Scanning Microscopy

Volume 4 | Number 3 Article 9

8-2-1990

Reproducible Method for Calibrating the Backscattered Electron Signal for Quantitative Assessment of Mineral Content in Bone

T. M. Boyce
Bone & Joint Research Laboratories

R. D. Bloebaum
Bone & Joint Research Laboratories

K. N. Bachus
Bone & Joint Research Laboratories

J. G. Skedros
Bone & Joint Research Laboratories

Scanning Microscopy, Vol. 4, No. 3, 1990 (Pages 591-603)
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

REPRODUCIBLE METHOD FOR CALIBRATING THE BACKSCATTERED ELECTRON SIGNAL FOR QUANTITATIVE ASSESSMENT OF MINERAL CONTENT IN BONE

T.M. Boyce,* R.D. Bloebaum, K.N. Bachus, and J.G. Skedros

Bone & Joint Research Laboratories (151F)
VA Medical Center, 500 Foothill Blvd.
Salt Lake City, UT 84148 U.S.A.

(Received for publication April 25, 1990, and in revised form August 2, 1990)

Abstract

Backscattered electron (BSE) imaging shows promise for orthopaedic and bone research. BSE images of bone may be captured on-line directly from the scanning electron microscope (SEM), and then analyzed to produce a backscattered electron profile (BSEP), a modified image graylevel histogram which is representative of the mineral content in bone. The goals of this work were 1) develop a reproducible graylevel calibration technique for bone specimens, and 2) determine a conservative time interval during which SEM operating conditions would remain stable.

Calibration standards containing pure aluminum and pure magnesium wires were placed in the SEM with human cancellous bone. Baseline imaging conditions were first established by adjusting the SEM until the bone image displayed good resolution and graylevel separation between regions of different mineral content. Microscope brightness and contrast controls were randomly changed to initiate the new operating conditions of another imaging session, and graylevel values from the calibration metals were used to readjust the microscope back to baseline operating conditions. Weighted mean graylevel values of the BSEPs from calibration trials were compared to those of the baseline. Data showed that bone images could be reproduced within 1.2 percent. It was also concluded that our equipment required calibration checks at 20 minute intervals.

KEY WORDS: Backscattered Electron (BSE) Imaging; Backscattered Electron Profile (BSEP); Bone; Atomic Number; Image Analysis; Operating Conditions; Graylevel Calibration; Calibration Standard; Mineral Quantitation; Filament Fatigue

*Address for correspondence:

Todd M. Boyce
Bone & Joint Research Laboratories (151F)
VA Medical Center, 500 Foothill Blvd.,
Salt Lake City, UT 84148 U.S.A.
Phone : (801)-582-1565, Ext. 1198

Introduction

Researchers have been attempting to understand the mechanisms of bone mineralization and the changes that occur with disease, prosthetic implantation, weightlessness, and aging. Many experimental methods, including ashing[4, 11], chemical assays,[4] and photon absorptiometry [7, 8, 19, 21, 26, 29] have provided essential information about gross mineralization changes in bone. While the value of these methods in bone research is indisputable, there is a need to understand the mineralization responses occurring at the microscopic level.

Microradiographic densitometry [30] has proven useful in the study of cortical bone, but is inadequate for application to cancellous bone. The process is plagued by two limitations. First, inconsistency in the exposure and development of the radiographic film leads to variations in the optical density of the resulting microradiograph. This can be corrected to a certain degree by the use of an aluminum step wedge at the time of exposure. However, projection errors which occur as a result of x-ray projection through a three-dimensional cancellous structure [28] cannot be corrected, and can alter bone estimates significantly [6, 24].

With the recent application of backscattered electron (BSE) imaging and image analysis to the study of bone, a microscopic technique is now available which allows the study of localized mineral variations in both cortical and cancellous bone. The physical basis for the ability of the BSE signal to distinguish between materials of different average atomic number is described by Castiang [10] Robinson et al. [25], and Ball and McCartney [2]. Their work showed that the strength of the BSE signal, as determined by the magnitude of the detector output voltage, was dependent on the elements present in the specimen and the relative weight fraction of each element.

Computer Assisted Graylevel Analysis of BSE Bone Images

Currently, BSE imaging has two principal applications in the study of bone, which have their origin in the early leadership of Holmes [17] and Boyde

and Jones [9]. When used to quantify the amount of bone present in an image (histometry) [28, 17, 18], BSE images have much higher accuracy than other methods because they provide a nearly ideal stereological plane. This improvement in accuracy results from the limited penetration (1-5 μm) into the specimen by the electron beam [9, 23].

The utility of BSE imaging methods in bone research has recently been extended beyond histometry to include mineral content analysis, by the incorporation of on-line automated image analysis methods for graylevel assessment [23, 24]. Boyde and Jones [9] have utilized BSE images from a cancellous specimen to demonstrate that bone consists of multiple mineral "phases," and that qualitative differences in mineralization can be observed. The term "phase" refers to variations in graylevels, presumably associated with regions of differing mineralization, which can be observed at microscopic distances from each other.

Utilizing an on-line computer system for the direct capture and analysis of BSE bone images, Reid and Boyde [24] used cadaveric human rib tissue (2 months to 59 years) to demonstrate age-related mineral changes. These changes were determined according to the relative graylevel shifts produced in the BSE images, and were reported to relate to differences in mineral density, as expressed in grams per cubic centimeter. Each image was represented by a histogram which described the area of the image covered by pixels from each of eight graylevel ranges. By overlaying histograms for bone of different ages, it was demonstrated that the graylevel distributions shifted toward a whiter (higher graylevel) image with age, as opposed to the darker graylevels present in younger specimens.

While the techniques introduced by Reid and Boyde held excellent potential for increasing the understanding of bone mineralization, they were limited by an inability to accurately compare images that were collected during different operating sessions. Without a calibration technique to provide consistency in the image graylevels, the usefulness of the captured images to the understanding of mineralization dynamics in bone is limited. The BSE detector output voltage is influenced by a number of SEM controls (detector contrast and brightness, accelerating voltage, working distance, bias voltage) which may vary from imaging session to imaging session and detector-specific parameters (collection angle, efficiency) which will vary depending upon the particular equipment used [24, 25]. Additionally, other factors which cannot be controlled by the user (such as filament fatigue and electronic drift) may also influence the BSE signal and the graylevel data. All of these parameters combine to make it difficult to reproduce graylevel results, even from identical fields of the same specimen.

Based upon initial work conducted in our laboratory [5], we have hypothesized that pure metals such as aluminum and magnesium could be used as calibration standards in order to control changes in the operating conditions which may affect data. Since brightness adjustments are known to have an additive effect on the intensity value of each image pixel and contrast adjustments have a multiplicative effect [3], graylevel information from the pure metal calibration standards could be used to provide the proper brightness and contrast settings to achieve graylevel calibration. By adjusting the brightness and contrast controls of the BSE detector (to calibrate the processed analog signal) or the image analyzer (to calibrate the stored digital image), it would then be possible to correct for changes in the operating conditions.

If BSE imaging methods are to accurately quantify mineral content in bone, it is essential that image graylevel uniformity be maintained by calibrating the detector output with respect to a known standard. The calibration method employed should be universal, allowing many different types of microscopic and imaging equipment to be used without compromising the ability to compare calibrated graylevel data. Additionally, the frequency of calibration needed to maintain stable operating conditions within a system which is experiencing electronic drift and filament fatigue must also be determined. The goal of this communication is to demonstrate that a calibration method has been developed and that guidelines for recalibration frequency have been established.

Methods

Analysis System

The imaging system used in this calibration study is diagrammed in Figure 1. The unit is a general-purpose hardware-based image analysis system (CRYSTAL, Link Analytical, Redwood City, CA), and analysis may be controlled by an IBM-compatible microcomputer (80386, Zenith Data Systems, St. Joseph, MI). This system allows for the capture of high resolution (512 x 570 pixel), noise-reduced digital images directly from the scanning electron microscope (JEOL JSM-T330A, JEOL USA, Inc., Peabody, MA). The image is generated by scanning the microscope beam across the specimen surface, then translating the analog output of the BSE detector (solid state annular quadrant detector, GW Electronics, Inc., Norcross, GA) into a series of discrete pixel-sized regions with graylevel values based upon the detector voltage level.

Graylevel analysis of the BSE image is performed to distinguish the degree of mineralization in the bone. Each pixel in the image is associated with an integer value between 0 and 255 which designates its intensity on the black and white monitor. Black pixels on the image have a value of 0, while bright white pixels have the value 255. The remaining values between (1-254) represent discrete shadings of gray from dark to light. Rather than selecting each of the 256 graylevels individually to determine its

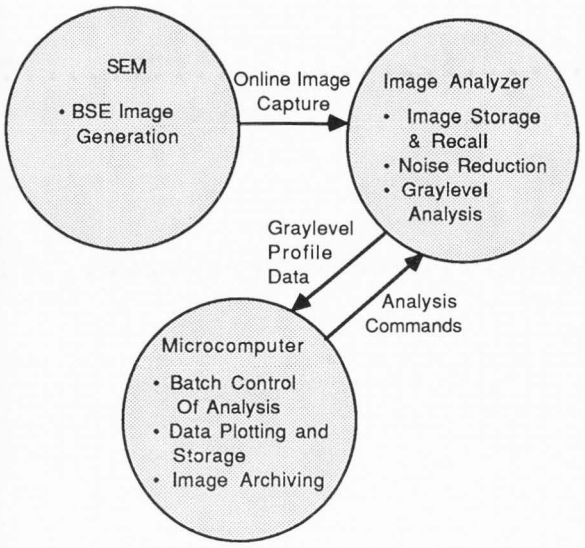

Figure 1 - Diagram describing the functions of the SEM, image analyzer, and microcomputer, as well as the flow of data.

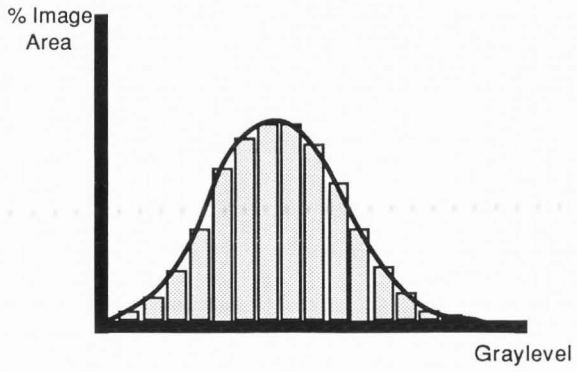

Figure 2 - An image graylevel histogram and the resulting Backscattered Electron Profile (BSEP).

representation in the image, the graylevels were grouped into 51 smaller subranges or "bins" (i.e. graylevels 0-5, 6-10, 11-15..., 251-255). By doing this, sufficient graylevel sensitivity could be maintained while significantly reducing the analysis time.

Analysis of the image was controlled by a microcomputer software routine. This program used an image analysis technique known as "thresholding" [12] in the CRYSTAL system. For each of the 51 bins, the program selected all image pixels which had graylevel values within the range of the bin. The cumulative area of the selected pixels was recorded as a percentage of the analysis region. The final result is a series of 51 numbers representing the proportion of the selected region occupied by pixels of each graylevel bin. These numbers may be plotted against their graylevel to produce the outline of an image graylevel histogram, which we have termed the "backscattered electron profile" (BSEP) for the image. (Figure 2).

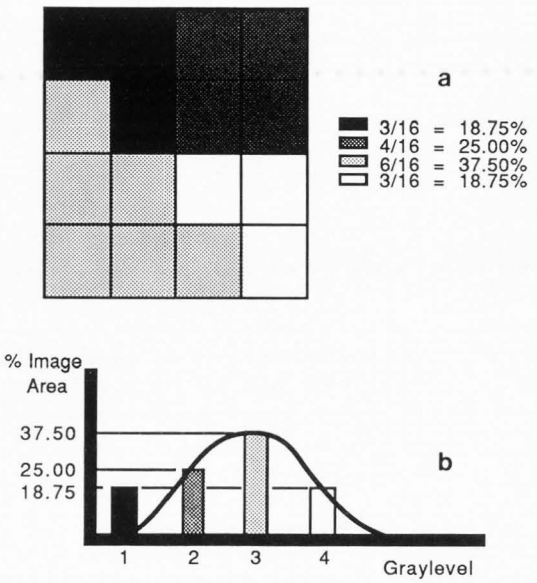

Figure 3 - a) 16 pixel "image" with 4 graylevels and b) the image graylevel histogram and resulting BSEP.

The distribution of the BSEP may be skewed or multi-modal, and therefore the use of mean and standard deviation as descriptors is not always justified. Since no single number could properly describe both the location and shape of the distribution, images were described by plots, range of the distribution, and average graylevel according to the weighted equation:

$$WMGL = \sum_{i=2}^{51} \left[\frac{(A_i)(GL_i)}{A_t} \right] \qquad -(1)$$

where WMGL is the weighted mean graylevel of the BSEP, A_i is the percent area measured for graylevel subrange i, GL_i is the lower graylevel defining subrange i, and At is the total area of the analysis region. The percent area representing the polymethyl methacrylate (PMMA) and non-mineralized tissues (bin 1) varies significantly between bone images, especially in cancellous bone specimens. These variations are due to differences in the volume fraction of PMMA-filled marrow spaces, lacunae, Haversian canals and other void spaces between specimen regions. If the PMMA and non-mineralized tissue regions are not excluded from analysis, they can artificially influence the weighted mean graylevel results by changing the shape of the BSEP and decreasing the weighted mean graylevel. Therefore, the black pixels (i = 1) were not included in the calculation of the weighted mean graylevel .

Interpretation of the Backscattered Electron Profile

To better clarify the steps involved in the graylevel analysis of a bone image, a hypothetical

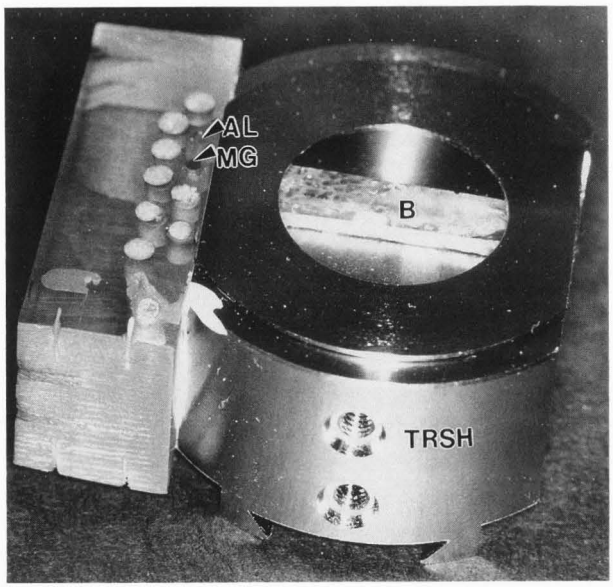

Figure 4 - Top referencing specimen holder (TRSH), human cortical bone specimen (B), aluminum (AL) and magnesium (MG) calibration standards used for graylevel calibration.

image consisting of only 16 "pixels" and using only four graylevels is shown in Figure 3a. The number of pixels of each graylevel are counted and expressed as a fraction of the total number of pixels in the analysis region. In this hypothetical image, 18.75% of the region is black (graylevel 1), 25.00% is dark gray (graylevel 2), 37.50% is light gray (graylevel 3), and 18.75% is white (graylevel 4). These percentages are plotted as a histogram (Figure 3b). The BSEP is represented as a smooth curve through the histogram peaks, excluding the black or lowest graylevel. Applying equation 1, the weighted mean graylevel in the BSEP is

$$WMGL = \sum_{i=2}^{4} \left[\frac{(A_i)(GL_i)}{A_t} \right]$$

$$= \frac{(25.00)(2) + (37.50)(3) + (18.75)(4)}{25.00 + 37.50 + 18.75}$$

$$= 2.92$$

and the range is from graylevel 2 to 4.

Graylevel Calibration

A calibration block was made by drilling holes and inserting 1.52 mm diameter aluminum wire (99.99 % pure, Cominco Electronic Materials, Spokane, Washington) and 1.59 mm diameter magnesium wire (99.8% pure, Johnson Matthey/Aesar, Seabrook, New

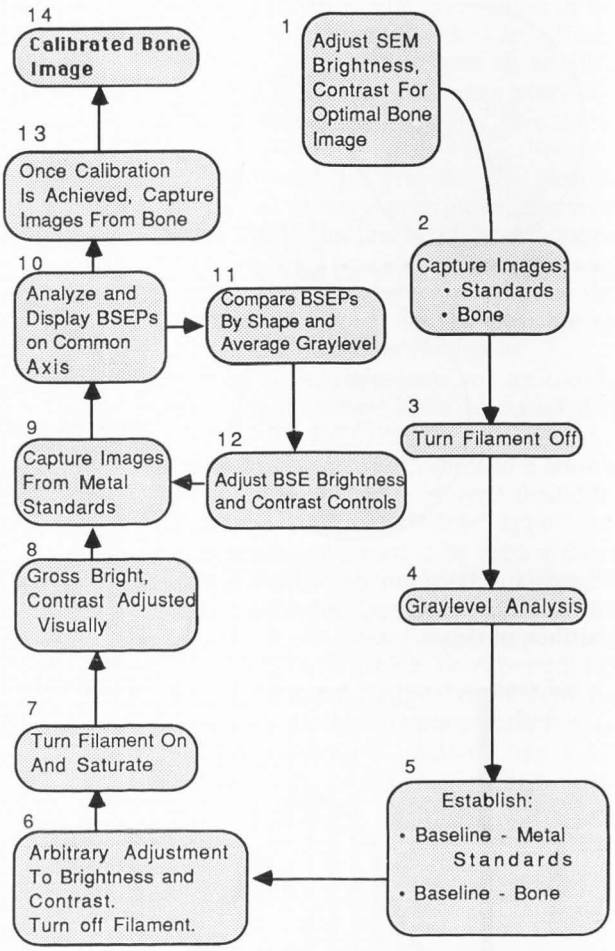

Figure 5 - The steps involved in determining baseline BSEPs (steps 1-5) and performing a trial calibration (steps 6-14).

Hampshire) into a plastic block (Figure 4). The block containing the metal standards was ground on a rotary grinding wheel using graded sandpapers (wet/dry silicon carbide 60, 100, 240, 400 grits, S.L. Fusco, Inc., Compton, CA) and polished using alpha alumina (1.0 micron, Fusco Abrasives, Tempe, AZ) with a polishing cloth (Buehler Ltd, Lake Bluff, IL) until an optically smooth finish was achieved. The calibration block was placed into a sputter coater (Hummer Model III, Technics, Inc, Alexandria, VA) and coated with a thin layer of gold for 3.5 minutes at a pressure of 55 mTorr and 10 mA current.

A polymethyl methacrylate-embedded specimen of human cancellous bone was also prepared by grinding and polishing as above, and placed in a top-referencing specimen holder (JEOL U.S.A., Peabody, MA). The calibration block was attached to the specimen holder at the same level as the bone specimen (Figure 4). This configuration was then placed in the SEM. All images were captured at 200x, using 30 keV accelerating voltage and 15 mm specimen-detector working distance.

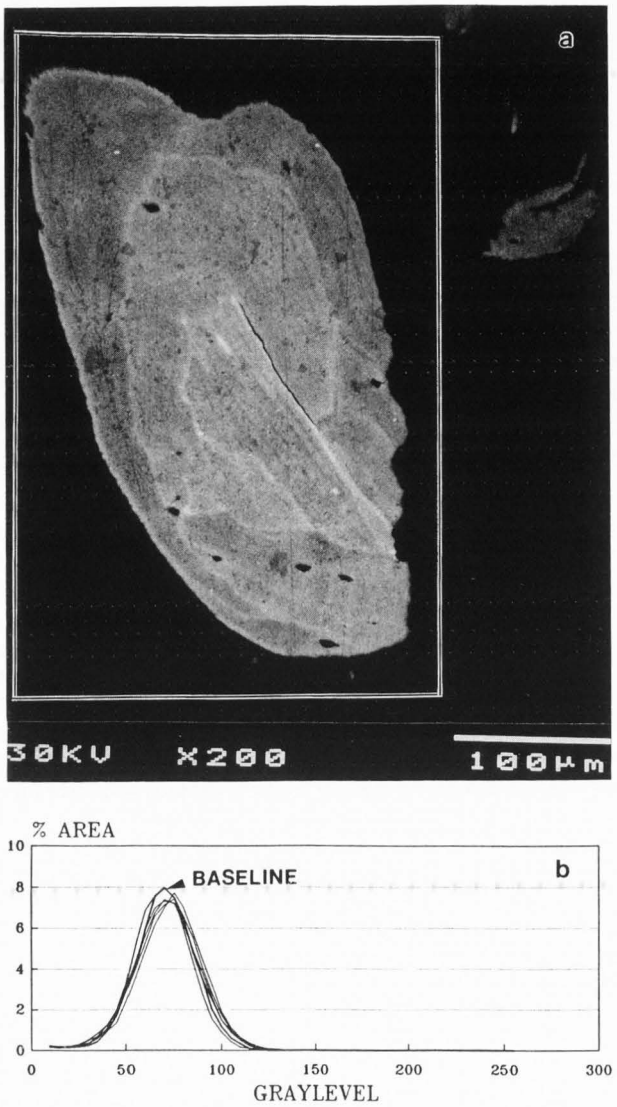

Figure 6 - a) Spicule of bone used for determining the effectiveness of each calibration trial. The box indicates the region of analysis. b) Backscattered electron profiles of baseline and five calibration trials for the bone specimen.The peaks representing polymethyl methacrylate and soft tissues (black, 0 graylevel) have been discarded.

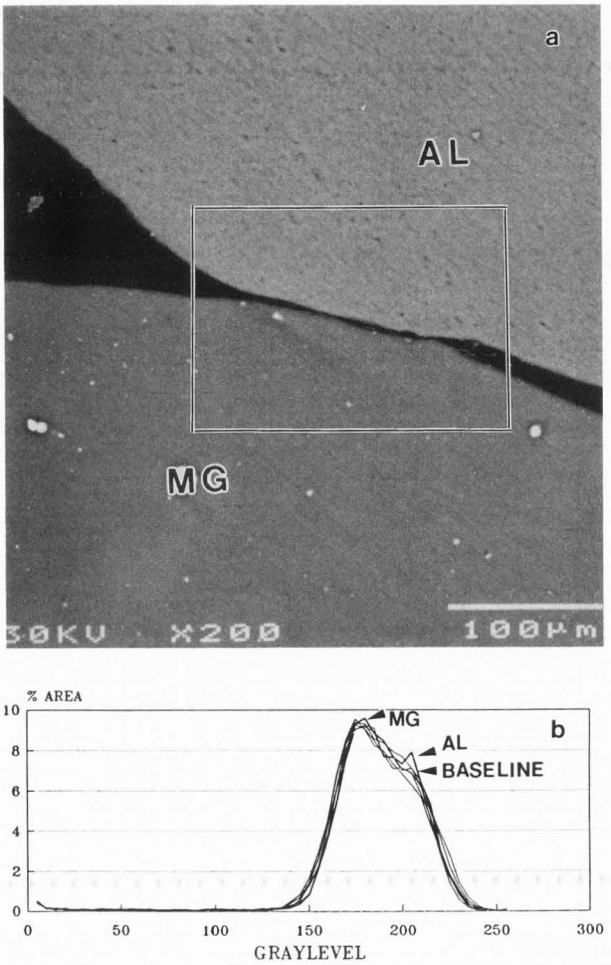

Figure 7 -a) BSE photomicrograph of calibration metals, Aluminum (AL) and Magnesium (MG).The box indicates the region of analysis. b) Backscattered electron profiles of baseline and five calibration trials for the standards.The peaks representing polymethyl methacrylate and soft tissues (black, 0 graylevel) have been discarded.

The steps involved in establishing the baseline BSEPs and calibrating the microscope are described in Figure 5. Bone, the material of interest, was used to initially establish the best visual graylevel range for the BSEP baseline conditions (Step 1). This was accomplished by adjusting the BSE brightness and contrast controls to give a visually optimal image of an isolated cancellous bone spicule. The resulting image had good graylevel separation, making it possible to observe the mineral phases in the bone (Figure 6a). Using only the x- and y-stage controls to shift the specimen, an image of the aluminum and magnesium standards (Figure 7a) was then captured (Step 2). The images of the metals and bone were then analyzed (Step 4), and the BSEP of each image was plotted, forming the baseline profiles (Step 5) against which the five subsequent calibration attempts would be compared.

The isolated cancellous bone spicule was chosen in order to allow the same bone to be analyzed during each calibration trial. Although the exact placement of the analysis region might have changed slightly between calibration trials, it always fully enclosed the bone spicule. This assured that the BSEP would not be altered, since the surrounding polymethyl methacrylate was always black (bin 1) and its contribution was always discarded. The image of the metal standards (Figure 7a) was too large at this magnification (200x) to allow the analysis box to enclose both metals fully.

These constraints required that the analysis region be placed according to minute, but easily identifiable marks at the perimeter of the analysis region.

The sequence of events involved in producing a calibrated bone image is given by steps 6 through 14 of Figure 5. Once the baseline BSEPs had been established (Steps 1-5), the BSE detector brightness and contrast controls were arbitrarily adjusted away from their original positions, and the filament current was slowly reduced to zero, then switched off (Step 6). After waiting a few minutes, a new operating session was initiated by switching on and saturating the filament (Step 7).

It has been noted that the human eye can distinguish up to 120 different graylevels between black and white [22] under good conditions, making it an efficient tool for coarse calibration. Coarse brightness and contrast settings (Step 8) were adjusted visually by the operator using the bone specimen, in order to provide a starting point for the graylevel calibration procedure. Next, a series of steps (9-12) was repeated until calibration was achieved. An image of the pure metal standards was captured and analyzed (Steps 9, 10). The resulting BSEP and the baseline profile (from Step 5) were both displayed on the microcomputer screen using a common graylevel axis. By comparing the two profiles, the operator could estimate the BSE detector brightness and contrast adjustments needed to align the two profiles. Calibration was achieved when both profiles were matched on the screen. Once calibrated, images of the calibration metals and bone were analyzed to determine the accuracy of the graylevel calibration procedure.

Filament and Electronics

As part of another study which required calibration of the BSE signal, images were captured from several specimens and stored during two 8 hour working periods and one 24 hour working period. Recalibration between specimens was performed as described above, except that only the magnesium calibration standard was used. During these working periods, we recorded observations in order to investigate the possible influence of electronic drift, filament fatigue, and other factors which combine to affect the stability of the whole SEM - image processing system. Frequent calibration attempts were made so that the most conservative time interval of stable SEM operating conditions could be determined. The system stability was tested by comparing the magnesium BSEP of the calibration attempt to a pre-established baseline BSEP for magnesium.

Results

Results of the calibration experiment are shown in Figures 6 and 7, and summarized in Table 1. The metal standards displayed a baseline BSEP with an weighted mean graylevel value of 186.2. The average of the WMGL values for five calibration trials was 185.6 $\pm$ 1.4 graylevels, indicating an average of 0.3 percent error per trial. All of the curves (Figure 7b) displayed a graylevel range between 150 and 255. It was possible to distinguish the magnesium peak, centered at approximately graylevel 175, from the aluminum peak centered around graylevel 205.

The bone spicule produced a baseline BSEP with a lower WMGL (70.4) than the pure metals. Using the operating conditions established by calibrating the pure metal standard images as previously described, the average of the WMGL values of bone for five calibration trials was 71.0 $\pm$ 1.9 graylevels. The average error over the five trials was 0.1 percent. The range of all the bone curves (Figure 6b) was between 40 to 105 graylevels.

Filament and Electronics

During 40 hours of SEM operation, there were times when power fluctuations in the main supply were erratic, causing large graylevel deviations in the image. Additionally, the filament failed once and was replaced. In both cases, it was possible to re-establish graylevel calibration by referring back to the baseline BSEP of the pure magnesium.

Two criteria were used to determine when recalibration was required. If the weighted mean graylevel of the BSEP deviated by more than 1.5% from the weighted mean graylevel of the baseline BSEP, or if visual inspection revealed that the BSEP curves did not match well, then graylevel calibration was performed. Even during periods in which operating conditions appeared relatively stable, it was determined that the graylevel calibration should be checked every 20 minutes.

Discussion

The utility of the calibration method developed in this study is dependent upon the choice of calibration standards. Ideally, the calibration material should have four characteristics. First, it should be homogeneous in composition to eliminate any variations in the intensity of the BSE signal which can be related to specimen location. The pure magnesium and aluminum standards that were chosen meet this criterion, since they are elemental materials, with purity documented by the suppliers. Second, the standards should be readily obtainable by researchers who intend to calibrate images using this method. Magnesium and aluminum specimens of certified purity are available from a variety of chemical supply companies. Third, the standard material should have an atomic number and WMGL value which approximates that of bone, without saturating the analog to digital converter. This provides the maximum graylevel range for imaging nearly the entire spectrum of mineral content in bone. Both magnesium ($Z = 12$) and aluminum ($Z = 13$) have graylevel ranges which approximates bone ($Z \approx 9$ to 11). Finally, the calibration standards should be stable under normal specimen processing and imaging conditions. Both aluminum and magnesium are known to form surface oxide layers when exposed to the

TABLE 1 - Results of five graylevel calibration trials using pure aluminum and pure magnesium as standards. Graylevel range includes all subrange "bins" with >1% image area. % Error = (Baseline WMGL - Trial WMGL / 255) x 100%.

		Baseline	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	AVE	SD
Calibration Metals Graylevel Profile	WMGL	186.2	184.8	184.8	184.2	187.5	186.5	185.6	1.4
	Range	155-205	150-255	150-255	150-230	155-230	150-230	N/A	N/A
	% Error	N/A	0.6	0.6	0.8	-0.5	-0.1	0.3	0.6
Bone Graylevel Profile	WMGL	70.4	68.4	70.4	70.9	73.4	72.1	71.0	1.9
	Range	45-100	40-95	40-100	40-100	45-105	40-105	N/A	N/A
	% Error	N/A	0.8	0.0	0.0	-1.2	0.7	0.1	0.8

atmosphere which help to stabilize them.

Aluminum forms an oxide film which is reported to be 5 to 10 nm in thickness [2, 16]. This surface oxide layer is stable, protecting the pure aluminum beneath from chemical changes which might otherwise occur. Magnesium surfaces exposed to the atmosphere also produce a thin oxide film. In contrast to the aluminum oxide layer, the magnesium oxide layer only decreases the rate of chemical change rather than halting it [1]. Thus, both the aluminum and magnesium wires used in this study have all four required characteristics. Aluminum may be slightly advantaged over magnesium for BSE signal calibration purposes because of this greater stability.

The depth dimension for the interaction volume created when 30 keV electrons strike an aluminum specimen has been calculated by the Kanaya-Okayama equation to be 8.3 μm [14], or approximately one thousand times the thickness of the oxide layer. The air-formed oxide layer has been reported to alter the BSE signal from an aluminum specimen by less than 1 percent [2].

Grain size, grain orientation and crystal structure are material properties which may vary significantly between pure metal specimens that have different processing histories. Since these may affect the BSE signal, the calibration standard should be standardized according to ASTM or other published material specifications. In this way, investigators can be assured that the same material is used in each study, permitting comparisons between data.

The metal calibration materials used in this study were in wire form. The standard extrusion and drawing methods that were used in their manufacture ensured that the wire was polycrystalline, with grains oriented along the length dimension. Since the angle of beam incidence relative to the specimen crystal structure is the basis for electron channeling contrast [20, 27], the size, number, and orientation of grains can be significant factors in determining the magnitude of the BSE signal which comes from the detector. Ball and McCartney [2] state that beam convergence of greater than 5 x 10^{-3} radians will average out any signal variation due to crystal orientation, making the electron channeling effects negligible. For the microscope used in this study with a 150 μm radius aperture, the convergence angle θ is given by

$$\theta = \tan^{-1} (150 \times 10^{-6}/WD)$$

where WD is the working distance to the specimen

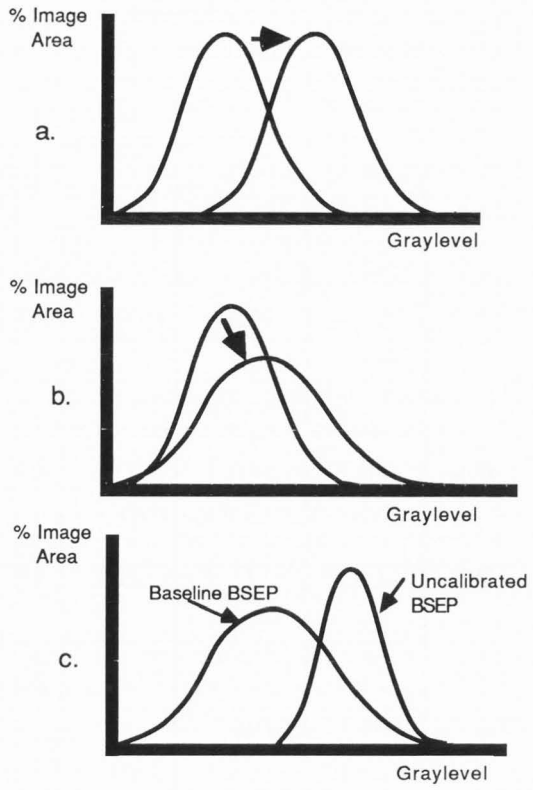

Figure 8 - a) Adjustments to the SEM brightness control produce a shifting of the BSEP along the horizontal axis. b) Adjustments to the SEM contrast control change the spread or base width of the image backscattered electron profile. c) Calibration is achieved by adjusting the microscope controls until the BSEP of the uncalibrated image is aligned with the backscattered electron profile of the baseline image.

surface (Personal Communication, Tony LaDotte, JEOL USA). Using this SEM, the required $\theta = 5x10^{-3}$ radians can be achieved, in theory, by limiting the working distance to less than 30 millimeters.

The heterogeneous nature of bone, with its non-uniform distribution of mineral produces BSEPs with a relatively large range. The local variations in mineral phase and composition can easily account for these graylevel variations within the specimen. It was surprising, however, that the pure calibration metals also displayed BSEPs with significant spread, rather than narrow peaks. This could be the result of a combination of several factors. The high contrast setting which was required to produce the bone mineral phase separation also tended to increase the number of graylevels in the metals. Surface contaminants present on the surface at the time of calibration might contribute to deviations in the BSEP of the metals by the introduction of topography effects and atomic number differences. Also, slight surface scratches and irregularities from the grinding process which were not macroscopically visible could be detected by the image analyzer and might also contribute to the spread of the peaks.

BSE brightness and contrast adjustments alter the image as illustrated in Figure 8. Brightness adjustments cause the digital image to appear brighter or darker by adding a constant value to or subtracting a constant value from each pixel [3], causing the BSEP to shift along the graylevel axis (Figure 8a). Contrast adjustments change the graylevel separation (dynamic range) in the image by multiplying or dividing pixel values by a constant [3]. This produces an increase or decrease in the BSEP spread (Figure 8b). Calibration of an image is accomplished when the image BSEP (Figure 8c) is closely matched with a pre-determined baseline BSEP, indicating that the correct brightness and contrast had been achieved. The principle of the BSE calibration technique was based upon first establishing optimal visual conditions for distinguishing the maximum number of graylevels in bone. Our starting point was to establish a bone image with good graylevel separation, and this dictated the Mg = 175 and Al = 205 weighted mean graylevel values. In a bone mineral study, each additional specimen would then be calibrated by matching the Mg and Al peaks from the calibration standard with the profiles obtained earlier.

The bi-modal shape of the BSEP for the metal standards is the result of using two calibration materials. While this bi-modal shape aids the operator in the process of calibration by providing additional visual landmarks, it is our experience that a single metal may also be used with reproducible results.

From another study conducted in our laboratory, we have observed that electron bombardment induces changes at the specimen surface of PMMA embedded bone specimens within 60 minutes of continuous scanning. A pronounced "bleaching" or whitening of the bone occurred over time. This "bleaching effect" is thought to be the result of changes that occur in the PMMA when it is bombarded with high energy electrons, as described elsewhere [13, 15]. Reid and Boyde [24] have attempted to reduce the degradative changes produced by the electron beam through the introduction of styrene into their embedding media. The PMMA used in this study was not similarly stabilized, because we were using pure metals rather than the bone specimen itself for calibration purposes and the period of actual beam interaction with the bone specimen was limited. Pilot studies from this laboratory have shown that the pure metal standards used in this study remained stable after more than six hours of continuous scanning.

Checking calibration every 20 minutes was considered to be a conservative interval since the operating conditions often remained stable for 30 to 40 minutes before calibration was actually necessary (>1.5% change in the weighted mean graylevel). Longer periods of stable operation may be obtained by isolating the SEM power lines or providing a voltage regulator to minimize power fluctuations.

This study has demonstrated that reproducible calibration of the BSE signal can be accomplished for bone mineral analysis using two pure metals as calibration standards. By selecting standards which are near to the range of atomic numbers for bone, all three materials could be maintained within the 0-256 graylevel scale using the same SEM settings. In the future, this choice of standards may permit average atomic number measurements by relating the BSE detector output (graylevel) of bone to an average atomic number, obtained by linear extrapolation from the standards in a manner similar to that used by Robinson for compounds [25].

The procedure described should be adaptable to any equipment which has the capability for BSE image generation, direct capture, and standard image analysis functions. Although this method is currently time-consuming, a real-time method is now being developed to increase the speed and ease of the calibration procedure. At this time, a 20 minute interval is recommended as a guideline for periodically checking calibration, when performing bone mineral analysis using the BSE technique. Results of this study have shown that the reproducibility of the calibration technique is excellent, even with the limitations noted.

Acknowledgments

This work was supported by the Department of Veterans Affairs Medical Center in Salt Lake City, Utah, the Whitaker Foundation, and the University of Utah Division of Orthopedic Surgery. The authors wish to thank Dr. A.U. Daniels of the University of Utah Division of Orthopedic Surgery for providing helpful comments about aluminum and magnesium metals, and Ms. Gwenevere Shaw for her patience and skill in typing this manuscript.

References

1. ASM Committee on Magnesium and Magnesium alloys (1979) Corrosion resistance of magnesium and magnesium alloys. In: Metals Handbook Ninth Edition, Baker H, Benjamin D (eds) American Society for Metals, Metals Park, OH, 596-600.
2. Ball MD, McCartney DG (1981) The measurement of atomic number and composition in an SEM using backscattered detectors. J Microscopy. 124(1), 57-68.
3. Baxes GA (1984) Digital image processing: A practical primer. Prentice-Hall, Englewood Cliffs, N.J., 32-37, 132-137.
4. Biltz RM, Pellegrino ED (1969) The chemical anatomy of bone. J Bone Joint Surg. 51-A(3), 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(1) 56-85.
6. Bloebaum RD, Reid SA, Bachus KN (1989). Backscattered electron imaging in the scanning electron microscope for bone-implant interface analysis. Trans 35th Orthop Res Soc. 14, 399.
7. Bohr H, Schaadt O (1983) Bone mineral content of femoral bone and the lumbar spine measured in women with fracture of the femoral neck by dual photon absorptiometry. Clin. Orthop. 179, 240-245.
8. Bohr HH, Schaadt O (1987) Mineral content of the upper tibia assessed by dual photon absorptiometry. Acta Orthop. Scand. 58, 557-559.
9. Boyde A, Jones S (1983) Backscattered electron imaging of skeletal tissues. Metab Bone Dis Rel Res. 5, 145-150.
10. Castaing R (1960) Electron probe microanalysis. In: Adv. Electron Phys., Marton LL, Marton, C (eds) Academic Press, New York, N.Y., 317-386.
11. Currey JD (1984) Effects of differences in mineralization on the mechanical properties of bone. Philos Trans R Soc Lond (Biol). B(304), 509-518.
12. Dougherty ER, Giardina CR (1987) Matrix structured image processing. Prentice-Hall, Englewood Cliffs, NJ, 37-69.
13. Everhart TE, Herzog RF, Chang MS, DeVore WJ (1972) 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.
14. Goldstein JI, Newbury DE, Echlin P, Joy DC, Fiori C, Lifshin E (1981) Scanning electron microscopy and x-ray microanalysis. A text for biologists, materials scientists, and geologists. Plenum Press, New York, 72-74.
15. Goldstein JI, Newbury DE, Echlin P, Joy DC, Fiori C, Lifshin E (1981) Scanning electron microscopy and x-ray microanalysis. A text for biologists, materials scientists, and geologists. Plenum Press, New York, 59-60.
16. Hollingsworth EH, Hunsicker HY (1979) Corrosion resistance of aluminum and aluminum alloys. In: Metals Handbook Ninth Edition, Baker H, Benjamin D (eds) American Society for Metals, Metals Park, OH, 204-599.
17. Holmes RE (1979) Bone regeneration within a coralline hydroxyapatite implant. Plast Reconstr Surg. 63, 626-633.
18. 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.
19. Lessig HJ, Meltzer MS, Siegal JA (1987) The symmetry of hip bone mineral density. A dual photon absorptiometry approach. Clin. Nucl. Med. 12, 811-812.
20. Lloyd GE (1987) Atomic number and crystallographic contrast images with the SEM: a review of backscattered electron techniques. Mineralogical Magazine. 51, 3-19.

21. Meade JB (1989) The adaption of bone to mechanical stress: Experimentation and current concepts. In: Bone Mechanics, Cowin SC (ed) CRC Press, Boca Raton, FL, 212-251.
22. Newbury D, Joy D, Echlin P, Fiori C, Goldstein J (1986) In: Advanced scanning electron microscopy and x-ray microanalysis., Plenum Press, New York, 183-186.
23. Reid SA (1986). A study of human skeletal maturation using the scanning electron microscope. PhD. Dissertation, University College London, UK.
24. 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.
25. Robinson VNE, Cutmore NG, Burdon RG (1984) Quantitative composition analysis using the backscattered electron signal in a scanning electron microscope. Scan Electron Microsc.;1984:II, 483-492.
26. Schaadt O, Bohr H (1982) Bone mineral by dual photon absorptiometry. Accuracy-precision-sites of measurement. In: Non-invasive bone measurements: Methodological problems, Dequeker J, Johnson C.C., Jr. (eds) IRL Press, Oxford, England, 59-72.
27. Schulson EM (1977) Electron channeling patterns in scanning electron microscopy. J. Mater. Sci. 12, 1071-1087.
28. 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.
29. Thompson DD (1980) Age changes in bone mineralization, cortical thickness, and Haversian canal area. Calcif. Tissue Int. 31, 5-11.
30. Vose GP (1962) Quantitative microradiography of osteoporotic compact bone. Clin Orthop. 24, 206-212.

Discussion With Reviewers

Alan Boyde: *We have experimented with exactly the same method of standardisation several years ago, and concluded that this is not a satisfactory means to ensure standardised running conditions. As is shown by the illustrations provided by the authors, the peaks from both Mg and Al are too diffuse and too close to assume that the method will work.*

A philosophical question is why, if the BSE method is to be regarded as useful, do pure metals give such broad histogram peaks? They have uniform atomic number and density.

Authors: We are aware that Dr. Boyde and his co-workers have attempted to calibrate the BSE signal. To our knowledge, this work has only been mentioned briefly (reference 24), but the methods and data are not given. We are unable to comment on their unpublished data. The authors of the present study as well as several medical students and fellows have routinely used this calibration technique for comparisons of bone mineral during the course of a number of studies conducted in our lab. We have now used this technique for more than one year and all users have found it to be very reproducible.

The widened bases of the BSEPs (figures 6b and 7b) are due to the increased SEM contrast setting, which is essential for visualizing the multiple mineral phases of the bone tissue. It would certainly be possible to obtain BSEPs with peaks falling in a single bin, or even a single graylevel. However, the detector contrast settings which would allow such an image are so low that they would be inadequate for bone mineral phase distinction and analysis, which is the object of of this method.

Alan Boyde: *The authors ignore channelling effects known to occur in metals. It would, therefore, be necessary to have an exact knowledge of the crystallographic orientation of the Mg and Al samples and to standardise upon the geometrical conditions for the collection of BSE in respect of the orientation of the crystallographic planes in the metal standards.*

Authors: Ball and McCartney (reference 2) estimate the effects of channeling at no more than 5% when the BSE Z-contrast mode is used. Even this value is an overestimate, however, due to the methods used in this study. As the specimens are ground and polished, the atomic structure in the crystals is changed to become more amorphous at the uppermost layers, thus reducing the effects of channeling. Ball and McCartney also stated that channeling effects were insignificant as long as the convergence angle of the beam was maintained sufficiently large (greater than $5x10^{-3}$ radians). By choosing our working distance to be 15 mm we have satisfied this requirement, and reduced the effects of electron channeling to negligible proportions. For an extrusion-formed metal wire which is always imaged on cross section, the calibration method described has proven to have a reproducibility error of less than 1% on average.

D.G.A. Nelson: *The pixel histogram spread of the pure metals also has a component from random fluctuations in the image digitization process. A good test for this, that will hopefully sharpen these histograms, is the use of image averaging algorithms during the digitization process.*

Authors: This is correct, and was accounted for in the study. During the image capture process, any fluctuations were minimized by the use of a frame integration formula which is built into the imaging system used for this study. Each image pixel value was obtained by summing the pixel values at its location from 8 scans, and then dividing by 8. Since the noise would be expected to be random in the image, the use of multiple image scans tends to minimize its presence. None of the images in this study showed any visual evidence of graylevel fluctuation.

D.G.A. Nelson: *In our experience, there is a large fluctuation at the individual pixel level when digitizing images from a video input. When looking at the value of one pixel during successive frame digitizations, its value can vary by 10-20%. This means that a "perfect" standard with a uniform greylevel value would have a BSEP with a width of 20-50 greylevels. In theory, adding a very large number of frames together will result in a BSEP with a width tending towards 1 without making any gain adjustments. Averaging 8 frames does not seem enough and that is why I believe you see broad histogram peaks for Al and Mg, as Alan Boyde suggests.*

Authors: The digital images that were used in this study are formed from the slow scan signals of the scanning electron microscope, not from video signals. Our experience has shown the frame-to-frame variability of individual pixels captured using this slow scan input to be approximately 11-22% of the 256 graylevel range, in agreement with your observations. After the image has been averaged over 8 frames, the pixel variability drops to 3-9%.

Regarding the number of frames to be averaged, the time required to capture 8 slow scan frames with our image analysis equipment is about 90 seconds. To produce an appreciable decrease in the noise (thereby decreasing the width of the BSEP as suggested) a very large number of frames would have to be averaged together. For the analog signal of the microscope (prior to digital conversion for the image analyzer), the signal to noise ratio (S/N ratio) increases in proportion to the square root of the number of signals averaged (See, for instance, R.A. Normann, Principles of Bioinstrumentation, J. Wiley and Sons, New York, 1988, 465-467). This means that improving the S/N ratio of the 8 frame image by a factor of 2 would require 32 frames, and over 5 minutes to capture one image. This is impractical, not only due to time constraints, but also because electronic drift or power fluctuations may change the SEM operating conditions during the actual image capturing events, and render the calibration ineffective.

In practice, the width of the BSEP does decrease as suggested (figure A), but the number of frames required for the BSEP to fall within a single graylevel is well over 256. As the number of frames is increased to 256, the WMGL value is changed by about 4 graylevels (figure B). The actual number of frames chosen is not as significant as assuring that all of the images from a study are averaged for the same number of frames.

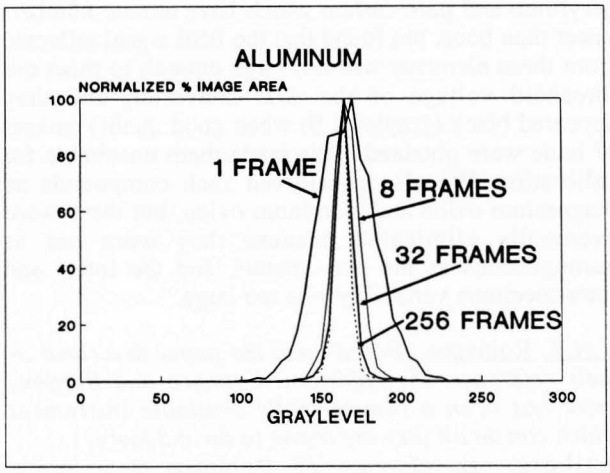

Figure A - Data showing how increased frame averaging reduces the BSEP width for pure aluminum.

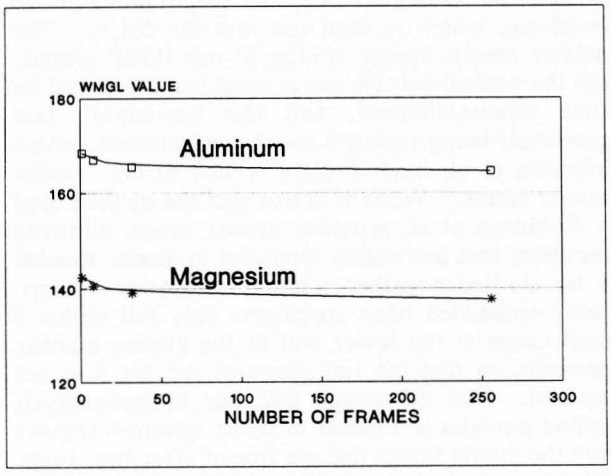

Figure B - Data showing the variation of the WMGL as aluminum and magnesium images are averaged for 1, 8, 32, and 256 frames. Note that the WMGL changes by only 4 graylevels for the pure metals.

D.G.A. Nelson: *The authors describe a calibration procedure for comparing BSE images of bone. However, this calibration procedure allows images to be compared, but does not describe the relationship between variation in Z and greyscale. This is because the authors only use two standards which lie outside the bone pixel greylevels. The use of a third standard with a low Z so that a calibration curve covering the bone pixel greylevels would allow the authors to relate greylevels to Z. A third standard would also eliminate the messy (and probably incorrect) procedure of matching the widths of pixel histograms during contrast adjustment.*

Authors: It has been well established (see, for example, references 2 and 20) that the BSE signal output for elements from approximately Z = 4 to Z = 40 closely approximates a straight line. Robinson (reference 25) extended this research to show that compounds also behave in a similar manner when subjected to BSE. Since two points define a straight line, the use of Mg and Al is sufficient to determine the average atomic number of the bone, using mean BSEP values.

The calibration materials that are used in this study were chosen carefully. We looked at pure

beryllium and pure carbon which have atomic numbers lower than bone, but found that the BSE signal reflected from these elements was not large enough to meet the threshold voltage of the A/D converter, and they appeared black (graylevel 0) when good quality images of bone were obtained. This made them unsuitable for calibration. We also considered such compounds as magnesium oxide and aluminum oxide, but these were eventually eliminated because they were not as homogeneous as the pure metals, and the inter- and intra-specimen variability was too large.

V.N.E. Robinson: *Based upon the paper described in their reference 25, Robinson, Cutmore and Burdon, there has been a commercially available instrument which can do all they are trying to do and more.*

Authors: In reference 25, Robinson et al. use a commercially available device called a "multi-channel analyzer" (MCA) (see also M.G. Hall and G.K. Skinner, J. Microscopy, v124, pp.69-75). This method translates the analog BSE detector output into a pulsed waveform, which is then fed into the MCA. The analysis results appear similar to our BSEP graphs, with the vertical axis (% image area) being replaced by pulse counts/channel, and the horizontal axis (graylevel) being replaced by channel number, which Robinson et al. have directly related to the "atomic number factor." While it is true that the method used by Robinson et al. provides greater range, allowing specimens that are widely separated in atomic number to be studied together without saturation effects, plastic-embedded bone specimens only fall within a small range at the lower end of the atomic number spectrum, so that the full range of the MCA is not required. The advantage that our image analysis method provides is a means to select specimen regions from the digital image that are free of scratches, voids, and other surface irregularities. If a high quality calibration standard with no surface irregularities were prepared, then we believe that the experiment described in this manuscript could be performed using either the MCA or the on-line image analysis equipment.

V.N.E. Robinson: *There are a number of problems with their methodology. First of all, both the magnesium and aluminum have an oxide layer on their surface. This means that it is not possible to get a good atomic number representation from looking at either of these metals. Secondly, they are far too close in atomic number to get any meaningful results. Thirdly, to suggest that the BSE signal between 4 and 40 closely approximates a straight line is an indication that the system being used is so inaccurate that it can not detect the curved nature of the relationship, see their reference 25 (Robinson et al.). This is supported by their curves figures 6b and 7b.*

Authors: As we described in the first three paragraphs of our discussion section, the surface oxide layers of the Mg and Al standards are relatively thin as compared to the interaction volume of the specimen,

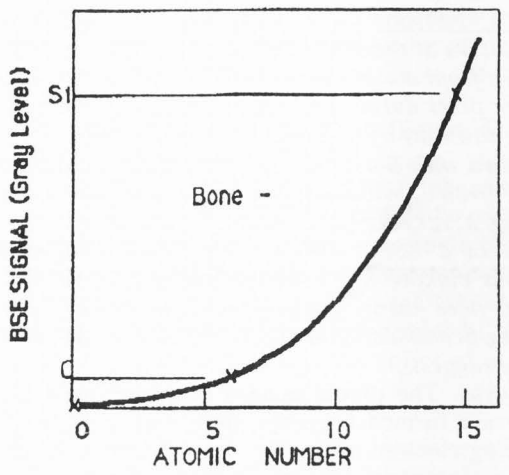

Figure C - Graph submitted by Dr. Robinson in his review.

and therefore they contribute very little (less than 1 percent for Al according to reference 2) to the BSE signal. Furthermore, even if the signal were affected significantly by the oxide layer, the fact that the oxide layers do not change appreciably throughout an imaging session makes them valid for the purpose of re-establishing the SEM operating conditions when imaging bone.

Our contention is that the region of the backscatter fraction vs. Z curve which includes bone can be modelled accurately with a straight line, not the whole curve. This is based upon figures 1,5, and A of reference 25, figure 4 of reference 20, and figure 1 of reference 2. Hall and Skinner (referenced above) have shown linearity for atomic numbers 6 through 32.

With respect to figures 6b and 7b of this paper, we do not understand what inaccuracies you attribute to our system or your basis for comparison, since these were intended to show the reproducibility of the technique.

V.N.E. Robinson: *What they should have done is as follows:*

a.) Set their zero standard, beam into a Faraday cage, signal output into Gray Level 1.

b.) Measure the Gray Level distribution for pyrolytic carbon (Z = 6) and also for pure silicon, (Z = 14). Neither of these oxidise significantly in times to conduct this experiment, i.e., over a few weeks. If these specimens were polished, this would remove any effect from Channelling contrast, due to crystallographic orientation. This gives three points, to which a freehand curve for the range 0-14 can be fitted, approximately as shown (Figure C).

c.) Determine the Gray Level distribution for the bone structure.

d.) Determine some mean atomic number, using the bone Gray Level and the curve.

A freehand curve is O.K., because the extrapolation is over a small range and between known points. It is a lot more accurate than a straight line of uncertain origin (Al and Mg oxides of unknown thickness) and close together. From the bone intensity, it would be possible to determine some sort of atomic number average for the bone, which then could be compared to the results obtained by others. As it is, a bone gray level of 71.0 (+/- 1.9) has no meaning to any other situation than the one they have used. Using their concept of a linear relationship between the Al, Mg and bone, Al = 205, Mg = 175, means that they have 30 gray levels per atomic number. Thus, 71 corresponds to an atomic number average of about 9. That at least enables some comparison with the results of others at some time in the future.

Authors: The methods that you describe are similar to those of this paper, except that we begin by first establishing a good quality image, showing the multiple mineral phases in our bone specimen. Carbon had been previously considered by our group, but when the proper operating conditions for bone are maintained, the BSE signal of carbon falls outside the sensitivity range of the analog to digital converter, making its graylevel indistinguishable from that of the embedding media. Silicon is currently being investigated.

The goal of this paper was to describe a calibration method, not to produce the average atomic number of bone. Your freehand curve suggests the obvious next step, which we have already taken (J.G. Skedros et al., Transactions of the Society for Biomaterials, volume XIII, 1990, 53)

V.N.E. Robinson: *The spread of their gray level curves is far too great to quote the accuracy of about 1% error.*

Authors: Referring to Table 1, the variation in the weighted mean graylevel is between 0 and 1.2 percent over five trials.