Scanning Microscopy, Vol. 9, No. 3, 1995 (Pages 741-755)
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

0891-7035/95$5.00+.25

CONSISTENCY IN CALIBRATED BACKSCATTERED ELECTRON IMAGES OF CALCIFIED TISSUES AND MINERALS ANALYZED IN MULTIPLE IMAGING SESSIONS

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

Bone and Joint Research Laboratory, Veterans Adminstration Medical Center (151F)
500 Foothill Blvd., Salt Lake City, UT 84148

(Received for publication April 26, 1995 and in revised form August 17, 1995)

Abstract

Pure metal standards have been used to calibrate the operating environment in quantitative backscattered electron (BSE) imaging of mineralized tissue, allowing comparisons to be made between various mineralization states of bone at the microscopic level. It has not previously been documented that calibration procedures produce consistent, reliable results over multiple imaging sessions. In this study, BSE images were obtained from bones, pure metals, and a naturally occurring mineral in multiple imaging sessions over a six day period. The graylevel histogram profile (GHP) from each specimen was analyzed for changes in the shape and relative placement on the graylevel spectrum. Computer controlled calibration and a retrospective calibration method using pure aluminum and pure magnesium-aluminum-zinc demonstrated consistency between imaging sessions. Calibrated weighted mean graylevels (WMGLs) for biological materials had an average standard deviation of 5.9 graylevels (2.4% variation) during the course of the study. WMGLs for inorganic materials had an average standard deviation of 0.9 graylevels (0.4% variation). A trend towards increased image brightness, due to specimen and/or embedding media degradation, was observed in the biological tissues. No increase in brightness was observed for the inorganic specimens. Kurtosis and skewness tests revealed a slight deviation from normality in all specimens, which remained consistent between multiple imaging sessions. These results demonstrate that BSE image analysis of bones and mineral can be calibrated with negligible precision error allowing comparisons between data within and between multiple imaging sessions.

Key Words: Backscattered electron imaging; bone; mineral quantitation; graylevel calibration; calibration standard.

*Contact for correspondence:
Roy D. Bloebaum, address as above.
Telephone number: (801) 582-1565 ext. 4607
FAX number: (801) 584-2533

Introduction

Boyde and Jones [9] introduced the use of the scanning electron microscope (SEM) in the backscatter mode as a tool for the qualitative study of microscopic mineralization differences in bone and dental tissues. Reid and Boyde [30] later developed a technique using backscattered electron (BSE) imaging for the quantitative analysis of bone mineral content. Boyce *et al.* [5] further advanced quantitative BSE image analysis by calibrating the SEM operating environment using pure elements as calibration standards, enabling the comparison of images captured at different times. The development of techniques to calibrate BSE images has allowed bone and mineral researchers to quantify relative differences in microscopic mineral density variations in mineralized tissue [2, 4, 11, 23, 34, 35, 39]. BSE imaging technology has also been used to study both mineral content variations and morphological parameters in cancellous bone [23, 36, 37].

Since BSEs interact within only a few micrometers of a typical bone specimen surface [24, 25], BSE images of bone have remarkably high resolution and contrast [1]. The graylevel intensity of a pixel in a digitized BSE image of bone is directly proportional to the backscatter coefficient η, which is primarily a function of the mean atomic number of the sampled electron interaction volume [22, 34]. Experimental studies have shown that as a result of atomic number contrast, mean graylevel variations in BSE images of bone are strongly correlated to mineral content (volume/volume, v/v; and weight/weight; w/w) and density (g/cc) [34, 35].

In many of the above applications of BSE imaging of bone, practical limitations often necessitate imaging over extended periods of time in several operating sessions. Large scale studies involving numerous samples [4, 11, 23] can require several days or even weeks. Calibration procedures that use pure standards were designed to enable researchers to maintain consistency in the BSE/SEM operating environment within and between imaging sessions. However, a review of the literature showed that the study of Boyce *et al.* [5] is the only one

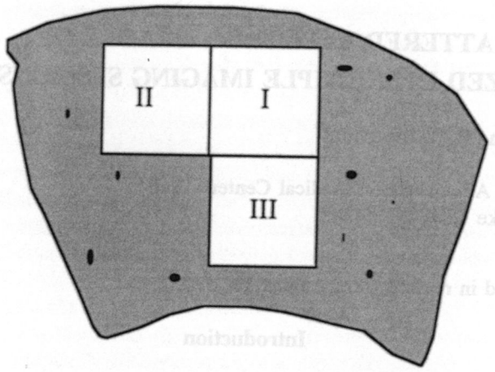


Figure 1. Graphic example of three adjacent fields captured with the SEM in a bone specimen. The three fields were assigned by location into one of three groups (I, II, III). The same pattern of three fields was captured from each of the seven specimens used in this study.

---

that experimentally demonstrates that pure materials can be used to calibrate the BSE/SEM operating environment with minimal precision error. They demonstrated that the relative placement along the graylevel spectrum of the graylevel histogram profile (GHP) of pure magnesium and aluminum standards could be maintained to within 0.8% error between 5 imaging sessions that were conducted at different times on the same day. Visual inspection showed that the shape of the profile remained consistent throughout the 5 imaging sessions.

Although Hofmann *et al.* [23] used this method to calibrate imaging sessions on different days, and both Torontali *et al.* [39] and Grynpas *et al.* [21] used pure standards and a modified technique to calibrate different imaging sessions, these investigators did not report the precision errors of their methods. Boyde *et al.* [7] consecutively re-imaged the same bone fields to calibrate images obtained during the same operating session, but did not report precision error. Documentation of precision error is an important step in validating the utility of any calibration technique, however, accomplishing this is not a straight-forward task in atomic-number-contrast BSE imaging. Error in these calibration methods can be the result of both brightness and contrast variations that are associated with surface contamination, filament fatigue, beam current changes, and other variables [2, 5]. Consequently, calibration of BSE images based on graylevels must include documentation that both brightness and contrast remain consistent [5]. This can be accomplished by verifying that both the size and shape, and relative placement of the GHPs of the calibration standards and experimental specimens remain consistent. It is not clear if these calibration procedures can maintain, with acceptable error, the consistency of the BSE/SEM operating environment between multiple imaging sessions on multiple days.

The objective of the present study is to use a modification of the calibration procedure of Boyce *et al.* [5] to test the following hypotheses: (1) sequential calibrations can be performed on pure metal standards with acceptable error in the size, shape, and relative placement of the GHP along the graylevel spectrum, between multiple imaging sessions on multiple days, and (2) acceptable consistency in the size, shape, and relative placement along the graylevel spectrum of mineral and bone tissue GHPs can be maintained in multiple calibrated imaging sessions on multiple days.

Materials and Methods

Specimen preparation

Both biological and inorganic specimens were obtained for this study. Biological specimens included: (1) bone from the femoral mid-diaphysis of a 17-day-old embryonic white leghorn chick; (2) cortical bone from the mid-femoral diaphysis of a previously healthy male human donor, 45 years of age; (3) cortical bone from the cranial ("compression") cortex at the proximal end of the calcaneus of a skeletally mature Rocky Mountain mule deer (*Odocoileus hemionus hemionus*); and (4) cortical bone from a mature antler of male Rocky Mountain mule deer. These mineralized tissues were selected because details of their remodeling rates [16, 27 (pages 214-217), 35, 36, 37] are known and they represent a broad range of mineralization [12]. Each bone specimen was manually cleaned of periosteum, where applicable, and other associated soft tissues. Marrow removal from porous spaces in the bone was facilitated with a high pressure stream of tap water (Water-Pik®, Teledyne Water Pik, Fort Collins, CO). The specimens were defatted by soaking for 20 days in a large volume of reagent-grade chloroform (Omnisolv, EM Industries, Inc., Gibbstown, NJ) under vacuum with constant stirring [15]. Residual chloroform was removed by placing the specimens in an 80°C oven for 3 days. The defatted tissues were then embedded in polymethyl methacrylate (PMMA) [13, 33] using conventional methods. After polymerization, the embedded tissues were cut into cubes with a band saw and glued together with a cyanoacrylate glue.

The inorganic materials included: (1) 99.8% pure magnesium-aluminum-zinc wire (93% Mg, 6% Al, 1% Zn) and 99.9999% pure aluminum wire (Johnson Matthey, Inc., Seabrook, NH), and (2) a naturally occurring feldspar (nominally: $KAlSi_3O_8$) crystal fragment (courtesy of F. DeCourten, University of Utah,

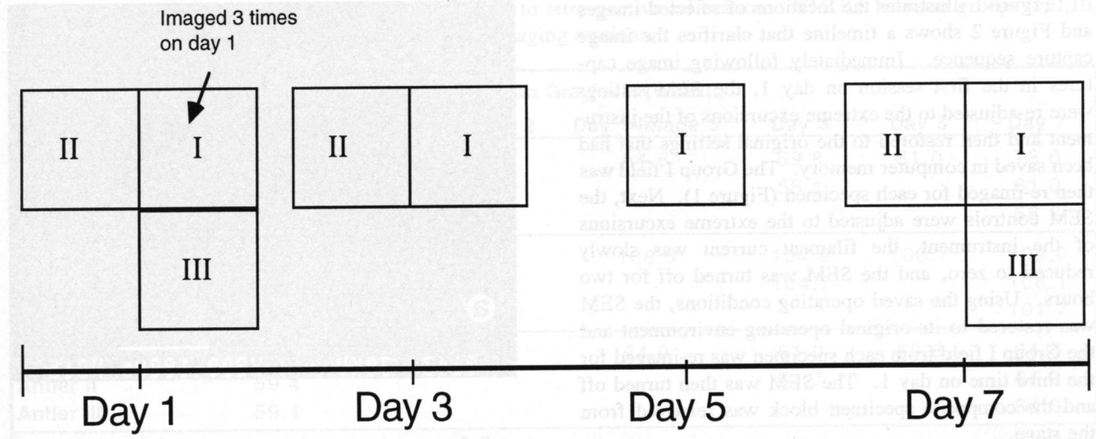


Figure 2. Time-line of captured images. The group numbers (I = Group 1; II = Group 2; and III = Group 3) refer to the image capture pattern for each specimen (see text and Fig. 1).

Department of Geology). For brevity, magnesium-aluminum-zinc will be referred to as magnesium. Cyanoacrylate glue was used to attach the inorganic materials directly to the block of embedded bone specimens.

The composite specimen block containing both the biological and inorganic materials was glued with an epoxy resin into a groove that had been previously milled in a Plexiglas® block. Following the protocol of Bloebaum *et al.* [2], the composite specimen block was milled, ground, and then buffed to achieve an optical finish. The block was lightly sputter coated with gold at 70 μm Hg and 10 mA for 75 seconds (Hummer Model VI-A Sputtering System, Anatech, Alexandria, VA).

Back scattered electrons (BSE) image capture

The composite specimen block was placed on the stage of a JEOL 6100 SEM (JEOL USA, Inc., Peabody, MA), with the polished surface perpendicular to the incident electron beam. Operating conditions of the SEM included an accelerating voltage of 30 kV, probe current of 0.750 nA, and working distance of 15 mm. BSEs were collected using a four quadrant, annular ring semiconductor BSE detector (Tetra, Oxford Instruments, Cambridge, UK). BSE images were captured onto magnetic media by a computer-controlled image capture and analysis system (eXL, Oxford Instruments) which allowed for image analysis at a later time. Image analysis was performed with a public domain computer software program (Image 1.55; copy available from Wayne Rasband, National Institutes of Health, Bethesda, MD, 20892; E-mail address: wayne@helix.nih.gov).

Three adjacent, non-overlapping 200X fields were imaged from each of the biological and inorganic specimens (Figure 1), for a total of 21 images. This approximately corresponds to a 450 μm x 450 μm square region analyzed in each image capture. Each field captured was assigned by location into one of three groups (I, II, or III) as depicted in Figure 1. The stage positions were recorded to allow the same field to be analyzed during different imaging sessions. The mechanized stage allowed for stage control accurate to $\pm$0.001 mm. Images were captured using nine slow scans and a Kalman frame averaging technique (eXL, Oxford Instruments) to increase the signal to noise ratio. Scanning time for each image capture was approximately 8 seconds. SEM settings, including brightness, contrast, objective lens strength, and condenser lens strength were computer controlled to maintain consistent BSE/SEM operating conditions.

Prior to each image capture, operating conditions stored in computer memory were restored to their original settings. Computer-controlled settings included: brightness, contrast, stage position, condenser lens strength, objective lens strength, and spot size. Additionally, an external probe current detector (model 485 autoranging picoammeter, Keithley Instruments, Inc., Cleveland, OH) was monitored to ensure that the probe current (corresponding to the number of incident electrons) remained consistent to within $\pm$ 0.001 nA. Prior to every image captured, probe current was measured and any deviation in probe current was manually corrected by fine adjustments in the condenser lens strength. The probe current has been observed to remain consistent, with only minor fluctuations, for at least twenty minutes [5]. This is well within the scanning time required for any individual image to be captured. This allowed for rapid on-line calibration of the BSE/SEM environment.

Figure 1 illustrates the locations of selected images and Figure 2 shows a timeline that clarifies the image capture sequence. Immediately following image captures in the first session on day 1, the SEM settings were re-adjusted to the extreme excursions of the instrument and then restored to the original settings that had been saved in computer memory. The Group I field was then re-imaged for each specimen (Figure 1). Next, the SEM controls were adjusted to the extreme excursions of the instrument, the filament current was slowly reduced to zero, and the SEM was turned off for two hours. Using the saved operating conditions, the SEM was restored to its original operating environment and the Group I field from each specimen was re-imaged for the third time on day 1. The SEM was then turned off and the composite specimen block was removed from the stage.

Two days later (Day 3) the specimen was placed on the stage and positioned using marks previously scribed on the specimen block. The SEM was again restored to its original operating conditions and the fields from Group I and Group II were re-imaged for each specimen. The SEM was turned off and an additional two days later (Day 5, or 4 full days after the first imaging session), the fields from Group I were re-imaged for each specimen using the saved operating conditions. On the final (6th) day of the experiment, the fields from all three groups were re-imaged for each specimen.

Backscattered electrons (BSE) image analysis

Pilot work showed that the aluminum and magnesium wires were too small to allow for the capture of three non-overlapping fields. Several of the captured images, therefore, included the interface between the metals and the PMMA. A noticeable decrease in pixel brightness of the metal near the interface was observed. The change in pixel intensity is likely the result of the altered electron interaction volume at the metal/PMMA [17 (pages 84-85, 92-93)] interface or possible topographical effects [25] due to the discontinuous nature of the interface. The metal/ PMMA interface was avoided by selecting analysis regions that were small enough to exclude the metal/ PMMA boundary on all captured images which included the interface. Additionally, on day seven, a large surface defect appeared on the feldspar specimen in region II. The cause of the defect was unknown and the area of the analysis region was modified to eliminate the contaminated area. For the remainder of the specimens, the analysis region consisted of the entire captured image.

The method of Bloebaum *et al.* [2] is used to calculate the weighted mean graylevel (WMGL) of each captured image. To reduce the time required to accomplish one analysis, this method, as originally described, sorts

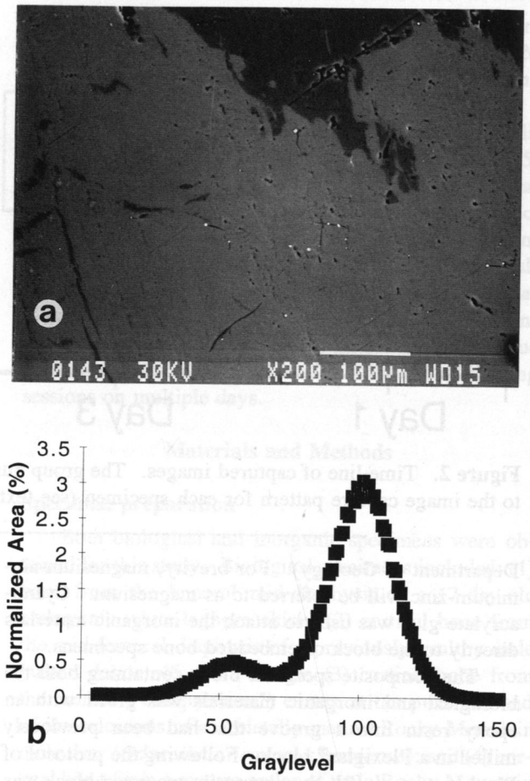


Figure 3. (a) BSE photomicrograph of feldspar and (b) the GHP generated from this image. The multiphasic nature of the feldspar is clearly seen in the photomicrograph and is seen in the GHP as a bimodal distribution.

---

the pixel graylevel data into 51 graylevel bins. The advanced analysis system used in the present study allowed all 256 bins to be used in the WMGL calculation:

$$WMGL = \sum_{i=6}^{255} \{(A_i GL_i)/A_t\} \quad (1)$$

where: A_i = area of ith graylevel, GL_i = ith graylevel, and A_t = total area imaged. WMGLs serve as indications of the relative placement of the graylevel histogram along the graylevel spectrum. GHPs were obtained from the graylevel frequency distribution of the image. Since alterations in the size and shape of the GHP are strongly influenced by image contrast [5], the GHP of each BSE image was analyzed for skewness and kurtosis [38]. Statistical analysis was performed with a commercially available computer software program (Number Cruncher Statistical System 5.02, Kaysville, UT). Each

Table 1. Calculated WMGLs for each specimen prior to retrospective calibration. WMGLs between groups (I,II,III) appear consistent. WMGLs are consistent between imaging sessions.

	Weighted Mean Graylevel (WMGL)					
	Day 1-Image 1	Day 1-Image 2	Day 1-Image 3	Day 3	Day 5	Day 7
Chick I	58.5	64.7	64.8	59.8	61.0	73.0
Chick II	57.7			58.1		61.0
Chick III	55.2					58.3
Human I	104.4	107.3	108.5	104.6	103.7	109.0
Human II	104.5			104.7		108.1
Human III	100.5					101.7
Antler I	58.0	63.4	64.0	61.0	58.8	64.6
Antler II	59.4			61.7		64.9
Antler III	59.1					64.8
Mule Deer I	111.9	119.3	119.7	115.3	115.5	117.1
Mule Deer II	113.3			115.3		116.5
Mule Deer III	113.8					117.2
Aluminum I	184.7	185.3	186.5	182.0	179.4	184.6
Aluminum II	190.8			186.5		188.9
Aluminum III	186.2					185.5
Magnesium I	148.3	147.2	146.9	144.4	142.6	144.3
Magnesium II	152.8			148.5		148.5
Magnesium III	150.8					146.8
Feldspar I	151.9	154.0	153.3	147.8	149.9	152.0
Feldspar II	153.8			151.1		153.9
Feldspar III	154.2					154.7

GHP was also visually examined for evidence of gross deviations from a unimodal, continuous distribution. Figure 3 displays the relationship between a BSE photomicrograph (Fig. 3a) and its associated GHP (Fig. 3b).

In previously published work, pure metal standards have been used as calibration standards [5, 21, 23, 34, 39] to compensate for fluctuations in the operating environment. The technique used by Boyce *et al.* [5] required a tedious process of superimposing the GHPs of magnesium and aluminum by adjusting the relative brightness and contrast settings during the image capture. In the present study, the WMGL from the magnesium and aluminum standards in region 1 were retrospectively used to normalize WMGL differences of the other materials in addition to calibration using the computer controls described above. The WMGL for the magnesium specimen from region I was arbitrarily defined as 150 for each imaging session. The WMGL for the aluminum specimen from region I was arbitrarily defined as 180 for each imaging session. The calibrated WMGL for the remaining specimens, and for the aluminum and magnesium from regions 2 and 3, was calculated using the following equation, so as to preserve the relationship between the specimen, magnesium standard, and aluminum standard:

$$\{(Al - Specimen) / (Mg - Specimen)\} =$$

$$\{(180 - Specimen_{calibrated}) / (150 - Specimen_{calibrated})\} \quad (2)$$

where: Al = Aluminum WMGL, Mg = Magnesium WMGL; Specimen = WMGL of specimen of interest; and $Specimen_{calibrated}$ = WMGL of specimen after calibration. Al, Mg, and Specimen were measured during each imaging session. Solving the equation for $Specimen_{calibrated}$ yields the new calibrated WMGL for Specimen.

Results

The WMGLs for each BSE image prior to the retrospective calibration are shown in Table 1. Over the 6 day experimental period, the WMGL of each BSE image of inorganic material remained consistent, although there were minor fluctuations with no obvious pattern. A

Table 2. Calculated kurtosis for each specimen. Kurtosis is consistent between imaging sessions.

	Kurtosis					
	Day 1-Image 1	Day 1-Image 2	Day 1-Image 3	Day 3	Day 5	Day 7
Chick I	-0.14	-0.02	-0.02	-0.09	-0.03	0.23
Chick II	0.65			0.85		0.96
Chick III	-0.39					-0.29
Human I	0.74	0.66	0.77	0.56	0.69	0.68
Human II	0.44			0.48		0.49
Human III	0.73					0.57
Antler I	0.50	0.57	0.55	0.48	0.47	0.60
Antler II	0.60			0.56		0.68
Antler III	0.19					0.30
Mule Deer I	9.47	10.47	10.69	10.13	9.52	9.60
Mule Deer II	9.87			9.24		9.06
Mule Deer III	10.24					9.47
Aluminum I	3.11	3.43	3.48	3.29	3.47	3.23
Aluminum II	3.19			2.84		2.75
Aluminum III	2.49					2.34
Magnesium I	6.19	6.02	6.39	6.33	6.30	6.31
Magnesium II	6.51			6.79		6.87
Magnesium III	6.70					6.87
Feldspar I	5.51	5.55	5.52	4.95	5.01	5.19
Feldspar II	4.94			4.93		3.47
Feldspar III	4.55					4.40

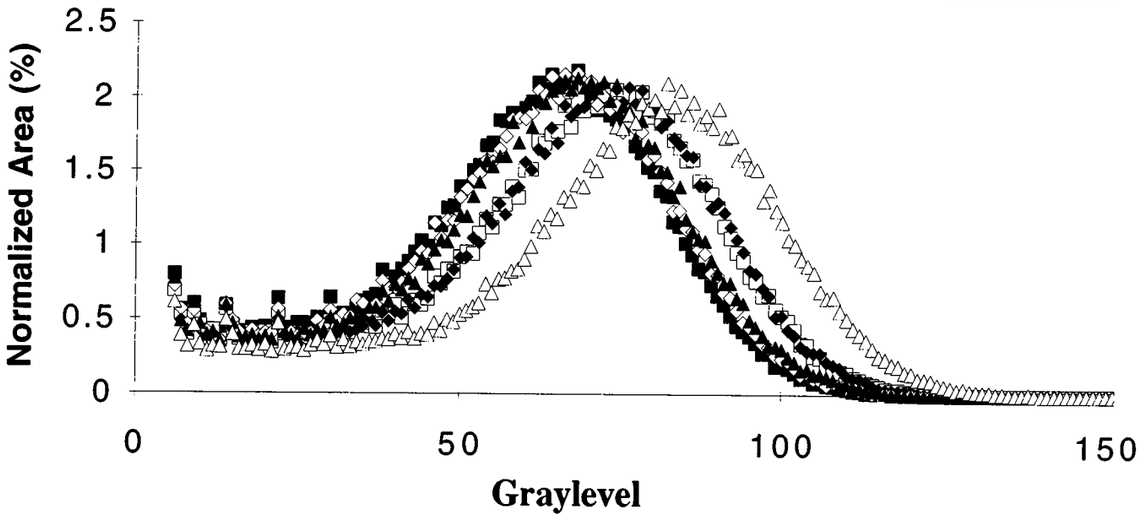


Figure 4. Embryonic Chick Femur GHPs from Group I. Graylevels 150-255 were not shown in this graph, so that the GHPs could be seen more clearly. The shape of the GHPs remains consistent between imaging sessions. Day 1 - Image 1: solid square; Day 1 - Image 2: open square; Day 1 - Image 3: solid diamond; Day 3: open diamond; Day 5: solid triangle; and Day 7: open triangle.

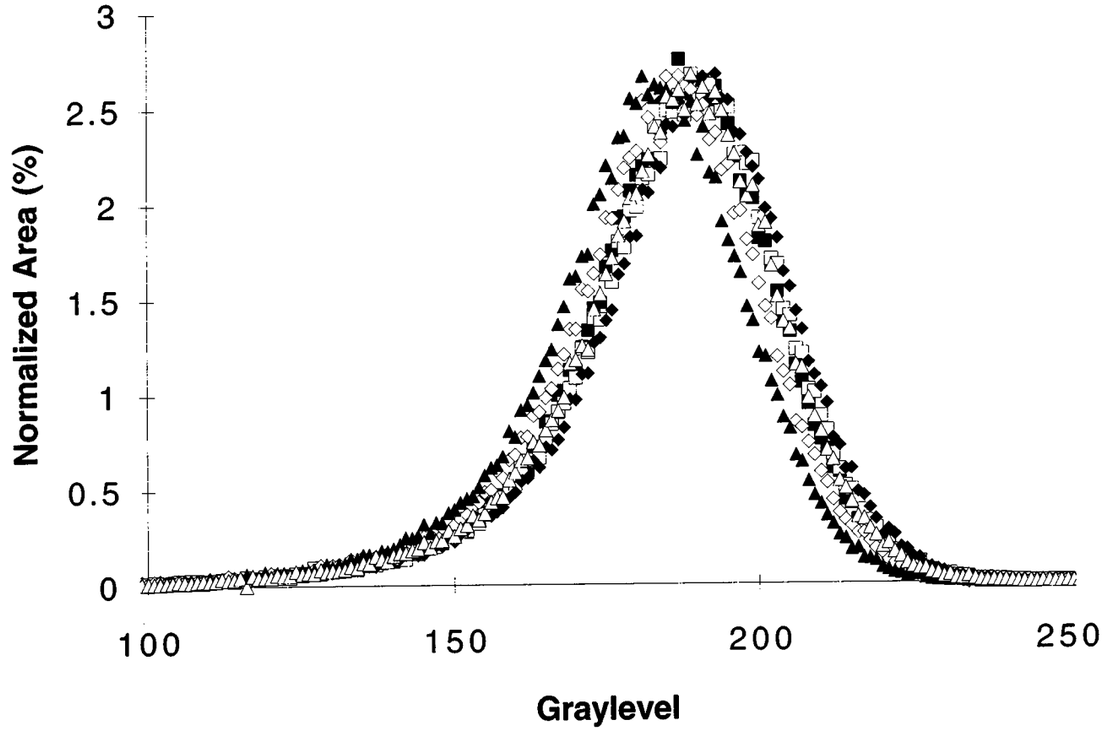


Figure 5. Aluminum GHPs from Group I. Graylevels 0-100 were not shown in this graph, so that the GHPs could be seen more clearly. The shape of the GHPs remains consistent between imaging sessions. Placement of the GHPs on the graylevel scale remained consistent between imaging sessions. Day 1 - Image 1: solid square; Day 1 - Image 2: open square; Day 1 - Image 3: solid diamond; Day 3: open diamond; Day 5: solid triangle; and Day 7: open triangle.

slight increase in WMGL with time occurred in the biological materials, but prior to the retrospective calibration, it is difficult to determine if it is a real trend or simply an artifact.

Kurtosis tests revealed that the GHPs typically, and consistently, deviated to a minor degree from an ideal normal distribution (Table 2). The shapes of GHPs between specimens ranged from very slightly platykurtic for the embryonic chick femur to moderately leptokurtic for the mule deer calcaneus. Kurtosis of the individual GHPs for each specimen, however, remained consistent throughout the course of the experiment. Figures 4 and 5 present examples of GHPs from selected specimens that demonstrate the consistency of the size, shape, and placement of the GHP over time in both biological and inorganic materials.

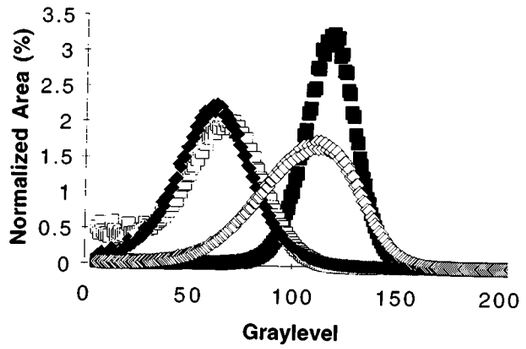
Tests for skewness revealed similar findings. A moderate range of skewness values were observed between specimens, but consistency between imaging sessions was observed for each region analyzed (Figs. 4 and 5, and Table 3). Most specimens exhibit a slight deviation from normality and are negatively skewed.

The GHPs of the bones and pure materials were observed to be continuous and unimodal. Feldspar was the only exception, with a GHP that tended towards bimodality. Figures 6 and 7 demonstrate the relative GHP shapes. For illustrative purposes, the figures represent cumulative GHPs for each specimen acquired over the six day interval [34].

The calibrated WMGLs for each BSE image are shown in Table 4. The calibrated WMGLs for aluminum and magnesium from Group I were, by definition, 150 and 180. The calibrated WMGLs for aluminum and magnesium from Group II and Group III were calculated with respect to Group I. After calibration, the consistency of the inorganic specimens was observed to improve. The average of one standard deviation for the inorganic specimens was 0.9 graylevels. Out of a possible 250 graylevels, 0.9 represents 0.4% of the entire graylevel spectrum. The calibrated WMGLs for the biological specimens were also consistent, but did show a slight trend towards increased brightness with time and/

Table 3. Calculated skewness for each specimen. skewness is consistent between imaging sessions. Most specimens are negatively skewed.

	Skewness					
	Day 1-Image 1	Day 1-Image 2	Day 1-Image 3	Day 3	Day 5	Day 7
Chick I	-0.49	-0.62	-0.63	-0.51	-0.53	-0.75
Chick II	-0.24			-0.23		-0.30
Chick III	-0.33					-0.40
Human I	-0.52	-0.53	-0.55	-0.51	-0.54	-0.53
Human II	-0.58			-0.59		-0.58
Human III	-0.52					-0.49
Antler I	0.09	-0.04	-0.06	0.03	-0.05	-0.04
Antler II	0.18			0.14		0.10
Antler III	0.07					-0.02
Mule Deer I	-2.05	-2.20	-2.24	-2.16	-2.09	-2.13
Mule Deer II	-2.11			-2.01		-2.00
Mule Deer III	-1.94					-1.80
Aluminum I	-1.03	-1.11	-1.14	-1.08	-1.08	-1.08
Aluminum II	-1.00			-0.95		-0.95
Aluminum III	-0.98					-0.94
Magnesium I	-0.15	-0.09	-0.13	-0.07	-0.04	-0.04
Magnesium II	0.28			0.35		0.31
Magnesium III	0.18					0.16
Feldspar I	-1.69	-1.71	-1.70	-1.62	-1.63	-1.66
Feldspar II	-1.49			-1.47		-1.26
Feldspar III	-1.26					-1.25



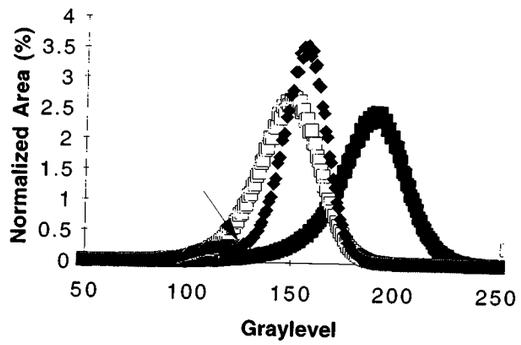


Figure 6 (at left). Cumulative GHPs for biological specimens. Open square: embryonic chick femur; solid diamond: mule deer antler; open diamond: human femur; solid square: mule deer calcaneus.
Figure 7 (at right). Cumulative GHPs for inorganic specimens. Open square: magnesium; solid diamond: feldspar; solid square: aluminum. Arrow indicates suggestion of bimodality in Feldspar GHP.

or number of times imaged. The average of one standard deviation for the biological specimens was 5.9 graylevels. This represents 2.4% of the entire graylevel spectrum. Analysis of variance (ANOVA) comparing the biological materials' calibrated WMGLs from Day 1 and Day 7 showed a statistically significant increase ($p < 0.01$). WMGLs from biological materials in Group I increased more than WMGLs from biological

Table 4. Calibrated WMGLs for each specimen. WMGLs for inorganic materials show excellent consistency between imaging sessions. WMGLs for biological specimens are consistent but show a trend towards increased brightness with increased imaging.

	Calibrated Weighted Mean Graylevel (WMGL)								
	Day 1-Image 1	Day 1-Image 2	Day 1-Image 3	Day 3	Day 5	Day 7	Mean	SD	
Chick I	76.0	85.1	87.8	82.4	83.6	96.8	85.3	6.9	
Chick II	75.4			81.1		87.8	81.4	6.2	
Chick III	73.3					85.8	79.5	8.9	
Human I	113.9	118.7	120.9	118.2	118.3	123.6	118.9	3.2	
Human II	113.9			118.3		123.0	118.4	4.5	
Human III	110.6					118.2	114.4	5.4	
Antler I	75.6	84.1	87.2	83.3	81.8	90.5	83.8	5.0	Avg. SD Biological 5.9
Antler II	76.7			83.9		90.8	83.8	7.0	
Antler III	76.5					90.6	83.6	10.0	
Mule Deer I	120.0	128.1	129.3	126.8	127.9	129.6	127.0	3.5	
Mule Deer II	121.1			126.7		129.3	125.7	4.2	
Mule Deer III	121.6					129.8	125.7	5.8	
Aluminum I	180.0	180.0	180.0	180.0	180.0	180.0			
Aluminum II	185.0			183.7		183.2	184.0	1.0	
Aluminum III	181.2					180.7	181.0	0.4	
Magnesium I	150.0	150.0	150.0	150.0	150.0	150.0			Avg. SD Inorganic 0.9
Magnesium II	153.7			153.3		153.1	153.3	0.3	
Magnesium III	152.1					151.8	152.0	0.2	
Feldspar I	153.0	155.4	154.8	152.7	155.9	155.7	154.6	1.4	
Feldspar II	154.5			155.4		157.1	155.7	1.3	
Feldspar III	154.8					157.7	156.3	2.0	

materials in Group III, although the difference was not statistically significant ($p > 0.05$). ANOVA for inorganic specimens comparing Day 1 and Day 7 showed no increase in WMGL with time ($p = 0.79$).

Discussion

The results of this study support the hypothesis that most modern SEMs with computer-controlled settings, calibrated with pure metals, can give reliable, consistent results over extended periods of time in studies dealing with calibrated BSE images of mineral and calcified tissues. Variations in the WMGLs and in the size and shape of the GHPs of each of the biological and inorganic materials remained consistent between operating sessions used in this investigation.

An increase in calibrated WMGLs of the bone tissues was observed with sequential imaging. Significant increases in WMGL from Day 1 to Day 7 were observed. Boyce *et al.* [5] described this phenomenon as "bleaching" and it has been attributed to specimen and/or embedding media degradation [7, 14] that is caused by prolonged scanning with a high energy electron beam. The data obtained from the biological specimens in Group I on Day 7 was an average of 13.8 graylevels higher than the data obtained on the same fields six days earlier. Data from the biological specimens in Group III on Day 7 was an average of 10.6 graylevels higher than the data obtained from the same fields six days earlier. This would indicate that the number of times imaged may affect the extent of bleaching, and a large portion of the bleaching occurs during the first or second image capture. It should also be noted that the bleaching was greatest in the embryonic chick femur and the deer antler (average WMGL increase: 16.7 and 14.5, respectively), both of which were more porous than the other specimens, have an inherently lower mineral content, and conversely a higher collagen content. These results may indicate changes in the embedding media and/or organic matrix lead to bleaching and an increase in WMGL.

The variation in calibrated WMGL for the pure metals between Day 1 and Day 7 was far less than that of the biological specimens. By definition, the change in WMGL for Group I was zero. The average change in WMGL for Groups II and III were -1.2 graylevels

and -0.4 graylevels, which validates their use as standards for calibrating the BSE/SEM environment between imaging sessions on different days. No signs of bleaching in any of the inorganic specimens were observed. In his review of work by Boyce *et al.* [5], Dr. V.N.E. Robinson had suggested that the use of pure metals as standards for long BSE imaging experiments would be hampered by the progressive increase in thickness of an oxide layer. The retrospective calibration method used in the present study would mask these changes, however, the results of the metal standards prior to calibration show that even if the oxidation layer increased on the magnesium and aluminum standards over the course of the 6 day experiment, the WMGLs remained consistent (Table 1). Additionally, the limited variability between calibrated WMGLs from Group I, II, and III in the metal standards suggests homogeneity within the specimens. This homogeneity is essential if metals are to be used for calibration standards.

The physical basis for the shape of a GHP obtained from BSE images of bone, minerals, and other materials is not completely understood [25, 34, 39]. Several investigators have interpreted these GHPs as being primarily a representation of a distribution of mineral densities [6, 8, 21, 29] corresponding to the sampled regions of the imaged material. Other investigators have suggested that the non-normal kurtosis and skewness of these distributions, although partially influenced by variations in mineral content, are also strongly affected by the interrelated variables of electronic noise and topographic irregularities [5, 25] on the specimen surface. Kurtosis and skewness data obtained in this study demonstrate consistent differences in the shape of the GHPs between species and between the biological and inorganic materials. The data show that these differences are not indicative of the Gaussian distribution that would be expected if the GHP is solely a representation of electronic noise (Tables 2 and 3, and Figs. 6 and 7). Therefore, to some extent, these differences are influenced by the characteristics of the material of each specimen.

Examination of mineral density distributions of bone powders ($\sim$ 50 μm diameter) obtained by density fractionation techniques helps to clarify the kurtosis and skewness data of the bone tissues listed in Tables 2 and 3. Published histogram distributions of density fractionated powders of bone from various species [3, 19, 20, 26, 28, 31] invariably show the presence of negative skewness. The absence of positive skewness, or curve asymmetry, can be attributed to the rather abrupt truncation of mineralization at higher bone mineral densities. This truncation is related to bone remodeling rates and to the primary and secondary phases of osteon mineralization [19]. During the initial 65% of mineralization (primary phase), mineral content of a bone progressively increases by the addition of mineral crystallites within and upon the collagenous matrix [16, 19, 35]. In the secondary phase of mineralization, increases in crystal size are thought to account for the remaining 35% of the final mineral content [19] in a specific bone type. In contrast to primary mineralization which can represent a broad, continuous range of mineralization from no mineral to approximately 65% of full mineralization [19], secondary mineralization affects a narrower range of mineralization which becomes relatively sharply limited as higher bone mineral densities are approached. In other words, the lower end of a mineral density distribution of fractionated bone powders appears quasi-asymptotic. In contrast, toward the higher end of the density distribution, the increase is not progressive, but becomes abruptly truncated. The negative skewness of GHPs obtained from the BSE images of bone tissues may be influenced by these differences. This interpretation does not, however, explain the negative skewness seen in pure aluminum. Pure aluminum does not have a mineral density distribution which would contribute to the negative skewness.

Residual topographic irregularity of the specimen surface of most of the biological and inorganic materials may contribute to the negative skewness. For example, the negative skewness of the aluminum standard, which was not seen in the magnesium standard, can be attributed to relatively more residual surface irregularities which were recognized in a retrospective examination of the imaged regions. We hypothesize that residual topographic irregularities, such as minute scratches, are more likely to occur in aluminum because it is softer than the magnesium composite used in this study. In bone tissues, it is also difficult to eliminate all surface topography [32]. As theorized by Howell and Boyde [25], topography effects may not symmetrically affect GHPs.

The progressive increase in kurtosis seen from the chick femur to the deer calcaneus (Table 2) can be explained in the context of differences in the remodeling rates of these bones. Remodeling in the cranial cortex of the deer calcaneus has been shown to be relatively inactive [36, 37]. Glimcher [16] suggests that bone remodeling in the chick embryo is so rapid that its entire bone volume may be renewed in only a few days. Additionally, a large percentage of the total bone volume would be expected to be composed of rapidly forming woven bone, which has a lower mineral content. Consequently, in embryonic bone, density fractionation data of bone powders shows a broad range of mineralization, which is seen as a negative kurtosis in histograms of density fractionation data [20, 31] and in the GHPs of BSE images [35]. In contrast, the narrower GHPs (higher kurtosis) of the BSE images of the cranial cortex

of the deer calcaneus reflect the fact that the bone, on average, is uniformly older, and hence a relatively larger percentage of the cortical volume of this region has undergone secondary mineralization and less residual woven bone is present. The bone is typically more uniform than an embryonic chick femur. The slight positive kurtosis exhibited by the GHPs of the BSE images of the middle aged human femur specimen partially reflects the moderate breadth of the distributions of bone mineralization that would be expected in this bone since it is likely to have a low to moderate remodeling rate [27 (pages 214-217)]. The kurtosis of the mule deer antler, and to a lesser extent the other bone specimens, is also influenced by regional heterogeneities in mineralization [4, 10, 11, 18, 27 (pages 20-52); unpublished observations of BSE images of deer antler cross-sections] which are seen as hypermineralization near lacunae, mineral variations across secondary osteons, and variable mineralization in cement-lines or cement-line-like structures.

However, because of the superimposed influences of electronic noise, surface topography, and electron energy distributions on BSE image GHPs, it is not possible to "dissect" out the relative contribution that mineralization has in affecting the graylevel of each image pixel. The results of this study for pure metals do not support the hypothesis that GHPs of bone are solely a representation of a mineralization profile. Pure metals would be expected to be confined to one graylevel corresponding to a uniform density and atomic number, rather than a broad spectrum of graylevels as is invariably observed. The relative breadth, or kurtosis, of the GHP cannot be interpreted as representative of the mineral density profile, as has been attempted by Mechanic *et al.* [29], Torontali *et al.* [39], and Boyde *et al.* [7], without accounting for other variables mentioned above. The authors of the present study caution against making this interpretation until the physical meaning of the GHP is more clearly understood.

Although electronic noise, surface topography, and various other interrelated variables can contribute to altering apparent atomic number contrast of a BSE image, the BSE/SEM imaging session can be calibrated to keep precision error at an acceptably low level. Progressive "bleaching" of bones with prolonged imaging is a potential problem which should be considered during image capture. Pure metal calibration standards do not display the "bleaching" phenomenon and produce extremely consistent results. More work is needed to advance current understanding of the physical basis of GHPs obtained from minerals, calcified tissues, and other materials imaged in the BSE mode of the SEM.

Acknowledgments

The authors would like to acknowledge the Department of Veterans Affairs Medical Research Fund, Salt Lake City, UT, for support of this project. We thank Rebecca Elliot for assistance in statistical analysis of data and Gwenevere Shaw for her assistance in manuscript preparation.

References

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

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

[3]. Bonar LC, Roufosse AH, Sabine WK, Grynpas MD, Glimcher MJ (1983) X-ray diffraction studies of the crystallinity of bone mineral in newly synthesized and density fractionated bone. Calcif Tiss Int **35**: 202-209.

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

[5]. 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.

[6]. Boyde A, Elliott JC, Bromage TG, Anderson P, Davis G, Collier J, Hausner WS, Mechanic GL, Katz EP, Arnaud SB, Durnova GN (1990) Quantitative mineralisation density mapping of space flight rat femoral bone at submicron resolution: BSE calibrated by μCT. Bone **11**: 217.

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

[8]. Boyde A, Howell PGT, Bromage TG, Elliott JC, Riggs CM, Bell LS, Kneissel M, Reid SA, Jayasinghe JAP, Jones SJ (1992). Applications of mineral quantitation of bone by histogram analysis of backscattered electron images. In: Chemistry and Biology of Mineralized Tissues. Slavkin H, Price P (eds.). Excerpta Medica, New York. pp. 47-60.

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

[10]. Burr DB, Schaffler MB, Frederickson RG (1988) Composition of the cement line and its possible mechanical role as a local interface in human compact bone. J Biomech **21**: 939-945.

[11]. Crofts RD, Boyce TM, Bloebaum RD (1994) Aging changes in osteon mineralization in the human femoral neck. Bone **15**: 147-152.

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

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

[14]. 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. pp. 81-86.

[15]. Folch J, Lees M, Stanley G (1957) A simple method for the isolation and purification of total lipids from animal tissues. J Biol Chem **226**: 497-509.

[16]. Glimcher MJ (1992). The nature of the mineral component of bone and the mechanism of calcification. In: Disorders of Bone and Mineral Metabolism. Coe FL, Favus MJ (eds.). Raven Press. Ltd., New York. pp. 265-286.

[17]. 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.

[18]. Goss RJ (1983) Deer Antlers. Regeneration, Function, and Evolution. Academic Press, New York. pp. 133-171.

[19]. Grynpas M (1993) Age and disease-related changes in the mineral of bone. Calcif Tissue Int **53**: S57-S64.

[20]. Grynpas M, Hunter G (1988) Bone mineral and glycosaminoglycans in newborn and mature rabbits. J Bone Miner Res **3**: 159-164.

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

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

[23]. 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.

[24]. 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.

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

[26]. Lian JB, Roufosse AH, Reit B, Glimcher MJ (1982) Concentrations of osteocalcin and phosphoprotein as a function of mineral content and age in cortical bone. Calcif Tissue Int **34**: S82-S87.

[27]. Martin RB, Burr DB (1989) Structure, Function and Adaptation of Compact Bone. Raven Press, New York.

[28]. Mbuyi-Muamba JM, Dequeker J, Gevers G (1989) Collagen and non-collagenous proteins in different mineralization stages of human femur. Acta Anat **134**: 265-268.

[29]. 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. The FASEB J **4**: 34-40.

[30]. 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.

[31]. Richelle LJ, Onkelinx C (1969). Recent advances in the physical biology of bone and other hard tissues. In: Mineral Metabolism: An Advanced Treatise. Comar C, Bronner F (eds.). Academic Press, New York. pp. 123-190.

[32]. Roschger P, Eschberger J, Plenk H Jr. (1993) Formation of ultracracks in methacrylate-embedded undecalcified bone samples by exposure to aqueous solutions. Cells Mater **3**: 361-365.

[33]. 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.

[34]. 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.

[35]. 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.

[36]. 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.

[37]. 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.

[38]. Sokal RR, Rohlf FJ (1995) Biometry. The Principles and Practice of Statistics in Biological Research. W.H. Freeman & Co., New York. pp. 111-115.

[39]. 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 Mater **4**: 125-134.

Discussion with Reviewers

P. Roschger: Because of the lacking online adjustment of reference graylevels (WMGLs), the authors introduced a useful method (equation 2) for retrospective calibration (correction) of specimen WMGLs, however, such corrections cannot be made for the GHPs. Therefore, if GHPs have to be combined and evaluated, a maximal accuracy in reference WMGLs in advance should be intended.

Authors: It is correct that the retrospective calibration proposed in this study does not calibrate GHPs, only WMGLs. However, as we stated at the end of Discussion, the interpretation of the GHP is still uncertain and we do not advocate the use of GHPs in the description of bone mineralization until a further experimentally based understanding is developed. WMGLs, on the other hand, have been described and correlated to the mineral content and density of bone tissues [34, 35].

P. Roschger: The WMGLs of the references Al and Mg are only 30 graylevels apart from each other and small deviations from the ideal value can introduce relatively large changes in BSE-image contrast. To use a second reference with a lower atomic number (Z) than that of Mg (Z = 12) would be advantageous.

Authors: Carbon (Z = 6) is a suitable pure element with Z less than bone (Z ranges from 9 to 11) that has been used in previous BSE studies as a calibration standard. Unfortunately, the working range to image both carbon and magnesium is six atomic numbers (12 - 6 = 6), which is less than that required to image both bone and aluminum (4 atomic numbers; 13 - 9 = 4). Therefore, the introduction of carbon necessitates a loss of sensitivity, because a larger atomic number range must be compressed into the same 256 graylevels. The 30 graylevel range between Al and Mg in this study could clearly be increased, to gain more sensitivity, by a suitable adjustment in contrast. The use of carbon would be advantageous, however, if low atomic number osteoid or unmineralized bone were to be analyzed in the same operating sessions as normal mineralized bone.

P. Roschger: This study was designed to evaluate the consistency of BSE images and the resulting GHPs from different standards and bone tissues. Thus, a discussion of the effects of (surface) artifacts etc. on the different biological specimen blocks seems appropriate. However, the influences of different mineral contents and compositions on the GHPs of these biological specimens, also with respect to their remodelling activities, are in the authors' opinion not yet so well understood to use GHPs for their description. So, why then are they discussed at such length in this methodological paper, without showing respective BSE-images? This should better go into a separate publication on this matter, also because GHP's seem well suited to characterize mineral distributions in different calcified tissues, as it was meanwhile demonstrated by the investigations of this BSE-method [41].

Authors: We do not believe that GHPs of bones are adequately understood at this point in time to interpret the height of a GHP at any one point as solely representative of the amount of tissue of a given mineral density, although this has been attempted in the literature [6, 8, 21, 29, 39]. Clearly, other artifacts affect the shape of the GHP as evidenced by the GHPs of pure elemental standards. However, this does not mean that mineral densities in bone have no influence upon GHPs and we would be remiss if we did not discuss this influence. Obviously, there are differences in the shapes of GHPs of different bones (Fig. 6) and we have attempted in this study to account for some of these differences to better understand the physical basis of the GHP. We have not provided BSE images of each specimen in this study because we felt it would be excessive as we have already provided GHPs (Figs. 6 and 7) for each specimen.

J. Hejna: Progressive "bleaching" of bones with prolonged imaging claimed in the paper is not fully confirmed by results of WMGL measurements in Table 4. There are large fluctuations of results in the course of the study.

P. Roschger: The "bleaching" phenomenon should be studied separately within one (not within multiple) calibrated BSE-imaging session by multiple prolonged scans of a defined bone area. It would provide useful information about the tolerable scanning time at a defined beam current and size of scanned area for quantitative BSE-imaging. However, the strategy for quantitative BSE-imaging should be to keep the exposure time of the analysed bone area to the electron beam as short as possible.

Authors: This study was not designed solely to examine the bleaching phenomenon in bone. In fact, bleaching proved to be a complication in our effort to show consistency in quantitative BSE graylevel analysis. The results in this study with regard to bleaching are not conclusive and Dr. Roschger has suggested the next logical study which we have already begun in our laboratory.

A. Boyde: Metals can be used to standardise one instrument provided that, as would be expected to be the case, working distance and repositioning of the standard field were exactly maintained. It cannot be used to calibrate the microscope because unpredictable channelling contrast effects would determine the exact mean value which is obtained and this will vary from sample to sample of the standard material. In our laboratories we used pure metals and minerals to calibrate the BSE signal in our earliest studies in our home-made histogram-only acquisition system described by Howell and Reid [40] and we have always been aware of the problems of anomalous histogram widths and shapes from so called pure standards. Metals and minerals have large crystals. Unless one assumes perfect randomisation of order within the layer depth sampled by the electron beam, then channelling contrast effects can neither be ignored, nor can they be dealt with. This is the reason not to choose them as standards.

Authors: You are correct in stating that pure metals can be used to standardise one instrument, as has been experimentally demonstrated in this study. Electron channelling patterns are typically very obvious when present, and were not observed in this study. This is not surprising, as the mechanical grinding and polishing methods used in specimen preparation produces a thin layer of highly stressed, disordered, and damaged material (the "Bilby layer") [17 (page 549)]. This effectively eliminates the production of channeling contrast patterns. We do not feel that the "anomalous histogram widths and shapes" are the result of channelling contrast.

In response to your remarks concering the variability of the graylevel mean value between pure samples, we have analyzed six pure (99.9999 %) aluminum standards. Three were prepared as previously described in this study and three were provided by Tousimis Research Inc. (Rockville, MD). Aluminum standards from Tousismis had been prepared by diamond polishing. BSE images were consecutively captured in one imaging session and WMGLs were calculated with the following results:

Al #1 WMGL	118.9
Al #2 WMGL	118.2
Al #3 WMGL	118.1
Tousimis Al #1 WMGL	124.9
Tousimis Al #2 WMGL	124.3
Tousimis Al #3 WMGL	124.5

Clearly, the variability between similarly prepared specimens is minor. The preparation methods, however, appear to influence WMGL. This suggests topographical effects may play a role. If it were indeed a result of electron channelling effects, consistency between similarly prepared specimens would not be expected as Prof. Boyde has mentioned. Future investigators should adequately describe the preparation methods to allow for reproduction of their work.

A. Boyde: Have the authors studied the surface profile of the material prepared by their protocol of milling, grinding and buffing?

Authors: We have analyzed the surface profile using optical phase shifting and vertical scanning interferometry (WYKO Corporation, Tucson, AZ) and have observed small (typically on the order of $\pm$ 0.5 μm) topographical irregularities. We are currently investigating the influence of this subtle topography upon BSE WMGLs.

M. Grynpas: In **Discussion**, you say: "it is not possible to "dissect" out the relative contribution that mineralization has in affecting the graylevel of each image pixel", but if various bone specimens are prepared in an identical manner, what else would the difference in graylevel histograms reflect beside differences in mineralization profiles?

Authors: Undoubtedly, electronic noise plays a role in shaping the GHP and is not necessarily consistent between similarly prepared specimens. This study and the report by Howell and Boyde [25] suggests that topography also influences GHPs. Because bone is extremely heterogenous, it cannot be assumed that topographical variations at the submicron level are consistent between similarly prepared specimens. The introduction of surface ultracracks [32] can be particularly problematic. Relative differences between GHPs **may** be the result of differences in mineralization profiles, but at this point, there is no experimental evidence that allows interpretation of the height of the GHP at any one point as representing the amount of calcified tissue of a given density.

J. Hejna: The authors sputter coated specimens with gold. Gold is not a good choice as a coating metal for material contrast studies, especially on low density specimens. Additionally, the thickness of the layer was not measured. I suggest to repeat experiments on specimens coated with low density material of known thickness. Layers of carbon or aluminum, or a very thin layer of chromium can be recommended. It would also be good to coat the calibrating standards.

Authors: It is true that the thin layer of gold used in this study is not ideal because it would **increase** the mean atomic number of all of the specimens. On the other hand, carbon, or any other low density coating, would **decrease** the mean atomic number of all specimens. Any coating material would affect the mean atomic number of a sample. This is an unfortunate consequence of making a specimen electrically conductive. If the coating is thin, even with a high atomic number element such as gold, Z contrast between bone specimens can still be observed and correlated to density and mineral content as was previously demonstrated by Skedros *et al.* [34, 35]. Because we were not interested in measuring the atomic number of the specimens in this study, it was not necessary to know the exact coating thickness. It was only necessary to insure that the coating was thin enough that Z contrast could still be observed. The calibrating standards used in this study were coated with gold as described in the **Materials and Methods**.

K. Murata: You scan the beam in an area of about 500 μm x 500 μm at a magnification of 200. Assuming 500 x 500 pixels for one frame scan, the area of one pixel is 1 μm^2. Is your probe size smaller than this area? What is the probe size?

Authors: If bone is assumed to have an atomic number intermediate to carbon and aluminum ($\sim$ 9.5) then the Kanaya-Okayama Range (R_{ko}) [17 (page 89)] for electron penetrations in bone would be 9.4 μm. Approximately 80 % of the backscattered electrons would be generated in a radial distance 0.46 R_{ko} [17 (page 104)] from the point of electron probe contact, which corresponds to a radial spread of 4.3 μm. This is a rough estimate of the interaction volume, but it demonstrates that the operating conditions used in this study did not provide optimal resolution. Resolution was sufficient, however, to distinguish morphologic features within the bone tissues, such as, individual lacunar spaces, lamellae, and canaliculi. Resolution could be improved in future studies, if necessary, by decreasing the accelerating voltage and to a lesser extent the probe diameter (increasing condenser lens strength).

K. Murata: Have you observed the GHP with secondary electrons (SEs)? I understand that backscattered electrons have a higher Z contrast. But SEs have a better topography contrast. So, it is interesting to see what type of GHP is obtained with the SEs.

Authors: The GHP of an SE image is largely the result of topography, and therefore, specimen preparation methods. A "rough" surface would have more topography and hence a broader GHP than a "smooth" surface. This would not, however, provide any insight into the actual height of these topographical irregularities.

K. Murata: Have you observed the GHP by keeping the beam stationary at a fixed position on the standard sample? As a matter of fact, the beam has to be digitized. The GHP thus obtained is free of the effect of topography although the contamination cannot be neglected.

P. Roschger: Have you observed fluctuations in the BSE-signal and then determined their pulse height distribution (i.e., GHP of the signal itself), when the beam is kept stationary at a fixed position on the standard sample?

Authors: We have not yet observed these fluctuations. This would provide insight to the noise inherent in this system, although it is unlikely to explain all of the variations in skewness and kurtosis that we have observed. This is an area that should be investigated in future experiments to help improve our understanding of GHPs.

K. Murata: Do the shape and broadness of the GHP change with the incident beam energy?

Authors: The breadth of the GHP changes with beam energy, the general "shape" or appearance of the GHP remains relatively consistent. The nature of the relationship between beam energy and GHP breadth has not been closely examined.

H. Plenk, Jr.: When the authors discuss the reasons for the negative skewness of "pure" aluminum wire and noticed surface irregularities as a possible reason, why did they not think of the metallic structure itself? The aluminum wire was most likely drawn, what was the metallurgical structure and crystallinity, what was the composition of oxides and other phases?

Authors: The broad histogram and negative skewness present in aluminum has been observed in many "pure" elements (Si, Cu, Au, Ti, V, C) imaged in our laboratory. The purity of these reference standards has been documented (Tousimis Research, Rockville, MD). The oxidation characteristics of these materials are greatly varied. Structure, crystallinity, oxidation, and impurities in the standards undoubtedly influence the GHPs, but they do not adequately explain the **consistently** broad and negatively skewed GHPs which are **virtually always** observed regardless of material or manufacturing method. We have provided hypotheses to explain the GHPs, but the physical basis of the histogram is still unclear to us.

Additional Reference

[40] Howell PGT, Reid SA (1986) A microcomputer based system for rapid on-line stereological analysis in the SEM. Scanning **8**, 139-144.

[41] Roschger P, Plenk Jr H, 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.