

# Quantitative Assessment of Bone Tissue Mineralization with Polychromatic Micro-Computed Tomography

Andrew J. Burghardt · Galateia J. Kazakia ·  
Andres Laib · Sharmila Majumdar

Received: 17 March 2008 / Accepted: 1 July 2008 / Published online: 7 August 2008  
© Springer Science+Business Media, LLC 2008

**Abstract** Micro-computed tomography ( $\mu$ CT) has become an important tool for morphological characterization of cortical and trabecular bone. Quantitative assessment of bone tissue mineral density (TMD) from  $\mu$ CT images may be possible; however, the methods for calibration and accuracy have not been thoroughly evaluated. This study investigated hydroxyapatite (HA) phantom sampling limitations, short-term reproducibility of phantom measurements, and accuracy of TMD measurements by correlation to ash density. Additionally, the performance of a global and a local threshold for determining TMD was tested. The full length of a commercial density phantom was imaged by  $\mu$ CT, and mean calibration parameters were determined for a volume of interest (VOI) at 10 random positions along the longitudinal axis. Ten different VOI lengths were used (0.9–13 mm). The root mean square error (RMSE) was calculated for each scan length. Short-term reproducibility was assessed by five repeat phantom measurements for three source voltage settings. Accuracy was evaluated by imaging rat cortical bone ( $n = 16$ ) and bovine trabecular bone ( $n = 15$ ), followed by ash gravimetry. Phantom heterogeneity was associated with  $<0.5\%$  RMSE. The coefficient of variation for five repeat measurements was generally  $<0.25\%$  across all energies and phantom densities. Bone mineral content was strongly correlated to ash weight ( $R^2 = 1.00$  for both specimen groups and both

threshold methods). Ash density was well correlated for the trabecular bone specimens ( $R^2 > 0.80$ ). In cortical bone specimens, the correlation was somewhat weaker when a global threshold was applied ( $R^2 = 0.67$ ) compared to the local threshold method ( $R^2 = 0.78$ ).

**Keywords** Micro-computed tomography · Computed tomography · Mineralization · Bone tissue · Density

Micro-computed tomography ( $\mu$ CT) has become an important tool for addressing a wide range of research questions related to the biology of bone and other calcified tissues. Quantitative measures derived from  $\mu$ CT have generally been limited to geometric and topological parameters describing the size and shape of trabecular and cortical bone [1–4]. These microstructural features have been shown to improve the prediction of bone strength independently of traditional measures of bone density [5, 6] and, therefore, represent a target for therapeutic manipulation. In addition to geometry, bone material composition is a determinant of bone strength [7] and can be positively modified pharmacologically [8]. Quantification of the mineral and organic properties of bone tissue has primarily been limited to two-dimensional (2D), destructive methods, including microradiography [9], infrared spectroscopic imaging [10], and backscatter electron imaging [11]. The ability to characterize tissue-level mineralization through 3D, nondestructive means would be an attractive complement to microstructural and biomechanical analyses already established for  $\mu$ CT.

Analogous to clinical quantitative computed tomography (QCT),  $\mu$ CT image intensity values (X-ray linear attenuation) reflect compositional properties of the imaged field of view (FOV). Where QCT gives apparent level measures of bone mineral density (BMD), the resolution of

A. J. Burghardt (✉) · G. J. Kazakia · S. Majumdar  
Musculoskeletal Quantitative Imaging Research Group,  
Department of Radiology and Biomedical Imaging, University  
of California, San Francisco, QB3 Building, Suite 203, 1700 4th  
Street, San Francisco, CA 94158, USA  
e-mail: andrew.burghardt@radiology.ucsf.edu

A. Laib  
Scanco Medical AG, Brüttisellen, Switzerland

$\mu$ CT is sufficient to potentially permit tissue-level characterization of the degree of mineralization. Recently, synchrotron radiation  $\mu$ CT (SR- $\mu$ CT) experiments have established the principle of densitometric  $\mu$ CT [12, 13] and have been applied to demonstrate pathological and pharmacological changes in tissue composition [14–17]. However, translation to widely available polychromatic  $\mu$ CT (X-ray tube desktop systems) faces several challenges. Specifically, polychromatic  $\mu$ CT is subject to beam-hardening effects, which result in spatial and geometric dependencies in the depiction of linear attenuation across the imaged FOV. Additionally, the signal-to-noise ratio (SNR) is typically lower for standard  $\mu$ CT, and commonly used cone-beam reconstruction methods are subject to geometric artifacts [18]. Nevertheless, a few commercial systems now provide protocols for calibrating reconstructed linear attenuation to mineral density.

The purpose of this study was to evaluate several important aspects of quantitative polychromatic  $\mu$ CT. Specifically, the following were characterized: (1) the error associated with sampling a relatively small section of a phantom comprised of hydroxyapatite (HA) and epoxy resin cylinders with known bulk concentrations but apparent microscale heterogeneity, (2) the short-term reproducibility of calibration measurements, and (3) the accuracy of mean measures of the degree of mineralization in trabecular and cortical bone through comparison to ash density reference measurements, with a focus on the effects of two different segmentation methods for identifying mineralized tissue.

## Materials and Methods

### $\mu$ CT

All imaging measurements in the present study were made on a commercial  $\mu$ CT system ( $\mu$ CT-40; Scanco Medical, Brüttiselen, Switzerland). This device consists of a micro-focus X-ray source with a 0.5-mm-thick aluminum filter that produces a narrow-angle cone beam, incident upon a  $1,024 \times 64$  element CCD detector. A previous generation of this scanner has been described in detail [19]. The reconstruction algorithm, beam-hardening correction, and HA phantom detailed below were provided by the manufacturer. Unless stated otherwise, all imaging measurements described below were acquired as follows: 1,000 half-field projection images were collected spanning 360 degrees, with each projection sampled for 200 milliseconds. The projections were then reconstructed across a  $1,024 \times 1,024$  matrix spanning a 36.9 mm FOV and resulting in an isotropic nominal resolution of 36  $\mu$ m. Projection frame averaging ( $n = 10$ ) was used for improved SNR for establishing the baseline calibration, and subsequently the

calibration was routinely monitored for quality-control purposes using single-frame acquisitions ( $n = 1$ ).

### Beam-Hardening Correction

To minimize geometric dependencies in reconstructed linear attenuation values, voltage-specific beam-hardening correction (BHC) factors were determined by the manufacturer. A wedge-shaped phantom ( $70 \times 35$  mm) composed of a 200 mg HA/cm<sup>3</sup> HA–epoxy resin mixture was imaged for three source voltages (45, 55, 70 kVp). This concentration was assumed to be a reasonable approximation of apparent level mineral density for cancellous bone biopsies/necropsies and whole-bone samples from small animals. This model for beam hardening assumes bone to be composed of two phases: a highly attenuating calcium phosphate mineral component and a soft tissue-equivalent, organic phase. The phantom geometry is shown in schematic form in Fig. 1a. For each voltage setting, the transmitted X-ray intensity through the phantom ( $I_{\text{meas}}$ ) and the unimpeded reference intensity ( $I_0$ ) were measured at 15 positions along the length of the wedge, corresponding to thicknesses of 0–7 mm. For unit attenuation, the deviation in the absorption curve ( $\ln[I_0/I_{\text{meas}}]$ ) from linearity— $\ln(I_0/I_{\text{theory}})$ —indicates the degree of beam hardening (Fig. 1b). A voltage-specific correction equation ( $\varepsilon_v$ , Fig. 1c) was derived using a third-order polynomial least squares fit of the difference— $\ln(I_0/I_{\text{theory}}) - \ln(I_0/I_{\text{meas}})$ —as a function of  $\ln(I_0/I_{\text{meas}})$ . Corrected absorption values ( $\ln[I_0/I_{\text{corr}}]$ ) are calculated at reconstruction time for tomographic acquisitions based on this correction equation for the corresponding voltage configuration:

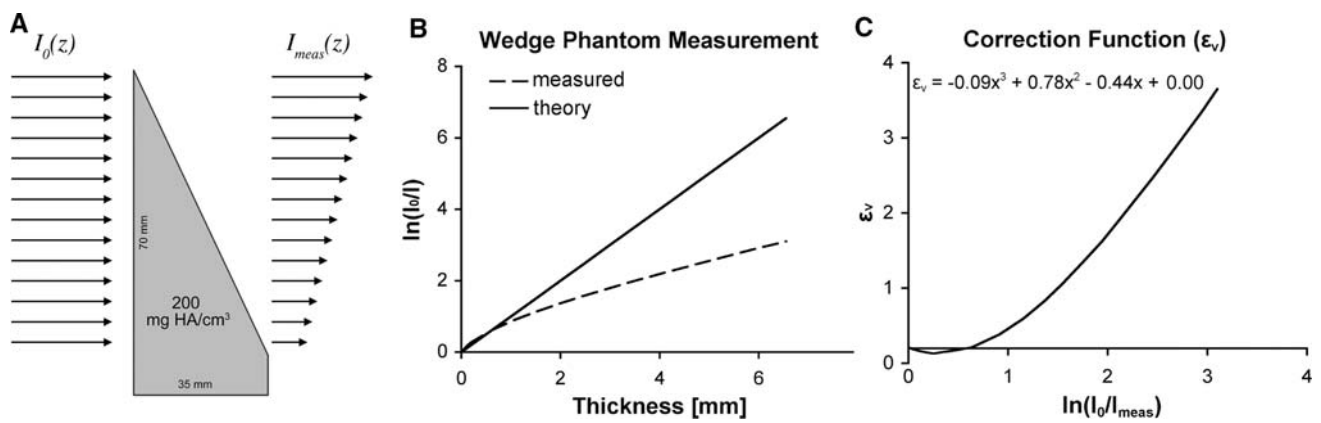
$$\ln\left(\frac{I_0}{I_{\text{corr}}}\right) = \ln\left(\frac{I_0}{I_{\text{meas}}}\right) + \varepsilon_v \left[ \ln\left(\frac{I_0}{I_{\text{meas}}}\right) \right] \quad (1)$$

### Density Calibration

Calibration of the reconstructed gray-scale attenuation values against mineral density was performed using a phantom composed of five cylinders of HA–resin mixtures with a range of mineral concentrations (0, 100, 200, 400, 800 mg HA/cm<sup>3</sup>), where 0 mg HA/cm<sup>3</sup> represents a soft tissue equivalent background devoid of mineral. For each anode voltage (45, 55, 70 kVp), 24 slices were acquired near the mid-length of the phantom. The mean attenuation value for each cylinder was calculated at each voltage setting, and individual linear relationships were determined against the known mineral concentrations:

$$[\text{HA}] = m\mu + b \quad (2)$$

where  $m$  and  $b$  represent the slope and intercept relating HA concentration ( $[\text{HA}]$  in mg HA/cm<sup>3</sup>) to linear



**Fig. 1** Schematic of BHC wedge phantom scan (a); representative wedge phantom measurement results for 70 kVp setting (b); plot of correction factor function,  $\epsilon_v$  (c)

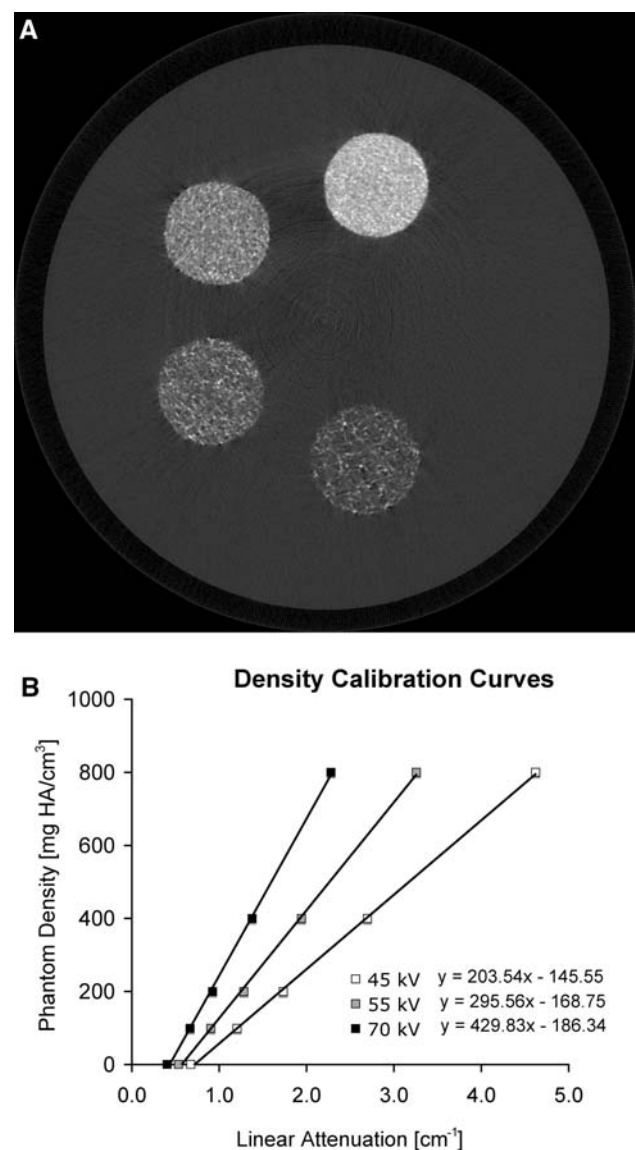
attenuation ( $\mu$ ). A representative tomographic image of the phantom and corresponding calibration plots are presented in Fig. 2. A simple script was developed to allow convenient calibration scan repetition at a fixed position for the reproducibility experiments described below and routine monitoring of system stability.

### Phantom Sampling

Due to the low solubility of HA, phantoms for bone mineral are typically composed of HA–epoxy resin composites. The granularity of such phantoms—readily apparent at  $\mu$ CT resolutions—is a potential source of error for the proposed calibration scheme [20, 21]. To characterize sampling limitations due to phantom heterogeneity, a single acquisition covering the full length of the phantom (16 mm) was acquired at 70 kVp. Mean attenuation values were determined for each HA cylinder in 10 regions of interest (ROI) randomly placed along the length of the phantom. This was repeated for six different ROI lengths ranging 0.9–13 mm (number of slices = 24, 48, 96, 182, 288, 360), which corresponded to even multiples of the detector stack size for this acquisition mode (24 slices). The root mean square error (RMSE) was calculated for each length against the mean attenuation of the full cylinder:

$$RMSE_{c,l} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\mu_{c,l,i} - M_c)^2} \quad (3)$$

where  $RMSE_{c,l}$  is the RMSE for a length  $l$  in the cylinder with concentration  $c$ ,  $\mu_{c,l,i}$  is the mean attenuation of the  $i$ th randomly placed ROI with length  $l$  in the cylinder with concentration  $c$ ,  $M_c$  is the mean attenuation for the full length of the cylinder with concentration  $c$ , and  $n$  is the number of regions sampled for each of the six lengths  $l$  (here  $n = 10$ ). RMSE has been reported as a percent of  $M_c$ . RMSE was also calculated for the calibration slope and intercept.



**Fig. 2** Representative slice from a phantom calibration scan (a) and typical calibration regression curves for each X-ray source voltage (b)

## Reproducibility

Short-term reproducibility of phantom calibration measurements was evaluated using the standardized calibration technique described above. Reproducibility was determined by acquiring five sets of calibration data for each voltage over the course of 3 days. Phantom repositioning was performed between repetitions, and scan localization was set to a fixed position in the scanner coordinate system. The mean linear attenuation coefficient was determined for each cylinder, for each repetition, and at each energy level. The coefficient of variation (CV%) was calculated for each cylinder at each energy level. Similarly, CV% was determined for the calibration slope and intercepts calculated for each repetition.

## Accuracy

The accuracy of  $\mu$ CT-derived measures of tissue-level mineral density was assessed by comparison to empirical measures of mineral content (ash gravimetrics) in two distinct groups of bone tissue: Cylindrical cores (8 × 18 mm) of bovine trabecular bone were acquired from the proximal femur ( $n = 8$ ) and proximal tibia ( $n = 7$ ). The mean morphological properties for this specimen group, calculated by direct 3D methods from  $\mu$ CT images described below [3], are listed in Table 1. A second set of specimens included mid-diaphyseal sections (4 mm in length) from the tibia ( $n = 9$ ) and femur ( $n = 9$ ) of adult Sprague-Dawley rats. All specimens were defatted using alternating cycles of deionized water and a mild biological detergent solution (1%, Tergazyme; Alconox Inc., New York, NY) under sonication and a mild water jet-wash. This process was necessary for gravimetry and was performed prior to imaging to avoid discrepancies in  $\mu$ CT and gravimetry data related to minor changes in bone mass as a result of the cleaning process. Specimens were stored at  $-20^{\circ}\text{C}$  when not being processed.

All specimens were imaged using a similar protocol as described earlier with two exceptions: the FOV was 16.4 mm (16  $\mu\text{m}$  nominal resolution) and no projection averaging was used in order to be consistent with a typical specimen protocol. Because the BHC and calibration

models assume a soft tissue-equivalent background, all bone specimens were immersed in deionized water and air bubbles were removed by repeat vacuum cycles at 28 psi. The SNR of the reconstructed images was calculated as the mean attenuation of the bone phase divided by the standard deviation of the water background.

Segmentation of the reconstructed gray-scale image into a binary bone/background image was carried out using two methods in an in-house customized version of IPL (IPL v5.01c-ucsf, Scanco Medical): (1) a global threshold automatically determined for each specimen using a common adaptive iterative method [22] and (2) a local adaptive threshold [23]. Briefly, the global adaptive iterative threshold method was implemented as follows: a 3D constrained gaussian filter ( $\sigma = 0.7$ , kernel = 3) was applied to remove high-frequency noise prior to the binarization step. Assuming a bimodal intensity distribution, the midpoint between bone and background peaks in the intensity histogram is iteratively determined according to Eq. 4:

$$T_{k+1} = \frac{\sum_{i=0}^{i=T_k} ih(i)}{2 \sum_{i=0}^{i=T_k} h(i)} + \frac{\sum_{i=T_{k+1}}^N ih(i)}{2 \sum_{i=T_{k+1}}^N h(i)} \quad (4)$$

where  $T_k$  is the trial threshold,  $h(i)$  is the histogram count for intensity  $i$ , and  $N$  is the maximum image intensity. When subsequent thresholds are identical, convergence has been reached, which effectively is equivalent to the midpoint between bone and background peaks. This value, determined for each specimen, was used as the global threshold. The mean global threshold values were equivalent to [HA] values of  $515 \pm 12$  and  $573 \pm 14$  mg HA/cm<sup>3</sup> for the trabecular and cortical specimens, respectively. The trabecular bone morphometric results presented in Table 1 were determined using global segmentation.

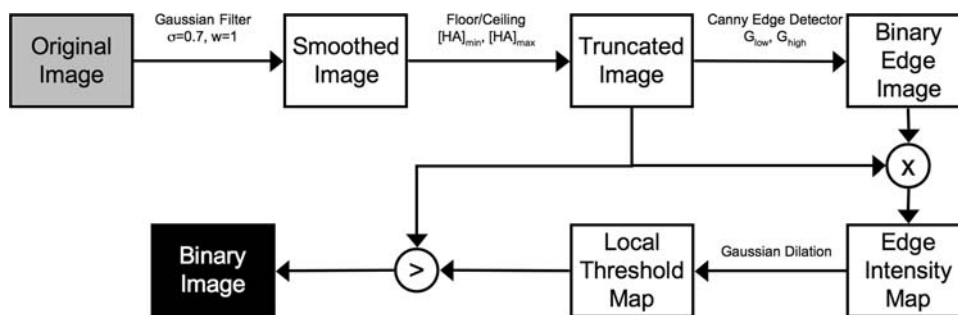
A summary schematic of the local threshold algorithm is presented in Fig. 3, while the complete algorithm details have been described elsewhere [23]. In this method, a Canny edge detector is used to locate bone marrow surfaces in the original gray-scale image. The intensity values of these edge voxels are then used as local threshold values for adjacent voxels. First, a 3D constrained gaussian filter ( $\sigma = 0.7$ , kernel = 3) is applied to remove high-frequency noise and texture. To further minimize false edge detection arising from noise or real image texture, the intensity values in the smoothed image were truncated using a floor and ceiling operator ( $[\text{HA}]_{\min} = 0$  mg HA/cm<sup>3</sup>,  $[\text{HA}]_{\max} = 1,050$  mg HA/cm<sup>3</sup>). A lower ( $G_{\text{low}}$ ) and upper ( $G_{\text{high}}$ ) gradient magnitude cut-off is required as input for the Canny algorithm to determine true edges. These parameters were automatically selected for each specimen based on the histogram of the gradient magnitude image.

**Table 1** Morphological properties of the bovine specimens (global segmentation)

	Femur ( $n = 8$ )	Tibia ( $n = 7$ )
BV/TV	$0.22 \pm 0.02$	$0.19 \pm 0.04$
Tb.Th ( $\mu\text{m}$ )	$154.9 \pm 4.5$	$173.3 \pm 10.9$
Tb.Sp ( $\mu\text{m}$ )	$637.8 \pm 51.4$	$761.5 \pm 123.7$
Tb.N ( $\text{mm}^{-1}$ )	$1.42 \pm 0.11$	$1.21 \pm 0.18$

BV/TV, bone volume/total volume; Tb.Th, trabecular thickness; Tb.Sp, trabecular separation; Tb.N, trabecular number

**Fig. 3** Summary flow diagram for local threshold algorithm



The same adaptive iterative approach used for the global segmentation was used to determine  $G_{low}$  from the gradient magnitude histogram [22]. As applied here, this technique iteratively selected the midpoint between the histogram peaks representing weak gradients (noise, texture) and strong gradients (bone–marrow interfaces). The histogram peak position corresponding to the strong gradient magnitude was taken as  $G_{high}$  (Fig. 4).

The mean tissue mineral density of the bone matrix (TMD), intraspecimen distribution of tissue mineral density (TMD.SD), and bone mineral content (BMC) were calculated in IPL based on the calibrated voxel attenuation values according to Eq. 2. To calculate TMD, the binary image from either segmentation process was used to mask the original gray-scale data. Each mask was eroded by two voxels to exclude partial volume. The mean linear attenuation value of voxels in the mask was converted to an equivalent HA concentration ( $mg\ HA/cm^3$ ) using the calibration slope and intercept for the appropriate source voltage. Additionally, the TMD.SD within each specimen was calculated as the standard deviation of TMD across all

bone voxels. BMC was calculated as the product of TMD and total bone volume.

Following the completion of imaging experiments, all specimens were transferred to preweighed porcelain crucibles, then placed inside a furnace (FB1300; Barnstead Thermolyne, Dubuque, IA) for 48 h at 600°C to remove all organic tissue components. The remaining ash mineral mass was weighed and corrected with the weight of the empty crucible to give ash weight. Bone volume determined by  $\mu$ CT (global threshold method) was used to calculate ash density. In this case, bone volume was determined prior to the voxel-peeling process used to calculate TMD.

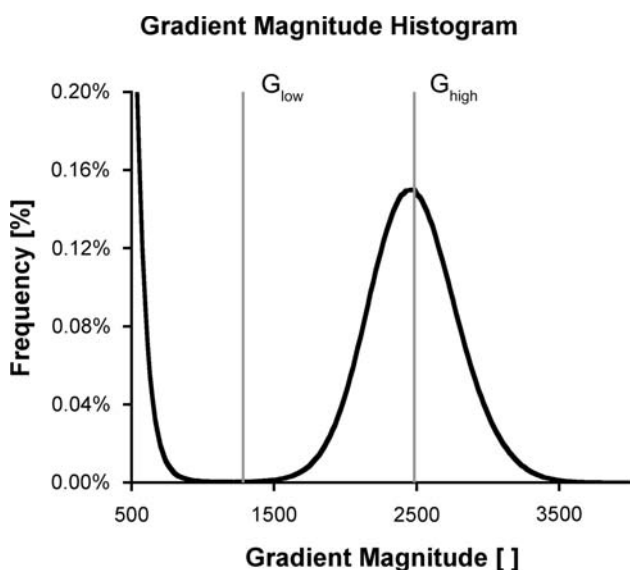
Statistics

The mean and standard deviation for all indices were calculated for trabecular and cortical specimens. A one-way analysis of variance (ANOVA) with Tukey’s honest significant difference (HSD) post hoc test was used to evaluate statistical differences for the reproducibility results. Regression analysis was performed to compare  $\mu$ CT-derived measures of mineralization to the gravimetric results. Linear correlations between BMC and ash weight were determined individually for trabecular and cortical specimens due to the large difference in weight between these groups. The correlation between ash density and TMD was also calculated separately for the trabecular and cortical specimen groups.

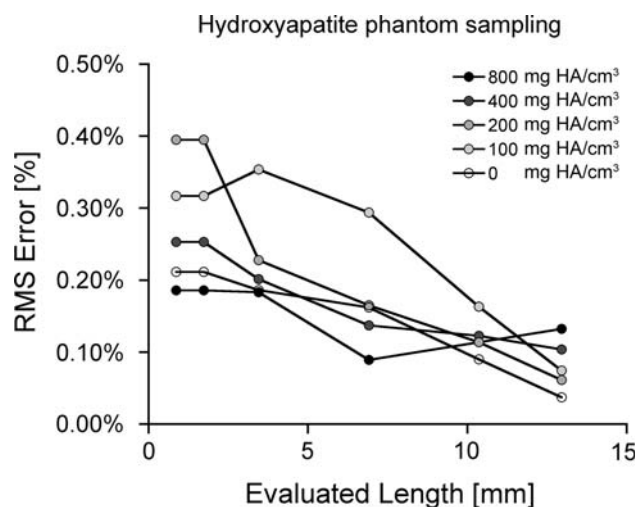
Results

Phantom Sampling

RMSE in the mean attenuation value for a phantom cylinder was found to decrease as the sampled length along the phantom increased (Fig. 5). For a single stack acquisition (24 slices, 0.86 mm) RMSE ranged 0.19–0.39%, while the longest sampled length (360 slices, 12.96 mm) yielded RMSE of 0.04–0.13%. The largest RMSE for attenuation was generally found in the low-density cylinders with non-zero HA concentrations and the lowest in the



**Fig. 4** Gradient magnitude histogram demonstrating the automated hysteresis threshold selection for the adaptive local threshold method



**Fig. 5** Error due to sampling a subsection of the whole phantom volume. RMSE was calculated for each concentration by measuring 10 randomly placed (in  $z$  direction) ROIs relative to the mean concentration for the entire phantom length

background (pure epoxy) and highest-density cylinders. The errors for the calibration ranged 0.15–0.21% for the slope and 0.29–0.50% for the intercept. The mean slope and intercept for the shortest and longest sampled lengths were not statistically different ( $P > 0.05$ ).

#### Reproducibility

Short-term reproducibility errors for phantom density calibration measurements were low (Table 2). The CV% of mean linear attenuation values calculated for repeat phantom scans ranged 0.03–0.21%. Reproducibility for the derived calibration slope and intercept was 0.09–0.20%. In general, CV% was highest for 70 kVp scans and lowest at 45 kVp. Results were analyzed using a one-way ANOVA between 45, 55, and 70 kVp acquisitions. This analysis revealed a significant effect for voltage ( $P < 0.0001$ ). Tukey's HSD test showed significant differences in CV% between each voltage configuration ( $P < 0.05$ ).

#### Accuracy

Representative images from the trabecular and cortical specimens are presented in Fig. 6, and mean densitometric

indices are given in Table 3. Using a typical acquisition protocol, a mean SNR of 27.9 was found for the trabecular bone images, while mean SNR was 28.9 for cortical bone images. The regression analysis results are plotted in Fig. 7. TMD ranged 900–1,100 mg HA/cm<sup>3</sup>, with the trabecular specimens spanning the lower half of the range and cortical specimens spanning the higher half. Global segmentation tended to yield a lower TMD value compared to local segmentation; however, this was only statistically significant for the cortical bone specimens ( $P < 0.01$ ), where local segmentation was more strongly correlated to ash density ( $R^2 = 0.78$  vs.  $R^2 = 0.67$ ). Qualitatively, cortical specimens appeared to have a greater range of mineralization levels. This was corroborated by a greater average TMD.SD for cortical specimens compared to trabecular specimens (111.4 vs. 93.2 mg HA/cm<sup>3</sup>,  $P < 0.0001$ ). BMC was very strongly correlated to ash weight ( $R^2 = 1.00$ ) for both specimen groups using either segmentation method. Local segmentation underestimated ash weight to a larger degree than did global segmentation (slope = 0.78 vs. 0.85, respectively).

#### Discussion

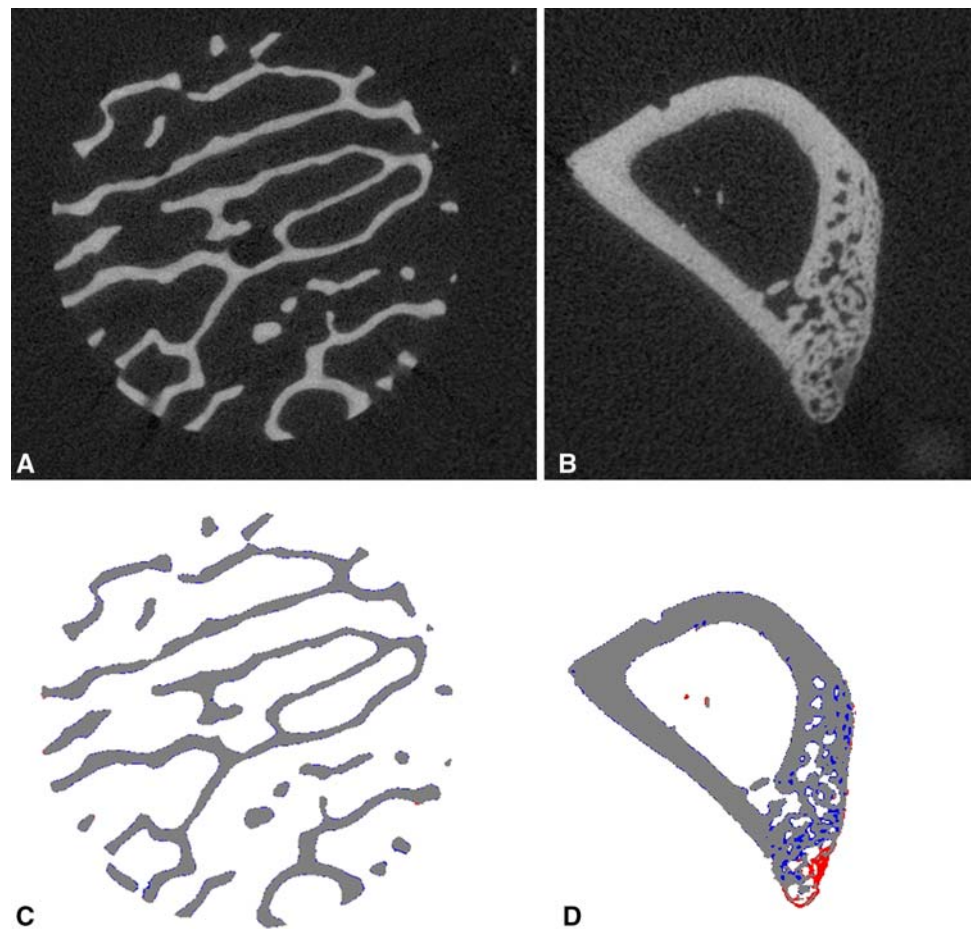
Quantitative characterization of tissue-level mineralization in a nondestructive, 3D fashion would be an attractive complement to bone morphometry and whole-bone measures of bone mass (BMD). In this study, the reproducibility and accuracy of a density calibration method for quantitative polychromatic  $\mu$ CT have been established. In addition, the accuracy of mean density parameters has been characterized in cortical and trabecular bone specimens by comparison to physical measures of mineral content.

Based on the results of phantom sampling experiments, a single stack acquisition (corresponding to 24 slices/0.864 mm) was deemed reasonable for routine calibration scans and, by extension, for evaluation of reproducibility and accuracy of the device for quantitative measures of tissue density. Errors related to reproducibility of calibration measurements at a fixed position in the phantom were low for all X-ray source voltages (CV < 0.25%), though CV tended to increase with increasing voltage. It is

**Table 2** Short-term reproducibility (CV%) of phantom calibration

Voltage (kVp)	Phantom cylinders					Calibration	
	0 (%)	100 (%)	200 (%)	400 (%)	800 (%)	Slope (%)	Intercept (%)
45	0.04	0.03	0.06	0.04	0.07	0.09	−0.20
55	0.08	0.07	0.10	0.13	0.13	0.15	−0.13
70	0.17	0.12	0.21	0.19	0.18	0.19	−0.10

**Fig. 6** Representative images for bovine trabecular specimens (a) and rat cortical specimens (b) with superimposed images comparing the global adaptive iterative and local adaptive segmentations (c, d), where gray represents common voxels, blue represents global-only voxels, and red represents local-only voxels



**Table 3** Mean and standard deviations for densitometric indices for trabecular and cortical specimens

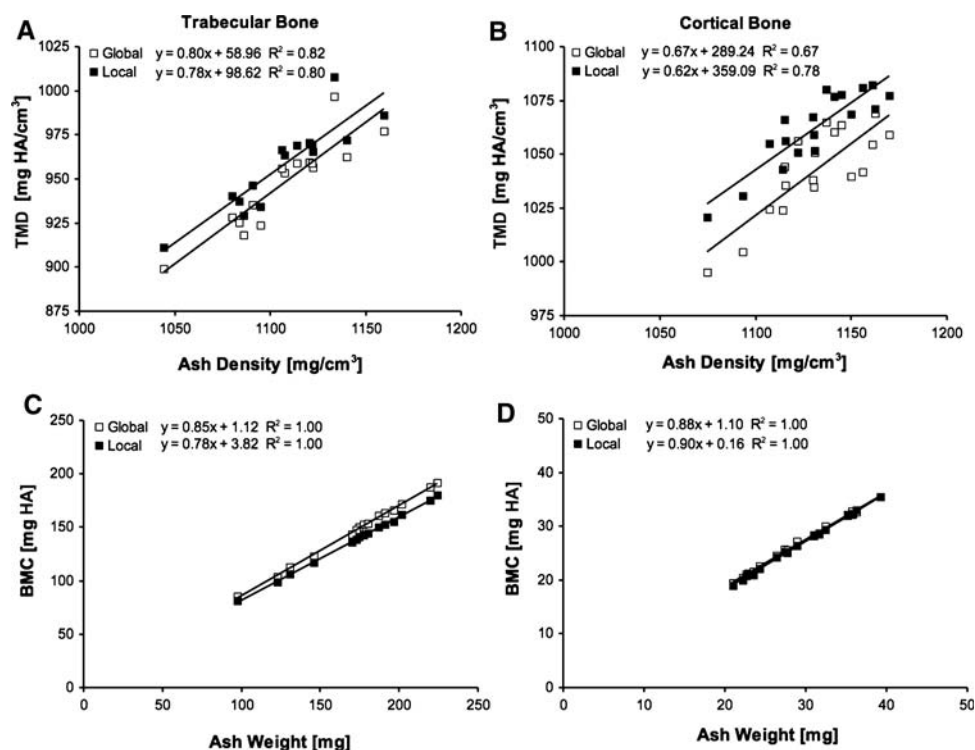
	Trabecular		Cortical	
	Global	Local	Global	Local
$\mu$ CT				
TMD (mg HA/cm <sup>3</sup> )	947.4 $\pm$ 25.2	958.1 $\pm$ 24.6	1,042.3 $\pm$ 20.4	1,062.0 $\pm$ 17.7
TMD.SD (mg HA/cm <sup>3</sup> )	99.1 $\pm$ 3.4	93.2 $\pm$ 2.5	129.3 $\pm$ 8.0	111.4 $\pm$ 5.5
BMC (mg HA)	147.8 $\pm$ 30.0	138.8 $\pm$ 27.6	26.8 $\pm$ 5.1	26.4 $\pm$ 5.2
Gravimetry				
Ash density (mg HA/cm <sup>3</sup> )	1,106.9 $\pm$ 28.4	–	1,130.8 $\pm$ 25.1	–
Ash weight (mg)	172.7 $\pm$ 35.3	–	29.1 $\pm$ 5.7	–

important to note that these results are specific to the phantom design, BHC method, and type of scanner used and may not be fully generalized across all  $\mu$ CT systems and phantom configurations.

An early study by Mulder et al. [21] presented a density calibration scheme for a similar polychromatic  $\mu$ CT system, which employed the manufacturer's generalized BHC used at the time and a polynomial calibration determined from sequential scans of liquid K<sub>2</sub>HPO<sub>4</sub> solutions. Accuracy was estimated by comparing mean density measures in a set of individual cylindrical phantoms to theoretical

attenuation values [24] based on the phantom concentration and mean effective energy (estimated from the same phantom scans). More recently, custom software BHC techniques have been introduced that are energy- and material-specific. Additionally, a standard density calibration phantom has been developed. The primary strength of this study is that the calibration methods applied include the new energy-specific BHC, optimized for bone tissue (using a 200 mg HA/cm<sup>3</sup> wedge phantom). Furthermore, in this study accuracy was determined in real bone specimens using empirical means (gravimetry) rather than estimated

**Fig. 7** Regression analysis results for TMD compared to gravimetric ash density for (a) bovine trabecular bone and (b) rat diaphyseal cortical bone. TMD was determined by segmenting the mineralized phase using either an adaptive iterative global threshold (*white*) or an adaptive local threshold (*black*). BMC was calculated as  $TMD \cdot BV$  and regressed against ash weight for trabecular bone (c) and cortical bone (d) specimens



comparisons to theory in homogenous solutions. Finally, the solid phantom used in this study is widely available and not subject to air bubble formation and the temporal instability associated with aqueous phantoms.

Two different segmentation techniques have been applied to identify bone voxels for calculating TMD in trabecular and cortical bone specimens: (1) a global method, where a single threshold level is automatically determined from the image intensity histogram, and (2) a local segmentation method, where voxel-specific thresholds are chosen based on local edge intensity values. The global iterative threshold method is advantageous in that it is fully automatic, is nonparametric, and requires no subjective input. Similar local adaptive threshold methods have been shown to provide improved segmentation under limited resolution scenarios and where bone tissue density is variable across the region of analysis [23, 25]. Visually, it is apparent that the cortical specimens ( $TMD.SD = 111 \text{ mg HA/cm}^3$ ) exhibit greater variability in TMD (Fig. 6b). Furthermore, the corresponding composite image (Fig. 6d) demonstrates that global segmentation excludes some significant regions of hypomineralization and does not capture some fine cortical pore structure. Our regression analysis results indicate that a more sophisticated segmentation can improve measures of TMD when a significant degree of variability in the degree of mineralization exists within a sample. Qualitatively, the discrepancy between the global and local segmentations was minimal for the trabecular specimens used in this

study (Fig. 6c). This coincided with a significantly lower average TMD.SD ( $93 \text{ mg HA/cm}^3$ ). These results suggest that simple global thresholding is sufficient for more homogenous specimens.

A limitation and potential source of error in the calibration scheme used here is the range of HA concentrations represented in the phantom ( $0\text{--}800 \text{ mg HA/cm}^3$ ). Because the biological range of TMD typically exceeds  $1,000 \text{ mg HA/cm}^3$ , extrapolation of the calibration curve is necessary for most applications in bone biology. Beyond these concentrations it is difficult to manufacture relatively homogeneous HA–epoxy resin mixtures. Recently, Schweizer et al. [26] proposed methods to manufacture pure HA-based phantoms with concentrations of approximately  $1,200\text{--}3,000 \text{ mg HA/cm}^3$ . While differences in background matrix between these phantoms must be taken into account, composite phantoms may overcome the need for extrapolated calibrations. It is not clear how multiple acquisitions for each density, as done by Schweizer et al., would differ from the calibration using a single phantom. Additionally, unlike clinical QCT, a limitation of this calibration method is that the calibration is performed as a separate acquisition from the specimen. While the calibration scan is highly reproducible over the short term, it is critical that stability is monitored over the longer course of an experiment to identify possible deviations.

The performance of the BHC has not been tested directly. Typically,  $\mu$ CT experiments are designed to compare specimens of comparable dimensions. For this reason, it was



deemed more relevant to consider the regression analysis independently for trabecular and cortical specimens. Nevertheless, it is important to note that there is a significant shift in the regression equations (i.e., the equation intercepts) between these two groups. This indicates a residual beam hardening-related geometric dependence in the depiction of mineralization. It is not clear to what extent these errors are affected by geometric and compositional variability; therefore, these results suggest that this BHC and calibration scheme are not well suited for densitometric comparisons across specimens with disparate geometries.

The results of this study recommend several important directions for future research. First, the accuracy results presented here only represent volume-averaged analysis of mean mineralization and do not demonstrate the local accuracy of mineralization depiction. Forthcoming studies presenting spatially resolved comparisons of polychromatic  $\mu$ CT-derived mineralization maps to SR- $\mu$ CT and Fourier transform infrared spectral imaging will address this question. Second, the resolution dependence of tissue density measures is unknown. The optical resolution of the system corresponds to approximately 1–2 voxels. For this reason, it is important to exclude two layers of voxels at the bone surface when calculating TMD because these voxels would be expected to contain significant partial signal from the marrow. Based on the morphology of our specimens (Table 1), mean trabecular thickness was spanned by approximately 10 voxels (16  $\mu$ m). Therefore, approximately 60% of the mean trabecular width was included in the density calculation following the peeling procedure. Accordingly, a greater fraction of bone volume would be excluded for specimens with higher surface areas (more rod-like structures); therefore, shape should be considered when establishing acquisition protocols. The results of this study do not necessarily extend to more limited resolution protocols. Finally, the accuracy results of this study were limited to two specimen types with moderate morphological variability. The feasibility of comparing densitometrics across specimens with more widely divergent geometries and compositions remains to be established.

In summary, several important aspects of quantitative polychromatic  $\mu$ CT have been characterized in this study. Errors due to sampling a small section of the calibration phantom do not exceed 0.5%, even when the sampled length is below 1 mm. Furthermore, calibration scans were found to be highly reproducible over the short term. Measures of bone mineral mass and density derived from  $\mu$ CT were moderately well correlated to gravimetric measures of mineral content when trabecular and cortical specimens were considered individually. This indicates that accurate measures of bone matrix mineralization are possible for sample populations with relatively consistent geometries. Finally, a simple global threshold technique is

preferred for moderately homogenous bone tissue. However, more sophisticated segmentation schemes can be beneficial when a significant degree of variability in mineralization is present.

**Acknowledgements** We thank Professor Tony Keaveny and the Orthopedic Biomechanics Lab at the University of California Berkeley for providing tissue and use of tissue-processing facilities. We also thank Margarita Meta of the University of California Los Angeles and Jennifer Schuyler of the University of California San Francisco Anesthesiology for providing and helping with the rat specimens. This study was supported with funds from the National Institutes of Health (RO1 AG17762 to S. M. and F32 AR053446 to G.J.K.).

## References

- Hildebrand T, Laib A, Muller R, Dequeker J, Ruegsegger P (1999) Direct three-dimensional morphometric analysis of human cancellous bone: microstructural data from spine, femur, iliac crest, and calcaneus. *J Bone Miner Res* 14:1167–1174
- Hildebrand T, Ruegsegger P (1997) Quantification of bone microarchitecture with the structure model index. *Comput Methods Biomech Biomed Engin* 1:15–23
- Hildebrand T, Ruegsegger P (1997) A new method for the model-independent assessment of thickness in three-dimensional images. *J Microsc* 185:67–75
- Odgaard A, Gundersen HJ (1993) Quantification of connectivity in cancellous bone, with special emphasis on 3-D reconstructions. *Bone* 14:173–182
- Ulrich D, van Rietbergen B, Laib A, Ruegsegger P (1999) The ability of three-dimensional structural indices to reflect mechanical aspects of trabecular bone. *Bone* 25:55–60
- Mosekilde L, Mosekilde L (1988) Iliac crest trabecular bone volume as predictor for vertebral compressive strength, ash density and trabecular bone volume in normal individuals. *Bone* 9:195–199
- Follet H, Boivin G, Rumelhart C, Meunier PJ (2004) The degree of mineralization is a determinant of bone strength: a study on human calcanei. *Bone* 34:783–789
- Boivin GY, Chavassieux PM, Santora AC, Yates J, Meunier PJ (2000) Alendronate increases bone strength by increasing the mean degree of mineralization of bone tissue in osteoporotic women. *Bone* 27:687–694
- Boivin G, Meunier PJ (2002) The degree of mineralization of bone tissue measured by computerized quantitative contact microradiography. *Calcif Tissue Int* 70:503–511
- Paschalis EP, Betts F, DiCarlo E, Mendelsohn R, Boskey AL (1997) FTIR microspectroscopic analysis of human iliac crest biopsies from untreated osteoporotic bone. *Calcif Tissue Int* 61:487–492
- Roschger P, Fratzl P, Eschberger J, Klaushofer K (1998) Validation of quantitative backscattered electron imaging for the measurement of mineral density distribution in human bone biopsies. *Bone* 23:319–326
- Nuzzo S, Peyrin F, Cloetens P, Baruchel J, Boivin G (2002) Quantification of the degree of mineralization of bone in three dimensions using synchrotron radiation microtomography. *Med Phys* 29:2672–2681
- Postnov AA, Vinogradov AV, van Dyck D, Saveliev SV, De Clerck NM (2003) Quantitative analysis of bone mineral content by X-ray microtomography. *Physiol Meas* 24:165–178
- Sone T, Tamada T, Jo Y, Miyoshi H, Fukunaga M (2004) Analysis of three-dimensional microarchitecture and degree of

- mineralization in bone metastases from prostate cancer using synchrotron microcomputed tomography. *Bone* 35:432–438
15. Nuzzo S, Lafage-Proust MH, Martin-Badosa E, Boivin G, Thomas T, Alexandre C, Peyrin F (2002) Synchrotron radiation microtomography allows the analysis of three-dimensional microarchitecture and degree of mineralization of human iliac crest biopsy specimens: effects of etidronate treatment. *J Bone Miner Res* 17:1372–1382
  16. Borah B, Ritman EL, Dufresne TE, Jorgensen SM, Liu S, Sacha J, Phipps RJ, Turner RT (2005) The effect of risedronate on bone mineralization as measured by micro-computed tomography with synchrotron radiation: correlation to histomorphometric indices of turnover. *Bone* 37:1–9
  17. Borah B, Dufresne TE, Ritman EL, Jorgensen SM, Liu S, Chmielewski PA, Phipps RJ, Zhou X, Sibonga JD, Turner RT (2006) Long-term risedronate treatment normalizes mineralization and continues to preserve trabecular architecture: sequential triple biopsy studies with micro-computed tomography. *Bone* 39:345–352
  18. Feldkamp LA, Davis LC, Kress JW (1984) Practical cone-beam algorithm. *J Opt Soc Am A* 1:612–619
  19. Ruegsegger P, Koller B, Muller R (1996) A microtomographic system for the nondestructive evaluation of bone architecture. *Calcif Tissue Int* 58:24–29
  20. Nazarian A, von Stechow D, Zurakowski D, Muller R, Snyder BD (2003) A quantitative technique to assess bone density from micro-computed tomography. *ORS Annual Meeting Transactions* ISSN 0149-6433
  21. Mulder L, Koolstra JH, Van Eijden TM (2004) Accuracy of microCT in the quantitative determination of the degree and distribution of mineralization in developing bone. *Acta Radiol* 45:769–777
  22. Ridler TW (1978) Picture thresholding using an iterative selection method. *IEEE Trans Syst Man Cybern* 8:630–632
  23. Burghardt AJ, Kazakia GJ, Majumdar S (2007) A local adaptive threshold strategy for high resolution peripheral quantitative computed tomography of trabecular bone. *Ann Biomed Eng* 35:1678–1686
  24. Berger MJ, Hubbell JH, Seltzer SM, Chang J, Coursey JS, Sukumar R, Zucker DS (1990) XCOM: photon cross sections database. National Institute of Standards and Technology, Gaithersburg, MD
  25. Waarsing JH, Day JS, Weinans H (2004) An improved segmentation method for in vivo microCT imaging. *J Bone Miner Res* 19:1640–1650
  26. Schweizer S, Hattendorf B, Schneider P, Aeschlimann B, Gauckler L, Muller R, Gunther D (2007) Preparation and characterization of calibration standards for bone density determination by micro-computed tomography. *Analyst* 132:1040–1045