natural History, Rockford, Illinois, U.S.A.; **CMMD**, Central Museum of Mongolian Dinosaurs, Ulaanbaatar, Mongolia; **CM**, Children's Museum of Indianapolis, Indiana, U.S.A.; **DMNH**, Denver Museum of Nature and Science, Denver, Colorado, U.S.A.; **FMNH**, Field Museum of Natural History, Chicago, Illinois, U.S.A.; **MOR**, Museum of the Rockies, Bozeman, Montana, U.S.A.; **MPC**, Mongolian Palaeontological Centre, Ulaanbaatar, Mongolia; **NMC**, Canadian Museum of Nature, Ottawa, Ontario, Canada; **NMMNH**, New Mexico Museum of Natural History and Science, Albuquerque, New Mexico, U.S.A.; **PIN**, Museum of Palaeontology, Moscow, Moscow Oblast, Russia; **ROM**, Royal Ontario Museum, Toronto, Ontario, Canada; **RSM**, Royal Saskatchewan Museum, Regina, Saskatchewan, Canada; **TMP**, Royal Tyrrell Museum of Palaeontology, Drumheller, Alberta, Canada; **UALVP**, University of Alberta Laboratory for Vertebrate Palaeontology, Edmonton, Alberta, Canada; **USNM**, National Museum of Natural History, Washington, D.C., U.S.A.; **ZPAL**, Institute of Palaeobiology, Polish Academy of Sciences, Warsaw, Poland.

4. Results

The linear regression formula produced for **APR** was:

$$y = 0.8886x + 0.6591 \text{ (Fig. 1)}$$

The linear regression formula generated for the **MLR** model was:

$$y = 0.959x + 0.5399 \text{ (Fig. 2)}$$

The multiple regressions performed for **ELLR** produced two equations, for use with anteroposterior diameters and mediolateral diameters respectively:

$$y = 0.9302x + 0.6324 \text{ (Fig. 3)}$$

$$y = 0.9572x + 0.5415 \text{ (Fig. 4)}$$

A two tailed student t-test was conducted on each set of results to assess any significant differences between each model's predictions and their corresponding known femoral circumferences (See Appendix 2 for full data). Four of the six models predicted values that were not significantly different from the true circumferences: the three regression models, along with **CML** ($p = 0.20$, t stat $= -1.29$), which had the lowest t-stat value. **ELLR** ($p = 0.79$, t stat $= -0.26$) had the highest p-value, with **MLR** ($p = 0.43$, t stat $= -0.80$) and **APR** ($p = 0.53$, t stat $= 0.63$) also showing no significant difference between predicted and true data. **CAP** and **ELL**, were the only models that produced predicted circumferences that were significantly different from the actual values (**CAP**: $p < 0.01$, t stat $= 5.13$; **ELL**: $p = 0.0002$, t stat $= -4.2556$). These results are summarised in Table 1.

Percent prediction error (PPE) was calculated for each model as a method of comparing the predictive strength of each set of results. Upper and lower confidence intervals of 95% were also produced, along with the standard deviation and range of PPE shown by a given model (Table 1). The models that incorporated the

mediolateral diameter had the lowest mean PPE and the narrowest range of 95% confidence intervals. As seen in Fig. 5, **CML** displayed a mean PPE of 6.55% and 95% confidence intervals of 4.6%–8.51%, whereas the mean PPE of **ELL** was 7.37% with 95% confidence intervals of 5.59%–9.14%. These error values were bettered only by the two mediolateral diameter based regression formulae. **MLR** had a mean PPE of 6.32% and a range of 95% confident intervals: 4.73%–7.88%, whereas **ELLR** had the lowest error rates of all, with a mean PPE of 4.16% and 95% confidence intervals of 2.65–5.67%. The two models, **APR** and **CAP**, which were fully reliant on anteroposterior diameter data, displayed the highest mean PPE and the widest range of 95% confidence intervals. **APR** produced a mean PPE of 11.63%, between 95% confidence intervals of 9.63% and 15.42%, whereas **CAP** performed as poorly as it did with the two tailed t-tests, with a mean PPE of 17.39% and 95% confidence intervals of 12.80%–21.98%. Although confidence intervals are heavily influenced by sample size, the models with the broadest range, **CAP** and **APR**, were also the two models with the largest sample sizes; therefore it is likely that these broader ranges reflect a greater variability of data.

Further to the Percent Prediction Error (PPE), the Standard Error of the Estimate (SEE) was calculated for each model (Table 1). The highest estimate of standard error, and therefore lowest assumed level of accuracy was produced by **CAP** (SEE = 68.13), and the other anteroposterior diameter based model, **APR**, displayed a similarly high value (SEE = 50.59). Again, **ELLR** displayed the least amount of error in its predictions (SEE = 26.52), however, in this test **ELL** had marginally lower levels of standard error (SEE = 33.09) than did **MLR** (SEE = 33.39) and **CML** (SEE = 35.32).

In order to assess the statistical significance of the variation in error across the five predictive models, two-tailed student t-tests were performed, which compared the results of PPE (Table A.1, Appendix 3) and SEE (Table A.2, Appendix 3) between the different models. In conjunction with these tests, false positives in the results were identified using the False Discovery Rate method (Benjamini and Hochberg, 1995). For the t-tests of PPE values, the cut-off (c) at $\alpha = 0.05$ was calculated as: 0.0238; any test which returned a q-value $\leq$ c was considered likely to represent a truly significant result.

ELLR again outperformed all the other models, as the only estimation method to have significantly smaller PPE values than three others: **CAP** ($p = 0.0001$, t stat $= 4.9607$, q $= 0.0004$), **APR** ($p = 0.0001$, t $= 4.1336$, q $= 0.0004$), and **ELL** ($p = 0.0068$, t $= 2.7918$, q $= 0.0126$). **MLR and CML** each had PPE values significantly smaller than two others: **APR** ($p = 0.0042$, t $= 2.9407$, q $= 0.011$; $p = 0.0062$, t-stat $= 2.8090$, q $= 0.013$ respectively) and **CAP** ($p = 0.0001$, t-stat $= 4.3287$, q $= 0.0004$; $p \leq 0.0001$, t-stat $= 4.2358$, q $= 0.0004$ respectively). **CAP** showed significantly higher levels of percent prediction error than all other models, except **APR** ($p = 0.037$, t-stat $= 2.1185$, q $= 0.0555$). Whereas **ELL** initially seemed to have significantly lower PPE levels than **APR**, this was rejected due to its high q-value (0.04), which suggested that the result of the t-test had been a false positive.

The student t-tests of SEE values among predictive models were less divisive, and only one model, when analysed using FDR ($\alpha = 0.05$, c = 0), displayed any sort of significant difference (Table A.2, Appendix 3). **ELLR**, the strongest model thus far, had significantly lower SEE values than **CAP** ($p = 0.0393$, t-stat $= 2.4601$, q $= 0.3702$), but no other combinations could be statistically separated by their SEE values.

Intertaxonomic comparisons were made within each model by calculating the mean PPE and 95% confidence intervals for each genus; these results are summarised in Fig. 6 (for full results see Appendix 2). The models incorporating only anteroposterior diameter data, **CAP** and **APR**, not only displayed higher mean error
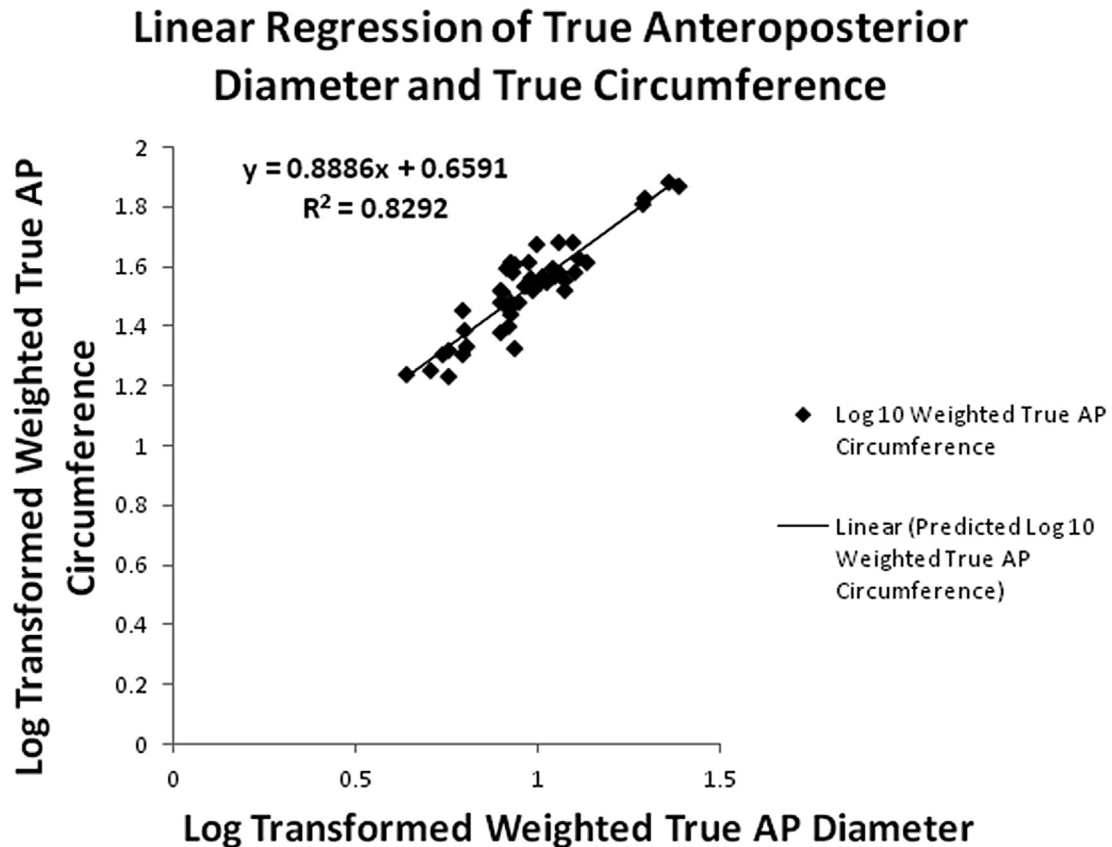
Linear Regression of True Anteroposterior Diameter and True Circumference

Fig. 1. Linear regression of true anteroposterior diameters of tyrannosaurid femoral shafts and their corresponding true circumference (**APR**). All data log transformed (Log10) and weighted.

levels (PPE and SEE), but also a greater variability of error values among the different taxa that made up the sample, despite having a greater overall sample size than the other models. The inter-taxonomic mean values for **CAP** ranged from 30.13% in *Alberto-saurus* to 5.09% in *Daspletosaurus*, and the 95% confidence intervals for *Albertosaurus* in this model do not overlap with the means of *Daspletosaurus* or *Gorgosaurus,* although its sample size was the same as the latter (n = 8) (Fig. 6). This high taxonomic variability in PPE is also witnessed in **APR**, which had a similarly wide range of mean PPE values compared to the mediolateral based models: 21.38% in *Albertosaurus* to 3.50% in *Daspletosaurus*. In this model, the 95% confidence intervals of *Tarbosaurus* do not overlap with the mean PPE of *Albertosaurus*, *Gorgosaurus* or *Daspletosaurus*. Its sample size is relatively similar to the first two genera (n = 10, n = 8, n = 8), but twice that of the latter (n = 5); its relationship with the albertosaurines is therefore more likely to represent true intertaxonomic error variability. The intervals of the two alberto-saurines do not overlap with the mean PPE of *Daspletosaurus* or *Tarbosaurus*; those of *Albertosaurus,* in fact, do not overlap with the means of any other taxon (Fig. 6). Although intertaxonomic varia-tion was strong, no individual genus had a lower mean PPE when using **CAP** or **APR**, as when using the four prediction models that incorporated mediolateral diameters, with the exception of *Das-pletosaurus*, which in the mediolateral models had a negligible sample size (n = 2). In both anteroposterior diameter models, *Albertosaurus* had the highest mean PPE and widest range of 95% confidence intervals.

Daspletosaurus expanded the 95% confidence interval range for PPE in **ELLR**, **CML**, **ELL** and **MLR**, due to its small sample size (n = 2). Whereas these two data points were correctly included in the

original prediction equations, in which each raw value was considered to be independent of all others, statistical results based on two values are representative only of the uncertainty caused by the impoverished sample. Removing *Daspletosaurus* from the sample for mediolateral diameter based models, the upper and lower 95% confidence intervals for **ELL** are restricted by 27.87% each, for **CML** by 60.48% (upper) and 45.45% (lower), for **MLR** by 148.28% (upper) and 88.29% (lower), and for **ELLR** by 38.2% each (Fig. 6). With these models, the ranges of mean PPE values among taxa were smaller than those produced by **CAP** and **APR**; **CML** (Range: 8.07% in *Tarbosaurus* to 5.12% in *Gorgosaurus*); **ELL** (Range: 8.92% in *Tyrannosaurus* to 3.98% in *Gorgosaurus*); **MLR** (Range: 7.52% in *Tarbosaurus* to 3.65% in *Gorgosaurus*); **ELLR** (Range: 5.01% in *Tyrannosaurus* to 2.14 in *Gorgosaurus*). No one genus dominated the highest mean error values across the four mediolateral diam-eter inclusive models, as *Albertosaurus* did with **CAP** and **APR**, however *Gorgosaurus* produced the lowest mean PPE values in each of these four models.

5. Discussion

There have been a number of different techniques employed within palaeontology to reduce the effect of missing data on ana-lyses and bolster sample sizes; each time in a manner appropriate and specific to the type of investigation being carried out. Grillo and Avezdo (2011) used regression equations to estimate the position of missing tail vertebrae in a specimen of *Staurikiosaurus* on the basis of measurements from the most anterior and posterior caudal vertebrae. Similarly, Funston et al. (2015) estimated the femoral dimensions of a large caenagnathid specimen from a single caudal
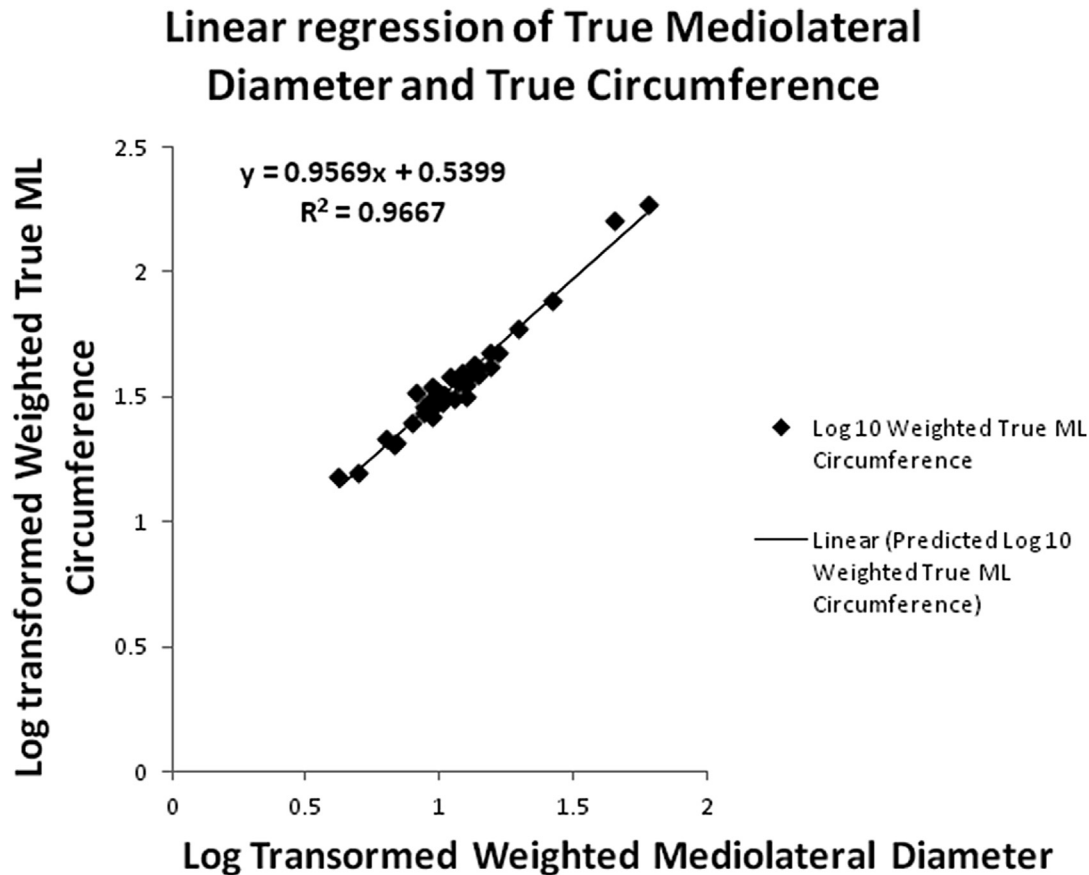
Linear regression of True Mediolateral Diameter and True Circumference

Fig. 2. Linear regression of true mediolateral diameters of tyrannosaurid femoral shafts and their corresponding true circumferences **(MLR)**. All data log transformed (Log10) and weighted.

vertebra, based on allometric equations between caudal vertebrae and femora for this group of theropods. On a histological scale, Cooper et al. (2008) used an elliptical equation (Ramanujan, 1914) with minor and major radii to find missing circumferences of concentric lines of arrested growth within femora of *Hypacrosaurus*. Through a novel study in missing data prevention, Arbour and Currie (2012) attempted to better understand the effects of taphonomic deformation in the skulls of ankylosaurids, using 3D retrodeformation techniques on C-T scan images, therefore minimising the confusion of morphological features for missing data points created by geological pressures.

Missing data estimation models have also been proposed for dealing with unrecordable values in larger, multivariate data sets. Norell and Wheeler (2003) dealt with the problem of missing data in phylogenetic analyses of fossils by introducing missing entry replacement data analysis (MERDA), in which they generate a host of possible phylogenetic outcomes and identify those which rely heaviest on missing data points. Brown et al. (2012) dealt with similarly large data sets as they tested the relative success of numerous missing data estimation techniques in large, multivariate, morphometric analyses. Measurements from the skull of an extant crocodilian were taken and analysed with PCA, before missing data values were introduced to the data set. Various missing data estimation models were employed and compared, with the complex Bayesian Principle Component Analysis missing value estimator (Strauss et al., 2003; Oba et al., 2003) producing the lowest error values in its estimations.

From the results of our study, there appear to be a number of viable models through which femoral diameter may be used to estimate missing femoral circumference values in tyrannosaurids, and that may be considered alongside previously proposed missing data estimation models. The two models based solely on the anteroposterior diameter of the femur, **CAP** and **APR**, consistently performed the poorest in the statistical analyses. This suggests that the relationship between anteroposterior diameter and circumference alone in tyrannosaurid femora is not strong enough to support missing data estimation models, and we advise against the use of **CAP** and **APR** for these taxa.

The models that incorporated mediolateral diameter data: **CML**, **ELL**, **ELLR** and **MLR**, performed significantly better in the statistical analyses than **CAP** and **APR**. Of these four methods, estimations using **ELLR**, performed marginally better than the rest. However, it could not be statistically separated on the basis of PPE and SEE from **MLR** or **CML**. Whereas **ELL** had fairly low PPE values, which were significantly better than **CAP**, it produced predictions that were significantly different from the true values. This may suggest that it had a relatively higher variance of predictions than other mediolateral models, or possibly that its statistical confidence was undermined by a lower sample size than was available for **CML** and **MLR**.

From these statistical analyses, it can be inferred that femoral circumference in tyrannosaurids is more closely related to its mediolateral rather than anteroposterior diameter, and **ELLR**, **CML**, and **MLR**, may be recommended as three acceptable missing data estimation methods for femoral circumferences across these taxa. **ELLR**, as the strongest statistical performer emerged as the best predictor of femoral circumference and should be favoured over the other five, although **MLR** and **CML** will also give low error

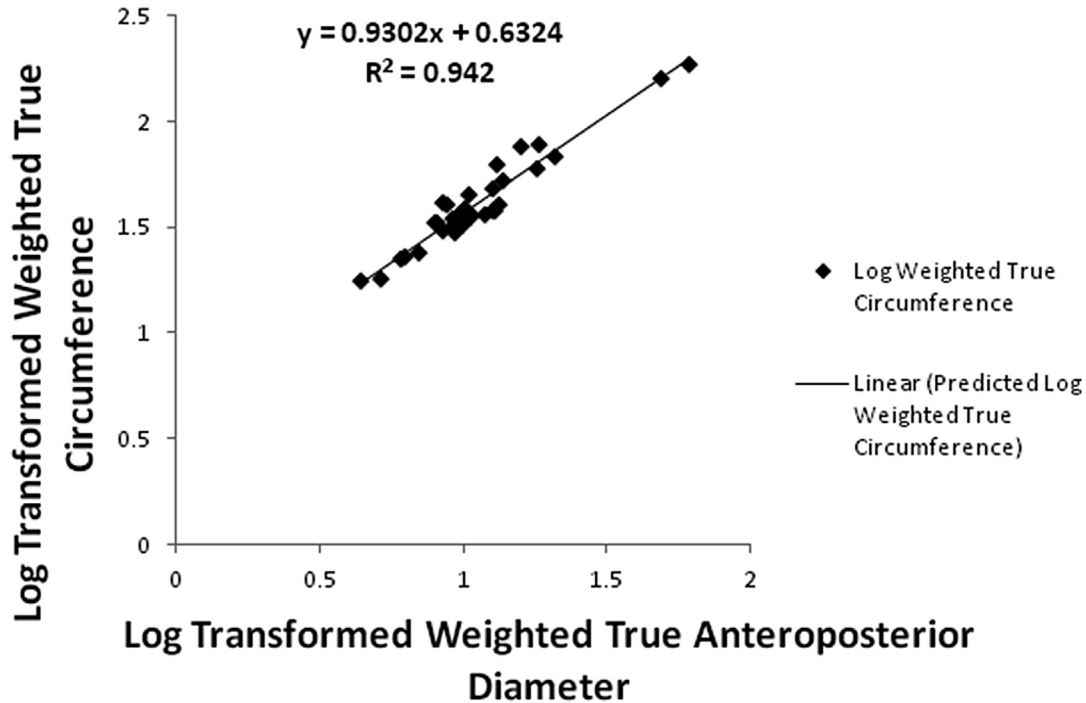## Multiple Linear regression of True Diameters vs True Circumference (Anteroposterior)

Fig. 3. Multiple Linear regression of true femoral diameters vs true femoral circumferences (**ELLR**-Anteroposterior). All data log transformed (Log10) and weighted.

predictive values, and are based off of larger data sets. Although **ELL** also displayed significantly lower error values than **CAP** in its predictions, this model requires two femoral measurements, anteroposterior and mediolateral diameter, rather than one, and therefore its effectiveness in increasing data sets will be diminished in comparison to **ELLR**, **MLR** and **CML**.

In order to demonstrate the utility of these models, two data sets were examined to see how much missing data the models could potentially recover: our own tyrannosaurid data set (Appendix 1) and Benson et al.'s (2014, Dataset S.1) data set for body mass estimations. Out of 71 tyrannosaurid specimens available in our original data set, confident femoral circumference measurements against which to test the five predictive models were only possible for 51, most likely due to taphonomic distortion. Using **ELLR**, **MLR** and **CML**, the most statistically successful models, the available sample size could be boosted by 35% (17 specimens) to 68 specimens.

There are only 11 specimens of the five tyrannosaurid genera in Benson et al.'s (2014, Dataset S.1) set, however **ELLR**, **MLR** and **CML** would still boost the sample by 9%. If the models, after additional testing, were extended to all theropod taxa within the same dataset, the available sample size would be increased by 62% (70 specimens). This result, although purely hypothetical, because **ELLR**, **MLR** and **CML** are thus far restricted to use on tyrannosaurid femoral data sets, highlights the utility of these types of missing data estimation models if they continue to be studied.

If adopted, the **ELLR**, **CML** and **MLR** models will increase the sample size for body mass estimation of tyrannosaurids, which is

one of the most common applications for femoral circumferences in palaeontology (Anderson et al., 1985; Christiansen & Farina, 2004; Campione & Evans, 2012; Campione, 2013; Campione et al., 2014). In order to illustrate one potential application of the **ELLR**, **CML** and **MLR** models, body mass estimations using the true femoral circumference of the largest adult specimen from each of the five genera were made using Campione et al.'s (2014) bipedal correction of the quadrupedal equation presented in Campione and Evans (2012) (Table 2). These were then compared to body mass estimates for the same specimens based on circumferences predicted using **ELLR**, **CML** and **MLR**, and also calculated with Campione et al.'s (2014) equation.

For each tyrannosaurid genus, the body mass values returned using **ELLR**, **CML** and **MLR** circumference estimations fell within the upper and lower boundaries of body mass values produced using true circumferences. In one example, *Tyrannosaurus*, the difference between the true circumference upper and lower body mass limit and the corresponding **ELLR** circumference body mass limits was only 32 and 19.6 kg respectively (True body mass:10592.9—6351.6; **ELLR** body mass:10560.2—6332.1 kg).

Aside from their utility in body mass estimations, which can give us insight into physiological, ecological and evolutionary questions (Farlow et al., 1995; Gillooley et al., 2001, 2006; Benson et al., 2014), femoral circumferences in tyrannosaurids have a wide range of applications from life history studies (Cooper et al., 2008; Woodward et al., 2011), to allometric scaling studies (Carrano, 2001; Bybee et al., 2006; Kilbourne & Makovicky, 2010). Some of these studies have even proven useful in examples of
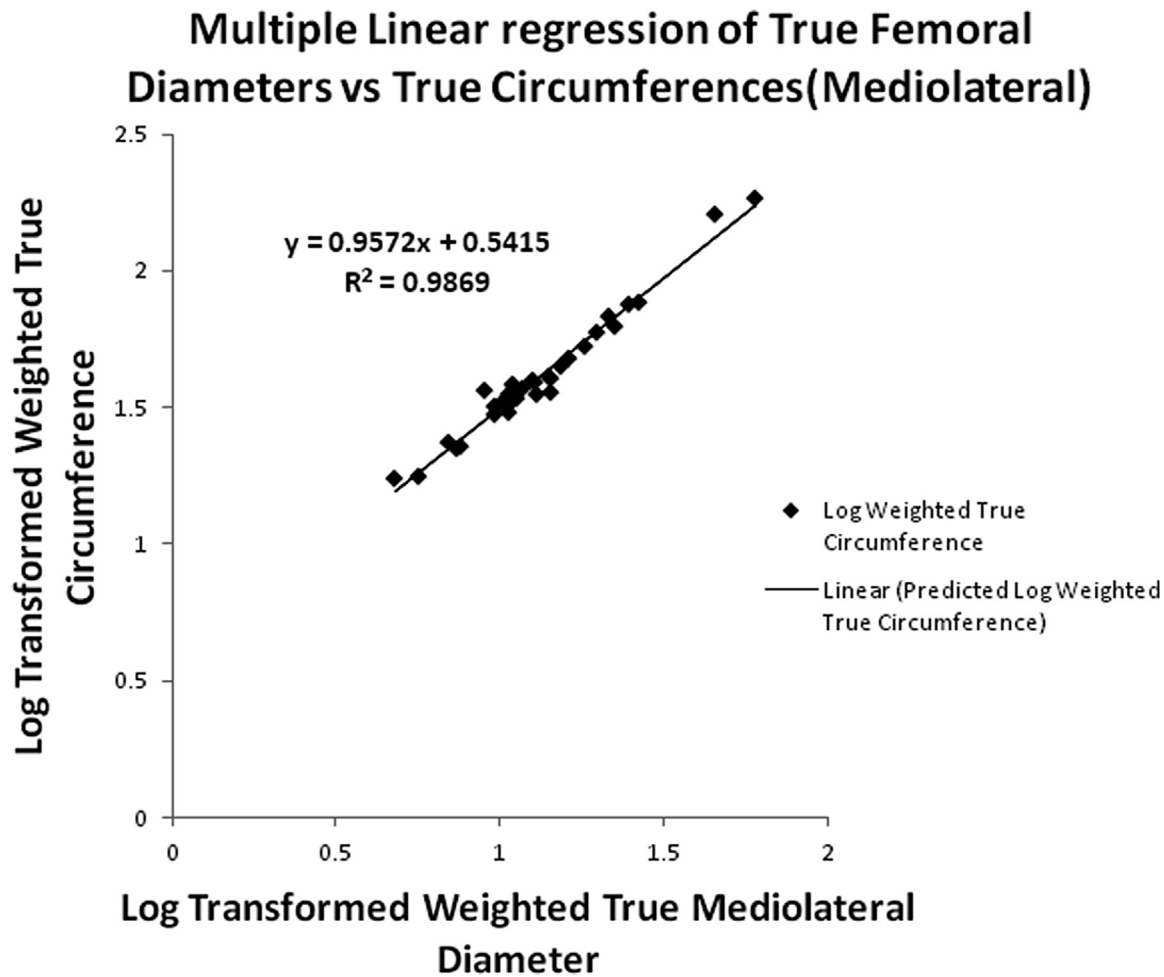
Multiple Linear regression of True Femoral Diameters vs True Circumferences(Mediolateral)

$$y = 0.9572x + 0.5415$$
$$R^2 = 0.9869$$

- ◆ Log Weighted True Circumference
- — Linear (Predicted Log Weighted True Circumference)

Fig. 4. Multiple Linear regression of true femoral diameters vs true femoral circumferences (**ELLR**-Mediolateral). All data log transformed (Log10) and weighted.

taxonomic disputes concerning tyrannosaurids such as *Gorgosaurus* and *Nanotyrannus* (Carr, 1999; Currie, 2003).

Despite the variety of their potential applications and predictive success in this set of statistical analyses, **ELLR, CML** and **MLR**, should be viewed as alternatives, rather than replacements for any of the missing data estimates that currently exist. Due to the restriction of the regression models within tyrannosaurid femora, they cannot be seen as comparable to large data set missing data recovery models such as BCPA (Brown et al., 2012) or MERDA (Norell et al., 2003), which are designed for use with multivariate morphometric and phylogenetic analyses respectively. A benefit of the currently narrow scope of predictions for which **ELLR, MLR** and

CML are appropriate is that they lack the digital complexity of models such as BCPA and MERDA, and can be calculated by simply substituting values into one of 4 simple equations. Another benefit is the minimal size and number of measurements required to produce a predicted value; a mediolateral radius from the minimum shaft width would suffice. Radii and diameters are often measurable even in taphonomically damaged fossils, and are usually available to access even on mounted specimens where full circumferences are not.

The strength of the regression based formulae, particularly **MLR** and **ELLR** is unsurprising given that they are based off true tyrannosaurid femoral data. However, that simple geometric models

Table 1
Summary statistics for model predictions vs. true femoral circumferences for five tyrannosaurid genera; two sample Student's t-test of model predicted vs. true femoral circumferences, percent prediction error (PPE) and standard error of the estimate (SEE).

| Model	CAP	CML	ELL	APR	MLR	ELLR
p-value (Predicted vs. True)	<0.01	0.20	0.0002	0.53	0.43	0.79
t-stat (Predicted vs. True)	5.13	−1.29	4.26	0.63	−0.80	−0.26
Mean PPE	17.39	6.55	7.37	11.63	6.32	4.16
Standard Deviation of Mean PPE	15.27	6.2	5.16	9.64	4.98	4.40
PPE Range	2.23–53.85	0.32–28.10	0.19–21.14	0.08–37.18	0.05–23.85	0.03–18.21
Confidence Intervals of PPE	Lower95: 12.80	Lower95: 4.6	Lower95: 5.59	Lower95: 9.63	Lower95: 4.73	Lower95: 2.65
	Upper95: 21.98	Upper95: 8.51	Upper95: 9.14	Upper95: 15.42	Upper95: 7.88	Upper95: 5.67
Sample Size	45	41	35	45	41	35
SEE	68.13	35.32	33.09	50.59	33.39	26.52
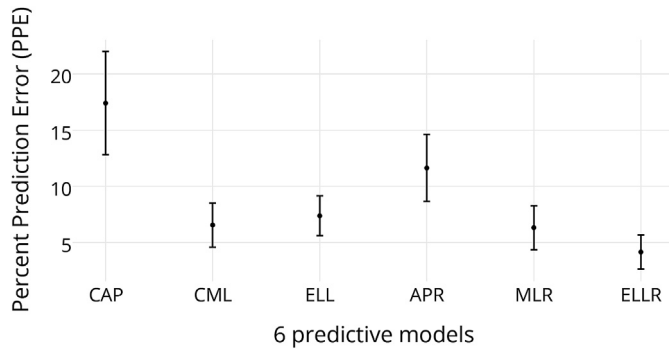
Fig. 5. Mean PPE values for six femoral circumference estimation models' predictions using data from five tyrannosaurid genera. Bars represent upper and lower 95% confidence boundaries for each model.

such as **CML** and **ELL** performed well under statistical examination bodes well for future projects. Unlike the regression models, these simple equations are not necessarily limited to tyrannosaurid femoral circumferences, and may hypothetically be used with any taxa. Although regression techniques may also be used with other

animals, they would need to be created anew using data sets from each new taxon, as was done in this study, whereas the simple equations are theoretically universal, and may be expanded with increased statistical testing across various taxa and limb bones.

6. Conclusions

The six predictive formulae, with varying levels of statistical success, display how diameter might be used in place of circumference in tyrannosaurid research, to fill voids left by missing data. Statistical testing suggests that three predictive models that incorporated mediolateral diameter measurements (**CML**, **MLR** and **ELLR**) are appropriate methods of estimating femoral circumference in five tyrannosaurid genera. Methods based exclusively on anteroposterior diameters, such as **CAP** and **APR**, showed significant levels of prediction error and should not be used for these taxa. **ELLR** performed best under statistical examination and can be used with both anteroposterior and mediolateral diameters, however, its predictions were not significantly more accurate than those of **CML** or **MLR**. Although **ELLR** and **MLR** provided the most accurate predictions in this study, they are restricted to tyrannosaurids; **CML**, **CAP** and **ELL**, as general geometric equations may theoretically be
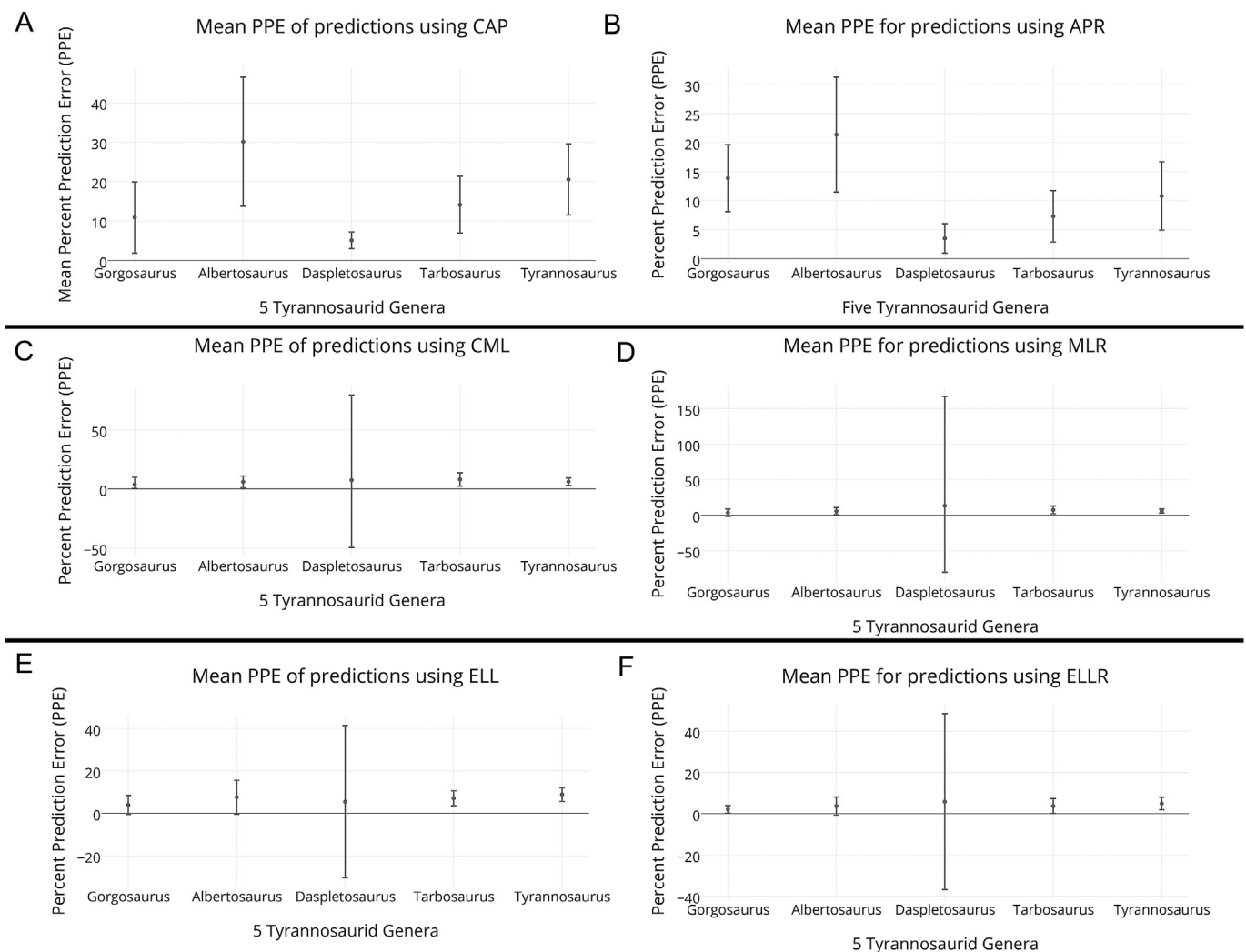

Fig. 6. Intertaxonomic mean PPE values for six femoral circumference estimation models' predictions using data from five tyrannosaurid genera. Bars represent upper and lower 95% confidence boundaries for each genus of tyrannosaurid in each particular model.

Table 2
Comparison of body mass estimations for five tyrannosaurid genera made with true femoral circumferences (Cf.), **ELLR**, **CML** and **MLR** estimated circumferences.

Genus and specimen	*Gorgosaurus* NMC 350	*Albertosaurus* TMP 1982.13.30	*Tarbosaurus* MPC-D107/02	*Daspletosaurus* AMNH 5434	*Tyrannosaurus* MOR 1128 "G-rex"
True Cf. (mm)	385	380	480	370	580
CML Cf. (mm)	414.69	388.30	518.36	376.99	584.34
MLR Cf. (mm)	400.10	382.97	509.89	351.70	583.09
ELLR Cf. (mm)	399.24	374.88	507.11	350.41	579.35
Body Mass (Kg)	3426.8–2054.8	3305.6–1982.1	6290.3–3771.8	3071.6–1841.8	10592.9–6351.6
Body Mass (Kg) **CML**	4204.8–2521.3	3508.3–2103.6	7773.7–4661.2	3234–1939.2	10812.6–6483.4
Body Mass (Kg) **MLR**	3809.8–2284.4	3377.3–2025.1	7428.9–4454.5	2761.1–1601.6	10749–6445.3
Body Mass (Kg) **ELLR**	3787.3–2270.9	3184.4–1909.4	7317.9–4387.9	2644.2–1585.5	10560.2–6332.1

used with other taxa, although their predictive power for each new taxon should be tested using methodology similar to that of this study.

Acknowledgements

References

Anderson, J.F., Hall-Martin, A., Russell, D.A., 1985. Long-bone circumference and weight in mammals, birds and dinosaurs. Journal of Zoological Society of London A 207, 53–61.

Arbour, V.M., Currie, P.J., 2012. Analyzing taphonomic deformation of ankylosaur skulls using retrodeformation and finite element analysis. PLoS One 7 (6), e39323.

Benson, R.B., Campione, N.E., Carrano, M.T., Mannion, P.D., Sullivan, C., Upchurch, P., Evans, D.C., 2014. Rates of dinosaur body mass evolution indicate 170 million years of sustained ecological innovation on the avian stem lineage. PLoS Biology 12 (5), e1001853.

Benton, M.J., Forth, J., Langer, M.C., 2014. Models for the Rise of the Dinosaurs. Current Biology 24 (2), R87–R95.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289–300.

Brown, C.M., Arbour, J.H., Jackson, D.A., 2012. Testing the Effect of Missing Data Estimation and Distribution in Morphometric Multivariate Data Analyses. Systematic Biology 61, 941–954.

Brusatte, S.L., Hone, D.W.E., Xu, X., 2013. Phylogenetic revision of Chingkankousaurus fragilis, a forgotten tyrannosauroid specimen from the Late Cretaceous of China. In: Parish, M., Molnar, R.E., Currie, P.J., Koppelhus, E.B. (Eds.), Tyrannosaurid Palaeobiology. Indiana University Press, Bloomington, Indiana, pp. 2–13.

Bybee, P.J., Lee, A.H., Lamm, E.T., 2006. Sizing the Jurassic theropod dinosaur Allosaurus: assessing growth strategy and evolution of ontogenetic scaling of limbs. Journal of Morphology 267 (3), 347–359.

Campione, N.E., Evans, D.C., 2011. Cranial growth and variation in edmontosaurs (Dinosauria: Hadrosauridae): implications for latest Cretaceous megaherbivore diversity in North America. PLoS One 6 (9), e25186.

Campione, N.E., Evans, D.C., 2012. A universal scaling relationship between body mass and proximal limb bone dimensions in quadrupedal terrestrial tetrapods. Bmc Biology 10 (1), 60.

Campione, N.E., 2013. MASSTIMATE: Body Mass Estimation Equations for Vertebrates. R Project for Statistical Computing. Comprehensive R Archive Network (CRAN).

Campione, N.E., Evans, D.C., Brown, C.M., Carrano, M.T., 2014. A mathematically derived equation for estimating body mass in terrestrial bipedal tetrapods. Methods in Ecology and Evolution 5 (9), 913–923.

Carr, T.D., 1999. Craniofacial ontogeny in tyrannosauridae (Dinosauria, Coelurosauria). Journal of Vertebrate Paleontology 19 (3), 497–520.

Carrano, M.T., 2001. Implications of limb bone scaling, curvature and eccentricity in mammals and non-avian dinosaurs. Journal of Zoology 254 (1), 41–55.

Cawley, G.C., Janacek, G.J., 2010. On allometric equations for predicting body mass of dinosaurs. Journal of Zoology 280 (4), 355–361.

Chapman, R.E., 1990. Shape analysis in the study of dinosaur morphology. Dinosaur systematics: approaches and perspectives, pp. 21–42.

Christiansen, P., 1998. Strength indicator values of theropod long bones, with comments on limb proportions and cursorial potential. Gaia 15, 241–255.

Christiansen, P., 1999. Long bone scaling and limb posture on non-avian theropods: evidence for differential allometry. Journal of Vertebrate Paleontology 19, 666–680.

Christiansen, P., Fariña, R.A., 2004. Mass prediction in theropod dinosaurs. Historical Biology 16, 85–92.

Cooper, L.N., Lee, A.H., Taper, M.L., Horner, J.R., 2008. Relative growth rates of predator and prey dinosaurs reflect effects of predation. Proceedings of the Royal Society B: Biological Sciences 275 (1651), 2609–2615.

Currie, P.J., 2003. Allometric growth in tyrannosaurids (Dinosauria: Theropoda) from the Upper Cretaceous of North America and Asia. Canadian Journal of Earth Sciences 40 (4), 651–665.

Dilkes, D.W., 2001. An ontogenetic perspective on locomotion in the Late Cretaceous dinosaur Maiasaura peeblesorum (Ornithischia: Hadrosauridae). Canadian Journal of Earth Sciences 38 (8), 1205–1227.

Erickson, G.M., Makovicky, P.J., Currie, P.J., Norell, M.A., Yerby, S.A., Brochu, C.A., 2004. Gigantism and comparative life-history parameters of tyrannosaurid dinosaurs. Nature 430 (7001), 772–775.

Farlow, J.O., Smith, M.B., Robinson, J.M., 1995. Body mass, bone "strength indicator," and cursorial potential of Tyrannosaurus rex. Journal of Vertebrate Paleontology 15, 713–725.

Funston, G.F., Persons, W.S., Bradley, G.J., Currie, P.J., 2015. New material of the large-bodied caenagnathid Caenagnathus collinsi from the Dinosaur Park Formation of Alberta, Canada. Cretaceous Research 54, 179–187.

Gillooly, J.F., Brown, J.H., West, G.B., Savage, V.M., Charnov, E.L., 2001. Effects of size and temperature on metabolic rate. Science 293 (5538), 2248–2251.

Gillooly, J.F., Allen, A.P., Charnov, E.L., 2006. Dinosaur fossils predict body temperatures. PLoS Biology 4 (8), e248.

Grillo, O.N., Azevedo, S.A., 2011. Recovering missing data: estimating position and size of caudal vertebrae in Staurikosaurus pricei Colbert, 1970. Anais da Academia Brasileira de Ciências 83 (1), 61–72.

Hedrick, B.P., Dodson, P., 2013. Lujiatun Psittacosaurids: Understanding Individual and Taphonomic Variation Using 3D Geometric Morphometrics. PLoS One 8 (8), e69265.

Heinrich, R.E., Ruff, C.B., Weishampel, D.B., 1993. Femoral ontogeny and locomotor biomechanics of Dryosaurus lettowvorbecki (Dinosauria, Iguanodontia). Zoological Journal of the Linnean Society 108 (2), 179–196.

Kearney, M., Clark, J.M., 2003. Problems due to missing data in phylogenetic analyses including fossils: a critical review. Journal of Vertebrate Paleontology 23 (2), 263–274.

Kilbourne, B.M., Makovicky, P.J., 2010. Limb bone allometry during postnatal ontogeny in non-avian dinosaurs. Journal of Anatomy 217 (2), 135–152.

Lee, Y.N., Barsbold, R., Currie, P.J., Kobayashi, Y., Lee, H.J., Godefroit, P., Chinzorig, T., 2014. Resolving the long-standing enigmas of a giant ornithomimosaur Deinocheirus mirificus. Nature 13874.

Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Systematic Biology 58 (1), 130–145.

Lipps, J.H., 2009. The protection and conservation of paleontologic field sites worldwide. In: Portland GSA Annual Meeting.

Miyashita, T., Arbour, V.M., Witmer, L.M., Currie, P.J., 2011. The internal cranial morphology of an armoured dinosaur Euoplocephalus corroborated by X-ray computed tomographic reconstruction. Journal of Anatomy 219 (6), 661–675.

McDonald, J.H., 2014. Handbook of Biological Statistics, 3rd ed. Sparky House Publishing, Baltimore.

Norell, M.A., Wheeler, W.C., 2003. Missing entry replacement data analysis: a replacement approach to dealing with missing data in paleontological and total evidence data sets. Journal of Vertebrate Paleontology 23 (2), 275–283.

Oba, S., Sato, M.A., Takemasa, I., Monden, M., Matsubara, K.I., Ishii, S., 2003. A Bayesian missing value estimation method for gene expression profile data. Bioinformatics 19 (16), 2088–2096.

O'Gorman, E.J., Hone, D.W., 2012. Body size distribution of the dinosaurs. PLoS One 7 (12), e51925.

Persons, W.S., Currie, P.J., 2011a. The Tail of Tyrannosaurus: Reassessing the Size and Locomotive Importance of the M. caudofemoralis in Non-Avian Theropods. The Anatomical Record 294 (1), 119–131.

Persons, W.S., Currie, P.J., 2011b. Dinosaur speed demon: the caudal musculature of Carnotaurus sastrei and implications for the evolution of South American abelisaurids. PLoS One 6 (10), e25763.

Ramanujan, S., 1914. Modular equations and approximations to π. Quarterly Journal of Mathematics 45, 350–372.

R Development Core Team, 2014. R: a language and environment for statistical computing, 3.1.2 ed. R Foundation for Statistical Computing, Vienna, Austria.

Seber, G.A., Lee, A.J., 2003. Linear regression analysis, 2nd Edition. John Wiley & Sons, Hoboken.

Seymour, R.S., Smith, S.L., White, C.R., Henderson, D.M., Schwarz-Wings, D., 2011. Blood flow to long bones indicates activity metabolism in mammals, reptiles and dinosaurs. Proceedings of the Royal Society B: Biological Sciences rspb20110968.

Strauss, R.E., Atanassov, M.N., De Oliveira, J.A., 2003. Evaluation of the principal-component and expectation-maximization methods for estimating missing data in morphometric studies. Journal of Vertebrate Paleontology 23 (2), 284–296.

Tschopp, E., Russo, J., Dzemski, G., 2013. Retrodeformation as a test for the validity of phylogenetic characters: an example from diplodocid sauropod vertebrae. Palaeontologica Electronica 16, 1–23.

Wiens, J.J., 2003. Incomplete taxa, incomplete characters, and phylogenetic accuracy: is there a missing data problem? Journal of Vertebrate Paleontology 23 (2), 297–310.

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39 (1), 34–42.

Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Systematic Biology 60 syr025.

Wilkinson, M., 2003. Missing entries and multiple trees: instability, relationships, and support in parsimony analysis. Journal of Vertebrate Paleontology 23 (2), 311–323.

Woodward, H.N., Rich, T.H., Chinsamy, A., Vickers-Rich, P., 2011. Growth dynamics of Australia's polar dinosaurs. PLoS One 6 (8), e23339.

Xu, X., Tan, Q., Wang, J., Zhao, X., Tan, L., 2007. A gigantic bird-like dinosaur from the Late Cretaceous of China. Nature 447 (7146), 844–847.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cretres.2015.02.010.