Intra-observer error of mouse long bone cross section digitization

Alena JINDROVÁ[1], Jan TŮMA[1,2] and Vladimír SLÁDEK[1,3]

[1] Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, 128 43 Prague 2, Czech Republic; e-mail: aja.jindrova@seznam.cz

[2] Department of Pathophysiology, Faculty of Medicine in Pilsen, Charles University, Lidická 1, 301 66 Plzeň, Czech Republic

[3] Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, v.v.i., Květná 8, 603 65 Brno, Czech Republic

Abstract. Digitization of periosteal and endosteal contours of bone cross section is one of several methods used for calculating long bone cross-sectional geometry (CSG). In this study, invasively obtained bone histological samples were used to calculate intra-observer measurement error for CSG parameters. Intra-observer error was measured based upon repeated measurements of cross-sectional areas (total area [TA], cortical area [CA]) and moments of area (I_{max}, I_{min}, I_{max}/I_{min}, J) in B6CBA mice ($n = 17$). Cross sections were cut at 50 % of the biomechanical length of the left tibia and the samples were further processed for CSG and histological analysis. Intra-observer error was measured to estimate the accuracy of the digitization method. Accuracy of the tested digitization method was expressed by mean difference (MD), mean absolute difference (MAD), and limits of agreement (LA). The results confirm our assumption that intra-observer error decreases with the number of repeated measurement events. Thus, the error can be minimized by acquiring experience in the section digitizing. Our results also show that TA, CA and polar moment of area (J) are more susceptible to intra-observer error than are I_{max}, I_{min} and I_{max}/I_{min}.

Key words: accuracy, cortical bone, CSG, digitization, tibia

Introduction

The biomechanical approach to cross-sectional geometry (CSG) analysis has a long tradition in biological research when studying an environmental impact (i.e., external mechanical stimuli) on bone tissues (for review, see Ruff 2007). The biomechanical approach is to analyse long bone cross sections with respect to engineering beam theory (Rybicki et al. 1972). Biomechanical parameters, such as areas and moments of area, are calculated from distribution and geometry of cortical tissue defined by periosteal and endosteal boundaries (Fig. 1).

It has been shown that the reliability of long bone CSG estimates used in biomechanical analysis is dependent not only on the technique used in taking the cross section but also upon the accurate determination of periosteal and endosteal contours in later processing (Stock 2002), when the periosteal contour of a bone is the most relevant to bone mechanical competence and the endosteal contour has less influence (Stock & Shaw 2007).

In deciding how to obtain the raw cross sections from long bone shaft, researchers may choose among non-invasive and invasive groups of methods (Ruff 2007; see summary of techniques in Table 1). Non-invasive methods estimate cross sections using external measurement (Pfeiffer 1980), moulding (Trinkaus & Ruff 1989, Trinkaus et al. 1999), Xray (O'Neill & Ruff 2004), computed tomography [CT] (Ruff et al. 1993, Ruff et al. 1994, Lieberman et al. 2004, Sládek et al. 2006a, b, 2007), micro-computed tomography [μCT] (Fritton et al. 2005), or peripheral quantitative

computed tomography [pQCT] (Sone et al. 2006). If available, CT is the foremost method of choice (Sumner et al. 1985), because the method is very rapid and produces accurate two-dimensional images of inner and outer bone contours. Standard CT cannot be used, however, when analysing bones of small size (e.g., mouse bone). Therefore, μCT is often used for evaluating bone morphology and microarchitecture *ex vivo* in mice and other small animal models (Bouxsein et al. 2010). pQCT uses a small, high-resolution CT machine that has builtin software for calculating section properties as well as bone density (Ferretti et al. 1996). There are alternative methods of CT technique, the most common using bi-planar radiography (Trinkaus & Ruff 1989), which can be used alone (the ellipse model method [EMM]; Stock 2002) or in combination with direct measurement of external bone contours (the latex cast method [LCM]; Trinkaus & Ruff 1989). LCM shows generally good correspondence to the true cross-sectional properties while EMM generally overestimates the true parameters (O'Neill & Ruff 2004). Previous research suggests that cross-sectional contours derived from LCM are within 5 % of those of the original bones (Trinkaus & Ruff 1989, Stock 2002).

Invasive methods visualizing periosteal and endosteal contours directly on natural or artificial cuts of the diaphyseal shaft include direct cutting of a section (Ruff & Hayes 1983), histological processing (Diab & Vashishth 2007), or using broken bone section (Ruff 2009). Cross sections can be either manually transformed into the image or scanned in using an optical scanner (Stock 2002). The choice between invasive and non-invasive techniques depends on the purpose of the given analysis. One of the main disadvantages of CT technique is an absence of histological information for cortical tissues between the periosteal and endosteal boundaries. Several techniques have been developed in recent years to obtain nondecalcified histological sections whereby both cross-sectional geometry as well as histology can be studied (Ciani et al. 2009).

What remains similar between invasive and non-invasive methods is how the raw periosteal and endosteal contour is later processed. In each case, the contour must be transferred in order to be useful for calculating cross-sectional properties. Several software applications have been developed for deriving cross-sectional geometric properties either from section images directly (Ruff 2007) or from digitized pictures [CT-i] (Sailer et al. 2003), [MomentMacro for ImageJ] (Ruff 2006). These all work according to

basic engineering principles (for a description of the algorithms, see Nagurka & Hayes 1980). The most reliable transfer of periosteal and endosteal contour is direct digitization by setting points on the periosteal and endosteal boundaries. Experience shows that it is difficult to produce reliable estimates using an automated digitization approach (i.e., whereby contours are estimated automatically using pixel values for boundaries between 'air' and 'bone' that are defined a priori). The raw cross sections are usually accompanied by artefacts introduced by technique as well as preservation. Since cross-sectional properties are calculated as 'ideal' geometry of cortical tissue between periosteum and endosteum, the artefacts are not of interest per se. The points must therefore be carefully located manually on the cross-sectional boundary, with the researcher deciding whether is artefact unimportant for estimated values, what is the effect of preservation, where is the natural boundary of the cross section, and how to set up the points to best reflect the natural geometry of the cross section. It is clear that part of variability differences in later biomechanical analysis of CSG properties is due to intra- and inter-observer error introduced by digitizing. The goal of this paper is to determine intra-observer measurement error in calculating CSG parameters based on the digitizing of periosteal and endosteal
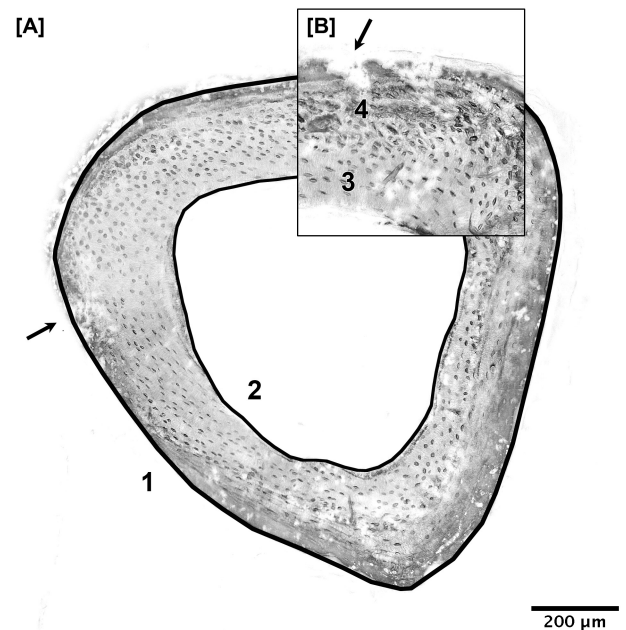

Fig. 1. *The left tibia cross-section at 50 % biomechanical length [A] with the detail of cortical bone [B]. The periosteal (1) and endosteal (2) digitalized margins are distinguished by black lines. The artefacts (3, 4) are pointed out in the detailed picture.*

341

Table 1. Summary of techniques used for cross-sectional geometry analysis.

Methods	Periosteum	Endosteum	References
Non-invasive			
External diameters[1]	non-directly estimated	non-directly estimated	Pfeiffer 1980
Bi-planar X-ray EEM	non-directly estimated	non-directly estimated	O'Neill & Ruff 2004
Periosteal mould[2]	direct (mould)	not calculated	Trinkaus & Ruff 1989
Bi-planar X-ray LCM	direct (mould)	non-directly estimated	Ruff & Hayes 1983, Stock 2002
CT, μCT, pQCT	direct digitizing	direct digitizing	Sumner et al. 1985, Ruff & Leo 1986, Ruff et al. 1993, 1994, Fritton et al. 2005, Sládek et al. 2006a, b, Sone et al. 2006, Sládek et al. 2007, Sabsovich et al. 2008, Bouxsein et al. 2010, Rittweger et al. 2010
Invasive			
Histology processing	direct digitizing	direct digitizing	Sheng et al. 1999, Diab & Vashishth 2007
Direct sectioning (DSM[5])	direct (natural contour)	direct (natural contour)	Ruff & Hayes 1983, Lieberman et al. 2004
Natural break	direct (natural contour)	direct (natural contour)	Lovejoy & Trinkaus 1980, Trinkaus & Ruff 1999

[1] *Only external diameters; anteroposterior and mediolateral breadths are used.*
[2] *Only periosteal mould is used.*
CT, computed tomography; DSM, direct sectioning method; EMM, eccentric elliptical method; LCM, latex cast method; μCT, micro-computed tomography; pQCT, peripheral quantitative computed tomography.

contours. It is assumed that the intra-observer measurement error of the digitization method tested in this study will be similar to that of CT digitization (Sailer et al. 2003) and can be minimized by the observer's acquiring experience in the cross-sectional digitizing.

Material and Methods

Laboratory mice (*Mus musculus*) of the B6CBA strain were used (males: $n = 10$; females: $n = 7$) for the CSG parameters analysis. The animals were nine months old and kept in standard laboratory conditions. All procedures involved in obtaining bone samples were in compliance with EU guidelines for scientific experimentation on animals and with the permission of the Ethical Commission of the Faculty of Medicine in Pilsen, Czech Republic.

Preparation of the histological samples
Cross sections were obtained from the left tibia at 50 % of the biomechanical length of the bone. The histological staining process was taken from the study of Ciani et al. (2009) and used the fluorescent properties of fluorescein isothiocyanate (FITC, Sigma-Aldrich Co.). An advantage of this method is the ability to analyse both CSG parameters and microstructure of non-decalcified bone tissue. All the sections were imaged using a Leica TCS SP2 confocal microscope with the AOBS (Acousto-Optical Beam Splitter) system having the following parameters: resolution (2048 × 2048), pixel size (0.732 μm),

distance between scans (1.79 μm), 10× lens. Using ImageJ software, the following CSG properties of the tibia were analysed: total area (TA), cortical area (CA), moments of area (I_{max}, I_{min}), and polar moment of area (J).

MomentMacro for NIH Image and ImageJ (Ruff 2006) was used to calculate the geometric properties. MomentMacro had to be modified according to the method used and size of the specimens analysed in this study (http://www.natur.cuni.cz/biologie/servisni-laboratore/laborator-konfokalni-a-fluorescencni-mikroskopie/imagej-macros/moment-macro-j/view). On each scan, the borders of cortical bone had to be estimated by manually defined periosteal and endosteal lines (the region of interest). The image of a cross section had to be zoomed in when estimating the cortical borders. The manual settings of both periosteal and endosteal contours was completed when the last point reached the first. From the given region of interest, the basic geometric parameters were calculated (TA, CA, I, I_{max}, I_{min}). These values were then transferred to Excel spreadsheets and the index of circularity (circularity index $= I_{max}/I_{min}$) and polar moment of area ($J = I_{max} + I_{min}$) were additionally calculated. Cross-sectional geometric properties (areas and moments of area) measure the amount and distribution of skeletal tissue in a section (Larsen 1999). Areas include total area (TA) and cortical area (CA). CA is a measure of the amount of cortical bone in a cross section and is also an indicator as to strength of the long bone under pure axial loading.

Moments of area include second moments of area (I) and polar moment of area (J). Moments of area are geometric properties that are used to measure bending strength and torsional strength. Other values of I expressing the maximum and minimum bending strength in a cross section are referred to as I_{max} and I_{min}, respectively, where I_{max} measures the maximum strength in resistance of bone to bending and I_{min} measures the minimum strength in resistance of bone to bending (Larsen 1999). Values of I (I_{max} and I_{min}) are calculated as products of very small unit areas in the cross section and squared distances of the unit areas relative to the specific axis running through the cross section. Polar moment of area is equal to the sum of the values of I_{max} and I_{min}, which are perpendicular to each other (Larsen 1999). Because the cortical bone boundaries were manually defined by one individual with no previous experience with the digitization, the intra-observer error could be then measured to test the accuracy of the method used.

Intra-observer error measurement

Intra-observer error was tested in two different samples (set I and set II measurements). The set I measurement ($n = 10$; males) was performed by nine repeated measurements of CSG parameters. The first to eighth measurements were performed during one month's time. The ninth measurement was made six months later on the same sample. The eighth measurement was chosen as the reference, because it was assumed that the error of repeating measurements should be smallest between the seventh and eighth measurement and that the error decreases with the number of measurements performed. Intra-observer error of the set II measurements ($n = 7$; females) was tested one week later after the set I measurement by eight repeated measurements to determine whether the intra-observer error of the set I measurements could be influenced by memorizing of each cross section instead of from experiencing the section digitizing. Intra-observer error was tested according to the method used in a study by Sládek et al. (2012). Intra-observer error was calculated for TA, CA, I_{max}, I_{min}, I_{min}/I_{max} and J. The difference between each measurement was estimated by mean difference (MD). MD varies from negative to positive values. The MD was calculated by the following formula used in the study by Sládek et al. (2012):

$$MD = \frac{\sum_{i=1}^{n}(M1_i - M2_i)}{n}$$

Where $M1_i$ is the first compared (first-seventh; ninth) measurement, $M2_i$ is the eighth measurement and $n = 10$ (for set I measurements) and $n = 7$ (for set II measurements) is the sample size.

The overall error was expressed by mean absolute difference (MAD). MAD has only positive values and is computed as follows (Sládek et al. 2012):
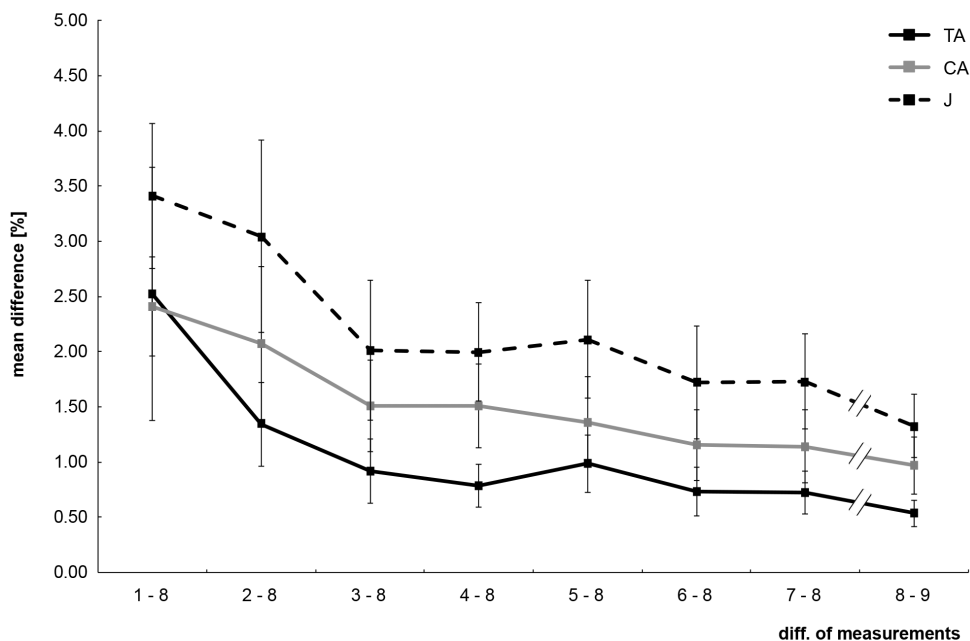

Fig. 2. *The mean error (%) from first to ninth measurements compared to eighth (reference) measurement for total area (TA), cortical area (CA) and polar moment of area (J).*

Table 2. *Total area (TA). Numbers 1–9 in column headings indicate measurement record for ten tested individuals. For computation and abbreviation, see Material and Methods.*

Individual	TA [mm²][1]	1–8 [mm²]	2–8 [mm²]	3–8 [mm²]	4–8 [mm²]	5–8 [mm²]	6–8 [mm²]	7–8 [mm²]	8–9 [mm²][2]
1	1.357	−0.034	−0.033	−0.029	−0.029	−0.030	−0.004	−0.002	0.005
2	1.273	0.014	0.001	0.005	0.018	0.011	0.009	0.005	−0.004
3	1.271	0.000	−0.008	0.000	0.009	−0.003	0.006	0.003	0.000
4	1.134	−0.009	−0.014	−0.003	0.006	−0.002	−0.002	0.000	−0.011
5	1.276	−0.004	0.005	0.002	−0.002	0.008	0.000	0.016	−0.015
6	1.206	−0.017	−0.007	0.002	0.000	0.000	−0.008	0.003	0.003
7	1.282	0.026	0.019	0.031	0.010	0.030	0.009	0.009	−0.013
8	1.050	0.016	0.011	0.021	0.010	0.009	0.011	0.009	−0.003
9	1.235	0.155	0.017	0.009	0.008	0.012	0.008	0.022	−0.005
10	1.128	−0.035	−0.047	0.011	−0.005	0.018	0.029	0.018	−0.007
MD		0.011	−0.006	0.005	0.003	0.005	0.006	0.008	−0.005
MAD		0.031	0.016	0.011	0.010	0.012	0.009	0.009	0.007
± 95 % LA		0.117; −0.095	0.036; −0.047	0.036; −0.026	0.028; −0.023	0.037; −0.026	0.026; −0.015	0.024; −0.007	0.008; −0.018

[1] *Values from the eighth measurement.*
[2] *6-month interval between eighth and ninth measurement.*

$$MAD = \frac{\sum\limits_{i=1}^{n}|M1_i - M2_i|}{n}$$

The symbols are described above.

The interval within which we would expect to find 95 % of differences between the first and second recordings is estimated by the limit of agreement (95 % LA). The 95 % LA was computed using the approach of Bland & Altman (1986):

$$LA = MD \pm 1.96 \times SD$$

Results

The results for the set I measurements are presented in Tables 2-4, Appendices 1-2 and Fig. 2. Each studied

Table 3. *Cortical area (CA). Number 1-9 in column headings indicate measurement record for ten tested individuals. For computation and abbreviation, see Material and Methods.*

Individual	CA [mm²][1]	1–8 [mm²]	2–8 [mm²]	3–8 [mm²]	4–8 [mm²]	5–8 [mm²]	6–8 [mm²]	7–8 [mm²]	8–9 [mm²][2]
1	0.880	−0.038	−0.034	−0.034	−0.037	−0.036	−0.003	−0.007	0.006
2	0.770	0.013	0.003	0.010	0.021	0.014	0.011	0.008	−0.006
3	0.781	−0.002	−0.008	0.001	0.008	−0.002	0.005	0.003	−0.002
4	0.679	−0.006	−0.008	0.003	0.011	−0.002	0.001	0.006	−0.012
5	0.839	−0.019	−0.002	−0.003	−0.006	−0.001	−0.004	0.014	−0.017
6	0.705	−0.022	−0.010	0.001	−0.002	−0.001	−0.010	0.002	0.001
7	0.856	0.014	0.008	0.024	−0.004	0.015	−0.002	0.002	−0.001
8	0.649	0.014	0.011	0.018	0.009	0.005	0.009	0.007	0.001
9	0.764	0.023	0.018	0.016	0.016	0.019	0.016	0.029	−0.016
10	0.793	−0.037	−0.061	0.009	−0.006	0.016	0.027	0.012	−0.014
MD		−0.006	−0.008	0.005	0.001	0.003	0.005	0.008	−0.006
MAD		0.019	0.016	0.012	0.012	0.011	0.009	0.009	0.008
± 95 % LA		0.038; −0.050	0.037; −0.054	0.036; −0.027	0.033; −0.031	0.034; −0.028	0.027; −0.017	0.026; −0.011	0.010; −0.022

[1] *Values from the eighth measurement.*
[2] *6-month interval between eighth and ninth measurement.*

Table 4. *Polar moment of area (J). Numbers 1-9 in column headings indicate measurement record for ten tested individuals. For computation and abbreviation see Material and Methods.*

Individual	J [mm⁴][1]	1–8 [mm⁴]	2–8 [mm⁴]	3–8 [mm⁴]	4–8 [mm⁴]	5–8 [mm⁴]	6–7 [mm⁴]	7–8 [mm⁴]	8–9 [mm⁴][2]
1	0.273	−0.016	−0.015	−0.014	−0.013	−0.014	−0.002	−0.002	0.002
2	0.232	0.007	0.001	0.003	0.008	0.006	0.005	0.003	−0.003
3	0.223	0.000	−0.004	0.000	0.004	−0.001	0.003	0.002	0.000
4	0.183	−0.003	−0.004	0.000	0.004	−0.001	0.000	0.001	−0.005
5	0.232	−0.004	0.002	0.000	−0.002	0.002	0.000	0.006	−0.007
6	0.211	−0.007	−0.003	0.001	0.000	0.000	−0.004	0.001	0.001
7	0.253	0.009	0.006	0.012	0.002	0.010	0.002	0.003	−0.003
8	0.154	0.006	0.004	0.007	0.004	0.003	0.004	0.003	−0.001
9	0.223	0.008	0.008	0.005	0.005	0.006	0.005	0.010	−0.004
10	0.197	−0.015	−0.019	0.004	−0.002	0.007	0.011	0.006	−0.004
MD		−0.002	−0.003	0.002	0.001	0.002	0.002	0.003	−0.002
MAD		0.007	0.007	0.005	0.004	0.005	0.004	0.004	0.003
± 95 % LA		0.016; −0.019	0.014; −0.020	0.015; −0.011	0.013; −0.011	0.015; −0.011	0.011; −0.006	0.010; −0.003	0.003; −0.008

[1] Values from the eighth measurement.
[2] 6-month interval between eighth and ninth measurement.

parameter showed decreasing error with increasing number of repeats. The greatest difference (MD) as well as overall error (MAD) between first and eighth measurements was found for I_{min} (3.63 %), J (3.41 %) and I_{max} (3.40 %). Compared to I_{max}, I_{min} and J, the I_{max}/I_{min} (1.68 %), CA (2.41 %) and TA (2.52 %) were less susceptible to error at the start of the digitization learning (Fig. 2). By the completion of eight consecutive measurements, each parameter was reduced below 2 % error: TA (0.73 %), CA (1.14 %), I_{max}/I_{min} (1.19 %), I_{max} (1.60 %), J (1.73 %) and I_{min} (1.94 %). The availability of a measurement after the interval of six months showed no effect in relation to the time that had passed since the digitization skill was acquired. However, the measurement errors for I_{max} (1.66 %) and I_{max}/I_{min} (1.42 %) were slightly increased while those for TA (0.54 %), CA (0.97 %), I_{min} (1.25 %) and J (1.33 %) were decreased.

The results from the set II measurements confirm learning of the digitization process, rather than simply learning of the given sample. The measurement error of the new sample was below 1 % for each parameter during each of the consecutive measurements. For the first measurement, the error was highest for I_{min} (0.48 %), I_{max} (0.46 %) and J (0.43 %) and lower for TA (0.18 %), I_{max}/I_{min} (0.21 %) and CA (0.32 %). In the last measurement, the intra-observer error was only slightly changed: I_{max} (0.62 %), J (0.47 %), I_{max}/I_{min} (0.37 %), CA (0.35 %), I_{min} (0.32 %), and TA (0.20 %). These results confirm our findings from the set I measurements that I_{max}, I_{min} and J are the most susceptible to measurement error in the first measurement on an unknown sample.

Discussion

The results of this study demonstrate that cross-sectional parameters derived from the tested digitization of periosteal and endosteal contours can be accurately used to estimate cross-sectional areas (TA, CA) and second moments of areas (I_{max}, I_{min}, I_{max}/I_{min}, J). Digitization of periosteal and endosteal cross-sectional contours of further use for histological processing is probably more difficult to be compared with methods using CT (e.g., Sailer et al. 2003). Histological samples are obtained invasively and undergo longer preparation processes that include cutting, grinding and histological staining of the cross section. The great advantage of the method used in this study, however, is the possibility to analyse both histological and CSG parameters of the long bones. Although we expected that intra-observer error of the tested method would be similar to that of CT, the results do not confirm our assumption. Intra-observer error from the digitization method used in this study was lower compared to the intra-observer error of the CT method used in Sailer et al. (2003). In the study by Sailer et al. (2003), however, the intra-observer error was calculated for measurements of human tibia cross sections. It remains unclear whether error differences between the two data sets could be influenced by the much different sizes of the measured cross sections. Accuracy of the tested method was estimated by intra-

observer error measurement according to the study by Sládek et al. (2012). Accuracy and precision can be interpreted only with respect to the level of error that is acceptable. The accepted error is frequently specified by a value derived from empirical observations (Hunter 1980, Meloun & Militký 1998) and is defined as the numerical accepted error (Sládek et al. 2010). A standard for acceptable error has not been developed for CSG comparison. Trinkaus & Ruff (1989) suggested a 5 % level of accepted difference to characterize reasonably accurate cross section measurements. Although the 5 % level of accepted error was later used in other methodological studies (Stock 2002, O'Neill & Ruff 2004), the established level was used only for human CT cross sections.

The intra-observer error measurement was tested in two samples. The first sample (set I) was digitized nine times. The last digitization was performed six months later to test the affect of placing a time interval between each measurement. Thus, this tested whether experience of a digitization process can be forgotten after some period of time. The results showed that the difference between intra-observer error for the eighth and ninth measurements was not influenced by the longer time interval between those measurements (see Fig. 2).

We assumed the possibility of memorizing each cross section during the digitization. According to this assumption, two different samples (set I and set II) were used to test whether the intraobserver error could be influenced by repeating and memorizing one sample digitization. Because intra-observer error of set II measurement was from the beginning smaller than that at the end of the set I measurement, we confirmed learning of the digitization process rather than learning of the sample.

Intra-observer measurement error could also be affected by the quality of the tested sample. The technique of obtaining the cross section and the process of histological staining can produce artefacts. Because the cross sections were calculated as an 'ideal' geometry of cortical tissue, the researcher had to decide which artefacts were unimportant for the estimated values. It is assumed that the intra-observer error could be affected by the frequency and size of artefacts in the raw cross sections. Learning of the digitization process could be slowed by a higher frequency of the cross-section artefacts.

Our results show that second moments of area (I_{max}, I_{min} and J) are more susceptible to intra-observer error than are cross-sectional areas (TA, CA). This could be caused by computational algorithms used for calculating second moments of area from defined periosteal and endosteal contour information. In our study, the intra-observer error in measuring total area was generally lower than was the intra-observer error in measuring cortical area. Cortical area is calculated using values from both periosteal and endosteal contours and total area is calculated using only values from the periosteal contour. These facts suggest that the digitization of endosteal contour might be more difficult to learn. Difficulties in digitizing of endosteal contour could be caused by a greater presence of artefacts on the endosteal boundary. While during the preparation of histological samples the soft tissue was removed from the periosteal and endosteal surface to reduce the presence of artefacts, it was more complicated to remove soft tissue from the endosteal surface.

Conclusion

Intra-observer error in the bone cross-section digitization technique used in this study was within 2 % for all the analysed CSG parameters (TA, CA, I_{max}, I_{min}, I_{max}/I_{min}, J). The intra-observer error attained cannot be interpreted as being either high or low, because the necessary level of accuracy and precision must be determined with reference to the specific research. Our results can serve the purpose of estimating accepted error for future studies. The results of this study confirm our assumption that the impact of intra-observer error during digitizing of periosteal and endosteal contour on CSG properties decreases with the number of repeated measurements. Thus, intra-observer measurement error for the digitization method used in this study can be minimized by acquiring experience in the section digitizing. This phenomenon is not caused by the memorization of measured bone cross sections and is not affected by the time interval between each measurement. Our results show that TA, CA and J are more susceptible to intra-observer error than are I_{max}, I_{min} and I_{max}/I_{min}. The results do not confirm our assumption that intra-observer measurement error for the tested digitization method is similar to that for CT digitization used in some previous studies. Even though digitization of histological sample can be more difficult to perform, the intra-observer error of the digitization method used in this study was lower than that reported earlier for CT digitization.

Acknowledgements

Literature

Bland J.M. & Altman D.G. 1986: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet 1: 307–310.*

Bouxsein M.L., Boyd S.K., Christiansen B.A., Guldberg R.E., Jepsen K.J. & Muller R. 2010: Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *J. Bone Miner. Res. 25: 1468–1486.*

Ciani C., Doty S.B. & Fritton S.P. 2009: An effective histological staining process to visualize bone interstitial fluid space using confocal microscopy. *Bone 44: 1015–1017.*

Diab T. & Vashishth D. 2007: Morphology, localization and accumulation of *in vivo* microdamage in human cortical bone. *Bone 40: 612–618.*

Ferretti J.L., Capozza R.F. & Zanchetta J.R. 1996: Mechanical validation of a tomographic (pQCT) index for noninvasive estimation of rat femur bending strength. *Bone 18: 97–102.*

Fritton J.C., Myers E.R., Wright T.M. & van der Meulen M.C. 2005: Loading induces site-specific increases in mineral content assessed by microcomputed tomography of the mouse tibia. *Bone 36: 1030–1038.*

Hunter J.S. 1980: The national system of scientific measurement. *Science 210: 869–874.*

Larsen C.S. 1999: Bioarchaeology: interpreting behavior from the human skeleton. *Cambridge University Press*, Cambridge.

Lieberman D.E., Polk J.D. & Demes B. 2004: Predicting long bone loading from cross-sectional geometry. *Am. J. Phys. Anthropol. 123: 156–171.*

Lovejoy C.O. & Trinkaus E. 1980: Strength and robusticity of the Neandertal tibia. *Am. J. Phys. Anthropol. 53: 456–470.*

Meloun M. & Militký J. 1998: Statistické zpracování experimentálních dat [Statistical analysis of experimental data]. *Plus, Praha. (in Czech)*

Nagurka M.L. & Hayes W.C. 1980: An interactive graphics package for calculating cross-sectional properties of complex shapes. *J. Biomech. 13: 59–64.*

O'Neill M.C. & Ruff C.B. 2004: Estimating human long bone cross-sectional geometric properties: a comparison of noninvasive methods. *J. Hum. Evol. 47: 221–235.*

Pfeiffer S. 1980: Age changes in the external dimensions of adult bone. *Am. J. Phys. Anthropol. 52: 529–532.*

Rittweger J., Goosey-Tolfrey V.L., Cointry G. & Ferretti J.L. 2010: Structural analysis of the human tibia in men with spinal cord injury by tomographic (pQCT) serial scans. *Bone 47: 511–518.*

Ruff C.B. 2006: MomentMacro for NIH Image and ImageJ. *Johns Hopkins University School of Medicine, Baltimore. Available from: http://www.hopkinsmedicine.org/fae/mmacro.htm*

Ruff C.B. 2007: Biomechanical analyses of archaeological human skeletons. In: Katzenberg M.A. & Saunders S.R. (eds.), Biological anthropology of the human skeleton. *John Wiley & Sons, Hoboken: 183–206.*

Ruff C.B. 2009: Relative limb strength and locomotion in *Homo habilis*. *Am. J. Phys. Anthropol. 138: 90–100.*

Ruff C.B. & Hayes W.C. 1983: Cross-sectional geometry of Pecos Pueblo femora and tibiae – a biomechanical investigation: I. Method and general patterns of variation. *Am. J. Phys. Anthropol. 60: 359–381.*

Ruff C.B. & Leo F.P. 1986: Use of computed tomography in skeletal structure research. *Am. J. Phys. Anthropol. 29: 181–196.*

Ruff C.B., Trinkaus E., Walker A. & Larsen C.S. 1993: Postcranial robusticity in *Homo*. I: Temporal trends and mechanical interpretation. *Am. J. Phys. Anthropol. 91: 21–53.*

Ruff C.B., Walker A. & Trinkaus E. 1994: Postcranial robusticity in *Homo*. III: Ontogeny. *Am. J. Phys. Anthropol. 93: 35–54.*

Rybicki E.F., Simonen F.A. & Weis E.B., Jr. 1972: On the mathematical analysis of stress in the human femur. *J. Biomech. 5: 203–215.*

Sabsovich I., Clark J.D., Liao G., Peltz G., Lindsey D.P., Jacobs C.R. & Yao W. 2008: Bone microstructure and its associated genetic variability in 12 inbred mouse strains: μCT study and in silico genome scans. *Bone 42: 439–451.*

Sailer R., Sládek V. & Berner M. 2003: Computer tomography and calculation of bone biomechanics in cross-sections of long bones. *Am. J. Phys. Anthropol. 120 (Suppl. 36): 182.*

Sheng M.-C., Bayling D., Beamer W., Donahue L., Rosen C., Lau K.-H. & Wergedal J. 1999: Histomorphometric studies show that bone formation and bone mineral apposition rates are greater in C3H/HeJ (high-density) than C57BL/6J (low-density) mice during growth. *Bone 25: 421–429.*

Sládek V., Berner M. & Sailer R. 2006a: Mobility in Central European late Eneolithic and early Bronze Age: femoral cross-sectional geometry. *Am. J. Phys. Anthropol. 130: 320–332.*

Sládek V., Berner M. & Sailer R. 2006b: Mobility in Central European late Eneolithic and early Bronze Age: tibial cross-sectional geometry. *J. Archaeol. Sci. 33: 470–482.*

Sládek V., Berner M., Sosna D. & Sailer R. 2007: Human manipulative behavior in the Central European late Eneolithic and early Bronze Age: humeral bilateral asymmetry. *Am. J. Phys. Anthropol. 133: 669–681.*

Sládek V., Berner M., Galeta P., Friedl L. & Kudrnová Š. 2010: Technical note: the effect of midshaft location on the error ranges of femoral and tibial cross-sectional parameters. *Am. J. Phys. Anthropol. 141: 325–332.*

Sládek V., Galeta P. & Sosna D. 2012: Measuring human remains in the field: grid technique, total station, or microscribe? *Forensic. Sci. Int. 221 (1–3): 16–22.*

Sone T., Imai Y., Joo Y.I., Onodera S., Tomomitsu T. & Fukunaga M. 2006: Side-to-side differences in cortical bone mineral density of tibiae in young male athletes. *Bone 38: 708–713.*

Stock J.T. 2002: A test of two methods of radiographically deriving long bone cross-sectional properties compared to direct sectioning of the diaphysis. *Int. J. Osteoarchaeol. 12: 335–342.*

Stock J.T. & Shaw C.N. 2007: Which measures of diaphyseal robusticity are robust? A comparison of external methods of quantifying the strength of long bone diaphyses to cross-sectional geometric properties. *Am. J. Phys. Anthropol. 134: 412–423.*

Sumner D.R., Mockbee B., Morse K., Cram T. & Pitt M. 1985: Computed tomography and automated image analysis of prehistoric femora. *Am. J. Phys. Anthropol. 68: 225–232.*

Trinkaus E. & Ruff C.B. 1989: Diaphyseal cross-sectional morphology and biomechanics of the Fond-de-Foret 1 femur and the Spy 2 femur and tibia. *Bull. Soc. Roy. Belg. Anthropol. et Préhist. 100: 33–42.*

Trinkaus E. & Ruff C.B. 1999: Diaphyseal cross-sectional geometry of Near Eastern Middle Palaeolithic humans: the tibia. *J. Archaeol. Sci. 26: 1289–1300.*

Trinkaus E., Stringer C.B., Ruff C.B., Hennessy R.J., Roberts M.B. & Parfitt S.A. 1999: Diaphyseal cross-sectional geometry of the Boxgrove 1 Middle Pleistocene human tibia. *J. Hum. Evol. 37: 1–25.*

Appendix 1. *Maximum second moments of area (I_{max}). Numbers 1-9 in column headings indicate measurement record for ten tested individuals. For computation and abbreviation see Material and Methods.*

Individual	I_{max} [mm⁴][1]	1–8 [mm⁴]	2–8 [mm⁴]	3–8 [mm⁴]	4–8 [mm⁴]	5–8 [mm⁴]	6–7 [mm⁴]	7–8 [mm⁴]	8–9 [mm⁴][2]
1	0.160	–0.008	–0.008	–0.008	–0.007	–0.007	–0.001	–0.001	0.002
2	0.129	0.006	0.000	0.002	0.006	0.005	0.004	0.003	–0.002
3	0.119	0.001	–0.002	0.001	0.002	0.000	0.002	0.001	–0.001
4	0.100	–0.002	–0.002	–0.001	0.002	0.000	0.001	0.000	–0.006
5	0.133	–0.004	0.001	0.000	–0.003	0.001	0.001	0.005	–0.004
6	0.131	–0.004	–0.002	0.000	0.000	0.000	–0.003	0.000	0.000
7	0.148	0.004	0.003	0.006	0.001	0.004	0.000	0.001	–0.001
8	0.081	0.003	0.003	0.004	0.003	0.002	0.003	0.002	–0.001
9	0.128	0.004	0.004	0.003	0.003	0.003	0.003	0.006	–0.002
10	0.113	–0.006	–0.006	0.002	–0.002	0.003	0.005	0.001	–0.002
MD		–0.001	–0.001	0.001	0.000	0.001	0.002	0.002	–0.002
MAD		0.004	0.003	0.003	0.003	0.003	0.002	0.002	0.002
± 95 % LA		0.009; –0.010	0.007; –0.009	0.008; –0.006	0.008; –0.007	0.008; –0.006	0.006; –0.003	0.006; –0.002	0.003; –0.006

[1] Values from the eighth measurement.
[2] 6-month interval between eighth and ninth measurement.

Appendix 2. *Minimum second moments of area (I_{min}). Numbers 1-9 in column headings indicate measurement record for ten tested individuals. For computation and abbreviation see Material and Methods.*

Individual	I_{min} [mm⁴][1]	1–8 [mm⁴]	2–8 [mm⁴]	3–8 [mm⁴]	4–8 [mm⁴]	5–8 [mm⁴]	6–7 [mm⁴]	7–8 [mm⁴]	8–9 [mm⁴][2]
1	0.114	–0.007	–0.006	–0.006	–0.006	–0.006	–0.001	–0.001	0.001
2	0.103	0.001	0.001	0.001	0.002	0.001	0.001	0.000	–0.001
3	0.104	–0.001	–0.003	0.000	0.002	–0.001	0.002	0.001	0.000
4	0.084	–0.001	–0.002	0.001	0.002	–0.001	–0.001	0.001	0.001
5	0.100	–0.001	0.001	0.001	0.000	0.001	–0.001	0.002	–0.002
6	0.079	–0.003	–0.001	0.000	0.000	0.000	–0.001	0.001	0.001
7	0.105	0.005	0.003	0.006	0.001	0.006	0.002	0.002	–0.003
8	0.073	0.002	0.002	0.003	0.001	0.001	0.001	0.001	0.000
9	0.095	0.004	0.003	0.002	0.002	0.003	0.001	0.005	–0.002
10	0.085	–0.009	–0.013	0.002	0.000	0.004	0.006	0.005	–0.002
MD		–0.001	–0.002	0.001	0.000	0.001	0.001	0.002	–0.001
MAD		0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.001
± 95 % LA		0.008; –0.010	0.008; –0.011	0.007; –0.005	0.005; –0.004	0.007; –0.006	0.005; –0.003	0.005; –0.002	0.002; –0.003

[1] Values from the eighth measurement.
[2] 6-month interval between eighth and ninth measurement.