Beam hardening artifacts in micro-computed tomography scanning can be reduced by X-ray beam filtration and the resulting images can be used to accurately measure BMD

Jeffrey A. Meganck\textsuperscript{a,b,1,2}, Kenneth M. Kozloff\textsuperscript{a,b,1}, Michael M. Thornton\textsuperscript{c,3}, Stephen M. Broski\textsuperscript{a}, Steven A. Goldstein\textsuperscript{a,b,*}

\textsuperscript{a} Orthopaedic Research Laboratories, Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI, USA
\textsuperscript{b} Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA
\textsuperscript{c} Imaging Research Laboratories, Robarts Research Institute, London, ON, Canada

Abstract

Bone mineral density (BMD) measurements are critical in many research studies investigating skeletal integrity. For pre-clinical research, micro-computed tomography (μCT) has become an essential tool in these studies. However, the ability to measure the BMD directly from μCT images can be biased by artifacts, such as beam hardening, in the image. This three-part study was designed to understand how the image acquisition process can affect the resulting BMD measurements and to verify that the BMD measurements are accurate. In the first part of this study, the effect of beam hardening-induced cupping artifacts on BMD measurements was examined. In the second part of this study, the number of bones in the X-ray path and the sampling process during scanning was examined. In the third part of this study, μCT-based BMD measurements were compared with ash weights to verify the accuracy of the measurements. The results indicate that beam hardening artifacts of up to 32.6% can occur in sample sizes of interest in studies investigating mineralized tissue and affect mineral density measurements. Beam filtration can be used to minimize these artifacts. The results also indicate that, for murine femora, the scan setup can impact densitometry measurements for both cortical and trabecular bone and morphologic measurements of trabecular bone. Last, when a scan setup that minimized all of these artifacts was used, the μCT-based measurements correlated well with ash weight measurements ($R^2 = 0.983$ when air was excluded), indicating that μCT can be an accurate tool for murine bone densitometry.

© 2009 Elsevier Inc. All rights reserved.

Introduction

The mechanical properties of bone tissue are determined by a variety of factors that range in size from the whole bone to the tissue ultrastructure. Bone mineral density (BMD) is one of these factors. BMD measurement methods continue to be an important area of research because ash content measurements, the gold standard for measuring mineral content in bone, are destructive and do not allow for measurement of site-specific mineral density patterns. To circumvent this, several non-destructive methods have been used to measure BMD in both clinical and basic science studies. While quantitative ultrasound methods have been used, the more typical methodologies use X-ray based imaging, such as single photon absorptimetry, dual photon absorptimetry, dual-energy X-ray absorptimetry (DXA), peripheral quantitative computed tomography (pQCT), and micro-computed tomography (μCT) \cite{1,2}. DXA measurements have been a valuable screening tool for bone diseases, but the accuracy of DXA measurements has been questioned and DXA cannot account for the three-dimensional architectural properties of bone \cite{2,3}. To overcome the challenges of DXA, peripheral quantitative computed tomography (pQCT) has been used to separate trabecular bone from cortical bone and estimate mechanical strength \cite{4,5}. While pQCT can assess both mineral content and structural properties in three dimensions from the same scan, its relatively low resolution can lead to errors when scanning small specimens \cite{6,7}. The resolution of micro-computed tomography (μCT) images is superior to clinical pQCT and, as a result, μCT has become the standard for accurate morphological and mineral density measurements in many pre-clinical studies \cite{7}.

The micro-radiographic techniques that form the foundation for quantitative μCT-based density measurements of bone were published
nearly two decades ago [8]. At that time, only relative densities were reported because these values were not calibrated or validated against a standard. This relative density is difficult to verify because the CT image acquisition process is subject to artifacts from partial-voluming, photon starvation, photon scatter, under-sampling, and beam hardening [9]. Beam hardening is arguably the most problematic for accurate BMD quantification and is caused by a preferential absorption of low-energy photons. This results in artifacts that appear as cupping, streaks, dark bands, or flare artifacts [9–11]. Because most laboratory and clinical CT systems use sources that generate polychromatic X-ray spectra, beam hardening artifacts must be taken into account for accurate quantitative imaging.

Corrections for beam hardening can be applied during the image acquisition process, during image reconstruction, or as empirical corrections. Procedures that are applied during the image acquisition process may require dual-energy imaging. This corrects for beam hardening and can be used to minimize cupping, streak, and flare artifacts [10,12,13]. Reconstruction based approaches to prevent beam hardening artifacts in the 3D image can also be used in both dual-energy and single-energy imaging if the input spectrum is known [14], or may require a thresholding step so that path lengths can be estimated [11,15,16]. Last, iterative reconstruction approaches based on Poisson distributions have also been proposed, both with and without the need for segmentation [17–19]. Despite the availability of sophisticated reconstruction algorithms, empirical corrections are arguably the most widely used class of beam hardening corrections. These can be applied prior to image reconstruction [20,21], or applied to the reconstructed image by applying polynomial basis functions, linearization procedures, calibration curves, or conversion tables [22–26]. These empirical approaches have been used in laboratory desktop μCT systems with some success, but the polynomial corrections that were used were not perfect and could not completely remove beam hardening artifacts for all cases [27–29].

Despite the wealth of possibilities to correct for beam hardening artifacts that result from the use of polychromatic X-ray spectra, it would be preferable to avoid beam hardening artifacts altogether. Monochromatic synchrotron radiation can be used for μCT and can allow for accurate BMD assessments, but limited synchrotron access can make studies difficult [30–32]. While a crystal monochromator or band-pass filters can be used to convert a polychromatic spectrum into a monochromatic or quasi-monochromatic spectra, [33,34] filtration is typically used to pre-harden the X-ray spectrum by removing low-energy X-rays. This common filtration approach can be enhanced by using water to ensure that the path lengths of the X-ray beam are approximately equivalent as they pass through the object being imaged [35]. Because beam hardening affects BMD measurements, the purpose of this study was to assess beam hardening artifacts associated with μCT imaging and ensure that accurate BMD measurements can be obtained. This was accomplished in a three-part study. In the first part, we investigated X-ray filtration in conjunction with beam flattening as a method to reduce cupping artifacts in bone-like materials. In the second part, we performed an investigation to determine if beam hardening affects μCT-based density measurements of murine femora. In the last part, we compared μCT-based measurements with ash weights of murine vertebrae to assess the accuracy and effectiveness of quantitative μCT as a method of bone densitometry.

Methods

Animal use

Bones were harvested from mice for the second and third portions of this study. These mice were primarily utilized in other experiments that were performed under approval of the University Committee on the Use and Care of Animals (UCUCA) at the University of Michigan. For the second part of this study, femora were dissected from 12 mice ranging in age from approximately 1 month to 10 months. These mice were maintained in colonies to investigate the effects of Thrombospondin 2, Thrombospondin 3, LRP5, and LRP6. While this study was not specifically designed to look at how these skeletal phenotypes change with age, these bones were chosen to minimize the effect of a particular bone phenotype or age on the results [36–39].

For the third part of this study, the 10th caudal vertebra was dissected from mice at 1, 2, 6, and 12 months of age from Brtl/ + and WT mice to obtain tissues with a wide variation in mineral content. Brtl/ + mice are heterozygous for a point mutation in col1a1 and have been used as a model for type IV osteogenesis imperfecta [40]. Prior to testing, endplates were removed at the growth plate using a diamond cutoff wheel as required for a separate study. A total of 8 vertebrae were analyzed (one per age group/genotype).

Part 1: Assessment and quantification of beam hardening-induced cupping artifacts

Phantom design

The ability of different filter materials and a beam flattener to reduce beam hardening-induced cupping artifacts was assessed using a tower phantom design with 11 separate circular tiers combined into one object (Fig. 1). The circular geometry was chosen because cupping artifacts are most prominent in circular sections. One phantom was made from a material that mimics cortical bone (SB3; physical density of 1.82 g/cm3) [41], and a second phantom was made from CB2-50% (physical density of 1.56 g/cm3) to represent bone with lower densities (Gammex RMI, Middleton, WI, USA) [42].

Image acquisition protocols and X-ray beam filtration

A commercially available μCT system was used (eXplore Locus SP, GE Healthcare Pre-Clinical Imaging, London, ON, Canada). This system uses a micro-focus source with an 8 μm focal spot size and a tungsten

---

**Fig. 1.** Phantom design for beam hardening assessments. This schematic demonstrates the design for the phantom that was used to assess beam hardening. There are 11 different thicknesses (see Table 2 for dimensions) that are circular in shape.
a source voltage of 80 kVp was used for this study to generate a spectrum that is primarily in the energy range of photoelectric absorption after filtration. In this system, five choices are available for beam filtration as X-rays exit the source: no added filtration, 0.254 mm aluminum (Al), 0.508 mm Al, 1.016 mm Al, and 0.254 mm Al followed by 0.254 mm copper (Cu) (Table 1). The specimens were immersed in distilled water and an acrylic beam flattener was used to equalize the beam path length within the field of view (Supplementary Fig. S1). Each of these filters affects the X-ray spectrum (Supplementary Fig. S2), so the effectiveness of each filter in reducing cupping was investigated. Current and integration times were selected to ensure that the photon statistics reaching the detector used approximately 75–85% of the dynamic range of the detector (Table 1). The scan setup utilized a magnification of 2.60 with 2 × 2 detector binning, resulting in an acquired pixel size of 18 μm. 720 projections were acquired for each individual scan over 360° of rotation. These projections were corrected using low-end and high-end outlier replacement in conjunction with a sinogram based long-term trend correction and reconstructed using a filtered cone-beam backprojection algorithm with a Ram-Lak filter to generate images with an isotropic voxel size of 18 μm. Noise measurement

X-ray beam filtration is known to affect the signal to noise ratio (SNR) and the contrast to noise ratio (CNR) [45,46]. To determine how phantom material influences noise measurements, a cube region of interest that was 25 × 25 × 25 pixels in size was placed in 9 separate locations of pure water at regions adjacent to tiers 2–10 of the tower phantom. Tiers 1 and 11 were not used due to proximity to the top and bottom of the scan, and to alleviate partial volume artifacts at these tower levels. This process was repeated for each filter for both the SB3 and CB2-50% phantoms, with and without the beam flattener. At each water location, standard deviation of the voxel grayscale values was calculated to estimate the noise level [47].

Beam hardening quantification

To quantify the amount of beam hardening that occurred as a function of beam filtration, specimen thickness, and specimen material, two-dimensional slices were taken for each combination of these variables. Histograms were used to select global threshold ranges to delineate the specimen from water; one range was chosen for the Al and Cu filter and another range was chosen for the remaining filters (Fig. 2). Single slice images were plotted and grayscale values were mapped onto a color scale using the limits determined from the global thresholding procedure. A line plot across the center was then created and a 30 pixel wide moving average filter was applied to reduce noise so that cupping could be visually detected.

Beam hardening effects will be most apparent when comparing voxels near the tower edge, which should be relatively unaffected, to voxels near the tower center, which will be most affected. A stochastic sampling approach was used to quantify the amount of beam hardening. A lognormal distribution was defined for the outer portion of the phantom to avoid partial volume artifacts (Supplementary Fig. S3A). A similar distribution was defined for the central portion of the phantom. These distributions both had values for the cumulative density function (CDF) of 0.995 at the half-tier radius. An outer voxel value was selected by defining a random radius based on the outer sampling distribution, and a pseudorandom angle selected from a uniform distribution over the interval [0,2π]. This process was

Table 1

The source current and detector integration times used for the image acquisition protocols.

<table>
<thead>
<tr>
<th>Supplementary material</th>
<th>Flattener</th>
<th>Source current (μA)</th>
<th>Integration time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Y</td>
<td>65</td>
<td>1600</td>
</tr>
<tr>
<td>0.254 mm Al</td>
<td>N</td>
<td>50</td>
<td>1600</td>
</tr>
<tr>
<td>0.508 mm Al</td>
<td>N</td>
<td>70</td>
<td>1600</td>
</tr>
<tr>
<td>1.016 mm Al</td>
<td>Y</td>
<td>80</td>
<td>1800</td>
</tr>
<tr>
<td>0.254 mm Al and 0.254 mm Cu</td>
<td>N</td>
<td>80</td>
<td>2100</td>
</tr>
<tr>
<td>0.254 mm Al and 0.254 mm Cu</td>
<td>N</td>
<td>80</td>
<td>6000</td>
</tr>
<tr>
<td>0.254 mm Al and 0.254 mm Cu</td>
<td>N</td>
<td>80</td>
<td>5400</td>
</tr>
</tbody>
</table>

Fig. 2. Extensive beam filtration results in a decrease in contrast. Histograms for the thickest tier in the phantoms made from (A) SB3 and (B) CB2-50% when scanned without the flattener. In all cases, the peak centered around the value of 0 represents water in the background of the image. The use of the 0.254 mm Al/0.254 mm Cu filter resulted in a shift downward in the voxel HU values, indicating less contrast in these images. When beam hardening was present, as seen most severely for the data obtained with no filtration, there was an alteration in the shape of the histogram peak for the phantom. Thresholds were also chosen based on these histograms. For SB3, the values 2200–3200 HU were used for the 0.254 mm Al/0.254 mm Cu filter and the values 2000–4300 were used for all other filters. For CB2-50%, the values 1100–2400 HU were used for the 0.254 mm Al/0.254 mm Cu filter and the values 1600–3200 HU were used for all other filters.
repeated using the central voxel distribution. In this manner, voxels were sampled with replacement for 10⁶ iterations, and the mean grayscale differences and percent differences were calculated between the sampled outer voxel population and inner voxel population. Beam hardening artifacts were considered significant when the difference between central and edge population means was greater than the baseline noise level.

**Statistical analysis**

To quantify how cupping artifacts affect BMD measurements, commercially available software was used to quantify the mineral density (MicroView 2.2 Advanced Bone Analysis Application, GE Healthcare Pre-Clinical Imaging, London, ON, Canada). In this software, the voxel grayscale value for SB3 in a manufacturer-provided phantom is correlated to a physical mineral density of 1073 mg/cc and voxel mineral contents are calculated using a linear correlation. The mineral densities for all other tissues are either interpolated or extrapolated based on this point and water at 0 HU with 0 mg/cc of mineral. For our own phantoms, these relationships resulted in an estimated mineral density of 1056 mg/cc for SB3 and 695 mg/cc for CB2-50%. The slight difference in estimated SB3 mineral densities was verified in a side-by-side comparison and may represent manufacturing inhomogeneities because these two materials were obtained at different times and came in slightly different forms. Global threshold levels were chosen based on the histograms of the phantom material to make tissue specific measurements.

**Image acquisition**

Femora were scanned on the same μCT system used in this first part of this study. The scan protocol entailed the use of the beam flattener with a 0.508 mm Al filter, the source set at 80 kVp and 80 μA, a magnification of 2.60, an exposure time of 1600 ms and an increment angle of 0.5°. The images were reconstructed using a Feldkamp cone-beam backprojection algorithm with a Ram-Lak filter to obtain an isotropic voxel size of 18 μm. To determine if artifacts affect mineral density measurements and common morphometric parameters in setups used to increase throughput, these bones were scanned using four methods. In the first method, 4 bones were simultaneously scanned using acquisitions limited to 200° of rotation to represent the shortest scan time. To examine artifacts in the reconstructed images caused by scanning multiple bones simultaneously, the same bones were then scanned individually over 200° of rotation. In this scanner, 200° of sample rotation was used because it is 180° plus the cone angle, representing the minimal complete data set for a reconstruction [48]. To help elucidate artifacts which may result from this minimal data set, these bones were scanned 4 at a time over 360° of rotation. As a relative gold standard image that avoided these limitations, each bone was then individually scanned over 360° of rotation. In every scan setup, each bone was placed in a sample holder away from the center of rotation (Supplementary Fig. 54). When 4 bones were scanned simultaneously, this configuration resulted in X-rays transmitting through two separate samples for some projection images. The scanner was calibrated once daily using a three point calibration of water, air, and SB3 to account for underlying day to day variation in the system stability.

**Image analysis**

A standard image analysis procedure was used to analyze the morphologic and mineral density measurements from these bones. Briefly, the images were reformatted using tricubic interpolations to align the long axis of the bone with a principal axis of the image. Images for each common bone were then registered using a rigid-body transform (translation and rotation) based on the selection of 4 sets of fiducial points. The femoral length was measured on the image of the bones scanned individually over 360°, and the regions of interest (ROIs) were normalized to this length. For the cortical bone, a ROI that was 20% of the femoral length was placed in the mid-diaphysis. For the trabecular bone, a ROI that was 10% of the femoral length was placed in the distal metaphysis. Trabecular bone was semi-automatically segmented from cortical bone by defining splines along the cortical–trabecular interface no more than 10 CT slices apart followed by linear interpolation between these selections. Because the images were registered, only one cortical and one trabecular ROI were defined for each bone. Based on these ROIs, the morphology and mineral density of the cortical bone and trabecular bone (using standard stereological techniques) were measured using commercially available software (MicroView 2.2 Advanced Bone Analysis, GE Healthcare Pre-Clinical Imaging). The same set of global thresholds was used for each image (one threshold for the cortical bone and one threshold for the trabecular bone).

**Statistical analysis**

To analyze the data, a subset of the variables was chosen to capture the morphologic and density properties of interest while limiting dependencies between the data points. To assess cortical bone morphology, the cortical thickness, moment of inertia, outer perimeter, and cross-sectional areas were measured. To assess the trabecular bone morphology, only the BV/TV ratio and trabecular number were analyzed because these are the only two independent measures using stereologic approaches. For densitometry assessments, both the tissue mineral content and tissue mineral density values were examined, even though they are mathematically related, because both have a unique physiologic interpretation. A repeated measures ANOVA was used for each variable to compare the data between the four scan protocols using a mixed linear model (SPSS 16, SPSS Inc., Chicago, IL, USA). Post-hoc tests were performed to determine pairwise differences using data for the bones scanned individually over 360° as a reference. Bonferroni corrections were used to adjust for multiple comparisons (α = 0.05 after adjustment).

**Part 3: Comparison of μCT-based density measurements to ash weights**

Vertebral specimens were placed in a custom polycarbonate scanning holder, allowing for simultaneous acquisition of data from all 8 specimens in one scan. As in the second part of this study, X-rays transmitted through up to 2 bones in this setup and a beam flattener was used. This setup was similar to that used when multiple bones were scanned in part 2 of this study. The small size allowed the vertebrae to be stacked in the sample holder. Scanning was performed using a commercially available μCT system (EVS [now GE Healthcare Pre-clinical Imaging] MS-8, London, ON, Canada). Samples were scanned using 495 projections at 80 kVp, for a total exposure of 88 mA*s, and the images were reconstructed at an 18 μm isotropic voxel size using a Feldkamp cone-beam backprojection algorithm. A 0.5 mm aluminum filter, which is very close to the 0.508 mm Al filter used in the prior studies, was used to minimize beam hardening effects.

**Image analysis**

Apparent BMD (BMDapp) was measured using manufacturer-provided software (MicroView 1.1, GE Healthcare Pre-Clinical Imaging, London, ON, Canada), representing the bone mineral content...
(BMC) of the specimen normalized by a parallelepiped region of interest enclosing the vertebra. A typical sampling of voxels consisted of two to three peaks: a high intensity peak (mineralized tissue), a peak centered around zero (water), and a peak around $-1000$ (air trapped within the specimen or reconstructed volume). A global threshold was applied to each volume and the bone volume fraction ($\text{BV/TV}$) of the region of interest was obtained. Volumetric BMD ($\text{vBMD}$), representing average tissue mineralization of each specimen, was calculated by normalizing $\text{BMD}_{\text{app}}$ by $\text{BV/TV}$:

$$\text{vBMD} = \frac{\text{BMD}_{\text{app}}}{\text{BV/TV}}.$$  

Because this calculation included a summation operation of all grayscale values in the region of interest, two parameters were

---

Fig. 3. Noise levels increased with extensive beam filtration and use of a beam flattener. Mean noise levels for water adjacent to the tiered phantoms made from (A) SB3 and (B) CB2-50%. ANOVA analyses indicated that use of a beam flattener results in an increase in the noise level. In addition to this, filtration also affected the baseline noise level. * indicates significance in comparison to the 0.254 mm Al/0.254 mm Cu filter and + indicates significance in comparison to the 1.016 mm Al filter. Data are presented as the mean ± one standard deviation.
calculated for BMC and vBMD. In the first parameter, a full range (BMCfull; every voxel in the region of interest) was used, and the second parameter (BMCexclude) a value of −500 HU, halfway between air (−1000 HU) and water (0 HU), was chosen to exclude voxels with values less than −500 to prevent voxels that may have included air from biasing the mean after summation.

Ash content assessment
To validate BMC measures, the vertebrae were dried overnight at 110 °C and the dry tissue weight was measured after completion of the μCT scanning. Specimens were then placed in individual ceramic crucibles and ashed at 800 °C for 4 h. At the end of this period, the ash weight was measured (BMCash) and the calcium content was quantified using a standard colorimetric assay (Sigma 587). A linear regression was then performed to assess the ability of the μCT-based BMCfull and BMCexclude measurements to predict ash weight, dry weight, and calcium content.

Statistical analysis
The statistical significance of the slope of the lines from the linear regressions for BMCfull and BMCexclude was compared using commercially available software (GraphPad Prism 4.0a, GraphPad Software, La Jolla, CA). In order to compare the ability of the BMCfull and BMCexclude to predict ash weight, dry weight, and calcium content, a statistical test comparing the correlation coefficients between these two methods was used [49].

Results
Part 1: Assessment and quantification of beam hardening-induced cupping artifacts
Noise measurement
To characterize effects of tower thickness, material, and presence or absence of the beam flattener on baseline noise levels, cubic regions of interest in 9 locations of water were analyzed adjacent to 9 levels of both tower phantoms. In the presence of each material, both the beam flattener and filter affect baseline noise levels, but there is no interaction between the two (for the interaction term: p = 0.9868 for the CB2-50% phantom; p = 0.8342 for the SB3 phantom). Because no interaction was present, the effects of the flattener and the filters were interpreted separately. In the presence of both materials, there was a statistically significant increase in the amount of measured noise when the flattener was used (Figs. 3A and B; p = 0.0009 for the CB2-50% phantom; p = 0.0276 for the SB3 phantom). When the noise levels were compared across the filters, results of the ANOVA analysis indicated that there were some statistically significant differences (p < 0.0001 for the CB2-50% phantom; p = 0.0001 for the SB3 phantom). Post-hoc tests indicated that there was more noise with the 0.254 mm Al/0.254 mm Cu filter for all cases except when compared to the 1.016 mm Al filter with SB3. In the presence of CB2-50% there was also more noise with the 1.016 mm Al filter than with the 0.254 mm Al filter (Fig. 3). To determine a cutoff point for beam hardening quantification (see the section on Beam Hardening Quantification below), the noise measurements for no filtration and the three Al filters were averaged resulting in baseline noise levels of 116 HU and 109 HU in the presence of CB2-50% with and without the flattener, respectively, and 124 HU and 119 HU in the presence of SB3 with and without the flattener, respectively. Data from the 0.254 mm Al/0.254 mm Cu filter was not included in these calculations because no measurable cupping occurred with this filter.

Beam hardening quantification
Histograms were used to establish global threshold levels (Fig. 2). Lower threshold ranges were required for the filter with Cu, coinciding with a decrease in contrast as beam filtration increases. In addition, there was a change in the histogram peak shape for thicker portions of the phantom due to cupping artifacts. To determine the specimen thickness where beam hardening artifacts begin to occur, a stochastic approach was used to find the difference in grayscale values between the outer and central portions of the

<table>
<thead>
<tr>
<th>Phantom material</th>
<th>Test condition</th>
<th>Filtration</th>
<th>Phantom thickness (mm)</th>
<th>CB2-50%</th>
<th>SB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>With flattener</td>
<td>None</td>
<td>3.4%</td>
<td>5.3%</td>
<td>9.3%</td>
<td>8.9%</td>
</tr>
<tr>
<td></td>
<td>0.254 mm Al</td>
<td>1.2%</td>
<td>4.0%</td>
<td>7.5%</td>
<td>7.5%</td>
</tr>
<tr>
<td></td>
<td>1.016 mm Al</td>
<td>1.4%</td>
<td>3.8%</td>
<td>7.3%</td>
<td>7.0%</td>
</tr>
<tr>
<td></td>
<td>0.254 mm Al and</td>
<td>−1.9%</td>
<td>0.1%</td>
<td>3.5%</td>
<td>3.5%</td>
</tr>
<tr>
<td></td>
<td>0.254 mm Cu</td>
<td>−7.1%</td>
<td>−7.3%</td>
<td>−4.6%</td>
<td>−5.5%</td>
</tr>
<tr>
<td>Without flattener</td>
<td>None</td>
<td>5.4%</td>
<td>9.0%</td>
<td>12.5%</td>
<td>12.9%</td>
</tr>
<tr>
<td></td>
<td>0.254 mm Al</td>
<td>5.0%</td>
<td>6.8%</td>
<td>10.8%</td>
<td>10.2%</td>
</tr>
<tr>
<td></td>
<td>0.508 mm Al</td>
<td>1.5%</td>
<td>4.3%</td>
<td>8.6%</td>
<td>8.3%</td>
</tr>
<tr>
<td></td>
<td>1.016 mm Al</td>
<td>−2.3%</td>
<td>0.6%</td>
<td>4.6%</td>
<td>4.6%</td>
</tr>
<tr>
<td></td>
<td>0.254 mm Al and</td>
<td>−9.0%</td>
<td>−8.8%</td>
<td>−5.5%</td>
<td>−6.0%</td>
</tr>
<tr>
<td></td>
<td>0.254 mm Cu</td>
<td>−10.5%</td>
<td>−8.3%</td>
<td>−5.3%</td>
<td>−6.5%</td>
</tr>
</tbody>
</table>

|                  | None           | 5.0%       | 7.8%   | 12.2%  | 12.3%  | 14.4%  | 18.1%  | 20.7%  | 23.6%  | 26.0%  | 30.7%  | 32.6%  | 35.1%  |
|                  | 0.254 mm Al    | 1.8%       | 4.8%   | 7.8%   | 8.3%   | 9.4%   | 12.2%  | 15.9%  | 17.4%  | 19.6%  | 22.5%  | 25.2%  | 29.5%  |
|                  | 0.508 mm Al    | −0.3%      | 1.9%   | 5.3%   | 6.0%   | 6.9%   | 9.9%   | 12.8%  | 14.8%  | 16.3%  | 19.1%  | 20.9%  | 23.3%  |
|                  | 1.016 mm Al    | −2.8%      | −1.6%  | 2.2%   | 2.8%   | 3.2%   | 6.5%   | 8.1%   | 9.6%   | 11.1%  | 13.5%  | 14.8%  | 15.0%  |
|                  | 0.254 mm Al and | −10.1%   | −7.7%  | −7.2%  | −6.6%  | −6.7%  | −5.0%  | −4.5%  | −3.2%  | −1.8%  | −0.6%  | −0.8%  |
|                  | 0.254 mm Cu    | −10.0%     | −7.7%  | −7.2%  | −6.6%  | −6.7%  | −5.0%  | −4.5%  | −3.2%  | −1.8%  | −0.6%  | −0.8%  |

Measurements indicate the percentage difference between voxel grayscale values at the inner and outer portions of the phantom. Shaded fields indicate areas where beam hardening artifacts occurred.
phantom. We assumed that significant beam hardening artifacts occurred when this difference was greater than the baseline noise level. The results indicate that both filtration and use of the beam flattener affect the onset of beam hardening artifacts. Table 2 summarizes the percentage changes between the inside and outside of the phantom for all cases and Fig. 4 demonstrates the results graphically for the case of SB3 with the flattener (the actual quantifications of the HU changes are included in the Supplementary Table S1 and visualizations of the remaining data are included in Supplementary Fig. S5 through S7). The thickness where cupping artifacts becomes significant increases as the amount of filtration increases. No significant cupping artifacts occurred for any thickness tested for either material when the filter containing Al and Cu was used. After cupping began to occur, the magnitude of the change increased as the phantom became thicker.

BMD estimation

The ultimate utility of μCT is bone mineral density measurements rather than beam hardening quantifications. We calculated the tissue mineral densities (TMD) for the SB3 and CB2-50% phantoms from scans that were obtained both with, and without, the beam flattener (Fig. 5). Here, we define TMD as the Bone Mineral Content of a region of interest normalized by bone volume of that region, creating a true volumetric measure of mineral density independent of bone size and shape. There was no measurable change in the TMD measurements from the expected density values when the filter of 0.254 mm Cu and 0.254 mm Al was used for either material in either scan condition. However, for the other filter materials, the least amount filtration corresponded with the greatest difference from the expected TMD value for both materials. For the test conditions without the beam flattener, there was a deviation from the expected TMD values even for the smallest diameter.

Part 2: Assessment of scan protocol parameters that contribute to accurate density measurements

Quantification of the cortical thickness, moment of inertia, outer perimeter indicated that there are some statistically significant differences when the specimens are scanned 4 at a time in comparison to when the bones were scanned individually (data not shown). However, there were no differences in the measured cross-sectional areas, the value for the mean thickness difference was less than the size of a voxel, and the thickness difference was within the range of error previously reported for a similar methodology [50]. Because of

![Fig. 4](image-url)
the mathematical relationship between thickness, moment of inertia, perimeter and cross-sectional area, we believe that these sub-voxel differences in morphology do not represent meaningful changes that occurred because of the scan protocol. When the trabecular bone morphology was examined, the results also indicated that there was a statistically significant decrease in the measured bone volume fraction with multiple bones in the path of the X-ray beam when the bones were scanned over 200° ($p=0.002$) and a trend toward a decrease when the bones were scanned over 360° ($p=0.080$) (Fig. 6E). However, the difference only represented a 2% change in the volume fraction. Therefore, this may not represent a physiologically meaningful change. Likewise, analysis of the trabecular number indicates that scanning 4 bones simultaneously over 360° results in fewer measured trabeculae than when the bones are scanned individually over 360° (Fig. 6F). For the densitometry measures, the results indicated that scanning the bones individually over 200° results in a statistically significant ($p<0.001$) and meaningful increase in the measured mineral content and mineral density of cortical bone (Figs. 6A and B). The results from the mineral content and density measurements of trabecular bone were slightly different. In these analyses, simultaneously scanning over 200° results in a significant underestimation of the tissue mineral content (TMC) ($p=0.042$), and a trend toward a decrease over 360° ($p=0.058$) (Fig. 6C). When these measurements are normalized to the volume of bone, both protocols that used a scan angle of 200° resulted in a slight overestimation of the TMD ($p=0.001$ for 4 bones simultaneously,
p = 0.011 when bones were individually scanned). The protocol where 4 bones were simultaneously scanned over 360° resulted in an underestimation of the TMD (p = 0.045, Fig. 6D).

**Part 3: Comparison of μCT-based density measurements to ash weights**

Results from the regression analyses indicated that BMCexclude predicted both the ash weight and dry weight with higher $R^2$ values, respectively, than BMCfull (Table 3). Calcium content did not correlate well with either BMCexclude or BMCfull. BMC values assessed by μCT were not significantly different than BMC measured by ash weight, as determined by paired t-tests between BMCash and BMCexclude and between BMCash and BMCfull. A linear regression revealed a near one-to-one relationship between the measurements (Fig. 7, slope $\pm$ 95% confidence intervals: BMCexclude vs. BMCash 0.990 ± 0.13; BMCfull vs. BMCash 1.05 ± 0.22). The difference between these slopes was not statistically significant, although the R value for BMCexclude was significantly higher, implying that this approach provided a better fit.

This strong correlation between the mineral content measurement methodologies, taken in conjunction with the beam hardening artifacts, implies that it is possible to visualize spatial patterns of mineralization. These patterns can be demonstrated using grayscale values or representative color maps (Supplementary Fig. S8). Visualization of mineralization patterns in vertebrae correlate with the ability to reduce beam hardening artifacts. This strong correlation between the mineral content measurement methodologies, taken in conjunction with the beam hardening artifacts, implies that it is possible to visualize spatial patterns of mineralization. These patterns can be demonstrated using grayscale values or representative color maps (Supplementary Fig. S8). Visualization of mineralization patterns in vertebrae correlate with the ability to reduce beam hardening artifacts. Furthermore, the ability to correct a skewed histogram [28], similar to what was seen in this study using filtration. Similarly to the filtration in this study, errors of up to approximately 45% and 60% occurred with inadequate correction and could be reduced to less than 5% with the correct density-specific linearization [29,31]. While the data in this study could be used for this type of linearization correction, we have chosen to use filtration and an acrylic beam attenuator to minimize the fundamental problem. These data show that the beam attenuator has the ability to reduce beam hardening artifacts. Furthermore, increasing the amount of filtration can minimize or reduce these artifacts. In general, the bone-like material SB3 was more prone to beam hardening artifacts than a similar material with a lower radiodensity (CB2-50%). The magnitude of the artifacts increased with thickness and corresponded to decreases in the measured BMD. Interestingly, the acrylic beam attenuator did not seem to have as much of an effect in SB3 as it did in CB2-50%. The acrylic is similar to water and provided a similar reduction in beam hardening artifacts than a similar material with a lower radiodensity (CB2-50%). The magnitude of the artifacts increased with thickness and corresponded to decreases in the measured BMD. Interestingly, the acrylic beam attenuator did not seem to have as much of an effect in SB3 as it did in CB2-50%. The acrylic is similar to water and provided the same path length effect that water bags provided in early CT scanners [35], so it was a reasonable first choice as a material. However, choosing a material that has a lower radiodensity match to SB3 may reduce the artifacts more.

The differences between the filter that used 0.254 mm Al and 0.254 mm Cu and the filter that used 0.101 mm Al were also evident. The filter that used Cu reduced cupping to an undetectable level for the thicknesses and materials tested in this study. This can be attributed to the difference in the X-ray spectrum after transmission through the filter with Cu (Supplementary Fig. S2). In fact, previous data have shown that the spectrum transmitted by a 0.10 mm Cu filter is very close to the spectrum transmitted by a 3.7 mm Al filter [52]. Filtration can be used to reduce the radiation dose for in vivo imaging [45,51,53], although the results of this study indicate a reduction in contrast and increase in noise with filtration. In fact, the decrease in contrast with the 0.254 mm Al and 0.254 mm Cu filter in this study required the use of an entirely separate color map to visualize this

**Discussion**

The overall goal of this study was to investigate interactions between beam hardening and mineral density measurements in μCT imaging. As a first goal, we attempted to determine when beam hardening-induced cupping artifacts occur in μCT imaging, how these artifacts impact BMD measurements, and determine the efficacy of beam filtration and beam flattening to reduce these artifacts. This theoretical example may not translate directly to research projects, so the possibility that artifacts may bias morphological and mineral density measurements was then investigated in a typical murine phenotyping study. Finally, μCT-based measurements of the BMC were compared to ash weights to verify the accuracy of μCT-based measurements.

The quantitative results of these studies are specific to the μCT system in this study, but the principles still apply to all scanners that use a polychromatic X-ray tube. In fact, it is possible to make meaningful μCT-based mineral density measurements using other scanners [27,28,29,31,32,51]. These systems utilized linearization procedures based on step wedge calibration or polynomial based approaches to correct the beam hardening artifacts, but even these corrections may be limited. The data in one of these studies indicates the scan setup can affect bone densitometry and trabecular morphology measurements. Comparisons of measurements on murine cortical bone for the (A) TMC of diaphyseal cortical bone, (B) TMD of diaphyseal cortical bone, (C) TMC of trabecular bone, (D) TMD of trabecular bone and (E) bone volume fraction of the trabecular bone. The results are presented as paired comparisons to the scanning condition where each bone was scanned individually over 360°. An asterisk indicates a statistically significant difference (p < 0.05 unless indicated).
data. Last, as a practical issue, using the filter composed of 0.254 mm Al and 0.254 mm Cu required a substantial increase in the integration time to obtain adequate photon statistics at the detector, resulting in scans that were approximately 3.5 to 3.75 times longer in comparison to no filtration. Therefore, many investigators use synchrotron based systems with monochromatic spectra. This has spurred interest in comparing synchrotron based systems with the more readily available laboratory systems [30,31,54,55].

While the X-ray spectrum is arguably the most important consideration to control when trying to minimize beam hardening artifacts, it is also possible that other aspects of the image acquisition and/or reconstruction processes can influence the measurement results. Charge integrating detectors are less prone to beam hardening artifacts than photon counting detectors [56]. In addition, increasing the number of views can increase the SNR [57]. These facts led us to the second portion of this study where we investigated limitations that arise when imaging multiple samples with a short scan. Four mouse femurs were simultaneously scanned over 200° for the high throughput approach, and scanning each specimen individually over 360° was used as the relative gold standard. Cortical bone densitometry results suggest that imaging four bones simultaneously does not bias these measurements. This is not surprising since the longest path length for two sections of mouse cortical bone is smaller than the first tier of our SB3 phantom where no significant cupping was detected. However, limiting the number of views resulted in a statistically significant and meaningful increase in measurements of TMC and TMD in cortical bone when the bones were scanned individually. This may indicate that the ‘front’ bone did still act as a low level filter when the bones were scanned 4 at a time. We have seen evidence that this type of effect occurs in another study where only two mice tibias were simultaneously scanned (Supplementary Fig. S9).

This possibility is also supported by the densitometry measurements for trabecular bone. The beam path length for trabecular bone will be longer because of an increased amount of tissue. This would increase the likelihood for Compton scattering that may occur for some of the higher photon energies in this X-ray beam. Cone-beam effects may also occur because the metaphysis is further from the center of the field of view. This could theoretically result in a lower density than expected for the second specimen in the beam path and, in fact, this occurred when four bones were simultaneously scanned over 360°. The trabecular bone data when the specimens were scanned individually shows the same overestimates for the mineral content and mineral density that were seen in cortical bone, further reinforcing the bias induced by limiting the number of views. Just as for cortical bone, there is an interaction between limiting the number of views and potential beam hardening when four specimens were scanned simultaneously. The TMC measurements for these data indicate that the subtle beam hardening effect may be prevalent, whereas TMD measurements indicate that limiting the number of views may dominate, so the underlying cause is still difficult to discern since the TMD change can be mathematically attributed to the decrease in BV/TV. These distinctions may be avoidable in the future by increasing the amount of beam filtration used to scan mice bones, but they may also be attributed to the inherent difficulties in assessing mouse trabecular bone structures. Limiting the number of views will reduce the SNR, and this may be particularly troublesome for trabecular bone due to resolution limitations that have been previously reported [50,58–60]. Only one of these studies is for mouse bones [50], and the voxel size used in this study was smaller than the 20 μm voxel size that correlated well to their gold standard image [50]. It is still possible, however, that those images have a better spatial resolution than reported in this study because those images were obtained by digital downsampling that may not represent the actual resolution if the bones were scanned at that voxel size [58].

There may still be other limitations in our assessment of the trabecular bone morphology and mineral density. First, we used a global thresholding approach that may limit the capability to obtain accurate quantifications. Partial volume artifacts around the edges of trabeculae can make thresholding difficult [61], so several studies have previously attempted to validate different thresholding or segmentation algorithms [61–67]. Most of these studies validate the thresholding algorithm by comparing the μCT images to a histological reference standard. While we could have performed a similar comparison in this study to use as a gold standard, our main goal in the second part of this study was to understand the limitations of different μCT protocols. Using the protocol that generates the best image quality obtainable was adequate for the comparisons of interest.

The end goal for these analyses is to have accurate measures of the mineral density. The only way to verify these measurements is to compare them with physical measurements which, for the case of mineral content data, are typically performed by ashing. Therefore, in the third study, we performed this type of analysis using mice of different ages both with and without an osteogenesis imperfecta-mimicking mutation to obtain a range of mineral densities. There was a high correlation between the ash weights and μCT-based densitometry measurements for these specimens, verifying that this setup of a μCT system can be used to obtain accurate mineral density measurements for murine vertebra. This data compares well to other studies indicating correlation coefficients greater than 0.9 after correcting for beam hardening, although those studies are for bovine or human samples [31,32,51]. While this data is only on murine vertebra, the correlation was performed on measurements of an entire vertebra so it was a statistical average for a large number of bone voxels containing both cortical and trabecular bone. Femora may be slightly different than vertebra, but the TMC and TMD data for the femora were also a statistical average, so we believe the data are similar. Furthermore, we cannot make mineral density estimates on a voxel-by-voxel level in our system because this type of measurement would be prone to image quality issues such as noise, partial volume artifacts, center of rotation artifacts and motion artifacts. However, it is possible to visualize patterns of mineralization that are based on the grayscale values.

There are also other limitations of this work. The studies are presented here in the logical order of understanding how beam hardening affects BMD measurements, scan setup on BMD measurements, and finally verifying the accuracy of BMC measurements. Despite this presentation, the chronological order of these studies was the exact opposite of this presentation. The scanner used in the third study irreparably failed and was replaced, precluding a side-by-side comparison of the systems. The X-ray source, system setup, and detector are identical between the two systems (with the exception of a field of view increase due to a change in CCD size), so we believe the data for these three studies are cohesive. The use of 0.508 mm Al filtration in the first study compares directly to the amount of filtration used in the second and third studies. The path lengths where artifacts began to occur with this amount filtration were longer than typical path lengths for mouse femora or vertebrae. Therefore, since beam hardening was avoided in the third study, this data still represents a valid accuracy test. Increasing the amount filtration increases noise, so the precision of densitometry measurements could plausibly decrease, but the accuracy should be consistent as long as beam attenuation is dominated by photoelectric absorption. It is also possible that adjusting the scan protocols to use more projections, such as done in the second study, could help account for variation that was not accounted for in study 3. The few percent variation in densitometry that was detected in the second study is consistent with unmeasured error in the third study.

In conclusion, filtering the X-ray beam can reduce beam hardening-induced cupping artifacts in bone–like materials at thicknesses relevant to typical μCT studies. Filtration does not necessarily require software based beam hardening corrections, but it does decrease contrast,
increase the baseline noise, and decrease throughput. These artifacts do not significantly impact estimates of the mineral density in cortical bone, but may be concerning for trabecular bone. Beam hardening induces less artifact for morphology than densitometric measurements. Comparing the ash weights of vertebra to μCT-based mineral density measurements showed a strong correlation, suggesting that this setup of a μCT system can be used to obtain densitometry data that is both precise and accurate.

Acknowledgments

The authors are indebted to Charles Roehm for his skillful machining of the SB3 and CB2-50% phantoms used to assess the magnitude of cupping artifacts. The authors would also like to thank Dr. Joan Marini and Dr. Tom Uveges for their assistance with the Brlt/+ mice, Avantika Varma for her assistance with the study, Kristi Ovgaard for her critical review of this manuscript, Jaroslav Tuszynski for providing the PhotonAttenuation2 script on MATLAB, and Dr. Paul Picot and Dr. Tim Morgan for their insight during the course of this study. We would also like to acknowledge the support of the National Science Foundation Graduate Research Fellowship program, NIH grants P30-AR46024 and T90-DK070071, and the University of Michigan Medical Student Biomedical Research Program for supporting this work.

Appendix A. Supplementary data


References


