Cells and Materials Vol. 6, No. 1-3, 1996 (Pages 79-92)
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

1051-6794/96$5.00+.25

VALIDATION OF ENERGY DISPERSIVE X-RAY SPECTROMETRY AS A METHOD TO STANDARDIZE BACKSCATTERED ELECTRON IMAGES OF BONE

E.G. Vajda, R.D. Bloebaum* and J.G. Skedros

Bone and Joint Research Laboratory, VA Medical Center, Salt Lake City, UT

(Received for publication May 9, 1996 and in revised form September 23, 1996)

Abstract

The use of backscattered electron (BSE) imaging as a tool for the qualitative measurement of mineral content in bone has been well documented. The challenge still remains to develop BSE imaging as a tool for quantitative mineral content analysis in bone. The limiting factor has been the ability to standardize the BSE signal within and between laboratories. Energy dispersive x-ray spectrometry (EDX) has been proposed as a method to standardize the BSE operating environment. The goal of this study is to investigate the relationship between EDX-determined wt.% Ca measurements and BSE graylevels. A comparison with traditional ash content measurements will indicate the validity of the procedure. A comparative study was performed on a series of bones representing a broad range of mineralization. Results confirmed a high correlation between BSE graylevels and wt.% Ca measured with EDX. However, the BSE method consistently underestimated the mineral content of bone determined by traditional ash measurements. The results suggest that quantitative BSE imaging can be standardized by EDX measurements, but an empirically determined correction may be necessary if comparisons with known and accepted mineral content measurement techniques are to be performed. Further investigation into the nature of this empirical correction is warranted before the procedure can be universally applied to bone mineral content analysis.

Key Words: Backscattered electron imaging, energy dispersive X-ray spectrometry, bone, mineral content, standardization.

*Address for correspondence:
Roy D. Bloebaum
Bone and Joint Research Laboratory
VA Medical Center (151F)
500 Foothill Blvd.
Salt Lake City, UT 84148
Telephone Number: (801) 582-1565 ext. 4607
FAX Number: (801) 584-2533

Introduction

Backscattered electron (BSE) imaging in the scanning electron microscope (SEM) provides a new approach to the microscopic measurement of mineral content variations in bone and mineralized tissues. BSE imaging is based upon the measurement of high energy, elastically scattered electrons. The fraction of incident electrons which are returned as backscattered electrons is a monotonically increasing function of the sampled region's mean atomic number [27]. Boyde and Jones first realized that this feature could be exploited in the study of bone tissue [13]. Since that time, experimental studies have demonstrated that the BSE signal is linearly proportional to mineral content [49, 50] in bone and simulated bone tissues. In both experimental [3, 54] and theoretical [30] studies, BSE imaging has demonstrated superior spatial resolution to the densitometric evaluation of microradiographs, which is the standard in the measurement of microscopic variations in bone mineral content [2, 45]. The combination of high resolution and proportionality to mineral content in BSE images of bone permits the use of BSE imaging technology for analysis of mineral variations within individual morphological features in addition to traditional histometry. In bone research, quantitative BSE imaging has been developed into a tool to analyze growth- and age-related variations [10, 14, 24, 41] structural/material adaptation and remodeling [35, 51], influences of therapeutic regimes [22, 25] and bone tissue reaction to orthopaedic and dental implants [6, 7, 8, 28, 29].

For quantitative BSE imaging to fulfill its potential as an analysis tool, proper calibration is essential. Calibration involves the control of two distinct aspects of any SEM/BSE operating session. First, instrument stability must be maintained and periodically measured to verify stability. The authors refer to this problem as instrument calibration. Secondly, quantitative data must be obtained in a form that is universally accepted and allows comparison of data between laboratories. This problem will be referred to as standardization.

Investigators have attempted instrument calibration by imaging and re-imaging regions of bone during an operating session to verify beam stability [12, 37].

Unfortunately, embedded bone specimens are not stable for prolonged periods of electron bombardment and investigators have documented the potential pitfalls of specimen degradation [11, 57]. Published reports have also indicated that the embedding media can be particularly unstable under the electron beam [21, 26, 48]. Boyde *et al.* [12] introduced the use of styrene to stabilize polymethylmethacrylate embedding media, but quantitative results have not been published. Due to these limitations this method of calibration has not gained wide acceptance.

Reid and Boyde [41] and Robinson [42] proposed a method of instrument calibration using pure elemental standards. This technique was further developed by Bloebaum *et al.* [5] and Boyce *et al.* [11] using pure Al and pure Mg alloy reference standards. Boyce *et al.* [11] experimentally demonstrated that the average graylevel intensity, or weighted mean graylevel (WMGL), of BSE images of bone remained consistent between consecutive imaging sessions using Mg and Al as reference calibration standards. Boyce's method was enhanced and simplified by Vajda *et al.* [57] using computer controls and a retrospective calibration method, dramatically decreasing the time required for instrument calibration. Vajda *et al.* [57] analyzed a selection of various bones and inorganic materials and demonstrated consistency not only in the WMGL, but also in the shape of the graylevel histogram profile, in a series of images collected over a one week period. Instrument calibration with pure elements is now being used [25, 44, 55] and provides a simple method of ensuring beam stability during the course of an experiment. The use of pure elements has allowed for accurate calibration of a single instrument and the measurement of *relative* differences in mineral content.

Published reviews of the paper by Vajda *et al.* [57] suggested that crystallographic variations, oxidation changes, and manufacturing methods can result in pronounced differences between pure elemental standards. The authors acknowledged this point and provided data indicating manufacturing methods and/or specimen preparation of pure elements can influence WMGLs in BSE images [57] . This emphasizes the difficulty obtaining *absolute* measurements of mineral content with backscattered electron imaging and the need for standardization as well as instrument calibration.

Roschger *et al.* [44] attempted to standardize the BSE signal by correlating the WMGL of BSE images to wt.% Ca measurements. On the same region of bone, the investigators captured BSE images and calcium Kα-line intensities with energy dispersive x-ray spectrometry (EDX) and found a very high correlation between WMGL and wt.% Ca. Although the linear relationship they observed is encouraging, validation of this technique requires correlation with known and accepted analysis methods. Roschger *et al.* [44] attempted to obtain absolute measurements of calcium, and therefore the mineral content, of bone without any *a priori* knowledge of the "true" calcium content. Although these results are a promising advance in the development of quantitative BSE analysis, they do not solve the standardization issue. An ideal standardization method would be able to convert BSE graylevels into weight percent mineral measurements, which are analogous to traditional ash fraction measurements. This would allow comparisons between laboratories and with previous work correlating ash fraction and mechanical properties [15, 16, 17, 59].

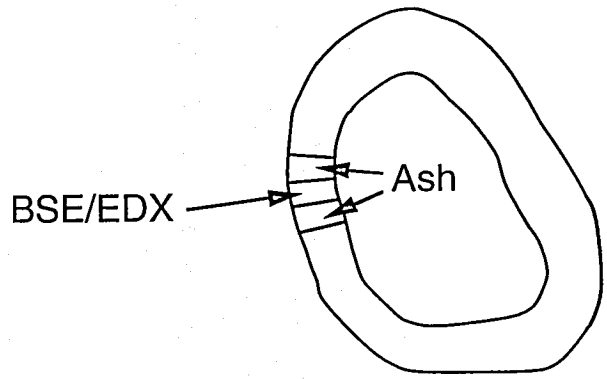

Figure 1. Diagram of a typical sectioning pattern used. The ash percent values reported in this study are the mean values of the two segments obtained from each bone.

The objective of the present study is to use the instrument calibration technique of Vajda *et al.* [57] and the standardization method of Roschger *et al.* [44] to: 1) verify that BSE graylevels are correlated to EDX-determined wt.% Ca measurements, 2) establish standardization equations between BSE graylevels and wt.% Ca measurements, and 3) use standardization equations and BSE graylevels to predict traditional ash measurements.

Materials and Methods

Specimen preparation

Thirteen bones were obtained from nine species (Table 1). Bones were selected to represent a broad range of mineralization [15]. Specimens were stored in 70% ethyl alcohol for a minimum of two weeks, after which time each bone was removed from alcohol and sectioned into three adjacent segments (Fig. 1). Two segments from each bone were used for ash percent measurements and the third segment from each bone was

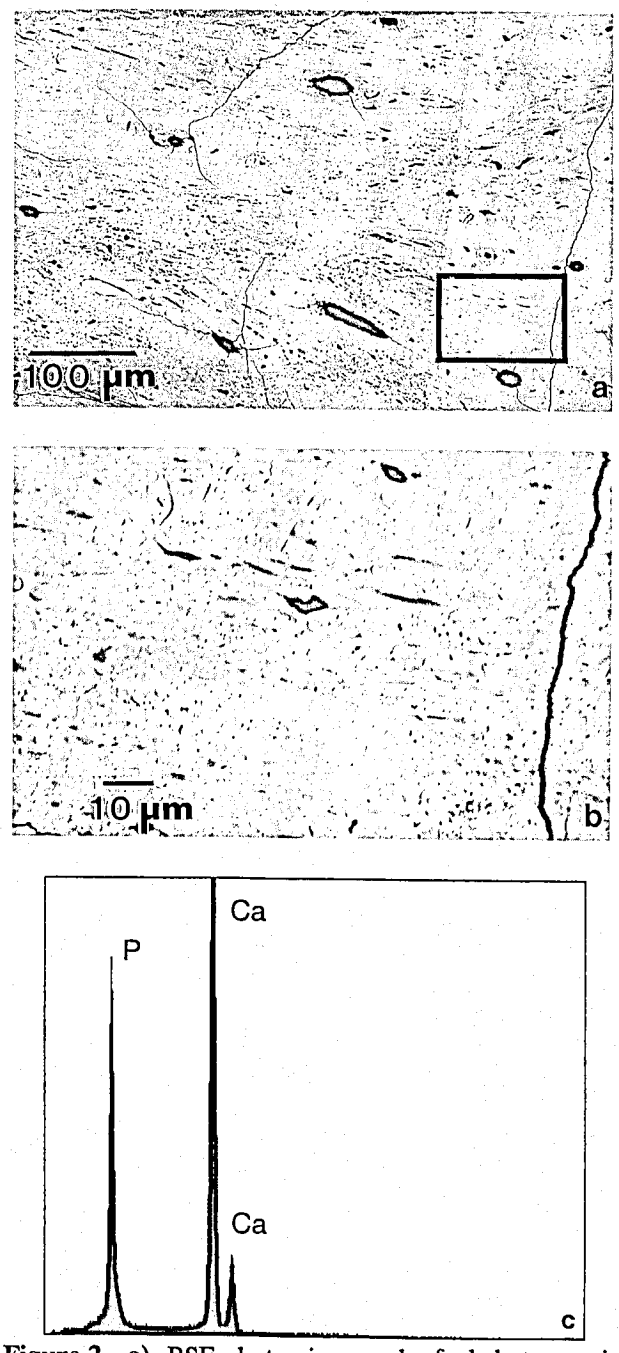

Figure 2. a) BSE photomicrograph of whale tympanic bulla. The box in the lower right corner depicts a small region, appearing relatively homogenous, to be selected for BSE standardization. b) High magnification BSE photomicrograph of outlined region in Fig. 2a. The weighted mean graylevel (WMGL) of the image is calculated as previously described [57]. c) EDX spectrum captured from same region as Fig. 2b. Ca and P peaks are clearly visible. ZAF corrections are used to calculate wt.% Ca measurements. Wt.% Ca measurements are then correlated with WMGL.

used for BSE analysis. Each bone segment was manually cleaned of soft tissue and periosteum where applicable. Marrow and soft tissue removal was facilitated by a high pressure stream of water (Water-Pik®, Teledyne Water Pik, Fort Collins, CO). Following soft tissue removal, each bone was defatted in a large volume of reagent-grade chloroform (Omnisolv, EM Industries, Inc., Gibbstown, NJ) under vacuum for 20 days. Chloroform was replenished at regular intervals to ensure that it did not completely evaporate.

Bone segments for BSE analysis were placed in ascending concentrations of alcohol and cleared in xylene following previously published protocols [19]. BSE bone segments were subsequently embedded in polymethyl methacrylate [19, 46, 47]. Embedded bones were cut into cubes and glued together with a cyanoacrylate glue. The composite structure was glued into a groove previously milled in a Plexiglas™ block. A 99.9999% pure aluminum wire and a 99.8% pure magnesium-aluminum-zinc (93% Mg, 6% Al, 1% Zn) wire (Johnson/Matthey, Inc., Seabrook, NH) were glued into drill holes in the specimen block. For brevity, magnesium-aluminum-zinc will be referred to as magnesium. The specimen block was milled, ground, and polished to an optically smooth finish [5]. The polished surface was lightly sputter-coated with a thin layer of gold for 75 seconds at 70 μm Hg and 10 mA (Hummer Model VI-A Sputtering System, Anatech, Ltd., Alexandria, VA).

Ash percent measurements

Bone segments used for ash measurements were placed in an oven at 80°C for 4 days to remove residual chloroform. Segments were immediately placed in a desiccator containing anhydrous $CaSO_4$ (Drierite, W.A. Hammond Drierite Co., Xenia, OH) for 6 hours and allowed to cool to room temperature. Bone segments were weighed to the nearest 0.00001 g on a precision analytical balance (Mettler H51, Mettler Instrument Co., Hightstown, NJ). Bone segments were subsequently placed in a furnace at 550°C for 24 hours. The bone segments were removed from the furnace, placed in a desiccator for 6 hours, and then re-weighed. Ash percent was calculated as 100 times the ratio of weight of ashed bone (W_{AB}) to weight of dry, defatted bone (W_{DB}) [($W_{AB} \div W_{DB}$)*100]. The value reported in this study is the mean of the two bone segments used from each bone. It has been experimentally demonstrated that small regional variations in mineral content do occur within one bone [35, 52]. Because the segments chosen for ash measurements were directly adjacent to the segments used for BSE analysis (Fig. 1), the mean value of the two measurements should be an accurate measurement with minimal error due to regional variations in

mineralization.

SEM operating conditions

The composite specimen block was placed on the SEM stage with the polished surface perpendicular to the incident electron beam. Working distance was set at 15 mm, accelerating voltage at 20 kV, and probe current at 2.50 nA. BSEs were collected by a four quadrant, semiconductor BSE detector (Tetra, Oxford Instruments, Cambridge, UK) which was configured around the electron beam. Digital BSE images with a resolution of 512 x 512 pixels and 8 bits/pixel (256 distinct graylevels) were stored on magnetic media using a computer-controlled image capture and retrieval system (eXL, Oxford Instruments). All BSE images were captured using nine scans and a Kalman frame averaging technique (Oxford Instruments).

EDX spectra were captured using a Link Pentafet system (Oxford Instruments) under the same operating conditions, with a live capture time of 130 seconds. The take off angle of the detector was set at 35°. Commercially available software (Link ZAF-4, Oxford Instruments) was used to calculate atomic number (Z), absorbance (A), and fluorescence (F) correction factors. Pure $CaCO_3$ and InP (Tousimis Research Inc., Rockville, MD) were used as reference standards for the major constituents of bone -- Ca and P. Additionally, Au, NaF, KCl, and MgO (Tousimis Research) were used as reference standards for the conductive coating and other minor elemental constituents [32]. The major light elements consist of carbon, hydrogen, nitrogen, and oxygen. For the purpose of this investigation carbon was treated as the dominant matrix element and was the unanalyzed element used in ZAF corrections, following the procedure of Roschger *et al.* [44] A pure Cu reference standard was used for gain calibration of the EDX spectra.

BSE Standardization with Ca Kα-lines

BSE brightness and contrast controls were adjusted to simultaneously provide an optimum image of both high and low mineral content bones. The brightest (highest mineral content) bone was found and a BSE image was captured as described above. Immediately following image capture an EDX spectra was obtained from the same region (Fig. 2). This was repeated at four different locations, to yield a total of five BSE images with five corresponding EDX spectra from regions of highly mineralized bone. Next, five BSE images and five EDX spectra were obtained from the darkest (lowest mineral content) bone. Image magnification was selected such that each BSE image appeared as a homogenous field with minimal visible graylevel variations. Magnification varied from 500X to 1000X.

At twenty minute intervals, the BSE image and EDX spectra capture sequence was interrupted and a BSE image of the pure aluminum and magnesium reference standards was obtained. This allowed for beam stabilization corrections, following the protocol of Vajda *et al.* [57]. Gain calibration of the EDX signal was performed after every second EDX spectra capture.

The weighted mean graylevel (WMGL) of each BSE image was calculated using public domain image analysis software (NIH Image 1.55, available from: wayne@helix.nih.gov) following previously published protocols [5, 11, 57]. Wt.% Ca was calculated for each EDX spectrum using the Ca Kα-line and ZAF corrections as described above. Linear regression analysis was performed to obtain the equation relating wt.% Ca and WMGL. The relationship between wt.% Ca and WMGL is highly linear [44], therefore any two points adequately determine the linear equation relating the two variables. Roschger *et al.* [44] used pure gelatin to determine the 0% Ca point, however this is not necessary in the calibration routine described in this study. The selection of five dark regions and five bright regions proved more than sufficient to determine the standardization equation.

Correlating BSE WMGL and mineral content measurements

Immediately following the standardization routine described above, BSE images of all of the bone tissues were captured. Five 200X images (corresponding to approximately 450 μm x 450 μm per image) were captured from each bone segment. Beam stabilization was performed at twenty minute intervals [57] by imaging the magnesium and aluminum standards. WMGL was calculated for each image [5, 11, 57]. The mean value of the five WMGL calculations from each bone is the value reported in this study.

Using the standardization equation determined above, WMGL values were converted to wt.% Ca values. Neglecting minor constituents and phase differences, bone mineral is a form of apatite, specifically hydroxyapatite (HA) $[Ca_{10}(PO_4)_6(OH)_2]$. Calcium occupies 39.9% of stoichiometric hydroxyapatite by weight. Wt.% Ca can therefore be converted to wt.% HA by the multiplicative factor 2.51 (100% HA ÷ 39.9% Ca = 2.51).

The entire data collection routine, including BSE and EDX data collection from bright and dark regions, calculation of the regression equation, and BSE image capture from all bone specimens, was repeated twice after re-adjusting the brightness and contrast controls on the SEM. This was done in order to determine if changes in brightness and contrast would significantly alter the relationship between WMGL and wt.% Ca.

Table 1. List of the bone specimens used and their anatomical provenances.

Species		Anatomical Location	Comments
human (S162)	*Homo sapiens*	midshaft, anterior cortex, femur	female, 25 years of age
human (S141)	*Homo sapiens*	midshaft, anterior cortex, femur	male, 64 years of age
cow	*Bos taurus*	midshaft, anterior cortex, femur	skeletally mature, domestic
fin whale	*Balaenoptera physalus*	tympanic bulla	skeletally mature, wild
rocky mountain mule deer	*Odocoileus hemionus hemionus*	antler	mature (no velvet present)
rocky mountain mule deer	*Odocoileus hemionus hemionus*	proximal shaft, cranial cortex, calcaneus	skeletally mature, wild
king penguin	*Aptenodytes patagonica*	midshaft, cortex, humerus	skeletally mature, captive
sheep	*Ovis aries*	midshaft, anterior cortex, tibia	skeletally mature, domestic
greyhound dog (D56)	*Canis familiaris*	midshaft, anterior cortex, femur	skeletally mature, domestic
greyhound dog (D282)	*Canis familiaris*	midshaft, anterior cortex, femur	skeletally mature, domestic
greyhound dog (85)	*Canis familiaris*	midshaft, medial cortex, tibia	skeletally mature, domestic
rabbit	*Lepus sp.*	midshaft, cortex, femur	skeletally mature, domestic
semi-aquatic turtle	*Chelydra serpentina*	midshaft, cortex, tibia	skeletally mature, captive

Table 2. Table of wt. % Ca measurements using EDX and corresponding WMGL values from the same region.

	Trial 1		Trial 2		Trial 3	
Specimen	WMGL	Wt.% Ca	WMGL	Wt.% Ca	WMGL	Wt.% Ca
High Mineral 1	152.8	29.7	159.4	28.9	214.7	29.8
High Mineral 2	152.0	25.9	164.8	29.9	211.0	29.6
High Mineral 3	157.1	28.8	175.7	30.4	213.1	30.3
High Mineral 4	163.2	31.2	171.4	30.1	213.1	30.0
High Mineral 5	149.6	29.6	166.8	29.7	213.9	30.0
Low Mineral 1	60.5	19.0	73.1	20.4	164.0	20.1
Low Mineral 2	61.8	20.1	73.8	21.1	165.4	20.2
Low Mineral 3	65.8	20.3	72.5	17.2	164.9	20.3
Low Mineral 4	64.3	20.4	68.4	19.9	164.8	20.7
Low Mineral 5	65.3	19.7	75.8	18.0	162.6	19.4

These data were used to establish regression equations relating wt. % Ca and ash percent.

All BSE WMGL measurements obtained with the new brightness and contrast settings were converted to wt. % Ca measurements using the new regression equations.

Results

Standardization equations

The standardization equations demonstrated highly correlated relationships between BSE WMGL and wt. % Ca (Fig. 3 and Table 2). Changing brightness and contrast resulted in a change in both the slope and intercept of the regression line, but did not influence the strength of the correlation. R^2 values were 0.942, 0.956, and 0.997, respectively, for the three trials. This is consistent with results reported by Roschger *et al.* [44], where they demonstrated the highly linear relationship between WMGL and wt. % Ca measurements.

Wt. % HA and ash percent correlation

Wt. % HA, calculated using BSE WMGLs and the standardization equations, showed a very high positive correlation to ash percent measurements (r^2 = 0.932, r^2 = 0.949, and r^2 = 0.926 for the three trials respectively). Fig. 4 shows that the data are well fit by a linear regression line. Changing brightness and contrast between the three trials did not alter the strength of the correlations between wt. % HA and ash percent. The slopes of the lines are nearly identical to the slope of a theoretical perfect correlation (Fig. 4). The y-intercepts are substantially lower than the theoretical correlation, however, indicating the wt. % HA calculations consis-

Table 3. Table of ash percent data, WMGL data, and wt.% HA data.

	Trial 1		Trial 2		Trial 3		Wt.% HA		Ash %
	WMGL	Wt.% HA	WMGL	Wt.% HA	WMGL	Wt.% HA	Mean	S.E.	
Bulla	151.8 ±1.5	72.2 ±0.4	169.0 ±2.1	75.1 ±0.5	205.5 ±3.1	71.3 ±1.6	72.9	0.5	85.8 ±2.5
Human (S141)	89.1 ±3.6	56.3 ±0.9	106.4 ±5.0	57.8 ±1.3	167.6 ±6.0	52.2 ±3.1	55.4	0.8	70.1 ±0.1
Cow	108.4 ±1.6	61.2 ±0.4	125.3 ±1.2	63.1 ±0.3	182.1 ±1.4	59.5 ±0.7	61.2	0.4	73.5 ±0.2
Human (S162)	90.8 ±2.5	56.7 ±0.6	106.7 ±1.9	57.9 ±0.5	172.0 ±3.0	54.4 ±1.5	56.4	0.5	70.7 ±0.0
Penguin	103.4 ±1.4	59.9 ±0.4	119.9 ±1.3	61.6 ±0.3	175.7 ±4.0	56.3 ±2.0	59.2	0.7	74.7 ±1.1
Dog (D56)	98.7 ±4.9	58.7 ±1.2	113.3 ±1.8	59.8 ±0.5	174.1 ±3.9	55.4 ±2.0	58.0	0.6	73.4 ±0.2
Sheep	100.4 ±2.6	59.2 ±0.6	116.8 ±1.6	60.7 ±0.4	175.6 ±1.9	56.2 ±1.0	58.7	0.5	74.4 ±0.5
Dog (D85)	79.3 ±1.8	53.8 ±0.5	98.3 ±2.5	55.6 ±0.6	166.9 ±3.8	51.8 ±1.9	53.7	0.5	70.5 ±0.5
Antler	47.5 ±2.8	45.8 ±0.7	69.9 ±1.4	47.8 ±0.4	155.2 ±2.6	45.9 ±1.3	46.5	0.3	61.6 ±0.4
Dog (D282)	91.1 ±1.3	56.8 ±0.3	106.6 ±2.3	57.9 ±0.6	170.4 ±2.6	53.6 ±1.3	56.1	0.5	72.2 ±0.3
Deer Calcaneus	92.1 ±2.9	57.1 ±0.7	113.0 ±3.9	59.7 ±1.0	172.8 ±2.3	54.8 ±1.1	57.2	0.6	71.4 ±0.7
Turtle	93.3 ±2.2	57.4 ±0.5	104.5 ±3.4	57.3 ±0.9	165.9 ±13.9	51.3 ±7.0	55.3	1.2	68.3 ±1.1
Rabbit	115.1 ±2.1	62.9 ±0.5	130.8 ±3.7	64.6 ±0.9	185.0 ±1.7	61.0 ±0.9	62.8	0.4	75.4 ±0.2
Mean							58.0		72.5

Wt.% HA was calculated using the regression equations (Fig. 3) relating WMGL and wt.% Ca with the multiplicative factor 2.51 (see text). Results from three repeated trials are displayed. Data is reported as Mean ± Standard Deviation. Smaller standard deviations indicate greater homogeneity at the microscopic level. The two columns titled "wt.% HA" are the mean and standard error of the mean based upon data collected from all three trials. Small standard errors suggest the analysis has acceptable precision. Wt.% HA values are consistently low, however, with respect to ash measurements.

tently underestimate the ash percent measurements by an average of 14.5% in this investigation (Table 3).

Discussion

The three standardization equations obtained for the three data sets in this study demonstrate a high correlation between wt.% Ca and WMGL. Changes in brightness and contrast resulted in significant changes in both the slope and intercept of the regression equations, but had no observable influence on the strength of the correlation (Fig. 3). Roschger *et al.* [44] reported similar results with a correlation coefficient (r) of 0.998 (r^2 = 0.996), which is similar to the correlations found in this study.

Ten data points were collected for standardization in the present study. The high r^2 values indicates that a correlation between WMGL and wt.% Ca does exist. However, the data points were collected only at the extreme range of the mineralization spectrum. It can not, therefore, be concluded that the relationship is linear, because a polynomial, exponential, or other relationship could also explain the observed correlation. The previous study by Roschger *et al.* [44] did include data points with varying mineral content and linearity was observed. The data obtained in the present study is consistent with this finding, suggesting linearity, although it is not conclusive due to the method of data collection.

The high correlation of the three regression lines obtained during three SEM/BSE operating sessions along with the previously published work by Roschger *et al.* [44] suggests that acceptable precision of these measurements might be attainable when this analysis is performed by different laboratories, regardless of brightness and contrast settings. These results are particularly surprising when the interaction volumes of the two analysis methods are compared. Characteristic X-rays sample a much larger volume than backscattered electrons [23 (pages 146-147)] and it would therefore not be anticipated that such a strong correlation would exist. Boyde expressed this concern in his published review of the study by Roschger *et al.* [44]. However, the high correlations observed in the present study and in that of Roschger *et al.* [44] suggests that regional heterogeneities in bone composition and differences in sampling volume have only a minor influence upon data variation.

When the standardization equations were used to convert BSE graylevels to wt.% HA in the remaining bone tissues, a highly linear correlation between wt.% HA and ash percent was observed (Fig. 4 and Table 3). The strength of the correlation was consistent between

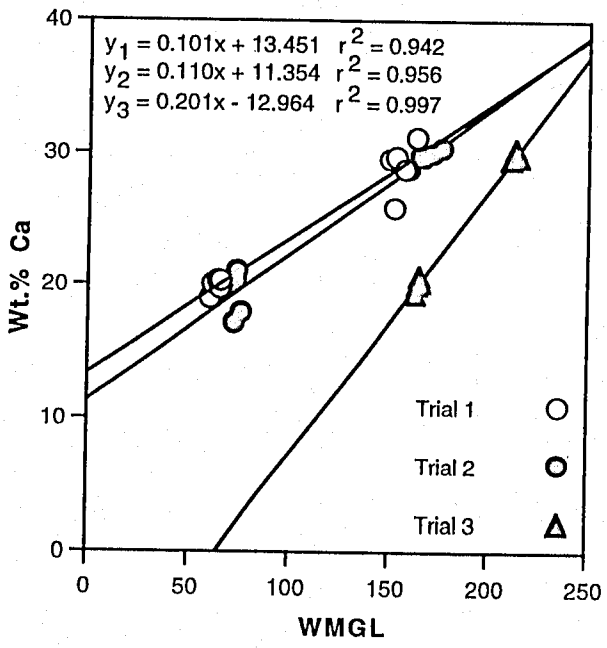

Figure 3. Scatterplot of wt.% Ca versus WMGL. Three trials, with varying brightness and contrast settings, are displayed. A strong correlation was observed in all three trials. The slope and intercept of the regression lines varied between trials. The regression lines y_1, y_2, and y_3 correspond to Trial 1, 2, and 3 respectively.

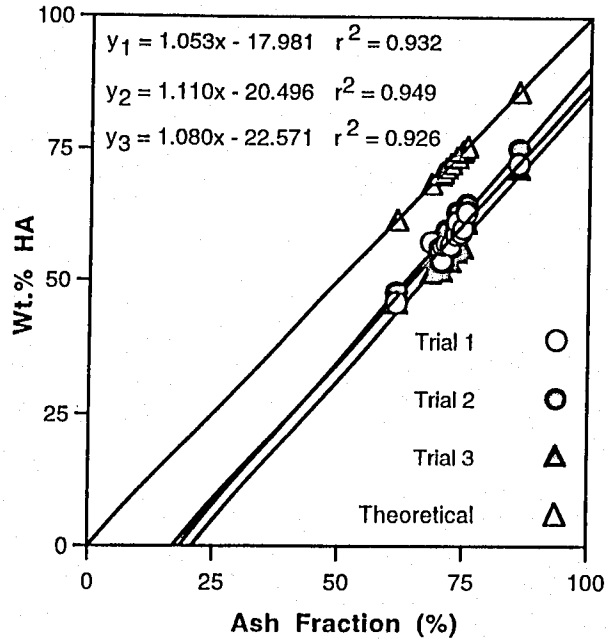

Figure 4. Scatterplot of wt.% HA versus ash percent. Wt.% HA was calculated using WMGL measurements (Table 3) and regression equations (Fig. 3) with a multiplicative factor (2.51, see text). Highly linear relationships from three trials, corresponding to different brightness and contrast settings, are displayed. A theoretical, ideal correlation is also displayed. All four regression lines have similar slopes, close to 1. The y-intercepts of the three trials are significantly lower than for the theoretical regression line, indicating that this method consistently underestimates ash percent.

---

trials, even with substantially different brightness and contrast settings. Furthermore, the slope of the regression line is close to 1 in each trial, which indicates a 1 to 1 correlation between wt.% HA and ash percent, regardless of instrument settings. Analysis of covariance [53 (pages 499-521)] revealed no statistically significant difference ($p > 0.75$) between a theoretical perfect correlation with a slope of 1.0 and the slopes of the regression lines.

A small, but statistically significant difference ($p < 0.05$) in the y-intercept was observed among the three trials. Because this study seeks to compare several microscopic measurements to macroscopic ash measurements, the small variations between the three trials were anticipated. Regional heterogeneities could influence the standardization equations and all subsequent calculated wt.% Ca values. Precise determination of the standardization equations is essential for the proposed method to be successful. Furthermore, linear regression analysis is particularly susceptible to outliers [53 (pages 451-466)] making the measurement of the standardization equation more difficult. The use of more than ten data points for the standardization equation may help to reduce the potential influence of anomalous data and

---

improve the fit of data to a linear regression.

Nevertheless, the variation among the three trials is relatively small as is clearly seen in Fig. 4. The standard error of the mean is less than 0.9 in all bones except the turtle (Table 3) indicating good precision with 15 data points. A comparison of the precision observed in this study with microradiography would be beneficial, however, the authors have not found a similar study performed using microradiography. Some estimates of specific aspects of microradiographic precision have been reported. Several investigators using microradiography have reported the precision of specimen thickness measurements [20, 39, 45, 58], which typically introduce error on the order of 1-2 μm in a 100 μm thick specimen. Other investigators [36, 39, 40] have reported the precision of areal measurements, which are based upon graylevel density variations in microradiographs, to vary by 1%-4%. More recently, Martin *et al.* [34] used a computerized videodensitometry method to calibrate density variations in microradiographs.

They reported a 3% variation in the slope and a 5% variation in the y-intercept of the calibration line, but they did not report the variation in mineral content or compare results with accepted measurements of mineral content. The error introduced by these different sources makes it unlikely that microradiographic measurements of mineral content can be performed with substantially better precision than was obtained in this study.

The precision of the technique used in this study is within acceptable limits, however, the accuracy of the measurements is relatively poor. The y-intercept of the regression lines is consistently low with respect to the theoretical plot of wt.% HA vs. ash percent (Fig. 4). Analysis of covariance reveals a significant difference ($p < 0.01$) between the y-intercepts of each of the three trials with respect to the y-intercept of the theoretical correlation (Fig. 4). Consequently, this method consistently underestimates ash percent by an average of 14.5% (Fig. 4 and Table 3). Comparison with previous work by Roschger *et al.* [44] indicates that a similar underestimate existed in their study. Roschger *et al.* [44] reported that the histogram of wt.% HA in normal human bone ranged from 40%-70% HA with the peak near 57.3% HA. Visual examination of histograms and scatterplots published by Roschger *et al.* [44] suggest that the mean value may be even lower than 57.3%, although numerical values were not published. Previous studies using other analysis methods have found the mineral content in normal human bone to be much higher, ranging from 62%-70% ash by weight [4, 18, 56, 59], which is consistent with the ash values for human bone reported in this study.

The consistency of the deficit in this study suggests that an empirical correction could be used to adequately convert wt.% HA to ash percent. An equation of the form [ash % = wt.% HA + 14.5%] would allow the comparison of the current method with other accepted methods of measurement. The nature of the deficit is not understood, however, and it can not be assumed that the correction would be constant between microscopes, x-ray analyzers, and ZAF correction procedures. This would allow standardization to be achieved, but would require each laboratory to perform an analysis similar to the study reported here to determine the necessary empirical correction for their SEM and imaging equipment.

Clearly, the use of an empirical correction is not an optimal solution to the problem of BSE standardization. A better understanding of the observed deficit could potentially solve the problem. The study reported here attempts to compare EDX-determined wt.% HA measurements in embedded bone tissue to ash percent measurements in unembedded dry bone tissue. Dry bone tissue is composed primarily of collagen and HA. Obviously, after embedding bone, Ca as a function of the total weight (collagen + HA + PMMA) would decrease due to the addition of PMMA. This corresponds to the deficit observed. However, the consistent deficit in all bones of varied mineral content would not be anticipated. The weight fractions of the three major components of wet bone (HA, collagen, and H_2O) are nonlinearly related as a consequence of their different densities [33]. Similarly, the three major components of embedded bone (HA, collagen, and PMMA) would be expected to have a nonlinear relationship between weight fractions for the same reasons. As mineralization increased, this would result in a deviation of the regression slope between EDX-determined wt.% HA and ash percent, which was not observed in this study.

This argument, however, is based on the assumption that H_2O is replaced by PMMA in the embedding process. The extent to which this occurs is unknown, and the relationship between PMMA content and HA content is speculative. It should be observed that all three regression lines (Fig. 4) did have a slope greater than 1, although it was not statistically significant. Clearly, then the PMMA may be responsible for some or all of the observed deficit. Further studies are warranted. Even if the PMMA is responsible for the observed deficit, the problem of standardization still has not been solved. The use of a different embedding medium or slightly different polymerization techniques could change the nature of the deficit. This could potentially change the empirical correction between laboratories.

A second potential source of error in this study is the EDX measurement technique itself. Quantitative EDX analysis in thin biological specimens has been relatively well developed [31], but controversy still exists over the accuracy of quantitative EDX measurements in bulk biological samples [1, 9, 38, 43]. Payne and Cromey reported large errors in the EDX procedure [38]. They found no statistically significant differences in EDX-determined Ca/P molar ratios in a series of analytical standards with known Ca/P molar ratios ranging from 1.0-1.67. The authors attributed much of the error to inaccurate phosphorous measurements, implying that wt.% Ca measurements may be more reliable. A more recent comparative study by Åkesson *et al.* [1] found EDX measurements of wt.% Ca in bone to be slightly higher than measurements performed with neutron activation and chemical analyses. The authors provided plausible explanations for the minor differences, and reported values were consistent with previously published results [4]. It should be noted, however, that although good agreement was observed with respect to wt.% Ca measurements, Ca/P molar ratios were substantially higher than is typically observed [4, 60] in bone using chemical analysis methods suggesting potential errors. These previous

studies used different analysis methods and instrumentation making direct comparison with the present study difficult.

Due to limitations in the ZAF software, the methods employed by Roschger *et al.* [44] and duplicated in this study involve the substitution of one element, carbon, for the organic matrix in ZAF corrections. Boekestein *et al.* [9] report that this approach can lead to errors up to 15% and may explain the observed deficit. A more robust approach would allow the correct ratios between light elements. Yet, even this method would be problematic because the mass absorption coefficients for light elements have not been well characterized in the literature [23 (pages 503-504)]. Considerable errors could exist due to inaccurate ZAF corrections procedures. The present study is further complicated by the use of a thin layer of gold to make the bone electrically conductive. The influence of the conductive layer on EDX analysis is not well characterized and investigations into these aspects of EDX technology need to be addressed.

An alternative solution would involve the use of "windowless" or "thin" window technology, allowing for the direct measurement of low energy x-rays, corresponding to the light elements [23 (pages 346-347)]. Efforts in our laboratory to quantify the light elements with thin window EDX have been inconsistent (unpublished data). Thin window technology is still under development and may present the eventual solution to this dilemma, but at this point in time, reliable results have not been documented.

The study reported here confirms the findings of Roschger *et al.* [44]. BSE WMGLs in bone are highly correlated to EDX-determined wt.% Ca measurements. The strong positive relationship was observed regardless of SEM brightness and contrast settings. A consistent, highly linear correlation between EDX-determined wt.% HA measurements and ash percent was observed, suggesting EDX standardized BSE graylevels can be converted to ash measurements. However, EDX measurements consistently underestimated ash percent. Further investigation is needed to determine the nature of this deficit before EDX standardization of BSE images can be universally applied for the microscopic determination of mineral content in bone tissues.

Acknowledgments

The authors would like to acknowledge the Department of Veterans Affairs Medical Research Fund, Salt Lake City, UT, for their support of this project; Gwenevere Shaw for her assistance in manuscript preparation; Richard Kuenzler for his technical assistance; and Dr. Frank Whitmore (Smithsonian Institute), Dr. Dennis M. Bramble (University of Utah), and Lynn Barkley (Los Angeles County Museum of Natural History) for the donation of bone specimens.

References

1. Åkesson K, Grynpas MD, Hancock RGV, Odselius R, Obrant KJ (1994) Energy-dispersive X-ray microanalysis of the bone mineral content in human trabecular bone: A comparison with ICPES and neutron activation analysis. Calcif Tissue Int **55**: 236-239.

2. Amprino R, Engstrom A (1952) Studies on x-ray absorption and diffraction of bone tissue. Acta Anat **15**: 1-22.

3. Bachus KN, Bloebaum RD (1992) Projection effect errors in biomaterials and bone research. Cells and Mater **2**: 347-355.

4. Biltz RM, Pellegrino ED (1969) The chemical anatomy of bone. I. A comparative study of bone composition in sixteen vertebrates. J Bone Joint Surg [Am] **51-A**: 456-466.

5. Bloebaum RD, Bachus KN, Boyce TM (1990) Backscattered electron imaging: The role in calcified tissue and implant analysis. J Biomater Appl **5**: 56-85.

6. Bloebaum RD, Bachus KN, Rubman MH, Dorr LD (1993) Postmortem comparative analysis of titanium and hydroxyapatite porous coated femoral implants retrieved from the same patient. J Arthroplasty **8**: 203-211.

7. Bloebaum RD, Hofmann AA, Rubman MH (1991) Bone ingrowth into porous coated tibial components implanted with autograft bone chips: Analysis of eight consecutively retrieved implants. Trans Soc Biomat **XIV**: 99.

8. Bloebaum RD, Rhodes DM, Rubman MH, Hofmann AA (1991) Bilateral tibial components of different cementless designs and materials: Microradiographic, backscattered imaging, and histologic analysis. Clin Orthop **268**: 179-187.

9. Boekestein A, Stols ALH, Stadhouders AM (1980) Quantitation in X-ray microanalysis of biological bulk specimens. Scanning Electron Microsc 1980; II: 321-334.

10. Boyce TM, Bloebaum RD (1993) Cortical aging differences and fracture implications for the human femoral neck. Bone **14**: 769-778.

11. Boyce TM, Bloebaum RD, Bachus KN, Skedros JG (1990) Reproducible method for calibrating the backscattered electron signal for quantitative assessment of mineral content in bone. Scanning Microsc **4**: 591-603.

12. Boyde A, Elliott JC, Jones SJ (1993) Stereology and histogram analysis of backscattered electron images: Age changes in bone. Bone **14**: 205-210.

13. Boyde A, Jones SJ (1983) Back-scattered electron imaging of skeletal tissues. Metab Bone Dis Rel Res **5**: 145-150.

14. Boyde A, Jones SJ, Aerssens J, Dequeker J (1995) Mineral density quantitation of the human cortical iliac crest by backscattered electron image analysis: Variations with age, sex, and degree of osteoarthritis. Bone **16**: 619-627.

15. Currey JD (1984) Effects of differences in mineralization on the mechanical properties of bone. Philos Trans R Soc Lond (Biol) **B 304**: 509-518.

16. Currey JD (1988) The effect of porosity and mineral content on the Young's modulus of elasticity of compact bone. J Biomech **21**: 131-139.

17. Currey JD (1990) Physical characteristics affecting the tensile failure properties of compact bone. J Biomech **8**: 837-844.

18. Dickenson RP, Hutton WC, Stott JRR (1981) The mechanical properties of bone in osteoporosis. J Bone Joint Surg [Br] **63-B**: 233-239.

19. Emmanual J, Hornbeck C, Bloebaum RD (1987) A polymethyl methacrylate method for large specimens of mineralized bone with implants. Stain Technol **62**: 401-410.

20. Eschberger J, Eschberger DJ (1986) Microradiography. In: Handbook of Biomaterials Evaluation. Scientific, Technical, and Clinical Testing of Implant Materials. von Recum AF (ed). Macmillan Publishing Company, New York, 491-500.

21. Everhart TE, Herzog RF, Chang MS, DeVore WJ (1972) Electron energy dissipation measurements in solids. In: Proceedings of the 6th International Conference on X-Ray Optics and Microanalysis. Shinoda G, Kohra K, Ichinokawa T (eds). University of Tokyo Press, Tokyo, Japan, 81-86.

22. Fratzl P, Schreiber S, Roschger P, Lafage M-H, Rodan G, Klaushofer K (1996) Effects of sodium fluoride and alendronate on the bone mineral in minipigs: A small-angle X-ray scattering and backscattered electron imaging study. J Bone Miner Res **11**: 248-253.

23. Goldstein JI, Newbury DE, Echlin P, Joy DC, Romig AD, Jr., Lyman CE, Fiori C, Lifshin E (1992) Scanning Electron Microscopy and X-ray Microanalysis. A Text for Biologists, Materials Scientists, and Geologists. Plenum Press, New York.

24. Grynpas MD, Holmyard D (1988) Changes in quality of bone mineral on aging and in disease. Scanning Microsc **2**: 1045-1054.

25. Grynpas MD, Holmyard DP, Pritzker KPH (1994) Bone mineralization and histomorphometry in biopsies of osteoporotic patients treated with fluoride. Cells and Mater **4**: 287-297.

26. Hatzakis M (1971) New method of observing electron penetration profiles in solids. Appl Phys Les **18**: 7-10.

27. Heinrich KFJ (1966) Electron probe microanalysis by specimen current measurement. In: Proceedings of the Fourth International Congress X-Ray Optics and Microanalysis. Hermann, Paris, 159-167.

28. Hofmann AA, Bachus KN, Bloebaum RD (1993) Comparative study of human cancellous bone remodeling to titanium and hydroxyapatite coated implants. J Arthroplasty **8**: 157-166.

29. Holmes RE, Hagler HK, Coletta CA (1987) Thick-section histometry of porous hydroxyapatite implants using backscattered electron imaging. J Biomed Mater Res **21**: 731-739.

30. Howell PGT, Boyde A (1994) Monte Carlo simulations of electron scattering in bone. Bone **15**: 285-291.

31. Landis WJ (1979) Application of electron probe X-ray microanalysis to calcification studies of bone and cartilage. Scanning Electron Microsc 1979; II: 555-570.

32. LeGeros RZ (1981) Apatites in biological systems. Prog Crystal Growth Charact **4**: 1-45.

33. Lowenstam H, Weiner S (1989) On Biomineralization. Oxford University Press, New York, 162-163.

34. Martin RB, Papamichos T, Dannucci GA (1990) Linear calibration of radiographic mineral density using video-digitizing methods. Calcif Tissue Int **47**: 82-91.

35. Mason MW, Skedros JG, Bloebaum RD (1995) Evidence of strain-mode-related cortical adaptation in the diaphysis of the horse radius. Bone **17**: 229-237.

36. McQueen CM, Smith DA, Monk IB, Horton PW (1973) A television scanning system for the measurement of the spatial variation of micro-density in bone sections. Calcif Tissue Res **11**: 124-132.

37. Mechanic GL, Arnaud SB, Boyde A, Bromage TG, Buckendahl P, Elliott JC, Katz EP, Durnova GN (1990) Regional distribution of mineral and matrix in the femurs of rats flown on Cosmos 1887 biosatellite. FASEB J **4**: 34-40.

38. Payne CM, Cromey DW (1990) Limitations of ZAF correction factors in the determination of calcium/phosphorus ratios: Important forensic science considerations relevant to the analysis of bone fragments using scanning electron microscopy and energy-dispersive X-ray microanalysis. J Forensic Sci **35**: 560-568.

39. Phillips HB, Owens-Jones S, Chandler B (1978) Quantitative histology of bone: a computerized method of measuring the total mineral content of bone. Calcif Tissue Res **26**: 85.

40. Pugliese LR, Anderson C (1986) A method for the determination of the relative distribution and relative quantity of mineral in bone sections. J Histotechnol **9**:

91-93.

41. Reid SA, Boyde A (1987) Changes in the mineral density distribution in human bone with age: Image analysis using backscattered electrons in the SEM. J Bone Miner Res **2**: 13-22.

42. Robinson VNE (1987) Materials characterization using the backscattered electron signal in scanning electron microscopy. Scanning Microsc **1**: 107-117.

43. Roomans GM (1981) Quantitative electron probe X-ray microanalysis of biological bulk specimens. Scanning Electron Microsc 1981; II: 345-356.

44. Roschger P, Plenk H, Jr., Klaushofer K, Eschberger J (1995) A new scanning electron microscopy approach to the quantification of bone mineral distribution: Backscattered electron image grey-levels correlated to calcium Kα-line intensities. Scanning Microsc **9**: 75-88.

45. Rowland R, Jowsey J, Marshall J (1959) Microscopic metabolism of calcium in bone. III. Microradiographic measurements of mineral density. Radiat Res **10**: 234-242.

46. Sanderson C (1995) Polymerization of mineralized bone specimens embedded in methylmethacrylate using ultraviolet irradiation. J Histotechnol **18**: 323-325.

47. Sanderson C, Kitabayashi LR (1994) Parallel experience of two different laboratories with the initiator perkadox 16 for polymerization of methylmethacrylates. J Histotechnol **17**: 343-348.

48. Shimizu R, Ikuta T, Everhart TE, DeVore WJ (1975) Experimental and theoretical study of energy dissipation profiles of keV electrons in polymethylmethacrylate. J Applied Physics **26**: 1581-1584.

49. Skedros JG, Bloebaum RD, Bachus KN, Boyce TM (1993) The meaning of graylevels in backscattered electron images of bone. J Biomed Mater Res **27**: 47-56.

50. Skedros JG, Bloebaum RD, Bachus KN, Boyce TM, Constantz B (1993) Influence of mineral content and composition on graylevels in backscattered electron images of bone. J Biomed Mater Res **27**: 57-64.

51. Skedros JG, Mason MW, Bloebaum RD (1994) Differences in osteonal micromorphologies between tensile and compressive cortices of a bending skeletal system: Indications of potential strain-specific differences in bone microstructure. Anat Rec **239**: 405-413.

52. Skedros JG, Mason MW, Nelson MC, Bloebaum RD (1996) Evidence of structural and material adaptation to specific strain features in cortical bone. Anat Rec **246**: 47-63.

53. Sokal RR, Rohlf FJ (1995) Biometry. The Principles and Practice of Statistics in Biological Research. Wilson J, Cotter S (eds). W. H. Freeman & Co., New York.

54. Sumner DR, Bryan JM, Urban RM, Kuszak JR (1990) Measuring the volume fraction of bone ingrowth: A comparison of three techniques. J Orthop Res **8**: 448-452.

55. Torontali M, Koo I, Holmyard DP, Tomlinson G, Grynpas MD, Tenenbaum HC (1994) Backscattered electron image assessment of mineral density in bone formed *in vitro*. Cells and Mater **4**: 125-134.

56. Trotter M, Hixon BB (1974) Sequential changes in weight, density, and percentage ash weight of human skeletons from an early fetal period through old age. Anat Rec **179**: 1-18.

57. Vajda EG, Skedros JG, Bloebaum RD (1995) Consistency in calibrated backscattered electron images of calcified tissues and mineral analyzed in multiple imaging sessions. Scanning Microsc **9**: 741-755.

58. Vose GP (1962) Quantitative microradiography of osteoporotic compact bone. Clin Orthop **24**: 206-212.

59. Vose GP, Kubala ALJ (1959) Bone strength -- its relationship to X-ray determined ash content. Hum Biol **31**: 261-270.

60. Woodard HQ (1964) The composition of human cortical bone. Effect of age and some abnormalities. Clin Orthop **37**: 187-193.

Discussion with Reviewers

G.M. Roomans: The main reason for the "deficit" found is the presence of the PMMA. The authors could have investigated this by "tagging" the PMMA with an element not present in bone, e.g., an organic brominated compound. In this way, the PMMA content of the bone specimen could have been quantitatively determined. This content could then have been used to correct the data. I agree with the authors that the presence of an embedding medium probably makes it impossible to have a "universal" correction factor. Not only the embedding medium can vary, but penetration may vary even with the same embedding medium.

The authors mention as another potential source of error the use of C instead of N as matrix in the ZAF correction. I doubt whether for Ca, which is relatively little absorbed in either C or N the effect would amount to 15%. However, the authors could easily have checked the potential magnitude of the effect.

In my opinion, comparison of embedded specimens analyzed by EDX and/or BSE with unembedded specimens analyzed by bulk chemical methods poses a major, if not unsurmountable, limitation to what the authors are attempting to do. Although the use of unembedded specimens for EDX/BSE certainly will meet with considerable difficulties, the possibility should be considered. For EDX analysis, quantitative methods

based on peak-to-background (P/B) ratios rather than ZAF-corrections, should be used. It could also be considered to use frozen-hydrated specimens where a flat surface might be prepared. This may be a direction for future work.

Authors: We are currently investigating the influence of the matrix in ZAF corrections and the influence of PMMA on BSE/EDX measurements. Dr. Roomans has suggested a useful method for investigating the PMMA content of the embedded tissue which should be considered. Clearly, further research is needed. The routine use of unembedded specimens for BSE is not advisable, however, as this makes it extremely difficult to eliminate surface topography and limit the contamination and damage to the bone tissue. This surface topography may subsequently influence the BSE signal. Frozen-hydrated specimens are a potential solution if, in fact, PMMA is the limiting factor in standardization of BSE images.

G.M. Roomans: No commercial ZAF correction is equipped to deal with the extra complications introduced by a conductive coating in the various corrections forming part of ZAF. Likewise, the backscatter measurements must be affected by this gold layer, though probably not in the same way as the ZAF correction.

J. Mitchell: The use of a gold coating is a serious issue with this work. The gold conductive layer will attenuate the incoming electron energies and cause a much greater scatter than a coating of lower atomic material, such as carbon. Additionally, the gold film layer will tend to absorb the x-rays generated within the sample. This will not be a uniform circumstance, as the absorption will occur as a function of [exp(coating thickness)]. It is not indicated whether the authors were able to measure the thickness of the coating layer. Unless their quantitative standards were coated with an identical thickness of gold as all of the samples, the electron absorption and x-ray fluorescence will contribute uncontrollable errors to their results. In fact, this may well be the source of their technique underestimating the mineralized percentages.

There is yet another area of concern with the coating which the authors selected to use. In the collection of backscattered images from gold-coated samples, the major elastic electron scattering mechanism will not be the differences in the atomic number of the bone regions, but rather be a complex mixture of topographic contrast and coating thickness contrast.

Authors: As the reviewers have suggested, there are theoretical considerations that would suggest the use of a low atomic number conductive coating would be preferable in BSE analysis. Typically, carbon is used. However, both carbon and gold will influence the backscatter signal [64]. In our laboratory, we have tried to not only improve and quantify BSE technology, but also to make the procedure as simplistic as possible, allowing the technique to be widely applied to bone biology. Gold coating is less expensive and easier to work with than carbon - and this is the rationale for its use. In practice, we have experimentally demonstrated that it has negligible influence on BSE WMGLs of bone tissue and quantitative measurements of mineral content can still be performed [49, 50]. This is likely due to the extremely thin layers of material necessary to make the specimen electrically conductive.

It is correct, however, that the ZAF corrections used in the present study do not accurately control for surface coating. The calculations are based on the assumption that the material is homogeneously dispersed within the sampled volume. This problem would exist regardless of the conductive coating, although it may be anticipated that the errors would be more pronounced with a heavy element such as gold. The magnitude of the errors contributed to the technique due to conductive coatings and inaccurate ZAF corrections are unknown and need to be investigated. However, we feel it is unlikely that the 14.5% deficit is solely the result of a gold coating. As mentioned in the **Discussion** section, a similar deficit was observed in carbon coated specimens in the study by Roschger *et al.* [44]

P. Roschger: Can it be assumed that the bone ash is identical with HA or native bone mineral in its chemical composition? The calculation of wt.% HA from EDX-data assumes perfect HA-crystal, while physiologic crystals contain some carbonate groups replacing phosphate. What is the influence of this effect on the wt.% HA?

J. Mitchell: The authors make the assumption that the mineral composition of the bones used in the study was entirely hydroxyapatite. While it is known that this is the eventual major mineral deposit in bone, it is widely believed that the first deposits of mineral are amorphous $Ca_3(PO_4)_2$, and that many other Ca/P ratios may exist at different maturity stages. Can they give evidence (e.g. XRD data) to support their extrapolation of all of the diverse bone types all to 100% HA?

Authors: It is correct that bone mineral is not pure HA. Bone mineral is dahllite [33], a carbonated, calcium-deficient form of HA. Experimental data [4] indicates that a conversion factor of approximately 2.6 would convert wt.% Ca measurements to dahllite, although small variations exist between species. The assumption that bone is composed of HA, and the subsequent use of a conversion factor of 2.51, does not substantially improve or degrade the results reported in the present study. Similarly, while Ca/P ratios may change with bone maturity, the mineral of a mature bulk

specimen will be very close to HA [4]. The assumptions made in the present study with respect to mineral composition will undoubtedly introduce small errors. However, the errors are minimal in comparison with the large deficit observed between wt.% HA and ash fraction.

P. Roschger: The ash weight was determined using specimens from a slightly different bone region than the EDX-data. How large is the variability of mineral content within the bone sections used?
Authors: The standard deviations of the ash fraction data are included in Table 3. Typically, the values were very consistent within each bone section. The largest range observed in any one specimen occurred in the whale tympanic bulla (range: 84.00% ash to 87.57% ash).

Reviewer V: The images demonstrate the problems of excess penetration at edges such as cracks, and possible also charging: both are generic edge brightness artifacts which are unrelated to tissue composition. Further, their images show variations in brightness in the line scan direction, indicating system and/or sample instability. For a backscattered electron image to be measured, we must be sure that all instrumental and sample related artifact has been removed.
Authors: The reviewer is quite right that excess penetration at edges can not be interpreted as highly mineralized regions. This could be minimized by reducing accelerating voltage and electron dosage. Unfortunately, to achieve adequate overvoltage and electron dose to accurately measure EDX, this is not possible and edge effects will persist. Nonetheless, the strong positive correlation observed suggest that edge effects have had a minimal influence in the present study. We must politely disagree with the reviewer's suggestion that graylevel variations in the line scan direction exist in Figure 2a and 2b. The small variations in graylevel are the result of the rollers used to distribute the photographic emulsion on the Polaroid® film and do not appear in the digital images. As described in this paper, digital images were used for all analyses, and the photomicrographs are only used for publication purposes.

Reviewer V: From my experience, the use of 0.75nA at 30kV is at the upper end of the range of possibilities for most PMMA embedded bone samples. It will cause substantive damage to many samples.
Authors: The reviewer is mistaken. As stated in the **Materials and Methods**, a probe current of 2.5nA and an accelerating voltage of 20kV were used in the present study. While these settings will also undoubtedly cause specimen damage, a decrease in probe current will result in inadequate x-ray sampling for accurate EDX measurements. Similarly, a substantial decrease in accelerating voltage would provide inadequate overvoltage to sufficiently generate calcium x-rays. There is presently no adequate method to completely eliminate specimen damage. However, the strong correlation observed in the present study indicate that specimen damage was not the source of the observed error.

Reviewer V: Checking on instrument stability once every 20 minutes is probably not often enough.
Authors: We have found that our electron microscope is stable for 20 minute intervals as previously published [49, 50, 57].

Reviewer V: The SEM used in this study does not have a proper digital scan generator. This means that the beam is moving whilst data is acquired, and this defeats the fundamental stereological requirement that the measurement of the condition of the sample is made at a single point at a time, and that the points are discrete and do not overlap.
Authors: The reviewer is correct that it would be preferable to have non-overlapping regions sampled by the electron beam, although it is not a "fundamental stereological requirement". Overlapping regions leads to a degradation of resolution, which becomes more severe as the overlapping increases. However, the scanning mechanism used by this SEM is irrelevant with respect to the overlapping of sampled points. At 20 kV, the interaction volume of the electron probe in a low atomic number sample such as bone may be on the order of 5 μm or more in diameter [30, 61]. Therefore, to obtain a high magnification image (1000 X - approximately 80 μm x 80 μm) with 512 x 512 discrete pixels as shown in Figure 2b, necessitates an overlap of sampled regions due to the broadening of the electron beam within the bulk specimen, regardless of the scan generator.

Reviewer V: 93% Mg, 6% Al, 1% Zn has a mean atomic number calculated by the method of Lloyd [62] of 12.54. It is not Mg, 12, and it differs by not too much from Al, 13. Values for bone lie in the range 10.2 to 11.7. Pure Al is very difficult to polish. These peculiarities may be added to the list of reasons not to accept the Mg alloy and Al as calibration standards, which include channeling contrast in crystalline materials. Please comment.
Authors: We have been able to observe sufficient contrast to differentiate between Al and Mg alloy as previously published [57]. The reviewer may have misunderstood the use of Mg alloy and Al as calibration

standards in the present study. As stated in the **Introduction**, metal calibration standards are used to insure that the SEM operating environment remains consistent during imaging. For image calibration, it is not required that two different pieces of Al (or Mg alloy) have the identical backscatter signal. The only requirement is that the backscatter signal remains consistent in repeated imaging of the same field. This has been experimentally demonstrated in our previous publication [57]. Channeling contrast and polishing effects may cause variations between pure metals and this would be a reason to avoid their use to *standardize* the BSE signal, but they would be perfectly adequate to *calibrate* the BSE signal. This is the rationale for the current study as well as the study originally performed by Roschger *et al.* [44].

Reviewer V: The existence of regional variations in mineral content in bone is documented in a large literature and references to this should have been provided.

Authors: We are aware of the regional mineral variations in bone. Several papers from our laboratory have been published investigating mineral variations [35, 63]. We have attempted to minimize the influence of regional mineral variations by measuring ash content in two specimens immediately adjacent to the sampled BSE region as described in **Materials and Methods**. The mean value of these two ash measurements is the reported value which should closely correlate with the mineralization of the region used for BSE analysis. The goal of this study, however, was not to determine regional mineral variations, but to standardize the BSE signal. We have not, therefore, performed an extensive literature review of regional mineral variations.

Reviewer V: Please describe in detail what your weighted mean graylevel (WMGL) is as it is of such fundamental importance to the whole paper.

Authors: The calculation of the weighted mean graylevel has been previously described [5, 11, 57]. It is calculated using the following equation:

$$WMGL = \sum_{1-6}^{255} \frac{AiGLi}{At}$$

where: A_i = area of ith graylevel, GL_i = ith graylevel, and A_t = total area imaged. This provides a mean value for the backscattered signal, independent of porosity (black pixels 0-5 in the BSE image).

Additional References

61. Murata K (1974) Spatial distribution of back-scattered electrons in the scanning electron microscope and electron microprobe. J Appl Phys **45**: 4110-4117.

62. Lloyd GE (1987) Atomic number and crystallographic contrast images with the SEM: a review of backscattered electron techniques. Mineralogical Magazine **51**: 3-19.

63. Skedros JG, Bloebaum RD, Mason MW, Bramble DM (1994) Analysis of a tension/compression skeletal system: Possible strain-specific differences in the hierarchial organization of bone. Anat Rec **239**: 396-404.