Title

Multicenter precision of cortical and trabecular bone quality measures assessed by high-resolution peripheral quantitative computed tomography

Journal

Journal of Bone and Mineral Research, 28(3)

ISSN

0884-0431

Authors

Burghardt, Andrew J
Pialat, Jean-Baptiste
Kazakia, Galateia J
et al.

Publication Date

2013-03-01

DOI

10.1002/jbmr.1795

Peer reviewed

Published in final edited form as:
J Bone Miner Res. 2013 March ; 28(3): 524–536. doi:10.1002/jbmr.1795.

MULTI-CENTER PRECISION OF CORTICAL AND TRABECULAR BONE QUALITY MEASURES ASSESSED BY HR-PQCT

Andrew J. Burghardt[1], **Jean-Baptiste Pialat**[1,2], **Galateia J. Kazakia**[1], **Stephanie Boutroy**[2,3], **Klaus Engelke**[4,5], **Janina M. Patsch**[1,6], **Alexander Valentinitsch**[1,6], **Danmei Liu**[7], **Eva Szabo**[8], **Cesar E. Bogado**[9], **Maria Belen Zanchetta**[9], **Heather A. McKay**[7], **Elizabeth Shane**[3], **Steven K. Boyd**[8], **Mary L. Bouxsein**[10], **Roland Chapurlat**[2], **Sundeep Khosla**[11], and **Sharmila Majumdar**[1]

[1]Musculoskeletal Quantitative Imaging Research Group, Department of Radiology & Biomedical Imaging, University of California, San Francisco; San Francisco, CA USA

[2]INSERM Research Unit 1033 and Université de Lyon; Lyon, France

[3]Division of Endocrinology, Department of Medicine, College of Physicians and Surgeons, Columbia University Medical Center; New York, NY, USA

[4]Synarc Inc.; Hamburg, Germany

[5]Institute of Medical Physics, University of Erlangen; Erlangen, Germany

[6]Department of Radiology, Medical University of Vienna; Vienna, Austria

[7]Department of Orthopedics, Centre for Hip Health and Mobility, University of British Columbia; Vancouver, BC Canada

[8]Faculty of Medicine and Schulich School of Engineering, University of Calgary; Calgary, AB Canada

[9]Instituto de Investigaciones Metabólicas and Cátedra de Osteología y Metabolismo mineral, Universidad del Salvador; Buenos Aires, Argentina

[10]Endocrine Unit, Massachusetts General Hospital and Orthopedic Surgery, Harvard Medical School; Boston, MA USA

[11]Division of Endocrinology, Metabolism and Nutrition, Department of Internal Medicine, College of Medicine, Mayo Clinic; Rochester, MN USA

Abstract

High-resolution peripheral quantitative computed tomography (HR-pQCT) has recently been introduced as a clinical research tool for *in vivo* assessment of bone quality. The utility of this technique to address important skeletal health questions requires translation to standardized multi-center data pools. Our goal was to evaluate the feasibility of pooling data in multi-center HR-pQCT imaging trials.

Corresponding author: Andrew J. Burghardt, Musculoskeletal Quantitative Imaging Research Group, Department of Radiology & Biomedical Imaging, University of California, San Francisco, QB3 Building, Suite 203, 1700 4th St, San Francisco, CA 94158, Tel: +1 (415) 514-9658, Fax: +1 (415) 514-9656, andrew.burghardt@ucsf.edu.

Authors' Roles

Study design: AJB and SM. Study conduct: AJB. Data collection: AJB, JBP, SB, JMP, AV, DL, ESz, CEB, MLB. Data analysis: AJB. Data interpretation: AJB, JBP, GJK, SB, KE, SKB, HM, MLB, SM. Drafting manuscript: AJB. Revising manuscript content: AJB. Approving final version of manuscript: AJB, JBP, GJK, SB, KE, JMP, AV, DL, ESh, CEB, MBZ, HM, ESz, SKB, MLB, RC, SK, SM. AJB takes responsibility for the integrity of the data analysis.

Reproducibility imaging experiments were performed using structure and composition-realistic phantoms constructed from cadaveric radii. Single-center precision was determined by repeat scanning over short (<72hrs), intermediate (3–5mo), and long-term intervals (28mo). Multi-center precision was determined by imaging the phantoms at nine different HR-pQCT centers. Least significant change (LSC) and root mean squared coefficient of variation (RMSCV) for each interval and across centers was calculated for bone density, geometry, microstructure, and biomechanical parameters.

Single-center short-term RMSCVs were <1% for all parameters except Ct.Th (1.1%), Ct.Th.SD (2.6%), Tb.Sp.SD (1.8%), and porosity measures (6–8%). Intermediate-term RMSCVs were generally not statistically different from short-term values. Long-term variability was significantly greater for all density measures (0.7–2.0%; $p < 0.05$ *vs.* short-term) and several structure measures: Ct.Th (3.4%; $p < 0.01$ *vs.* short-term), Ct.Po (15.4%; $p < 0.01$ *vs.* short-term), and Tb.Th (2.2%; $p < 0.01$ *vs.* short-term). Multi-center RMSCVs were also significantly higher than short-term values: 2–4% for density and μFE measures ($p < 0.0001$), 2.6–5.3% for morphometric measures ($p < 0.001$), while Ct.Po was 16.2% ($p < 0.001$).

In the absence of subject motion, multi-center precision errors for HR-pQCT parameters were generally less than 5%. Phantom-based multi-center precision was comparable to previously reported *in vivo* single-center precision errors, although this was approximately 2–5 times worse than *ex vivo* short-term precision. The data generated from this study will contribute to the future design and validation of standardized procedures that are broadly translatable to multi-center study designs.

Keywords

HR-pQCT; osteoporosis; precision; bone; microstructure; bone strength; multi-center studies

Introduction

High-resolution peripheral quantitative computed tomography (HR-pQCT) is an emerging non-invasive method for *in vivo* 3D characterization of bone in the human peripheral skeleton (1,2). With a typical isotropic voxel size of 82μm, this device permits quantification of the geometric, microstructural, densitometric, and mechanical properties of human cortical and trabecular bone (3–5). Monitoring these bone quality features in a longitudinal fashion is of particular interest in clinical research, allowing patient-specific evaluation of disease progression, treatment response, and other manifestations of a change in bone metabolism (6,7).

To fully realize the potential of HR-pQCT to investigate the role of bone quality in fracture prediction and evaluate treatment efficacy, the technology must be translated to the multi-center setting. The first step towards establishing HR-pQCT as a feasible technique for large-scale multi-center clinical research studies is to characterize variability across different devices and imaging centers. A number of studies have reported the short-term, single-center reproducibility for measures of bone density, microstructure, and strength (1,2,8–11). One study has recently reported short-term *in vivo* precision data pooled from nine imaging centers participating in a multi-center trial (12). In addition, other studies have characterized various factors contributing to measurement variability including scan positioning (13,14), subject motion (15–18), and the image processing and analysis techniques applied (9,19,20). To date, no standard tools exist to assess long-term reproducibility of HR-pQCT bone quality measurements, or to investigate precision across multiple centers.

This study was designed to accomplish three specific goals: (I) to characterize the variation in noise and spatial resolution across multiple HR-pQCT systems; (II) to develop anthropomorphic microstructure-realistic imaging phantoms appropriate for long-term, multi-center quality control for HR-pQCT; and (III) to measure the precision in HR-pQCT derived measures of bone density, geometry, structure, and strength across imaging centers. To this end a phantom was designed to approximate peripheral skeletal anatomy, microstructure, and composition. Experiments to determine short-term, intermediate-term, and multi-center precision of HR-pQCT bone quality measures were performed by imaging these phantoms at nine different imaging centers.

Materials and Methods

Extremity Bone Structure (EBS) Phantom Design

The design goal established for our extremity bone structure (EBS) phantoms consisted of four critical criteria: that the phantoms (i) realistically model the microstructure, geometry, and composition of human bone from the peripheral skeleton, (ii) be structurally and materially stable over time for reliable long-term use, (iii) be easily portable for transfer between imaging centers, and (iii) facilitate standardized image acquisitions and analysis protocols. This goal was realized using cadaveric bone tissue embedded in a polymer resin with X-ray radio-opacity comparable to soft-tissue.

Specimens—Human cadaveric distal radius bone specimens (n=14) were acquired from the National Disease Research Interchange (NDRI) under a protocol that excluded donors with a known history of osteoporosis or other primary or metastatic diseases affecting bone. The tissue bank provided the bone specimens with a majority of the extra-osseal soft tissue dissected away, but with the bone marrow intact. The specimens were stored in a −20C freezer prior to the subsequent tissue processing steps. Descriptive information on the donors and the bone specimens, including bone density and structure measures determined using *ex vivo* micro-tomography (μCT-40, Scanco Medical AG, Brüttisellen, Switzerland) are provided in Table 1.

Tissue processing—Axial bone sections corresponding to the standard HR-pQCT imaging location for the distal radius were machined from the distal ends of the specimens. Each section was approximately 1-cm in length with the distal cut located 9-mm from the distal joint line. After sectioning, the bone marrow was removed using cycled treatments with a mild bio-detergent (Terg-A-Zyme, Alconox Inc.) in a sonicated water bath followed by a gentle wash by water jet.

Embedment—Each clean bone section was individually embedded in polymethyl methacrylate (PMMA) (Sigma-Aldrich, St. Louis, MO), and subsequently five sections were stacked for larger embedment in a polyethylene resin (Clear-Lite Casting Resin, Tap Plastics Inc., Dublin, CA) (Figure 1). Both plastics have effectively equivalent X-ray attenuation properties and approximate soft-tissue radio-opacity. The low viscosity of the monomer MMA allowed for robust infiltration of the marrow volume and complete displacement of air within the cancellous compartment under vacuum at 193 kPa (28 psi). Minimally sized container moldings (4-cm × 3-cm) were used for the individual embedments to minimize the risk of bubble formation during the exothermic polymerization process. The stacked sections were embedded in a 7-cm diameter by 20-cm long cylindrical molding (Proctor & Gamble, Cincinnati, OH). The cylindrical phantoms were machined to incorporate threaded inserts at either end for attaching mounting brackets (Scanco Medical AG, Brüttisellen, Switzerland) used for reproducible positioning and fixation within the scanner (Figure 1A). A total of three EBS phantoms, each with 5 bone sections, were constructed for use in this study. One

phantom included a single tibia section, which was not used for this analysis, therefore a total of 14 sections were evaluated.

Phantom Imaging Experiments

Imaging Centers—A consortium of nine academic imaging centers with HR-pQCT systems (XtremeCT, Scanco Medical AG, Brüttisellen, Switzerland) in North America, South America, and Europe was recruited to perform the phantom imaging experiments described below. A custom acquisition protocol was provided to each site by the organizing institution (UCSF), while the imaging experiments were performed by a local technician or researcher at each site. The phantom set was distributed to each participating site by a commercial freight service, while the image data was transferred via magnetic tape to the organizing institution for centralized processing. The arbitrary, de-identified label "site A" through "site I" will heretofore indicate the individual imaging center participating in this study.

Image Noise and Spatial Resolution—Two phantoms were imaged at each site to characterize the relative noise and spatial resolution performance of each HR-pQCT system. The noise performance was measured at each site using a single quality control (QC) phantom analogous to the phantom provided with each scanner by the manufacturer for density calibration and routine quality control. This phantom consists of a 7-cm diameter cylinder of a water-equivalent polymer resin with 1-cm diameter cylinders of hydroxyapatite mixtures (0, 100, 200, 400, 800 mg HA/cm^3) (Figure 2). The manufacturer's standard quality control protocol was adapted to image this phantom: a 73-slice stack of images (123μm isotropic voxel size) was acquired starting at a fixed offset from the end of the phantom (18mm) to ensure comparable positioning at each site.

Spatial resolution was measured using a custom-built wire phantom (QRM GmbH, Möhrendorf, Germany). The phantom consisted of an 8-cm diameter cylinder of soft-tissue equivalent polymer resin with five 25μm diameter tungsten wires embedded parallel to the long axis (Figure 3A). One wire was located at the center of the cylinder, while the remaining four wires were located at 90° intervals, 3-cm from the center. Tomographic images were acquired transverse to the long axis of the phantom, to provide an estimation of the in-plane point spread-function (PSF) (Figure 3B). The acquisition parameters for the resolution measurement were set to be equivalent to the standard *in vivo* imaging protocol; 750 projections images (1536×110 after 2×2 pixel binning of the detector) were acquired over 180° with a 100ms integration time. A total of 110 slices (9.02mm) were reconstructed across a 126mm field of view (FOV), resulting in 82μm isotropic voxels. To standardize the imaging location across HR-pQCT sites, a fixed offset (35.5mm) from the end of the phantom was used for positioning the first slice.

EBS Phantom Imaging—A standard acquisition protocol for imaging the EBS phantom was adapted from the manufacturer's standard *in vivo* protocol described in previous patient studies (1,2,8). A single projection radiograph was captured and subsequently used to prescribe scan positioning for the tomographic acquisition. The operator selected a single region that captured the five bone sections within the phantom (Figure 1B). The tomographic acquisition consisted of 750 projection images for each 180° rotation (110 slice stack spanning 9.02mm). Because the extent of the scan required coverage of approximately 75mm (915 slices), the entire acquisition required up to nine stacks. The X-ray tube voltage was set to 900μA at all sites except one site (site E) that was set to 1000μA. The integration time was set to 200ms to produce comparable noise performance in the phantom images to *in vivo* images of the forearm. The total imaging time per phantom was approximately 45

minutes. The projections were reconstructed across a 126mm FOV, discretized by a 1536×1536 image matrix, yielding 82μm isotropic voxels (Figure 1C).

Reproducibility Imaging—A series of imaging experiments were performed at nine imaging centers to assess multi-center reproducibility of the HR-pQCT technique. Additionally, short-, intermediate-, and long-term precision experiments were performed at a subset of centers. Due to practical considerations (scanner access, cost, and personnel time), repeat measurements were not acquired at all centers. Short-term reproducibility acquisitions were conducted at three imaging centers (sites A, D, and H). Three repeat acquisitions for each of the three EBS phantoms were performed within a maximum 72-hour window. Intermediate-term reproducibility was assessed for three centers following X-ray tube (sites A and H) and detector (site E) replacements. The interval between baseline and repeat measurements was between 3 and 5 months. Longer-term reproducibility was monitored for one center (site A) by imaging the EBS phantoms at 6-months, 21-months, and 28-months after the baseline measurements.

Image Analysis

Noise and Spatial Resolution Analysis—The signal-to-noise ratio (SNR) was calculated in the QC density phantom images (Figure 2). A concentric, cylindrical volume of interest (VOI) was placed in the 800 mg HA/cm^2 cylindrical region and the mean linear attenuation value was calculated for the signal measurement. A similar region of interest was placed in the 0 mg HA/cm^2 region (homogeneous polymer) and the standard deviation of the linear attenuation was calculated for the noise measurement. The ratio of these two measures was calculated to represent the system's SNR, with the mean and standard deviation reported for each imaging center based on three repeat acquisitions.

The spatial resolution was estimated from the PSF images of the wire phantom (Figure 3B). For each wire in the phantom image, a 32×32 neighborhood was identified around the approximate wire location in the central slice. A super-sampled PSF was generated by calculating the precise sub-voxel position of the wire in each slice according to the method described by Kwan *et al.* (21). A normalized modulation transfer function (MTF) was calculated from the Fourier transform of the PSF. The spatial resolution was defined as the spatial frequency corresponding to 10% of the MTF height and was reported in μm.

EBS Phantom Image Pre-processing—Semi-automated pre-processing scripts were developed in DIGITAL Command Language (DCL, Hewlett-Packard Company, Palo Alto, CA) and Image Processing Language (IPL v5.08b, Scanco Medical AG). For each 5-section phantom dataset, an operator identified the first complete slice of the first bone section. Because the length and relative position of each section is known and fixed, the remaining four sections could be automatically extracted and imported as unique measurements by the pre-processing software. In this way each phantom scan was deconstructed into its component sections for subsequent image analysis as individual pseudo-measurements.

EBS Phantom Image Analysis—Following the phantom image pre-processing, the individual bone sections were processed using an analysis protocol adapted from the manufacturer's protocol for standard *in vivo* acquisitions (1,2,22). Unlike the manufacturer protocol, which requires a trained operator to generate semi-automatic contours around the periosteal surface, a fully automatic contouring algorithm was used to identify the periosteal perimeter in the phantom images (11,23). This method was used to minimize operator bias in the phantom analysis. A simple 2-D registration procedure was utilized to constrain the volume of interest to a common region across all acquisitions in a set (22). The area enclosed by the automatically generated contours was calculated on a slice-by-slice basis

and cross-correlation was used to determine the slice offset for the maximal common VOI between datasets.

Standard densitometric and microstructural indices were automatically calculated for each co-registered VOI as described previously by Laib *et al.* (22). Trabecular bone volume fraction (BV/TV) was derived from the volumetric BMD of the trabecular compartment (Tb.BMD) using the assumption that compact bone has a matrix mineral density of 1200 mg HA/cm^3. Trabecular BMD was also calculated for two concentric sub-regions: a peripheral region adjacent to the cortex (pTb.BMD) and central medullary region (mTb.BMD) (24). The trabecular structure was extracted using a threshold-based binarization process (25). From the binary image, trabecular number (Tb.N) was measured using the direct 3D distance transform (DT) approach (26,27). Based on the densitometric BV/TV and direct Tb.N, trabecular thickness (Tb.Th) and trabecular separation (Tb.Sp) were derived using traditional plate model assumptions. In addition to the standard analysis, new methods were applied to measure cortical bone structure (11,28). A direct measure of cortical thickness (Ct.Th*) and the spatial variability in thickness (Ct.Th.SD) were calculated using the 3D DT approach (26). Similarly, the DT was used to measure the mean intra-cortical pore diameter (Po.Dm). Cortical porosity (Ct.Po) was measured by counting of intra-cortical void voxels in the cortical envelope (28,29). Finally, the structure model index (SMI) was calculated for the trabecular bone compartment (30).

μFE Analysis—Linear elastic micro-finite element analysis was performed for each co-registered VOI. For each model the binary image data set was converted to a mesh of isotropic hexahedral elements using a voxel conversion technique (31) and each element was assigned an elastic modulus of 6.829 GPa (4) and a Poisson's ratio of 0.3 (32). Cortical and trabecular bone were labeled as different materials, with identical material properties to facilitate calculation of compartmental load distribution. A uniaxial compression test in the axial direction (superior-inferior) was performed with an applied strain of 1%. An iterative solver (Scanco FE Software v1.12, Scanco Medical) was used to compute reaction forces at the superior and inferior ends of the sections for the prescribed boundary conditions. The model computations were performed at the UCSF/QB3 Shared Computing Facility – a mixed architecture linux cluster comprised of approximately 4500 processor cores. The apparent modulus and stiffness were calculated based on the reaction forces determined at the superior and inferior boundaries. Additionally, the cortical load fractions were determined at the proximal ($Ct.LF_{prox}$) and distal ($Ct.LF_{dist}$) boundaries of the VOI. Finally, the estimated failure load was calculated using the method of Pistoia *et al.* (33), based on criteria optimized for HR-pQCT: by back calculation of the reaction force at which 7.5% of the elements exceed a local effective strain of 0.7% (34).

Statistics

For all reproducibility experiments the standard deviation (SD) and coefficient of variation (CV) were calculated for each individual bone section and for each bone quality parameter. The precision was summarized for each parameter by calculating the least significant change (LSC) (35) and root mean square coefficient of variation (RMSCV expressed in %) (36). For short- and intermediate-term reproducibility studies performed at three different imaging centers, the mean LSC and RMSCV was calculated over the three sites, for each parameter. The long-term and multi-center reproducibility were reported as the LSC and RMSCV calculated over all time points and imaging centers, respectively. Analysis of variance (ANOVA) was performed using JMP version 8 (SAS Institute Inc., Cary, NC USA). A one-way ANOVA with Dunnett's *post hoc* test was used to compare the LSC for each parameter measured over intermediate, long-term, and across imaging centers to the LSC measured for the short-term experiments. A significance level of $p < 0.05$ was used to identify parameters

with significantly different intermediate- long-term, and multi-center reproducibility compared to short-term reproducibility.

For graphical visualization of the variability among imaging centers, the measured value of each parameter at each center was normalized to the mean across all centers for each bone section. The mean normalized value over all bone sections was calculated. Similarly, the progression of each parameter in the long-term reproducibility study was expressed as the mean of normalized values with respect to the individual baseline value for each bone section.

Results

Multi-center imaging performance

The noise performance for each site is presented in Figure 4A. The SNR ranged from 20.7 to 25.9 with a mean of 22.9±1.5. The spatial resolution (corresponding to the spatial frequency at 10%-MTF) at all positions is summarized in Table 2, with the resolution of the center wire presented in bar graph form in Figure 4B for comparison with the noise performance. The highest resolution was found at the location near isocenter (137.6±11.1 μm), while peripheral locations demonstrated systematically lower resolution performance. SNR was positively correlated with the 10% MTF (R^2=0.65, $p < 0.01$), indicating an inverse relationship between noise and resolution performance. Representative EBS phantom images for systems with image performance representing median noise/resolution, high noise/resolution, and low noise/resolution are illustrated in Figure 5.

Short-term single-center reproducibility

The short-, intermediate-, and long-term precision values of all bone quality measures determined for the EBS phantoms are reported in Table 3 and Figure 6A–D. The short-term reproducibility expressed as the average RMSCV ranged from 0.3% to 0.6% for bone density measures. The RMSCV for areal geometric measures were generally less than 1%, while Ct.Th estimated using an annular model or using the direct 3D approach was just over 1%. The heterogeneity of cortical thickness was somewhat less reproducible (2.6%). Among microstructure measures, Ct.Po and Po.Dm had the highest RMSCVs (7.9 and 5.9%, respectively), which corresponded to LSCs of 0.5% and 30μm, respectively.

Intermediate-term single-center reproducibility

In general the precision of HR-pQCT measures over the course of several months, including X-ray tube and detector repairs, tended to be slightly poorer compared to short-term precision. However, a statistically significantly decrease in precision was only observed for the peripheral trabecular BMD (pTb.BMD) and Tb.Ar (both $p < 0.01$ *vs.* short-term). While the intermediate-term LSC of these measures were approximately twice that of their short-term values, the RMSCV remained less than 1% for both. Systematic differences in the individual CVs (ANOVA with Tukey-Kramer *post hoc* test) between the two machines that had X-ray tube replacements and the machine that had a detector replacement were only observed for integral BMD (X-ray change: 0.4%; detector change: 1.0%; $p < 0.001$) and Tb.BMD (X-ray change: 0.6%; detector change: 1.4%; $p < 0.0001$).

Long-term single-center reproducibility

The evolution of bone quality measures in a single center over 28-months is illustrated in line-graph form in Figure 6E–H. While integral and Ct.BMD remained within 1% of baseline after 28-months, the lower numerical density parameters (Tb.BMD and its sub-compartment variants) were ±2–4% different at 28-months. The 28-month LSC of density measures ranged from 8.5 to 12.5 mg HA/cm^3, all significantly greater than the

corresponding short-term LSC ($p < 0.05$). Among structure measures, Ct.Po, Tb.Th, and SMI had significantly larger LSC over 28-months compared to their short-term LSC ($p < 0.01$). Cortical bone geometry decreased 2–6% after 28-months, with the long-term LSC of Ct.Th* significantly larger compared to short-term ($p < 0.01$). Biomechanical measures remained within ±1% of the baseline value on average, except $Ct.LF_{dist}$, which was 3% and 2% below baseline after 20- and 28-months.

Multi-center reproducibility

Compared to short-term single center precision (Figure 6A–D), bone quality parameters measured using the EBS phantom at nine different centers were generally less precise, with LSCs typically ranging from two-fold to eight-fold greater than their respective short-term single-center values. All densitometric parameters had significantly poorer multi-center precision (all $p < 0.0001$ *vs.* short-term single center LSC). The LSC of trabecular density (all regions) was less than 15 mg/cm^3, while integral and cortical BMD were somewhat greater (22 and 57 mg/cm^3, respectively). Among geometric and structure parameters, only Ct.Th.SD and Po.Dm did not have significantly worse precision across sites, compared to their respective short-term single center precision (22 μm and 24 μm, respectively). All biomechanical measures were found to have significantly lower multi-center precision ($p < 0.0001$, except $Ct.LF_{prox}$ $p < 0.001$ *vs.* short-term) while all RMSCVs were 2.5% or less, except $Ct.LF_{dist}$ (4.0%).

Multi-center variability

The variability of HR-pQCT-derived parameters among individual sites is illustrated in Figure 7, where the averaged normalized values (with respect to the mean for each bone section across all sites) are plotted by site. Densitometric measures spanned ±5% of the mean, while geometric measures were generally within ±3% of each respective mean. The exception was the standard measures of Ct.Ar and Ct.Th, where site E was a moderate outlier (6.1% and 6.4% greater than the mean, respectively). Among structure measures, Ct.Po showed the greatest relative variance, with site B +13.7% above and site G −23.9% below the mean. For the remaining structure measures all sites were generally within ±7% of the mean. As with the Ct.Po findings, the largest difference among sites for Tb.N was observed between sites B and G. The variability among sites of apparent biomechanical measures spanned ±2.5%, with site E being a positive outlier (+4.5%). Finally, cortical load fraction at the distal boundary was more variable (±4.2%) than at the proximal boundary (±2.0, $p < 0.05$ *vs.* $Ct.LF_{dist}$).

Discussion

Translation of HR-pQCT to long-term, multi-center clinical research studies of bone health requires the validation of standard techniques for the acquisition and analysis of images across time and across different devices. Our goal is to develop a phantom data-driven platform to optimize and validate standardized acquisition and post-processing procedures for reliable application of high-resolution CT in multi-center imaging studies. In this initial study, we have measured basic image performance properties of nine different first generation HR-pQCT devices and determined the long-term and inter-scanner precision of bone quality parameters – including bone density, geometry, microstructure, and strength.

The primary strength of this study is the use of a set of structure and material-realistic phantoms to measure precision. The use of these phantoms allowed the practical assessment of long-term and multi-center precision over 28-months and across nine different imaging systems. Previously, reproducibility experiments have been primarily conducted *in vivo* and limited to short- or intermediate-term scenarios in order to minimize confounding effects

due to biological change (1,2,8,9,11,12,17). Cadaveric specimens have also been used to assess short-term reproducibility (10), however these experiments, which used intact forearms, are not practical for long-term or multi-center use.

Long-term single-center precision is critical for the design of longitudinal studies. The long term LSC for density measures was approximately 10 mgHA/cm^3. To put that into context, the tolerance recommended by the manufacturer as part of the standard quality control (QC) procedures, as measured in a simple hydroxyapatite phantom, is ±8 mgHA/cm^3 (corresponding to 1% error for a cylindrical 800 mgHA/cm^3 insert). Over the course of the same 28-month period, the standard QC values were within this specified tolerance, except immediately prior to the failure of an X-ray source. The precision errors for non-densitometric parameters were not consistently poorer over the 28-month period compared to short-term precision errors. A notable exception was Ct.Po, where the long-term LSC was approximately double the short-term value. Changes in spatial resolution over time could possibly explain these findings, however we do not have concurrent MTF measurements to confirm this. Finally, it is worth acknowledging that long-term precision was only studied at a single center. A more rigorous investigation of precision over a greater interval and on more devices is an important priority of our ongoing effort.

Compared to single-center precision, essentially all bone quality parameters were significantly less reproducible across different scanners. LSCs ranged from approximately twice to five times as high as the short-term, single-center precision. While this indicates inter-scanner differences are an important source of error, the ideal multi-center precision errors (*i.e.* errors independent of subject motion) for non-densitometric parameters were comparable to short-term precision errors measured *in vivo* with minor to moderate motion (Scanco grade 1 and 2) (12,17). By contrast, inter-scanner precision errors for BMD parameters, which are less sensitive to subject motion, were approximately twice what has been reported *in vivo* (12). The comparability of microstructural parameters was particularly sensitive to the imaging performance of each system (see Figures 4 and 7C). Sites with relatively high SNR and low spatial resolution tended to have the lowest estimates of Tb.N, while sites with relatively low SNR and high spatial resolution had the highest estimates of Tb.N. Qualitatively, this is consistent with the variability in image quality observed for the different systems, as illustrated in Figure 5. Although the RMSCV for Ct.Po was relatively high compared to other parameters (RMSCV% > 15%), it is important to note that this corresponds to a least significant change of 1.7% (in units of relative pore volume). To put that into context, natural differences in porosity are quite large; in the distal radius Ct.Po has been found to be approximately 90% greater in post- *vs.* pre-menopausal women (28,37), 124% greater in a type II diabetic female population compared to healthy age-matched controls (38), and to increase by an estimated 300% from 20 to 90 years of age (39).

An important observation that becomes apparent when considering the variability among sites (see Figure 7) is that individual sites have varying degrees of deviation for different classes of parameters. For example, sites G and I were consistently outliers for densitometric parameters. This may indicate X-ray tube properties of these scanners have deviated to a greater extent from their baseline calibrations compared to other sites, warranting closer scrutiny and possibly correction using the concurrent density QC scans. In contrast, site E was a significant outlier for the standard cortical geometric measures, and to some extent the μFE parameters. Possibly this suggests a relative overestimation of the cortical bone volume during the thresholding process. That this site's BMD parameters are average indicates that the variability is not due to mis-calibration, but rather to variability in the attenuation values themselves due to the greater tube current used at this site (1000μA) and/or differences in the X-ray spectrum. Direct measurement of Ct.Th* (11), which has been shown to be more accurate (29) and less sensitive to bone tissue attenuation differences (19), normalized

cortical geometry for Site E. Finally, as discussed previously, the comparability of resolution-sensitive parameters such as Tb.N and Ct.Po are likely driven by spatial resolution and noise performance (sites B and G). Porosity measures derived from cortical BMD (40) may minimize error related to resolution variability, however it is not clear what other intrinsic sources of error would affect this approach. Collectively, these findings indicate that variability in specific intrinsic sources of error among devices is common, and that these properties affect precision in a parameter-specific fashion.

This study is also the first to report the impact of critical hardware replacement on the precision of bone quality parameters measured by HR-pQCT. It was encouraging to note that replacement of an X-ray source or detector did not significantly impact precision. This observation is important, as it is relatively common for an X-ray source to be replaced during the course of a typical two-year study.

An important caveat to our study is that the precision errors measured with the EBS phantom underestimate the absolute error expected for an *in vivo* measurement. Our data reflect the cumulative effect of intrinsic variability in the measurement (*e.g.* machine performance, geometric calibration) as well as the limitations and assumptions inherent to the reconstruction and post-processing procedures (*e.g.* the beam hardening-correction, segmentation algorithm, morphometry methods). However, *in vivo* measurements are additionally confounded by subject motion (12,17,18). This represents a significant and independent source of error that compounds the overall error in a clinical measurement. Variability introduced by subject motion is not inherently time or scanner dependent. Therefore, while the phantom-based design of this study has the advantage of isolating the intrinsic errors that are the most relevant to comparability across devices and time, this must be considered an estimate of the ideal precision under these circumstances.

Several limitations to this study should be acknowledged. First, the cadaveric specimens included in the EBS phantoms were from relatively old men and women (range 60–88 years). While this represents a clinically important population for age-related osteoporosis, the precision results may not translate to younger populations with different bone quality profiles. Furthermore, these phantoms only included distal radius specimens, and therefore are not suitable for estimating the precision error at the distal tibia. Collectively, the literature suggests that *in vivo* precision errors are possibly lower in younger populations compared to elderly, and higher at the forearm compared to the ankle (8,9,11,12). However, it is likely these trends are driven by the differences in the prevalence of subject motion rather than age or site-related differences in bone size or quality. Finally, it is not possible to rule out physical changes in the EBS phantoms over time. While nothing has led us to believe significant physical changes did occur, it was not possible to definitively evaluate this non-invasively.

The information from our multi-center precision results and the phantom image data collected will be critical for the development of targeted cross-calibration procedures. As discussed above, the precision results provide specific insight into important sources of error in bone quality measurements monitored over the long-term on a single scanner as well as pooled across multiple scanners. This knowledge can inform the direction of research into appropriate cross-calibration procedures and improved post-processing techniques that minimize these errors. Furthermore, the data itself could be exploited to derive parameter-specific numerical corrections to standard scales (41–43). Finally, the data can be used to optimize and validate *independently* developed cross-calibration and post-processing methods. This will likely be critical for the broader application of standardization procedures because the EBS phantom design (i.e. comprising unique cadaveric bone specimens) is not suitable for mass production. For example, an idealized phantom design, such as the

European Forearm Phantom (44,45), could potentially be used to routinely derive correction factors using procedures validated in EBS phantom data.

In summary, the results presented in this study are the first to describe the comparability of bone quality measures performed on different HR-pQCT scanners. Our principal finding is that intrinsic variability between scanners is comparable in magnitude to, and entirely independent of, error introduced by a mild amount of patient motion. This information alone could have immediate impact on the design of multi-center HR-pQCT studies currently being planned. Precision data derived from single-center reproducibility experiments likely leads to an underestimation of the number of subjects needed to achieve a given level of statistical power in long term and multi-center study designs. Finally, these image datasets and the knowledge garnered from the precision results of this study can form the fundamental basis for the development of cross-calibration and standardization procedures.

Acknowledgments

The authors would like to thank Bryan Hermannsson, Lucas Carvajal of UCSF, and Doug McVay of LabWorks Inc. for their contributions to constructing the phantoms used in this study. They also thank Dr. Andres Laib, formerly of Scanco Medical AG for technical consultations and feedback during the preparation of this manuscript. This publication was supported by the following grants: NIH R01 AR060700 (AJB), NIH R01 AG17762 (SM), NIH K01 AR056734 (GJK), NIH/NCRR S10 RR023405 (MLB), NIH P01 AG004875 and R01 AR027065 (SK), Synarc Inc. (SM), and Austrian Federal Bank 13468 (JMP), and France-Berkeley Fund (SM/RC).

Disclosures

Klaus Engelke is an employee of Synarc, Inc. Sharmila Majumdar received a research grant from Synarc, Inc. that partially supported the study presented in this paper.

References

1. Boutroy S, Bouxsein ML, Munoz F, Delmas PD. In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography. J Clin Endocrinol Metab. 2005; 90(12):6508–6515. [PubMed: 16189253]
2. Khosla S, Riggs BL, Atkinson EJ, Oberg AL, McDaniel LJ, Holets M, Peterson JM, Melton LJ 3rd. Effects of sex and age on bone microstructure at the ultradistal radius: a population-based noninvasive in vivo assessment. J Bone Miner Res. 2006; 21(1):124–131. [PubMed: 16355281]
3. Burghardt AJ, Kazakia GJ, Majumdar S. A local adaptive threshold strategy for high resolution peripheral quantitative computed tomography of trabecular bone. Ann Biomed Eng. 2007; 35(10): 1678–1686. [PubMed: 17602299]
4. Macneil JA, Boyd SK. Bone strength at the distal radius can be estimated from high-resolution peripheral quantitative computed tomography and the finite element method. Bone. 2008; 42(6): 1203–1213. [PubMed: 18358799]
5. MacNeil JA, Boyd SK. Accuracy of high-resolution peripheral quantitative computed tomography for measurement of bone quality. Med Eng Phys. 2007; 29(10):1096–1105. [PubMed: 17229586]
6. Seeman E, Delmas PD, Hanley DA, Sellmeyer D, Cheung AM, Shane E, Kearns A, Thomas T, Boyd SK, Boutroy S, Bogado C, Majumdar S, Fan M, Libanati C, Zanchetta J. Microarchitectural deterioration of cortical and trabecular bone: Differing effects of denosumab and alendronate. J Bone Miner Res. 2010; 25(8):1886–1894. [PubMed: 20222106]
7. Burghardt AJ, Kazakia GJ, Sode M, de Papp AE, Link TM, Majumdar S. A longitudinal HR-pQCT study of alendronate treatment in postmenopausal women with low bone density: Relations among density, cortical and trabecular microarchitecture, biomechanics, and bone turnover. J Bone Miner Res. 2010; 25(12):2558–2571. [PubMed: 20564242]
8. Kazakia GJ, Hyun B, Burghardt AJ, Krug R, Newitt DC, de Papp AE, Link TM, Majumdar S. In vivo determination of bone structure in postmenopausal women: a comparison of HR-pQCT and high-field MR imaging. J Bone Miner Res. 2008; 23(4):463–474. [PubMed: 18052756]

9. MacNeil JA, Boyd SK. Improved reproducibility of high-resolution peripheral quantitative computed tomography for measurement of bone quality. Med Eng Phys. 2008; 30(6):792–799. [PubMed: 18164643]
10. Mueller TL, Stauber M, Kohler T, Eckstein F, Muller R, van Lenthe GH. Non-invasive bone competence analysis by high-resolution pQCT: an in vitro reproducibility study on structural and mechanical properties at the human radius. Bone. 2009; 44(2):364–371. [PubMed: 19027092]
11. Burghardt AJ, Buie HR, Laib A, Majumdar S, Boyd SK. Reproducibility of direct quantitative measures of cortical bone microarchitecture of the distal radius and tibia by HR-pQCT. Bone. 2010; 47(3):519–528. [PubMed: 20561906]
12. Engelke K, Stampa B, Timm W, Dardzinski B, de Papp AE, Genant HK, Fuerst T. Short-term in vivo precision of BMD and parameters of trabecular architecture at the distal forearm and tibia. Osteoporos Int. 2012; 23(8):2151–2158. [PubMed: 22143491]
13. Boyd SK. Site-specific variation of bone micro-architecture in the distal radius and tibia. J Clin Densitom. 2008; 11(3):424–430. [PubMed: 18280194]
14. Mueller TL, van Lenthe GH, Stauber M, Gratzke C, Eckstein F, Muller R. Regional, age and gender differences in architectural measures of bone quality and their correlation to bone mechanical competence in the human radius of an elderly population. Bone. 2009; 45(5):882–891. [PubMed: 19615477]
15. Sode M, Burghardt AJ, Pialat JB, Link TM, Majumdar S. Quantitative characterization of subject motion in HR-pQCT images of the distal radius and tibia. Bone. 2011; 48(6):1291–1297. [PubMed: 21421091]
16. Pauchard Y, Ayres FJ, Boyd SK. Automated quantification of three-dimensional subject motion to monitor image quality in high-resolution peripheral quantitative computed tomography. Phys Med Biol. 2011; 56(20):6523–6543. [PubMed: 21937776]
17. Pialat JB, Burghardt AJ, Sode M, Link TM, Majumdar S. Visual grading of motion induced image degradation in high resolution peripheral computed tomography: impact of image quality on measures of bone density and micro-architecture. Bone. 2012; 50(1):111–118. [PubMed: 22019605]
18. Pauchard Y, Liphardt AM, Macdonald HM, Hanley DA, Boyd SK. Quality control for bone quality parameters affected by subject motion in high-resolution peripheral quantitative computed tomography. Bone. 2012; 50(6):1304–1310. [PubMed: 22445540]
19. Davis KA, Burghardt AJ, Link TM, Majumdar S. The effects of geometric and threshold definitions on cortical bone metrics assessed by in vivo high-resolution peripheral quantitative computed tomography. Calcif Tissue Int. 2007; 81(5):364–371. [PubMed: 17952361]
20. Valentinitsch A, Patsch JM, Deutschmann J, Schueller-Weidekamm C, Resch H, Kainberger F, Langs G. Automated threshold-independent cortex segmentation by 3D-texture analysis of HR-pQCT scans. Bone. 2012; 51(3):480–487. [PubMed: 22705149]
21. Kwan AL, Boone JM, Yang K, Huang SY. Evaluation of the spatial resolution characteristics of a cone-beam breast CT scanner. Med Phys. 2007; 34(1):275–281. [PubMed: 17278513]
22. Laib A, Hauselmann HJ, Ruegsegger P. In vivo high resolution 3D-QCT of the human forearm. Technol Health Care. 1998; 6(5–6):329–337. [PubMed: 10100936]
23. Buie HR, Campbell GM, Klinck RJ, MacNeil JA, Boyd SK. Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis. Bone. 2007; 41(4):505–515. [PubMed: 17693147]
24. Vico L, Zouch M, Amirouche A, Frere D, Laroche N, Koller B, Laib A, Thomas T, Alexandre C. High-resolution pQCT analysis at the distal radius and tibia discriminates patients with recent wrist and femoral neck fractures. J Bone Miner Res. 2008; 23(11):1741–1750. [PubMed: 18665795]
25. Laib A, Ruegsegger P. Comparison of structure extraction methods for in vivo trabecular bone measurements. Comput Med Imaging Graph. 1999; 23(2):69–74. [PubMed: 10227372]
26. Hildebrand T, Ruegsegger P. A new method for the model-independent assessment of thickness in three-dimensional images. J Microsc. 1997; 185:67–75.
27. Laib A, Hildebrand T, Hauselmann HJ, Ruegsegger P. Ridge number density: a new parameter for in vivo bone structure analysis. Bone. 1997; 21(6):541–546. [PubMed: 9430245]

28. Burghardt AJ, Kazakia GJ, Ramachandran S, Link TM, Majumdar S. Age- and gender-related differences in the geometric properties and biomechanical significance of intracortical porosity in the distal radius and tibia. J Bone Miner Res. 2010; 25(5):983–993. [PubMed: 19888900]
29. Nishiyama KK, Macdonald HM, Buie HR, Hanley DA, Boyd SK. Postmenopausal women with osteopenia have higher cortical porosity and thinner cortices at the distal radius and tibia than women with normal aBMD: an in vivo HR-pQCT study. J Bone Miner Res. 2010; 25(4):882–890. [PubMed: 19839766]
30. Hildebrand T, Ruegsegger P. Quantification of Bone Microarchitecture with the Structure Model Index. Comput Methods Biomech Biomed Engin. 1997; 1(1):15–23. [PubMed: 11264794]
31. Muller R, Ruegsegger P. Three-dimensional finite element modelling of non-invasively assessed trabecular bone structures. Med Eng Phys. 1995; 17(2):126–133. [PubMed: 7735642]
32. Van Rietbergen B, Odgaard A, Kabel J, Huiskes R. Direct mechanics assessment of elastic symmetries and properties of trabecular bone architecture. J Biomech. 1996; 29(12):1653–1657. [PubMed: 8945668]
33. Pistoia W, van Rietbergen B, Lochmuller EM, Lill CA, Eckstein F, Ruegsegger P. Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images. Bone. 2002; 30(6):842–848. [PubMed: 12052451]
34. Mueller TL, Christen D, Sandercott S, Boyd SK, van Rietbergen B, Eckstein F, Lochmuller EM, Muller R, van Lenthe GH. Computational finite element bone mechanics accurately predicts mechanical competence in the human radius of an elderly population. Bone. 2011; 48(6):1232–1238. [PubMed: 21376150]
35. Shepherd JA, Lu Y. A generalized least significant change for individuals measured on different DXA systems. J Clin Densitom. 2007; 10(3):249–258. [PubMed: 17616413]
36. Gluer CC, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporos Int. 1995; 5(4):262–270. [PubMed: 7492865]
37. Nicks KM, Amin S, Atkinson EJ, Riggs BL, Melton LJ 3rd, Khosla S. Relationship of age to bone microstructure independent of areal bone mineral density. J Bone Miner Res. 2012; 27(3):637–644. [PubMed: 22095490]
38. Burghardt AJ, Issever AS, Schwartz AV, Davis KA, Masharani U, Majumdar S, Link TM. High-resolution peripheral quantitative computed tomographic imaging of cortical and trabecular bone microarchitecture in patients with type 2 diabetes mellitus. J Clin Endocrinol Metab. 2010; 95(11):5045–5055. [PubMed: 20719835]
39. Macdonald HM, Nishiyama KK, Kang J, Hanley DA, Boyd SK. Age-related patterns of trabecular and cortical bone loss differ between sexes and skeletal sites: a population-based HR-pQCT study. J Bone Miner Res. 2011; 26(1):50–62. [PubMed: 20593413]
40. Zebaze RM, Ghasem-Zadeh A, Bohte A, Iuliano-Burns S, Mirams M, Price RI, Mackie EJ, Seeman E. Intracortical remodelling and porosity in the distal radius and post-mortem femurs of women: a cross-sectional study. Lancet. 2010; 375(9727):1729–1736. [PubMed: 20472174]
41. Hui SL, Gao S, Zhou XH, Johnston CC Jr, Lu Y, Gluer CC, Grampp S, Genant H. Universal standardization of bone density measurements: a method with optimal properties for calibration among several instruments. J Bone Miner Res. 1997; 12(9):1463–1470. [PubMed: 9286763]
42. Lu Y, Fuerst T, Hui S, Genant HK. Standardization of bone mineral density at femoral neck, trochanter and Ward's triangle. Osteoporos Int. 2001; 12(6):438–444. [PubMed: 11446558]
43. Shepherd JA, Cheng XG, Lu Y, Njeh C, Toschke J, Engelke K, Grigorian M, Genant HK. Universal standardization of forearm bone densitometry. J Bone Miner Res. 2002; 17(4):734–745. [PubMed: 11918231]
44. Ruegsegger P, Kalender W. A phantom for standardization and quality control in peripheral bone measurements by PQCT and DXA. Physics in medicine and biology. 1993; 38:1963.
45. Pearson J, Ruegsegger P, Dequeker J, Henley M, Bright J, Reeve J, Kalender W, Felsenberg D, Laval-Jeantet AM, Adams JE, et al. European semi-anthropomorphic phantom for the cross-calibration of peripheral bone densitometers: assessment of precision accuracy and stability. Bone Miner. 1994; 27(2):109–120. [PubMed: 7711519]

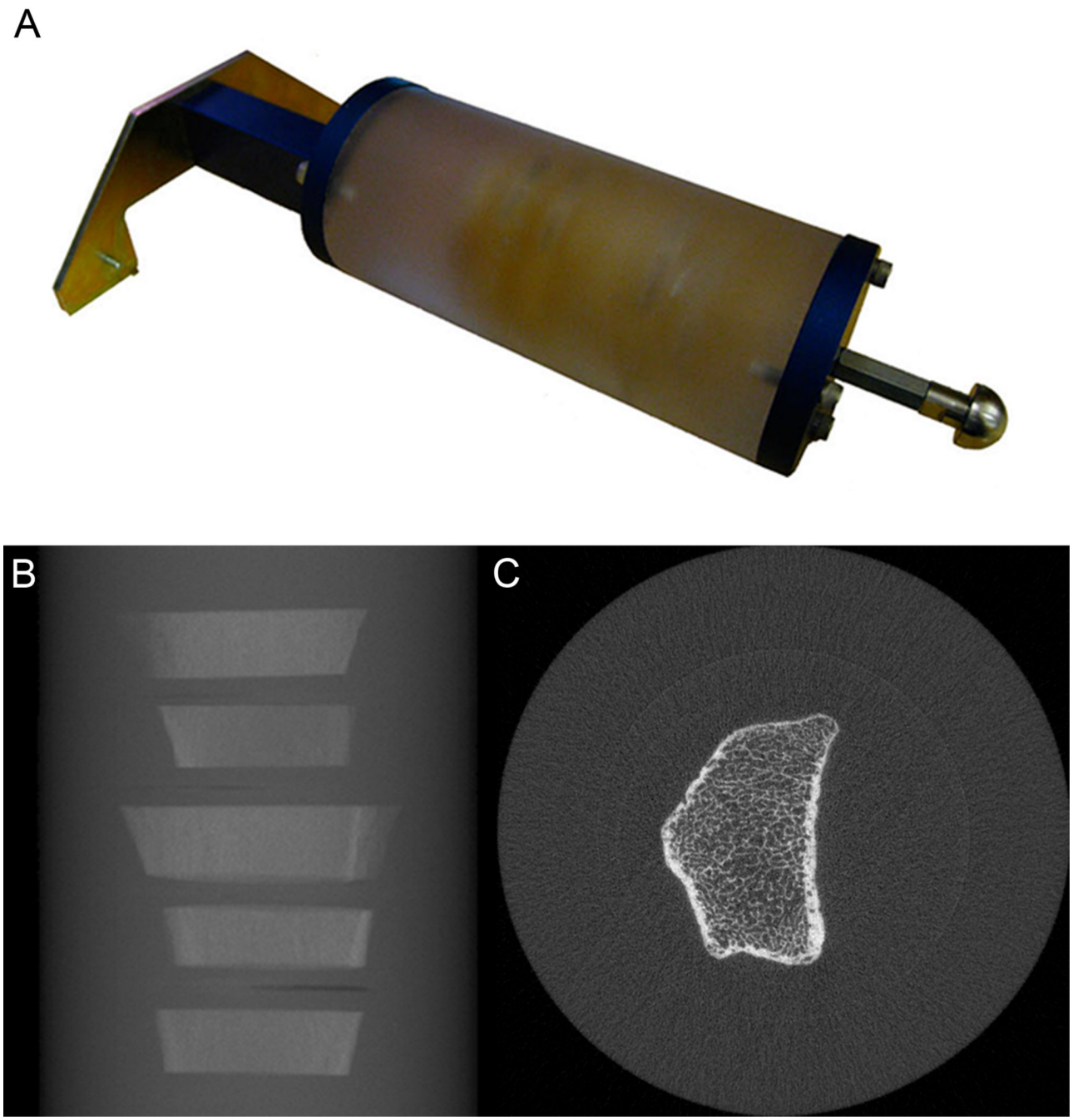
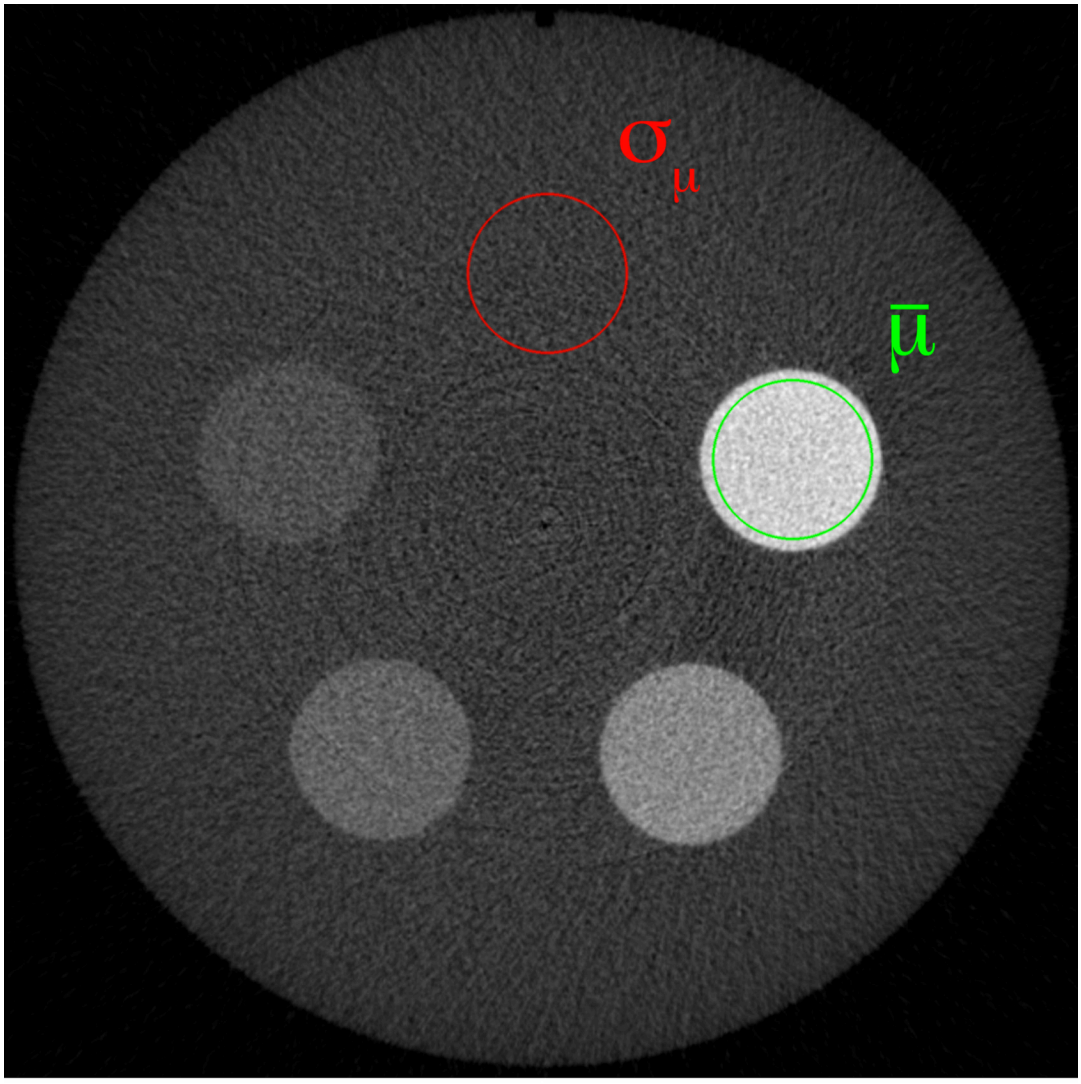


Figure 1.
The UCSF Extremity Bone Structure (EBS) phantom is shown in a photograph (A). Each phantom consists of a 7-cm diameter cylinder of soft-tissue equivalent polymer resin. Each cylinder contains five 1-cm thick cadaveric bone embedments (B), sectioned from the distal radius at a location corresponding to the standard location for an *in vivo* HR-pQCT exam (9.0-mm proximal to mid-point of the distal endplate). The phantom uses custom brackets for standardized mounting on the scanner gantry. A representative image for one bone section is shown on the right (C).



$$SNR = \frac{\bar{\mu}\ [800\ mg\ HA/cm^3]}{\sigma_\mu\ [0\ mg\ HA/cm^3]}$$

Figure 2.
The noise performance of each imaging system was assessed in a single phantom designed for routine density calibration and daily quality control. The signal to noise ratio (SNR) was calculated as the ratio of the mean attenuation of the high-density cylinder (μ, green region) to the standard deviation of the attenuation in the tissue-equivalent background (σ_μ, red region).

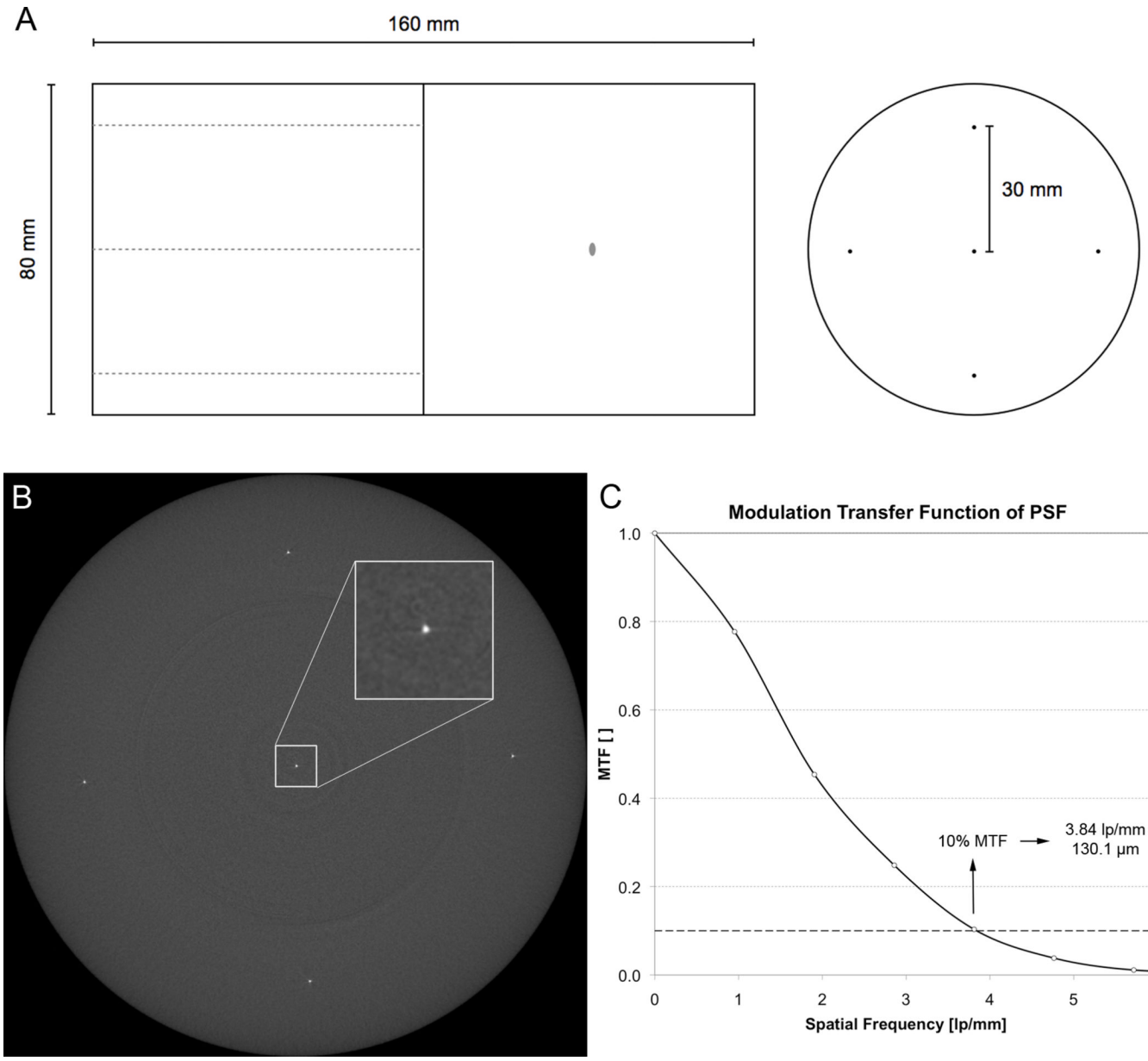


Figure 3.

The resolution performance of each imaging system was measured by estimating the point spread function (PSF) from images of a custom wire phantom (A). The phantom consisted of five 25-µm tungsten wires oriented parallel to the long axis of an 8-cm cylinder of tissue-equivalent polymer resin. One wire was located at the center, while the other four were located 3-cm from the center at 90° intervals. From the reconstructed image of the wire cross-sections (B), the PSF is estimated and then the resolution is calculated from the spatial frequency corresponding to 10% of the MTF (C).

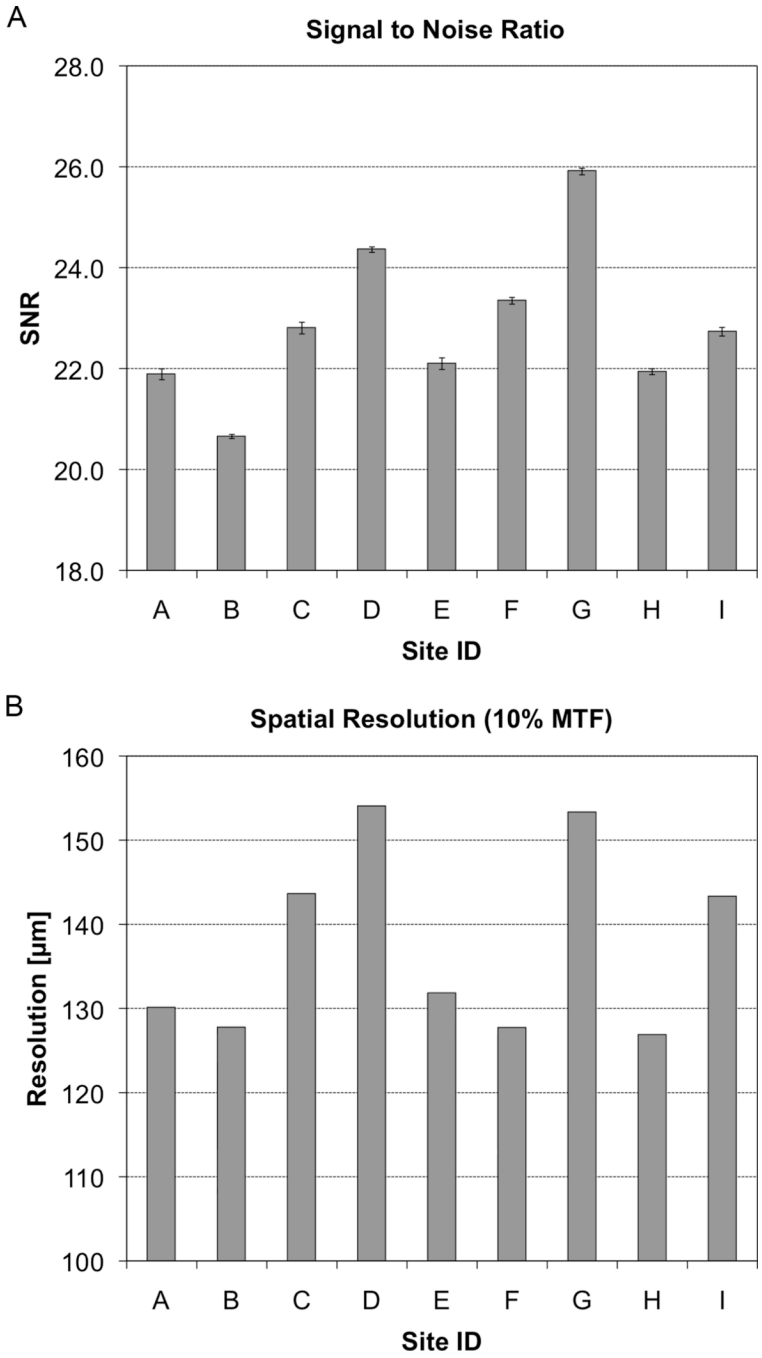


Figure 4.

The noise and spatial resolution of the individual HR-pQCT systems from the nine imaging centers participating in this study is presented in bar graph form. Above (A) the noise performance is characterized by a simple SNR measurement in the standard QC density phantom with error bars indicating the standard deviation of SNR for three repeat scans. Below (B), the spatial resolution is estimated to be the spatial frequency corresponding to the 10%-height of the modulation transfer function (MTF).

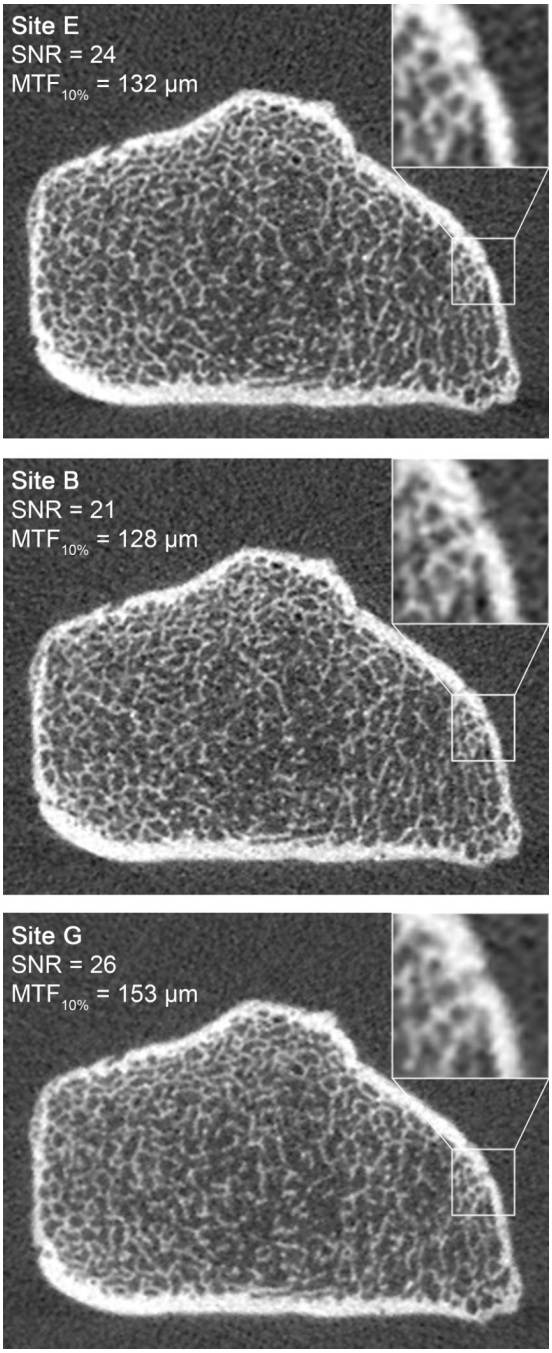


Figure 5.
HR-pQCT images of the UCSF EBS Phantom from three different imaging centers that represent the median SNR and resolution performance (site E, top), a high-resolution, high noise system (site B, middle), and a low resolution, low noise system (site G, bottom). For each image the mean SNR and spatial resolution (10% MTF) at the center of the FOV are indicated.

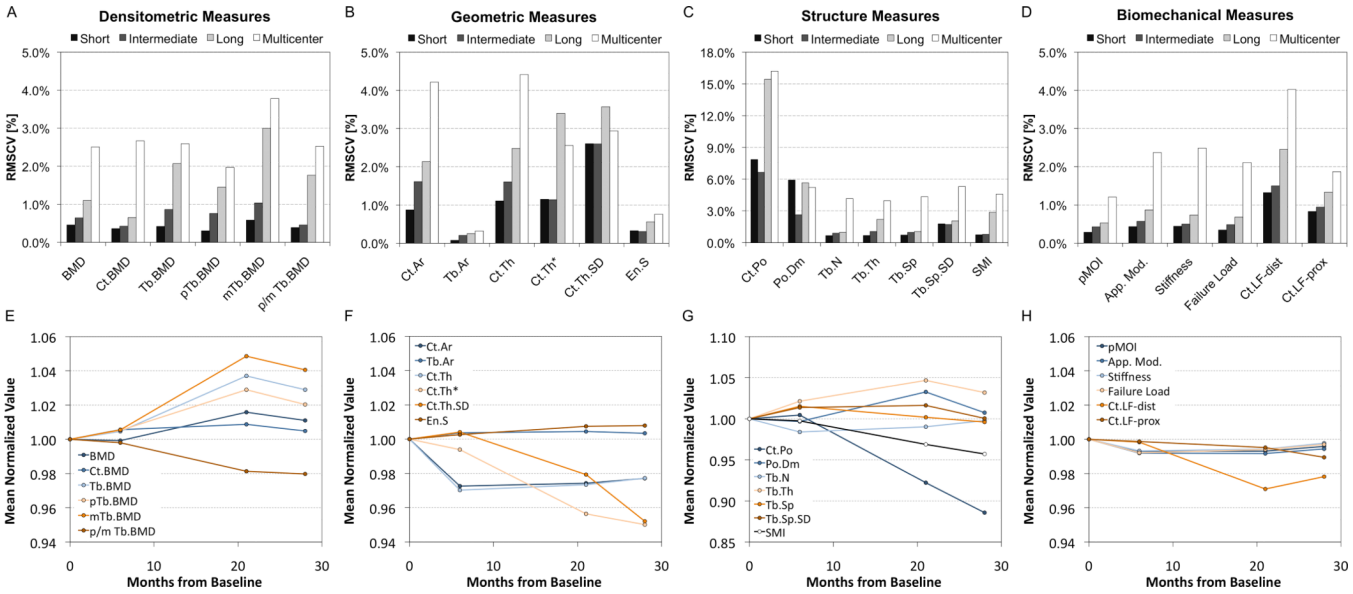


Figure 6.

On top (A–D) are summary bar graphs of the reproducibility of select HR-pQCT measures expressed as RMSCV [%] over short- (<72 hours), intermediate- (3–5 months), and long-term (28 months) for single centers (black, dark grey, light grey, respectively) compared to the multi-center reproducibility of nine different systems (white). On the bottom (E–H) are line graphs illustrating the stability of select HR-pQCT measures monitored at a single site over the course of 28 months. All values were normalized to the baseline for visualization purposes.

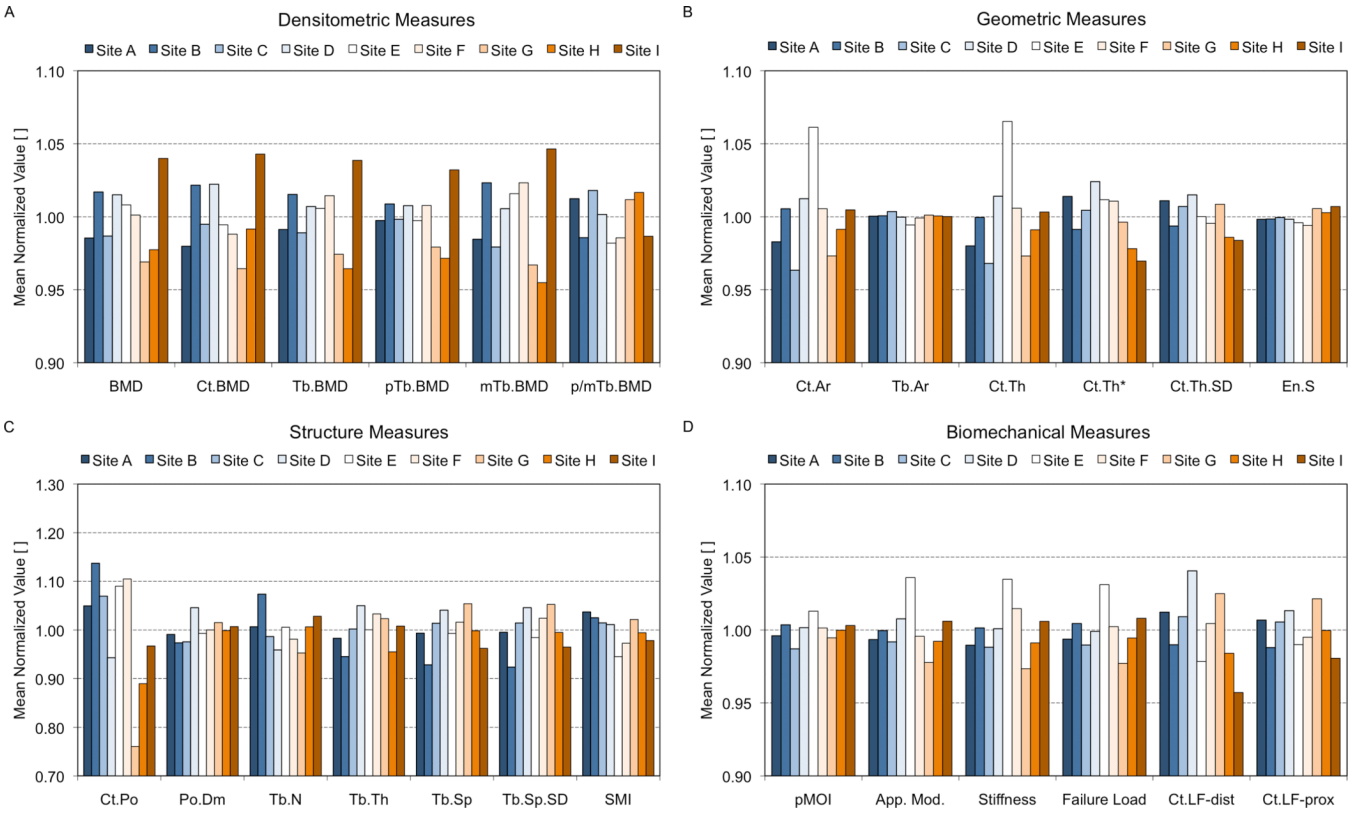


Figure 7.

The comparability of individual imaging systems is illustrated in this figure for the standard density (A), geometric (B), microstructural (C), and biomechanical (D) parameters. For each parameter the value measured for each 1-cm bone section was normalized to the mean value of that section across all nine systems. These graphs represent the mean normalized value for each parameter at a given center.

Table 1

Descriptive statistics for donor radius specimens

	units	avg ± sd
n	[ ]	14
n men/women	[ ]	7/7
Age	yrs	76.4 ± 8.6
BMD	mg/cm^3	327 ± 72
Ct.BMD	mg/cm^3	769 ± 130
Tb.BMD	mg/cm^3	214 ± 54
Ct.Th*	[mm]	0.64 ± 0.30
Ct.Po	[%]	3.9 ± 2.6
Po.Dm	[μm]	190 ± 31
Tb.N	mm^{-1}	1.63 ± 0.41
Tb.Th	μm	111 ± 14
Tb.Sp	μm	561 ± 253
Tb.Sp.SD	μm	286 ± 194

Table 2

Summary of spatial resolution measurements (10% MTF in μm) at center and 3-cm offset at different radial positions, 90° apart

Site	Center	0°	90°	180°	270°
A	130.1	146.3	160.6	142.0	167.9
B	127.8	132.5	172.9	146.0	182.2
C	143.6	139.1	164.1	135.6	168.6
D	154.0	134.3	169.8	152.2	187.4
E	131.8	155.0	154.9	149.0	170.6
F	127.7	135.9	167.8	134.1	164.6
G	153.3	147.8	166.5	131.2	180.1
H	126.9	151.5	164.0	171.3	168.4
I	143.3	144.9	147.9	135.2	154.4
avg ± sd	**137.6 ± 11.1**	**143.0 ± 7.9**	**163.2 ± 7.8**	**144.1 ± 12.5**	**171.6 ± 10.1**

Table 3

Precision results for short-term, intermediate-term, long-term, and multi-center phantom studies expressed as the least significant change (LSC) with the root mean square coefficient of variation (RMSCV in %) listed parenthetically. For each experiment, the number of HR-pQCT centers, the number of repeat scans at each site, and the number of degrees of freedom are also provided.

Parameter	Units	Short (< 72 hrs)	Intermediate (3–5 months)	Long (28 months)	Multi-center (9 sites)
centers/repeats		**3/3**	**3/2**	**1/4**	**9/1**
degrees of freedom		**84**	**42**	**42**	**112**
Density					
BMD	[mg/cm^3]	3.8 (0.5%)	6.0 (0.6%)	**8.5 (1.1%)**^{b}	**22.3 (2.5%)**^{d}
Ct.BMD	[mg/cm^3]	7.2 (0.4%)	8.9 (0.4%)	**12.5 (0.7%)**^{a}	**57.0 (2.7%)**^{d}
Tb.BMD	[mg/cm^3]	2.3 (0.4%)	5.0 (0.9%)	**10.0 (2.1%)**^{d}	**14.0 (2.6%)**^{d}
pTb.BMD	[mg/cm^3]	2.0 (0.3%)	**5.3 (0.8%)**^{b}	**9.5 (1.5%)**^{d}	**13.6 (2.0%)**^{d}
mTb.BMD	[mg/cm^3]	2.6 (0.6%)	5.0 (1.0%)	**10.4 (3.0%)**^{c}	**14.9 (3.8%)**^{d}
p/mTb.BMD	[ ]	0.02 (0.4%)	0.02 (0.46%)	0.11 (1.8%)	**0.16 (2.5%)**^{b}
Geometry					
Ct.Ar	[mm^2]	0.8 (0.9%)	1.9 (1.6%)	1.9 (2.1%)	**3.4 (4.2%)**^{c}
Tb.Ar	[mm^2]	0.6 (0.1%)	**1.5 (0.2%)**^{b}	**1.6 (0.2%)**^{d}	**2.4 (0.3%)**^{d}
Ct.Th	[μm]	12 (1.1%)	24 (1.6%)	26 (2.5%)	**41 (4.4%)**^{c}
Ct.Th*	[μm]	22 (1.1%)	23 (1.1%)	**75 (3.4%)**^{b}	**56 (2.6%)**^{d}
Ct.Th.SD	[μm]	21 (2.6%)	21 (2.6%)	27 (3.6%)	22 (2.9%)
En.S	[mm^2]	6.2 (0.3%)	5.3 (0.3%)	10.6 (0.6%)	**14.1 (0.8%)**^{c}
Microstructure					
Ct.Po	[%]	0.5 (7.9%)	0.7 (6.6%)	**1.3 (15.4%)**^{b}	**1.7 (16.2%)**^{c}
Po.Dm	[μm]	30 (5.9%)	14 (2.6%)	26 (5.6%)	24 (5.2%)
Tb.N	[mm^{-1}]	0.02 (0.7%)	0.04 (0.9%)	0.04 (1.0%)	**0.16 (4.2%)**^{d}
Tb.Th	[μm]	2 (0.7%)	3 (1.1%)	**7 (2.2%)**^{c}	**12 (4.0%)**^{d}
Tb.Sp	[μm]	20 (0.7%)	18 (1.0%)	22 (1.1%)	**102 (4.3%)**^{d}

Parameter	Units	Short (< 72 hrs)	Intermediate (3–5 months)	Long (28 months)	Multi-center (9 sites)
centers/repeats		**3/3**	**3/2**	**1/4**	**9/1**
degrees of freedom		**84**	**42**	**42**	**112**
Tb.Sp.SD	[μm]	30 (1.8%)	21 (1.7%)	30 (2.0%)	**71 (5.3%)**^{d}
SMI	[ ]	0.03 (0.7%)	0.03 (0.8%)	**0.09 (2.9%)**^{c}	**0.05 (4.6%)**^{d}
Biomechanics					
pMOI	[mm^4]	93.8 (0.3%)	162.9 (0.4%)	190.1 (0.5%)	**392.0 (1.2%)**^{d}
Apparent Modulus	[MPa]	16.5 (0.4%)	22.0 (0.6%)	33.5 (0.9%)	**80.9 (2.4%)**^{d}
Stiffness	[kN/mm]	1.0 (0.4%)	1.1 (0.5%)	1.5 (0.7%)	**5.3 (2.5%)**^{d}
Failure Load	[N]	35.3 (0.3%)	55.4 (0.5%)	66.9 (0.7%)	**199.7 (2.1%)**^{d}
$Ct.LF_{dist}$	[%]	1.4 (1.3%)	1.6 (1.5%)	2.2 (2.5%)	**3.8 (4.0%)**^{d}
$Ct.LF_{prox}$	[%]	1.4 (0.8%)	1.6 (0.9%)	2.2 (1.3%)	**3.2 (1.9%)**^{c}

*
Bold indicates statistically significant difference compared to short-term CVs:

$^{a}p < 0.05$,

$^{b}p < 0.01$,

$^{c}p < 0.001$,

$^{d}p < 0.0001$