Title

Operator variability in scan positioning is a major component of HR-pQCT precision error and is reduced by standardized training.

Journal

Osteoporosis International, 28(1)

Authors

Bonaretti, S
Vilayphiou, N
Chan, C
et al.

Publication Date

2017

DOI

10.1007/s00198-016-3705-5

Peer reviewed

HHS Public Access

Author manuscript

Osteoporos Int. Author manuscript; available in PMC 2018 January 01.

Published in final edited form as:

Osteoporos Int. 2017 January ; 28(1): 245–257. doi:10.1007/s00198-016-3705-5.

Operator Variability in Scan Positioning is a Major Component of HR-pQCT Precision Error and is Reduced by Standardized Training

Serena Bonaretti[1,a,*], **Nicolas Vilayphiou**[2], **Caroline Mai Chan**[3], **Andrew Yu**[1], **Kyle Nishiyama**[4], **Danmei Liu**[5], **Stephanie Boutroy**[6], **Ali Ghasem-Zadeh**[7], **Steven K. Boyd**[8], **Roland Chapurlat**[6], **Heather McKay**[5], **Elizabeth Shane**[4], **Mary L. Bouxsein**[9], **Dennis M. Black**[10], **Sharmila Majumdar**[1], **Eric S. Orwoll**[11], **Thomas F. Lang**[1], **Sundeep Khosla**[12], and **Andrew J. Burghardt**[1]

[1]Department of Radiology & Biomedical Imaging, University of California, San Francisco, CA, USA [2]Scanco Medical AG, Brüttisellen, Switzerland [3]University of California Berkeley, CA, USA [4]Division of Endocrinology, Department of Medicine, Columbia University Medical Center, New York, NY, USA [5]University of British Columbia, Vancouver, BC, Canada [6]INSERM UMR 1033, Université de Lyon, France [7]Department of Medicine, Austin Health, University of Melbourne, Melbourne, Australia [8]Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada [9]Center for Advanced Orthopaedic Studies, Beth Israel Deaconess Medical Center, Boston, MA, USA [10]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA [11]Division of Endocrinology, Bone and Mineral Unit, Oregon Health & Science University, Portland, OR, USA [12]Division of Endocrinology, Metabolism and Nutrition, Department of Internal Medicine, College of Medicine, Mayo Clinic, Rochester, MN, USA

Abstract

Introduction—HR-pQCT is increasingly used to assess bone quality, fracture risk and anti-fracture interventions. The contribution of the operator has not been adequately accounted in measurement precision. Operators acquire a 2D projection (“scout view image”) and define the region to be scanned by positioning a “reference line” on a standard anatomical landmark. In this study, we (i) evaluated the contribution of positioning variability to *in vivo* measurement precision, (ii) measured intra- and inter-operator positioning variability, and (iii) tested if custom training software led to superior reproducibility in new operators compared to experienced operators.

Methods—To evaluate the operator *in vivo* measurement precision we compared precision errors calculated in 64 co-registered and non-co-registered scan-rescan images. To quantify operator

Corresponding author (present address): Serena Bonaretti, Lucas Center for Imaging Stanford University 1201 Welch Road Stanford, CA 94305 USA Phone: +1-650-724-0361.

[a]Department of Radiology, Stanford University, Stanford, CA, USA (current affiliation)

Conflict of interest

Nicolas Vilayphiou is an employee of Scanco Medical AG. Serena Bonaretti, Caroline Mai Chan, Andrew Yu, Kyle Nishiyama, Danmei Liu, Stephanie Boutroy, Ali Ghasem-Zadeh, Steven K. Boyd, Roland Chapurlat, Heather McKay, Elizabeth Shane, Mary Bouxsein, Dennis M. Black, Sharmila Majumdar, Eric S. Orwoll, Thomas Lang, Sundeep Khosla, and Andrew J. Burghardt declare that they have no conflict of interest.

variability, we developed software that simulates the positioning process of the scanner's software. Eight experienced operators positioned reference lines on scout view images designed to test intra- and inter-operator reproducibility. Finally, we developed modules for training and evaluation of reference line positioning. We enrolled 6 new operators to participate in a common training, followed by the same reproducibility experiments performed by the experienced group.

Results—*In vivo* precision errors were up to three-fold greater (Tt.BMD and Ct.Th) when variability in scan positioning was included. Inter-operator precision errors were significantly greater than short-term intra-operator precision ($p<0.001$). New trained operators achieved comparable intra-operator reproducibility to experienced operators, and lower inter-operator reproducibility ($p<0.001$). Precision errors were significantly greater for the radius than for the tibia.

Conclusion—Operator reference line positioning contributes significantly to *in vivo* measurement precision and is significantly greater for multi-operator datasets. Inter-operator variability can be significantly reduced using a systematic training platform, now available online (http://webapps.radiology.ucsf.edu/refline/).

Keywords

HR-pQCT; operator reproducibility; precision; standardization; multicenter studies

Introduction

High-resolution peripheral quantitative computed tomography (HR-pQCT) is increasingly used [1] to assess bone quality [2,3], to investigate age-, sex-, and race-related differences [4,5], to evaluate bone diseases and monitor drug therapies [6–9], and to assess fracture risk [10,11]. HR-pQCT allows the acquisition of *in vivo* high-resolution images (isotropic voxel size 82μm, stack length = 9.02mm) of the distal radius and distal tibia with low radiation dose (3μSv). From the images, densitometric and structural parameters can be measured to characterize bone as a whole, and separately as cortical and trabecular compartments. In addition, the images can be used to perform finite element analysis to assess bone strength.

High measurement precision is fundamental to reliably calculate bone parameters [12]. In previous studies reported in the literature, precision errors for HR-pQCT were calculated from scan-rescan images with subject repositioning between the repetitions. Scan-rescan images are usually acquired longitudinally in short-, intermediate- and long-term repetitions. HR-pQCT precision errors are smaller *ex vivo* (<1.5%) than *in vivo* [13,14], for the tibia (<5.2%) than for radius (<6.3%) [3,15], and for densitometric (<1.5%) than for structural (<4.5%) measurements [2,4]. In general, precision errors are calculated as coefficient of variation of bone measurements extracted from co-registered scan-rescan images [16]. However, co-registration limits the measurement of bone parameters to a common volume between the scan-rescan images. Therefore this calculation of measurement precision includes the effect of factors such as inherent scanner precision, limb positioning in the scanner, and image quality degradation due to motion artifacts, but notably excludes the variability introduced by the operator when defining the anatomic region to be scanned.

To define the anatomic region to be scanned, the operator acquires a 2-D anterior-posterior projection of the proximal limb, commonly named “scout view image”. On the scout view image, the operator visually identifies an anatomic landmark and manually intersects it with a horizontal reference line (Figure 1). From the reference line, the anatomic volume to be scanned is offset by a standard distance. In the forearm, the anatomic landmark is traditionally an inflection in the curvature of the articular surface of the radius between the scaphoid and lunate fossae of the radiocarpal joint, whereas in the lower leg, the anatomic landmark is the apex of the distal articular plateau of the tibia at the tibiotalar joint. The anatomic landmark is often vague or absent in the forearm, whereas it is generally clearly visible in the lower leg. Identification of the anatomic landmark and positioning of the reference line may vary within and between operators, affecting bone parameter measurements especially in the radius, where morphology changes remarkably along the bone axis [17,18]. The variability in landmark identification is of paramount importance for cross-sectional or observational multicenter studies and when pooling cross-sectional datasets for retrospective analyses where multiple operators were involved in data collection. Furthermore, the reliability of using HR-pQCT to characterize an individual patient’s skeletal status in a clinical context hinges on minimizing operator variability in the measurement procedure and high quality reference data.

The goal of this study was to isolate and evaluate the contribution of operator reference line positioning to HR-pQCT measurement precision. Specifically, our first aim was to measure total precision error, including variability in scan positioning by the operator. We hypothesized that total precision error is significantly greater than what has been reported previously because the error due to positioning variability is removed by constraining the analysis to co-registered volumes of interest between repeat scans, which is the default behavior of the analysis output for follow up measurements. The second aim was to explicitly quantify the contribution of scan positioning variability to the measurement precision error within and between operators. To this end we developed software that simulates the acquisition interface of the HR-pQCT system and performed scan positioning reproducibility experiments. We hypothesized that operator variability in scan positioning is a major source of error for single and multicenter studies. Finally, the third aim was to evaluate the precision of new HR-pQCT operators who participated in a standard training curriculum. We developed training software for operators to perform scan positioning exercises on a collection of scout view images. We hypothesized that our training platform would lead to superior inter-operator precision errors.

Materials and methods

Study design

Our investigation of operator positioning variability and its contribution to measurement precision errors in HR-pQCT followed two complimentary study procedures (Figure 2). First, we performed an *in vivo* scan-rescan reproducibility study to compare the *true* total precision error for HR-pQCT bone measurements (including scan positioning variability) to precision error as it has previously been measured for HR-pQCT (excluding scan positioning variability). Specifically, we compared precision errors for bone parameters measured in the

entire volumes acquired for each repeat scan, to those measured in volume of interest constrained to overlapping slices between scan and rescan. Second, we isolated the error produced by operator positioning variability by calculating bone parameters in sub volumes corresponding to variable operator-defined scan positions. We performed retrospective positioning experiments using custom scan simulation software and the scout view images collected for this second study in order to evaluate: (i) short-term intra-operator, (ii) long-term intra-operator, and (iii) inter-operator positioning precision. Based on the positioning experiments, we quantified the absolute reference line position precision and the corresponding precision errors for bone parameters. This second study was performed in (i) a group of experienced operators, and (ii) a group of operators without previous HR-pQCT experience, who performed custom software training and evaluation exercises prior to their reproducibility experiments.

Operators

This study involved 14 HR-pQCT operators from different imaging centers situated in North America, Europe and Australia. Of these, 8 operators (S.B., S.B., N.P.D., A. G.-Z., M. H., D. L., K. N., C. D.) had significant previous research experience with HR-pQCT at their respective institutions, but heterogeneous training histories. These operators participated only in the reproducibility experiments (Figure 2). The remaining 6 operators (D.C., S.H., K.J., K.K., P.M., N.W.) were clinical research technicians employed to perform HR-pQCT scans for the Osteoporotic Fractures in Men (MrOS) study. These operators were licensed densitometry technicians without previous experience with HR-pQCT. Before participating in the reproducibility experiments, the operators trained on reference line positioning to receive certification. Certification consisted of passing evaluation modules on reference line positioning and performing 10 acceptable imaging exams following the MrOS study protocol. Five of the MrOS operators participated in centralized training at a single MrOS clinical center, while one operator received identical on site training by the same instructors (N.V. and A.J.B.) 4 weeks later.

Imaging

A total of 120 subjects were scanned at University of California, San Francisco (UCSF) and at Mayo Clinic (60 at each center) using each center's first generation HR-pQCT system (XtremeCT, Scanco Medical AG, Brüttisellen, Switzerland). Subjects comprised 56 men and 64 women, with age range of 65–80 years (mean age = 71 ± 4 years). At each imaging center, a single operator scanned all subjects, using an acquisition protocol provided by the organizing center (UCSF). After positioning the subject's non-dominant limb in a carbon fiber cast, the operator inserted the limb in the scanner and acquired standard dorsal-palmar (forearm) and oblique anterolateral-posteromedial (lower leg) scout view images. On this image, the operator identified the anatomic landmark and intersected the horizontal reference line with this position to define the tomographic scan region, according to the guidelines of the manufacturer. Volumetric images were acquired using standard *in vivo* settings (60 kVp, 900 μA, 100 ms integration time). For 34 subjects at UCSF and 30 subjects at Mayo Clinic, a single stack of slices was acquired (110 slices, 9.02 mm) at the standard fixed offset from the reference line (9.5 mm for radius and 22.5 mm for tibia) (Figure 3(a) and (b)). All subjects were scanned twice at the same location with repositioning performed

between acquisitions. These scan-rescan images were used to evaluate the contribution of operators in total acquisition reproducibility. For the remaining 26 subjects at UCSF and 30 subjects at Mayo Clinic, two sequential stacks (220 slices, 18.04 mm) were acquired. The double-stack volume was centered on the standard single-stack location. Therefore the offset from the reference line to the distal-most slice of the scan was 4.99 mm for the radius and 17.99 mm for the tibia (Figure 3(c) and (d)). For the double-stack protocol, all subjects were scanned once per anatomic site. These images were used to retrospectively investigate intra-operator and inter-operator positioning reproducibility, and to compare operators with and without standardized training.

At scan time the UCSF and Mayo Clinic operators assessed image quality by grading the central slice of a low resolution reconstruction from 1 (no motion artifacts) to 5 (severe motion artifacts) [19]. The operator repeated an acquisition when the grade was equal to or greater than 3. If the best scan acquired was greater than grade 3, the image was excluded from the analysis. All subjects signed informed consent to participate in the study, and the Committees on Human Research at UCSF and Mayo Clinic approved study procedures.

Image Analysis

The manufacturer's standard analysis protocol (IPL Version 5.08b, Scanco Medical) was used to semi-automatically segment all scans and calculate basic bone density and structure parameters. The software records mean values on a slice-by-slice basis, which is designed to facilitate future analysis of co-registered sub volumes for longitudinal analyses. We utilized this feature to derive mean values for sub-volumes corresponding to positions defined by each operator in our reproducibility study. In particular, we calculated volumetric bone mineral densities of total (Tt.BMD), cortical (Ct.BMD), and trabecular (Tb.BMD) bone [20]. We computed cortical thickness (Ct.Th) using an annular approximation [20,21], and cortical porosity (Ct.Po) using an extended cortical analysis [3,22,23]. Finally, we calculated trabecular number (Tb.N) and heterogeneity (Tb.Sp.SD) directly using the distance transform method [24].

Mechanical response was estimated using micro-finite element (FE) analysis. From the segmented images, we automatically generated FE meshes converting image voxels to isotropic hexahedral elements [25]. We assigned Young's modulus of 6.829 GPa [26] and Poisson's ratio of 0.3 [27] to all elements, but we labeled cortical and trabecular bone differently to calculate compartmental load distributions at the distal and proximal boundaries of the model. We applied a 1% uniaxial compressive strain to the nodes at the distal axial surface of the bone, and we constrained the nodes at the proximal axial surface. We solved the FE models using an iterative solver (Scanco FE Software v1.12; Scanco Medical AG). From the solutions, we extracted proximal cortical load fraction ($Ct.LF_{prox}$), distal cortical load fraction ($Ct.LF_{dist}$) and failure load ($L_{failure}$)[28].

Operator contribution to total acquisition reproducibility

We measured *in vivo* acquisition precision errors in standard-length stacks of scan-rescan images acquired with subject repositioning. From the 64 scan-rescan exams performed for the first part of our study, we analyzed 57 pairs of images of the radius and 63 pairs of

images of the tibia. We excluded 7 pairs of scan-rescan images of the radius and 1 pair of scan-rescan images of the tibia because of low image quality due to severe motion artifacts.

Scan and rescan images were co-registered by matching cross-sectional area of the bone, resulting in a slice-wise translation that provides an optimal overlap [29]. Positioning precision was calculated as root mean square of the slice shift (SD_{RMS}) between scan and rescan. The precision error for each bone parameter was calculated as root mean square of the coefficient of variation (CV_{RMS}, in %) for measurements computed from the entire 110-slice volumes for each paired scan (precision with operator positioning variability) and from sub-volumes constrained to overlapping slices between scan pairs (precision with operator positioning removed).

Operator positioning reproducibility

We directly measured precision errors contributed by operator variability in positioning the reference line by way of a retrospective reproducibility experiment. The 14 operators participating in this experiment performed reference line positioning on the set of scout view images that corresponded to our double-length acquisitions. The use of double-length stacks was critical to allow the analysis of 110-slice volumes that correspond to each operator's reference line positioning (Figure 3(e) and (f)). From the 56 subjects enrolled for the second study, the image quality of 50 radius and 55 tibia scans was acceptable for analysis. To facilitate the reproducibility experiments, we developed software that simulates the graphical user interface of the HR-pQCT acquisition software (μCT Tomography v5.4C, Scanco Medical AG, Brüttisellen, Switzerland) (Figure 4(b)). The interface includes a pixel-to-pixel reproduction of the scout view image presentation, i.e. one pixel of the image corresponded to one pixel on the scanner's display. All text and graphical elements in the acquisition software were reproduced with the same values, colors, and layout. Graphical control elements were embedded on the left side of the simulation interface to enable control of the reproducibility experiments. The software was created using Matlab (Matlab 2012b, The MathWorks, Inc., Natick, Massachusetts, United States). Scout view images were loaded automatically and sequentially for reference line positioning. During the experiments, operators identified anatomical landmarks on each scout view image and positioned the reference line, following their usual acquisition protocol. To optimally recreate the acquisition environment, operators performed the reproducibility experiments on the same video display used to perform HR-pQCT scans in their laboratory. The software was run from Apple MacBook or Windows PC laptops connected to the secondary video input of the scanner display. The three operator reproducibility experiments were performed in one single session that lasted approximately one hour.

Three specific reproducibility experiments were performed to quantify:

- Short-term intra-operator precision: all 14 operators positioned reference lines on a subset of scout view images that were randomly repeated twice (a total of 45 scout view images of the radius and 48 scout view images of the tibia);
- Long-term intra-operator precision: the two operators that acquired the original images at UCSF and Mayo Clinic repositioned the reference lines on all the

scout view images (50 for the radius and 55 for the tibia) 6–24 months after the original acquisition;

- Inter-operator precision: all 14 operators positioned reference lines on all scout view images (50 scout view images of the radius and 55 scout view images of the tibia).

Experienced operators were instructed to follow their local procedure for scan positioning, while the newly certified MrOS operators were instructed to follow a standard protocol provided to them as part of their training curriculum. Details of the positioning, training, and certification procedure are provided in detail in Appendix I. Short-term intra-operator and inter-operator precision errors were calculated for the experienced and new operators independently. The precision experiments for the MrOS operators were performed 8±2 weeks after training, and repeated to evaluate long term precision at 17 months. For each reproducibility experiment, the positioning precision was calculated as the root mean square of the standard deviation of the absolute reference lines positions (SD_{RMS}). Corresponding bone parameter precision errors were calculated as root mean square of the coefficients of variation (CV_{RMS}) based on measurements performed in the 110-slice sub-volume corresponding to each reference line positioning. Statistical comparisons between short-term, long-term, inter-operator precision errors, as well as between operator groups, were performed using Student's t-test with Bonferroni corrections applied to account for multiple comparisons.

Results

Total in vivo precision errors

Variability in reference line positioning was 0.38 mm for the radius and 0.20 mm for the tibia ($p<0.005$), corresponding to approximately 4.2% and 2.2% of the total stack length, respectively. *In vivo* precision errors were greater for non-co-registered than co-registered images, both for radius and tibia (Table 1). Differences were greater for densitometric and structural parameters than for mechanical parameters. *In vivo* precision errors were significantly higher for total, cortical and trabecular mineral density (Tt.BMD, Ct.BMD, Tb.BMD) and geometric (Ct.Th) measurements when variability in operator positioning was included in the measurement procedure (Non-match). Significant differences in precision errors were not observed for the microstructural parameters (Ct.Po, Tb.N and Tb.Sp.SD) or for FE-derived mechanical parameters. In general, precision errors were greater for radius than for tibia. Consequently, bone measurement errors were significantly greater for radius than for tibia ($p<0.005$), except trabecular number, inter-trabecular heterogeneity, failure load and proximal cortical load fraction.

Experienced Operator Reproducibility

Precision error measurements for operator variability in positioning are reported in Table 2. Short-term positioning precision was 0.3 mm. Short-term intra-operator positioning precision errors tended to be smaller than long-term intra-operator precision errors and significantly smaller than inter-operator precision errors ($p<0.001$), both for radius and tibia. Positioning precision errors were double for radius than for tibia. For both radius and tibia,

short-term intra-operator positioning errors were smaller than long-term intra-operator positioning errors and inter-operator positioning errors. The precision errors for bone parameter errors followed the trends of the positioning precision errors, with respect to short-term vs. long-term intra-operator and inter-operator precision, and radius vs. tibia. Among densitometric parameters, we observed the highest errors for Tt.BMD in the radius, where it reached 3.69% for experienced operators in the inter-operator experiments, and for Tb.BMD for the tibia, where it reached 0.75% for long-term intra-operator variability. Among geometrical and structural parameters, Ct.Th had the highest variations, especially in the radius where the inter-operator error reached 8.40%. In terms of μFE parameters, precision errors were greatest for $Ct.LF_{dist}$, where it reached 6.37% in the radius in the inter-operator reproducibility experiment. Inter-operator precision errors for $L_{failure}$ were modest (~1%).

New Operator Reproducibility

Short-term positioning precision for the group of newly trained MrOS operators was not statistically different from experienced operators at the radius or tibia. Accordingly, there were no significant differences in the short-term bone parameter precision errors between operator groups. In contrast, the inter-operator positioning precision at both radius and tibia for the newly-trained operators was significantly better than for the experienced operators with heterogeneous training histories (radius: 0.68 mm vs. 0.34 mm; tibia 0.30 mm vs. 0.16 mm; both $p<0.001$). Inter-operator positioning precision values for the newly trained operator group approached their respective short-term precision values (radius 0.34 mm vs. 0.28 mm; tibia: 0.16 mm vs. 0.11 mm, $p <0.001$), whereas the differentials between inter-operator and short-term positioning precision values were substantially greater for the experienced group (0.68 mm vs. 0.24 mm; tibia 0.30 vs. 0.13 mm; all $p < 0.001$). Long-term precision errors for the trainees were highly consistent with their initial performance. Intra-operator positioning SD_{RMS} was 0.25 mm for radius (vs. 0.28 mm for the initial experiments) and 0.11 mm for tibia (it was 0.11mm initially). Similarly, inter-operator positioning SD_{RMS} was 0.36 mm (compared to 0.34 mm) and 0.14 mm (vs. 0.16 mm) for the radius and tibia, respectively. Bone parameter precision errors followed the trends of positioning precision errors.

Discussion

In this study, we analyzed the contribution of the operator to HR-pQCT acquisition precision. In particular, we investigated the scan localization process that requires an operator to position a horizontal reference line over a specific anatomic landmark on a scout view image. Our specific aims were to analyze operator reference line positioning in a set of scan-rescan images; to separately investigate short-term and long-term intra-operator precision and inter-operator precision; and to evaluate the efficacy of a training tool for reference line positioning. Operator reference line positioning contributed significantly to acquisition precision. Precision errors, and consequently bone parameter and mechanical measurements, were larger for radius than for tibia, and for inter-operator than intra-operator cases. Short-term intra-operator errors were similar for experienced and trained operators,

whereas inter-operator errors for trained operators were smaller than those for experienced operators and tended to be similar to short-term inter-operator errors.

Our first significant finding is that operatory variability in reference line positioning is a major component of *in vivo* precision error, for both the radius and tibia. Our scan-rescan precision errors measured in co-registered volumes (which inherently excludes positioning variability) were on the same order of magnitude as previously reported in other reproducibility studies ([2,4,13,15,16]). The error we measured in unregistered repeat scans was somewhat greater than what was reported by Ellouz et al. [16], though both studies observed significantly higher precision errors in unregistered vs. co-registered volumes. In the radius, precision errors for densitometric parameters were two-fold greater, and Ct.Th three-fold greater when operator positioning variability was included. In fact, the error due to positioning variability is greater for these parameters than the error due to severe motion artifacts [15,19]. In general, precision errors for all bone parameters were significantly higher when variability in reference line positioning was included in the measurement, though the magnitude of the errors in trabecular structure and μFE parameters was less pronounced.

The precision errors measured by analyzing the full 110-slice volumes in our scan-rescan study are better estimates of true *in vivo* precision for cross-sectional studies [2–4,13,16]. On the other hand, positioning precision is less critical for longitudinal studies, where co-registration of baseline and follow up acquisitions - necessary to measure changes in the same volume of interest - inherently removes variability in scan positioning. Therefore, the precision errors measured by analyzing co-registered volumes in our scan-rescan study (and most previous HR-pQCT reproducibility studies) are reasonable estimates of precision for longitudinal studies that employ image registration.

The second main outcome of our study is the direct quantification of precision errors due to operator positioning variability. The targeted investigation of operator positioning revealed that inter-operator variability is significantly greater than intra-operator variability, both for radius and tibia. This underscores the critical need for standard training procedures for operators charged with collecting valuable reference data and for multicenter studies. Long-term variability was greater than short-term intra-operator variability – suggesting operator drift is a possible scenario that could be addressed with regular training refreshers. The parameters most affected by positioning variability were cortical thickness and integral bone density. The sensitivity of these parameters to operator positioning precision is consistent with the significant variance in these properties along the distal-proximal axis [17,18]. By the same token trabecular microstructure and cortical porosity, which do not vary as dramatically across the scan volume, were less sensitive to positioning variability. The precision errors for apparent mechanical properties derived by μFE were modestly sensitive to positioning variability (~1% or less), however load distributions between trabecular and cortical compartments were highly sensitive, especially in the radius. Finally, our operator reproducibility study demonstrated the relatively poorer reproducibility positioning the reference line in radius scout images, compared to tibia scout images, reflecting the challenge in landmark identification for the radius.

Our third principal finding is that inter-operator positioning precision errors were half for operators who used our training software to learn standard positioning procedures, compared to operators who had significant previous HR-pQCT experience, but heterogeneous individual training histories. In fact, the inter-operator precision values of our trainees approached their respective short-term intra-operator precision errors, whereas a significant discrepancy between intra- and inter-operator precision errors was found for the experienced group. In addition, we re-tested operator reproducibility of the MrOS trainees 17 months after certification. Encouragingly, the long-term performance indicated that the operators maintained or improved their personal precision performance, while comparability across operators was maintained over approximately 1.5 years. This interval corresponds to three quarters of the data collection period for these MrOS visits. This suggests that our training platform could be a useful tool to homogenize operator positioning, especially for multicenter studies. These results also illustrate the importance of providing operators standardized positioning criteria and training. In this study, we gave operators precise guidelines (as detailed in Appendix 1) that provided a more detailed description of the procedure recommended by the manufacturer. In addition, while training, operators had immediate feedback on their positioning performance, and had the opportunity to repeat training exercises until they achieved satisfactory scores. The current version of the software was specifically implemented for the MrOS study to certify the new operators. For general HR-pQCT operator training and certification, the software has subsequently been translated into a free web application (http://webapps.radiology.ucsf.edu/refline/). At the time of the paper publication, the webapp contains the same set of images used in this reproducibility study. In the future, we aim to develop modulus for other populations and anatomic sites, including protocols for pediatric and articular joint imaging [30–32].

The identification of the anatomical landmark was significantly more challenging for the radius compared to the tibia, with poorer positioning precision and larger resulting errors in bone parameter measurements. In fact the anatomical landmark recommended by the manufacturer (a peak representing the confluence of the scaphoid and lunate fossae on the radial articular surface) is often not visible, and therefore requires a secondary process for localizing the reference line. Measurement precision errors at the radius are compounded by the substantial variability in geometry and density along the distal-proximal axis [17,18]. In this study, positioning variability at the radius was significantly poorer than positioning variability for the tibia, demonstrating that the greater subjectivity involved in the landmark definition highly impacts positioning precision. Thus, improved landmark guidelines for the radius could significantly reduce measurement precision errors. The most straightforward solution would be to change the current anatomic landmark to a new anatomical landmark that is consistently visible in scout view images, less variable in its visual presentation, and features that are simple to precisely describe to operators. Possible new landmarks include the medial or the lateral margins of the radiocarpal articular surfaces of the radius (Figure 5). The medial margin has the advantage that it would exclude variability due to the length of the joint surface (Figure 5, length (d)), though it can be partially obscured by superposition of the ulna. On the other hand the lateral margin is likely very precisely identifiable, however it would include significant anatomic variability due to the length of the joint surface. One drawback of changing the landmark could be compromised comparability with

existing data; however adjusting the offset from the reference line to the scan region could be optimized to provide data comparability. An alternative would be software-based automatic positioning of the reference line at scan time [33]. Although we are actively working on this direction, automatic landmark detection involves significantly more hurdles to implement in a real-time product on the scanner. Furthermore this represents a more significant methodological transition that will require agreement by the user community and the manufacturer. The third solution could be to train operators to recognize anatomic landmarks with rigorously detailed and standardized protocols. In this study, we implemented this third solution and incorporated it into our training procedure and software. Operator training considerably decreased inter-operator positioning variability at the radius, which was half compared to experienced operators, though still greater than variability observed at the tibia. Therefore, changing the anatomic landmark combined with standardized operator training is warranted to improve the precision errors for HR-pQCT of the radius.

Strengths and limitations must be acknowledged to this study. This was the first study that systematically isolated the contribution of the operators in HR-pQCT reproducibility. Measurements were performed on the same dataset for different operators, using double-length images from which we could extract pseudo-measurements based on reference line positions set by the different operators participating in our reproducibility studies. By evaluating sub-volumes within each dataset we fixed all sources of acquisition variability except operator positioning. The experiments were performed under conditions closely approximating actual scanning, with software reproducing the XtremeCT interface and operators using their XtremeCT workstation's display to visualize the scout images. The dataset represents the main limitation of this study. For the new trained operators, images from the same dataset were used both for training and evaluation and for the precision study, albeit presented in different orders. Therefore, we did not test short- nor long-term precision in entirely independent datasets from the training data. However, trained operators participated in the precision experiments 33–75 days after their software training and following the initiation of MrOS study data collection. The number of operators in a training group was small, though compared to the total number of operational HR-pQCT scanners, this is not an insignificant number. It should also be noted that our tests did not account for small variations in limb orientation that may impact the radiographic presentation of the scout image, and therefore contribute to reference line placement variability. Future experiments could systematically examine the degree to which realistic variability in limb orientation impacts scan positioning precision. Finally, for practical considerations we were only able to perform reproducibility experiments after the MrOS operators completed training. In the future it would be worthwhile to test positioning precision in the same group of operators before and after completion of the training procedure. However, it is important for investigators to consider if "re-training" operators involved in on going data collection is advisable. It would be our recommendation that centers utilize the training platform for new technician hires, or for experienced technicians transitioning to new studies.

In conclusion, we demonstrated that scan positioning precision has a large impact on HR-pQCT measurement errors. Scan positioning has not been recognized as source of measurement error in previous reproducibility studies and not accounted for in published

precision data to date. This finding has significant implications particularly for cross-sectional study design. We also demonstrated that inter-operator precision errors are significantly lower for operators that participate in standardized training procedures compared to experienced operators without a standard training history. Our data also suggest that the selection of a new landmark for positioning distal radius scans should be considered to reduce the error introduced by anatomical variability. We released our scan positioning training tool as a web-based application and plan to extend its functionality to include training modules for pediatric protocols, and potentially protocols for other anatomic sites such as the hand, foot, and knee.

Acknowledgments

The authors thank Isra Saeed and Louise McCready for subject recruitment, Margaret Holets for data acquisition, James Peterson for data management, and Nicholas P. Derrico and the MrOS operators that took part in this study: Deborah Cusick, Shannon Hanson, Kristi Jacobson, Kyla Kent, Patricia Miller, and Nita Webb.

This study was funded by NIH/NIAMS R01 AR060700 and by The Osteoporotic Fractures in Men (MrOS) Study, which is supported by National Institutes of Health funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Advancing Translational Sciences (NCATS), and NIH Roadmap for Medical Research under the following grant numbers: U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, and UL1 TR000128.

References

1. Cheung AM, Adachi JD, Hanley Da, Kendler DL, Davison KS, Josse R, et al. High-resolution peripheral quantitative computed tomography for the assessment of bone strength and structure: a review by the Canadian Bone Strength Working Group. Curr Osteoporos Rep. 2013; 11:136–46. [PubMed: 23525967]
2. 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; 90:6508–15. [PubMed: 16189253]
3. 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; 47:519–28. [PubMed: 20561906]
4. Khosla S, Riggs BL, Atkinson EJ, Oberg AL, McDaniel LJ, Holets M, et al. Effects of sex and age on bone microstructure at the ultradistal radius: a population-based noninvasive in vivo assessment. J Bone Miner Res. 2006; 21:124–31. [PubMed: 16355281]
5. 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; 26:50–62. [PubMed: 20593413]
6. Kazakia GJ, Tjong W, Nirody Ja, Burghardt AJ, Carballido-Gamio J, Patsch JM, et al. The influence of disuse on bone microstructure and mechanics assessed by HR-pQCT. Bone. 2014; 63:132–40. [PubMed: 24603002]
7. Zhu TY, Griffith JF, Qin L, Hung VW, Fong T-N, Au S-K, et al. Alterations of Bone Density, Microstructure and Strength of the Distal Radius in Male Patients with Rheumatoid Arthritis: A Case-Control Study with HR-pQCT. J Bone Miner Res. 2014:1–42. [PubMed: 23712442]
8. Rizzoli R, Laroche M, Krieg M-A, Frieling I, Thomas T, Delmas P, et al. Strontium ranelate and alendronate have differing effects on distal tibia bone microstructure in women with osteoporosis. Rheumatol Int. 2010; 30:1341–8. [PubMed: 20512336]
9. Seeman E, Delmas PD, Hanley Da, Sellmeyer D, Cheung AM, Shane E, et al. Microarchitectural deterioration of cortical and trabecular bone: differing effects of denosumab and alendronate. J Bone Miner Res. 2010; 25:1886–94. [PubMed: 20222106]

10. Melton LJ, Christen D, Riggs BL, Achenbach SJ, Müller R, van Lenthe GH, et al. Assessing forearm fracture risk in postmenopausal women. Osteoporos Int. 2010; 21:1161–9. [PubMed: 19714390]
11. Christen D, Melton LJ, Zwahlen A, Amin S, Khosla S, Müller R. Improved fracture risk assessment based on nonlinear micro-finite element simulations from HRpQCT images at the distal radius. J Bone Miner Res. 2013; 28:2601–8. [PubMed: 23703921]
12. Glüer C, Blake G, Lu Y, Blunt B. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporos Int. 1995; 5:262–70. [PubMed: 7492865]
13. MacNeil JA, Boyd SK. Improved reproducibility of high-resolution peripheral quantitative computed tomography for measurement of bone quality. Med Eng Phys. 2008; 30:792–9. [PubMed: 18164643]
14. Burghardt AJ, Pialat J-B, Kazakia GJ, Boutroy S, Engelke K, Patsch JM, et al. Multicenter precision of cortical and trabecular bone quality measures assessed by high-resolution peripheral quantitative computed tomography. J Bone Miner Res. 2013; 28:524–36. [PubMed: 23074145]
15. Engelke K, Stampa B, Timm W, Dardzinski B, de Pappa E, Genant HK, et al. Short-term in vivo precision of BMD and parameters of trabecular architecture at the distal forearm and tibia. Osteoporos Int. 2012; 23:2151–8. [PubMed: 22143491]
16. Ellouz R, Chapurlat R, van Rietbergen B, Christen P, Pialat J-B, Boutroy S. Challenges in longitudinal measurements with HR-pQCT: Evaluation of a 3D registration method to improve bone microarchitecture and strength measurement reproducibility. Bone. 2014
17. Boyd SK. Site-specific variation of bone micro-architecture in the distal radius and tibia. J Clin Densitom. 2008; 11:424–30. [PubMed: 18280194]
18. Mueller TL, van Lenthe GH, Stauber M, Gratzke C, Eckstein F, Müller 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; 45:882–91. [PubMed: 19615477]
19. 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 MicroArchitecture. Bone. 2012; 50:111–8. [PubMed: 22019605]
20. Laib A, Häuselmann HJ, Rüegsegger P. In vivo high resolution 3D-QCT of the human forearm. Technol Health Care. 1998; 6:329–37. [PubMed: 10100936]
21. 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; 81:364–71. [PubMed: 17952361]
22. 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; 25:983–93. [PubMed: 19888900]
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; 41:505–15. [PubMed: 17693147]
24. Hildebrand T, Ru P. A new method for the model-independent assessment of thickness in three-dimensional images. 1997; 185:67–75.
25. Müller R, Rüegsegger P. Three-dimensional finite element modelling of non-invasively assessed trabecular bone structures. Med Eng Phys. 1995; 17:126–33. [PubMed: 7735642]
26. MacNeil J, 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; 42:1203–13. [PubMed: 18358799]
27. Van Rietbergen B, Odgaard A, Kabel J, Huiskes R. Direct mechanics assessment of elastic symmetries and properties of trabecular bone architecture. J Biomech. 1996; 29:1653–7. [PubMed: 8945668]
28. Mueller TL, Christen D, Sandercott S, Boyd SK, van Rietbergen B, Eckstein F, et al. Computational finite element bone mechanics accurately predicts mechanical competence in the human radius of an elderly population. Bone. 2011; 48:1232–8. [PubMed: 21376150]

29. Laib A, Hildebrand T, Häuselmann H, Rüegsegger P. Ridge number density: a new parameter for in vivo bone structure analysis. Bone. 1997; 21:541–6. [PubMed: 9430245]
30. Kirmani S, Christen D, van Lenthe GH, Fischer PR, Bouxsein ML, McCready LK, et al. Bone structure at the distal radius during adolescent growth. J Bone Miner Res. 2009; 24:1033–42. [PubMed: 19113916]
31. Burrows M, Liu D, McKay H. High-resolution peripheral QCT imaging of bone micro-structure in adolescents. Osteoporos Int. 2010; 21:515–20. [PubMed: 19322507]
32. Barnabe C, Feehan L. The Journal of Rheumatology High-resolution Peripheral Quantitative Computed Tomography Imaging Protocol for Metacarpophalangeal Joints in Inflammatory Arthritis: The SPECTRA Collaboration The Journal of Rheumatology is a monthly international serial edi. 2012; 39:7–9.
33. Carballido-Gamio J, Bonaretti S, Holets M, Saeed I, McCready L, Majumdar S, et al. Automated Scan Prescription For HR-pQCT: A Multi-Atlas Prospective Registration Approach. ASBMR. 2013

Appendix 1 – HR-pQCT operator training

To train new operators, we developed custom software and a detailed scan positioning protocol. The software comprised two modules, one to train reference line positioning and one to evaluate positioning performance for the purposes of operator certification. The graphical user interface reproduces the original acquisition software (Figure 1A(c)). Graphical control elements on the left side of the interface allowed the user to select training or evaluation modules, and from a series of scout view collections (8 scout images each) according to anatomic site, laterality, and difficulty. In training mode, the software provides operators both immediate and summary feedback on their positioning accuracy. The immediate feedback consists of a traffic light that is illuminated after each reference line is positioned. A green, yellow or red light is shown, corresponding to “exact”, “good”, or “out of range” positions, respectively. The final feedback consists of a separate window displaying a summary report of operator performance after completing the selected collection (Figure 4(d)). Positioning feedback was determined by comparing the user’s reference line position to a gold standard reference position. The gold standard positions were determined through independent consensus of three experienced operators at UCSF. “Exact” indicates that operator and reference positions coincided; “good” indicates that operator position was within 2 pixels from the reference position; and “out of range” indicates that operator position was greater than 2 pixels from the reference position. Two pixels corresponded to 0.34 mm, which was the short-term standard deviation among experienced operators.

We created training and evaluation collections for the radius and tibia from the double-length scans described in the Methods. A horizontal mirror image of each scout view image was created to have an equal number of “left” and “right” limb images, for a total of 112 scout view images per anatomic site. We divided the scout view images into two groups depending on the degree of difficulty in positioning the reference line. Difficulty was assigned based on the presence or absence of a clear anatomic landmark, as assessed by an experienced UCSF operator. From each of these two groups, 49 images were randomly selected to create the training collections, and we used the remaining 7 images to create the evaluation datasets. Both for training and evaluation modules, we created 15 collections of 8 images randomly selected within their visibility group. Before accessing the reproducibility

experiments, the 6 new operators had to complete all the datasets of the training module at least once, and had to succeed on at least 3 sets for the radius and 3 sets for the tibia. A set was considered successful when the sum of "*exact*" and "*good*" positions was 7 out of 8 positions.

Together with the training and evaluation software, we provided specific guidelines to identify the anatomic landmark and to position the reference line, as follow (Figure 1A). For the radius we defined the anatomic landmark as the peak of the radiocarpal articular surface of the radius, which separates the radiopaque subchondral bone (bright signal) from the radiolucent joint space (dark signal) (Figure 1A(a)–(b)). When the peak is clearly visible, the reference line must intersect the peak (Figure 1A(a)), whereas when the peak is not visible, the reference line must intersect the mid-point of the joint surface (Figure 5(b)). The hiker analogy in Figure 1A(e)–(h) illustrates the concept. Similarly, for the tibia the anatomic landmark is the peak of the edge of the tibial plafond, which separates the radiopaque subchondral bone (bright signal) from the radiolucent joint space (dark signal) (Figure 1A(c)–(d)). When the peak is clearly visible, the reference line must intersect the center of the peak (Figure 1A(c)), whereas when the peak is not visible, the reference line must intersect the flat plafond (Figure 1A(d)).

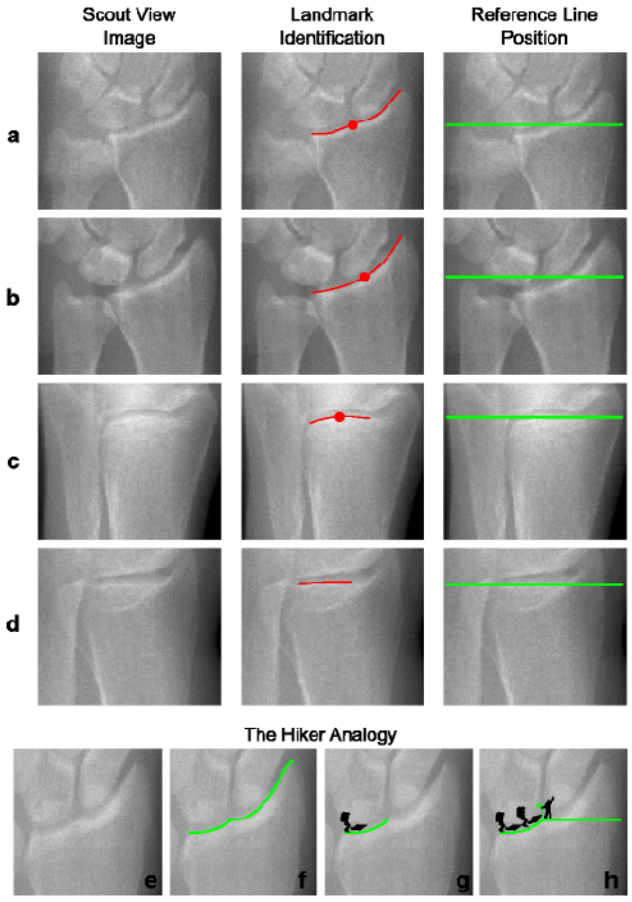

Figure 1A.

Guidelines for reference line positioning recommended for the training and evaluation software. In the radius, the reference line intersects (a) the peak when visible or (b) the mid-articular surface when the peak is not visible. Similarly, in the tibia, the reference line intersects (c) the peak when visible or (d) the flat plafond when the peak is not visible. In the radius, the hiker analogy easily explains the positioning of the reference line. The articular surface of the radius (e) can be considered as a hiking path (f). The hiker walks along the path (g) until he reaches the peak where he plants his flag, which coincides with the location where the reference line intersects the edge that represents the articular surface of the radius (h).

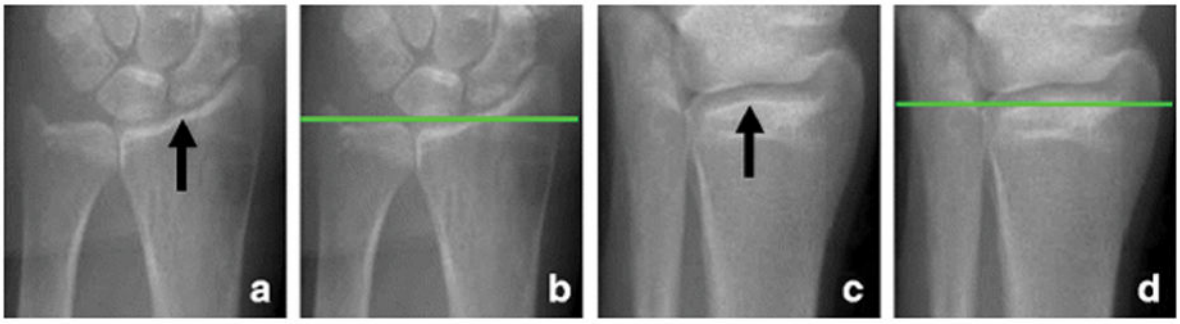

Figure 1.

Identification of the anatomic landmarks and positioning of the reference line in scout view images of radius and tibia. The anatomic landmarks are (a) an inflection in the curvature of the articular surface of the radius between the scaphoid and lunate fossae of the radiocarpal joint for the forearm scan, and (c) the apex of the distal articular plateau of the tibia at the tibiotalar joint for the lower leg scan. To define the scan region, the operator manually positions a horizontal reference line (b and d) to intersect the landmark. The volume of interest is offset from this reference line by a fixed distance.

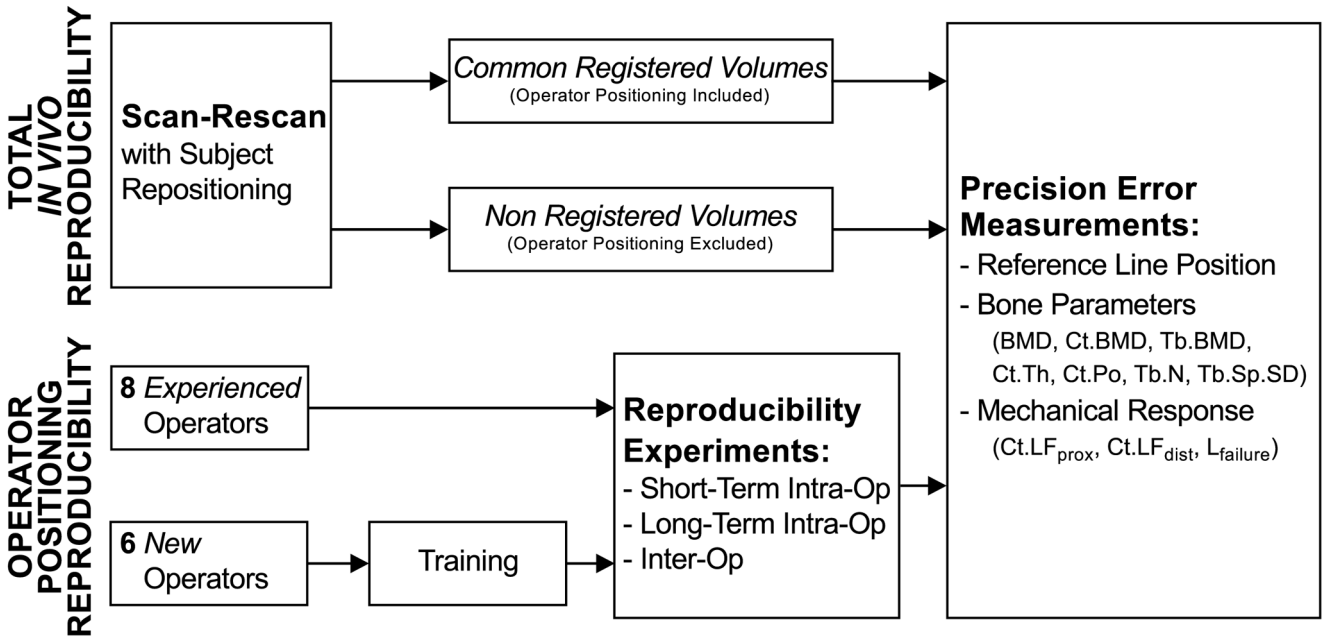

Figure 2.

Design of our study. First we investigated the contribution of HR-pQCT operators to total acquisition precision, and then we decomposed operator precision in intra- and inter-operator precision. In the last case, we also tested the efficacy of our training tool at reducing operator precision errors.

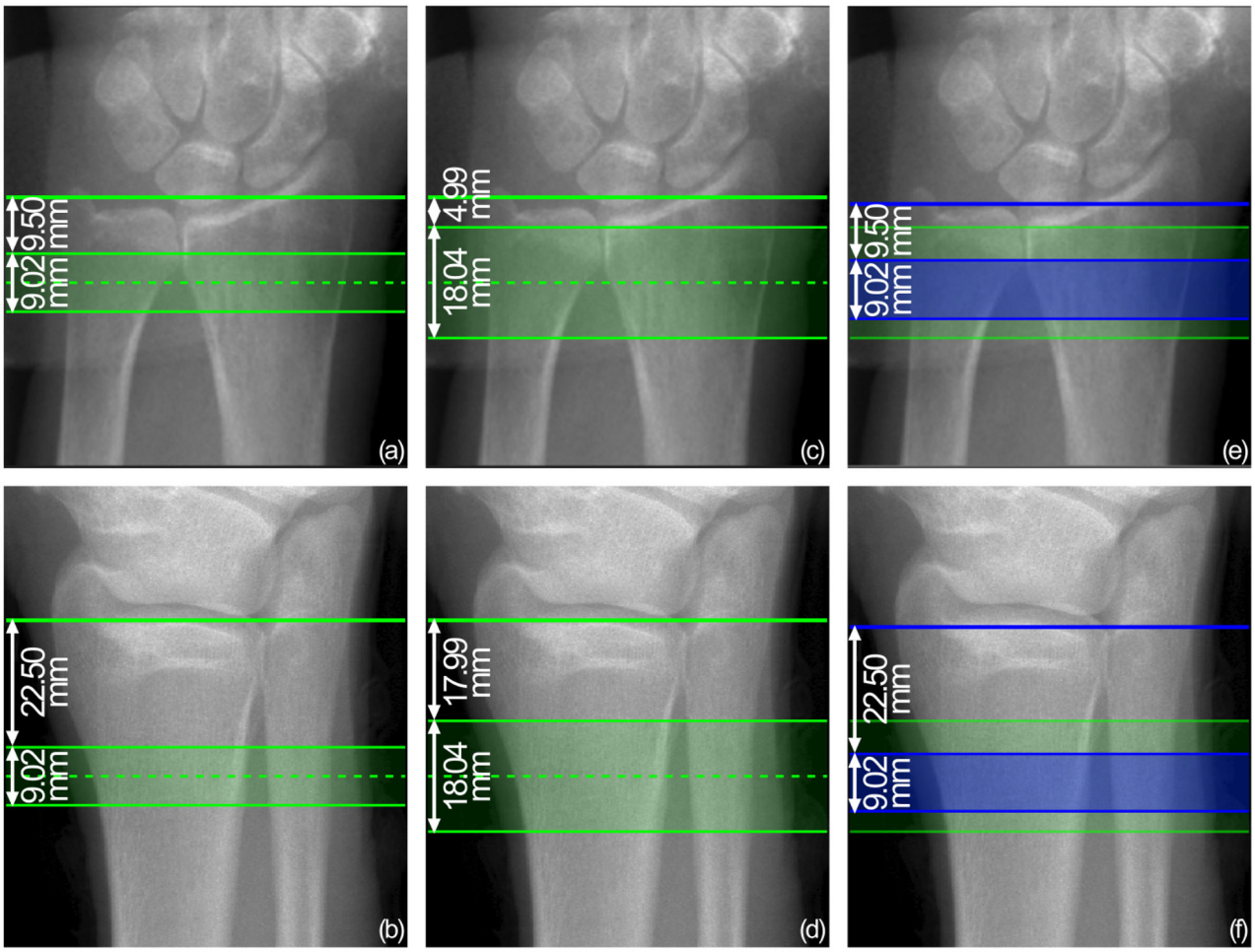

Figure 3.

Scout view images of radius and tibia, illustrating the regions scanned in this study. (a)–(b) Standard-length, single stack volume: scan coverage starts at a fixed offset (9.50 mm for radius and 22.50 mm for tibia) from the reference line and extends 9.02mm. (c)–(d) Double-length, two stack volume: scan coverage is centered on the standard stack position and extends 18.04 mm. (e)–(f) Standard-length sub-volume superimposed on the double-length volume, corresponding to retrospective operator reference line positioning: the standard stack (blue) starts at a fixed offset (9.50mm for radius and 22.50mm for tibia) from the reference line (blue) positioned by the operator in the scan simulation software, and extends for 9.02 mm within the double-length volume (green).

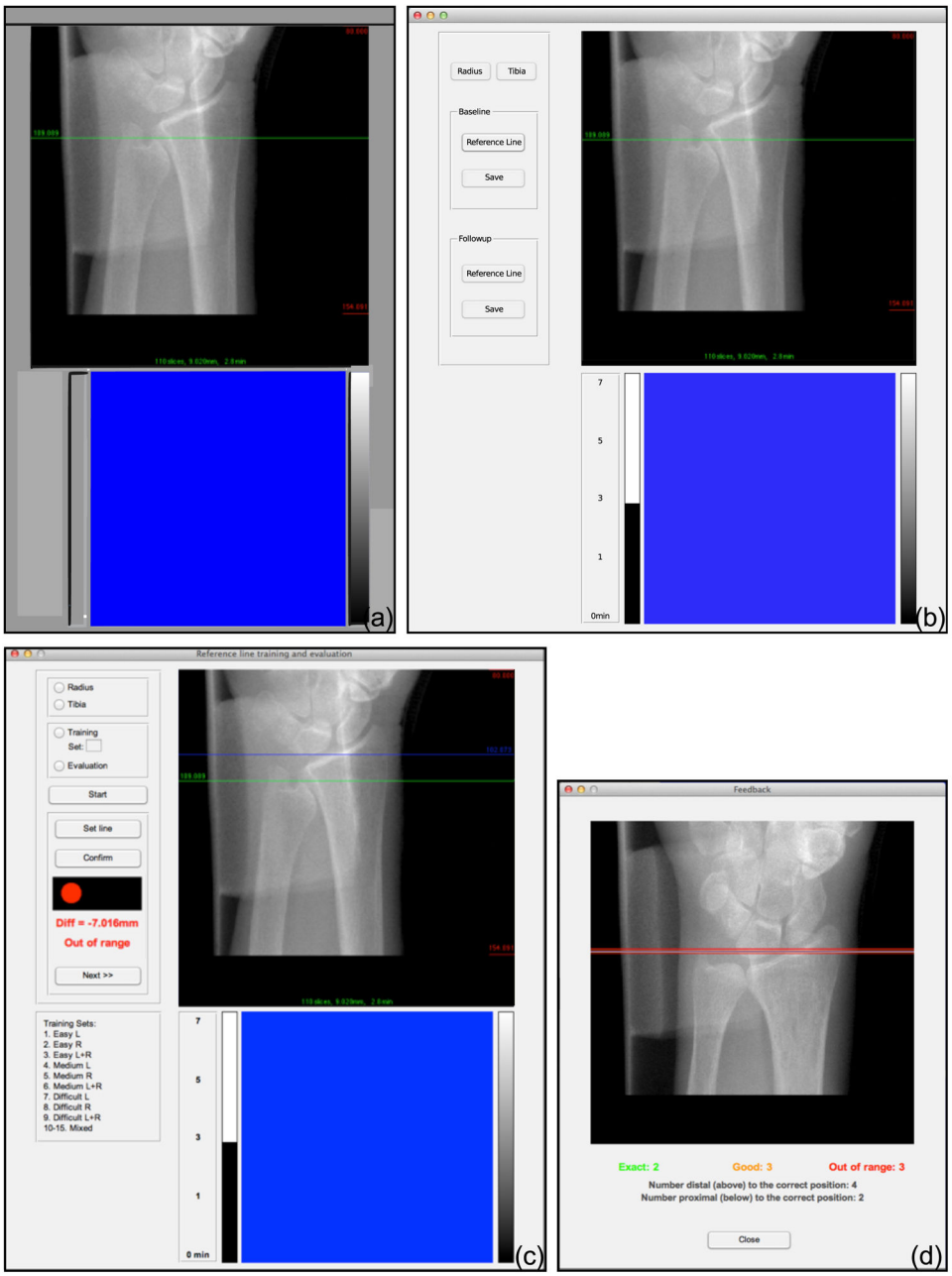

Figure 4.

The graphical user interface of the software used in this study. (a) Interface of the original acquisition software of the Scanco XtremeCT (μCT Tomography v5.4c). (b) Interface of the software created for the operator reproducibility study. The dimension of the top viewer was 512×512 pixels, and the dimension of the bottom viewer was 384×384 pixel. In the top viewer, the scout view is presented to the user as a pixel-level reproduction of the original acquisition software. (c) Interface of the training and evaluation software for reference line positioning. Operators could choose between training and evaluation modules, and limb and laterality on which to exercise and receive positioning feedback for each scout. (d) Summary feedback presented by the training software to the operator after completion of a scout view

collection. The green lines (“exact” positioning) shows the ground truth position on a representative scout image, whereas yellow lines (“good” positioning) and red lines (“out of range” positioning) report distances from the ground truth for the preceding exercise.

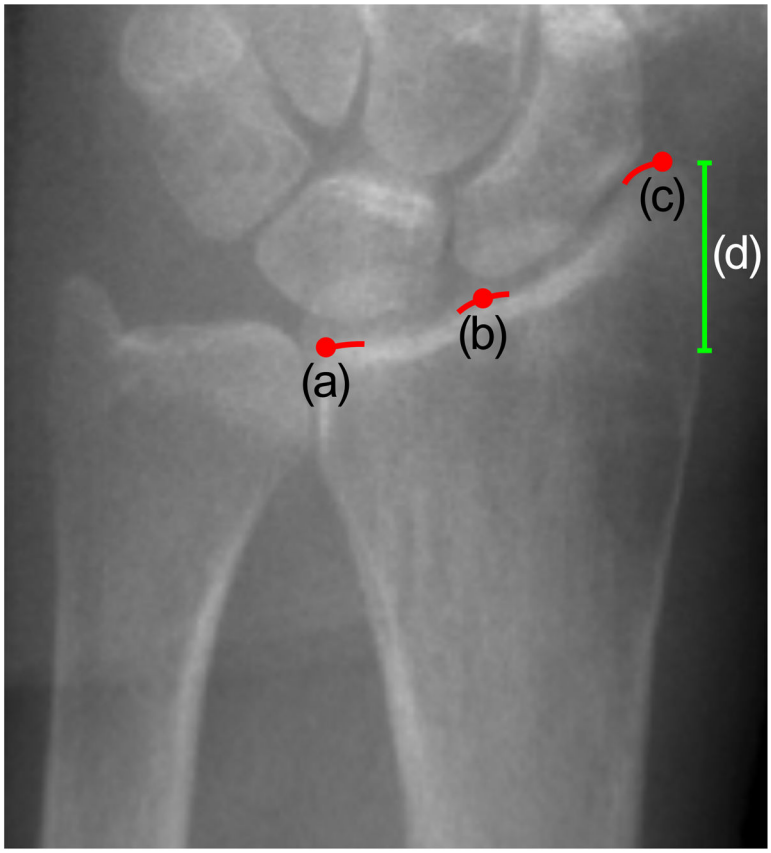

Figure 5.

Possible landmarks on the radius joint surface: (a) medial margin, (b) the notch defined by the scaphoid and lunate fossae and (c) lateral margin of the radiocarpal joint surface. Currently, HR-pQCT operators position the reference line that defines the region to be scanned at the notch between the scaphoid and lunate fossa of the radius (b). Medial (a) and lateral (c) margins represent the possible candidates for a more visible and easily identifiable anatomical landmark. The medial margin (a) might be preferable because it would exclude the subject-dependent joint height (d).

Table 1

Precision of reference line positioning and corresponding *in vivo* precision errors for bone parameter measurements in the full 110-slice scans (Non-match; includes operator positioning variability) and co-registered scan-rescan sub-volumes (Match; excludes operator positioning variability).

	Positioning Precision SD_{RMS} **[mm]**	**Bone Parameter Precision** CV_{RMS} **[%]**							**Biomechanics Precision** CV_{RMS} **[%]**		
		Tt.BMD	**Ct.BMD**	**Tb.BMD**	**Ct.Th**	**Ct.Po**	**Tb.N**	**Tb.Sp.SD**	$L_{failure}$	$Ct.LF_{prox}$	$Ct.LF_{dist}$
RADIUS *(n = 57; o = 2)*											
Non-match	0.38[b]	3.33[a,b]	2.35[a,b]	1.63[a,b]	7.88[a,b]	12.02[b]	4.86	6.70	2.88	2.17	5.45[b]
Match	/	0.89[a,c]	0.77[a,c]	0.73[a]	2.12[a,c]	8.85[c]	4.27	5.72	2.69	1.88	4.34
TIBIA *(n = 63; o = 2)*											
Non-match	0.20[b]	0.90[a,b]	1.54[a,b]	0.93[b]	2.52[a,b]	5.61[b]	5.71	5.61	2.76	1.96	3.95[b]
Match	/	0.43[a,c]	1.38[a,c]	0.76	1.30[a,c]	4.07[c]	5.82	6.33	2.73	1.87	3.73

[a] Non-match vs. match comparison; $p<0.001$.

[b] Non-match radius vs. tibia comparison; $p<0.005$.

[c] Match radius vs. tibia comparison; $p<0.005$.

n = number of pairs of scan-rescan images.

o = number of operators.

Table 2

Precision of reference line positioning and corresponding errors in bone parameter measurements due solely to operator positioning variability. Errors were calculated for short-term intra-operator, long-term intra-operator and inter-operator reproducibility. Operators were divided into groups: experienced operators without common training, and new operators with recent central training.

RADIUS	**Positioning Precision** SD_{RMS} **[mm]**	**Bone Parameter Precision** CV_{RMS} **[%]**							**Biomechanics Precision** CV_{RMS} **[%]**		
		Tt.BMD	**Ct.BMD**	**Tb.BMD**	**Ct.Th**	**Ct.Po**	**Tb.N**	**Tb.Sp.SD**	**$L_{failure}$**	**Ct.LF_{prox}**	**Ct.LF_{dist}**
Short-term intra-op (n = 15)*											
Experienced (o = 8)	0.24 ± 0.05^{a}	$1.39 \pm 0.32^{a,c}$	$0.93 \pm 0.19^{a,c}$	0.40 ± 0.10^{a}	$3.17 \pm 0.65^{a,c}$	$2.78 \pm 0.81^{a,c}$	0.47 ± 0.14^{a}	0.83 ± 0.26^{a}	$0.42 \pm 0.12^{a,c}$	0.58 ± 0.12^{a}	$2.65 \pm 0.62^{a,c}$
New (o = 6)	0.28 ± 0.08^{c}	$1.50 \pm 0.25^{a,c}$	1.26 ± 0.40	0.72 ± 0.57	$3.46 \pm 1.24^{a,c}$	3.19 ± 0.96^{c}	0.69 ± 0.56	1.00 ± 0.44^{c}	0.41 ± 0.11^{c}	0.67 ± 0.25	3.03 ± 0.90^{c}
Long-term intra-op (n = 50; o =2)	0.35	2.09^{c}	1.41^{c}	0.70	4.76	4.29^{c}	1.05	1.77	0.71^{c}	0.83	3.94
Inter-operator (n = 50)											
Experienced (o = 8)	$0.68^{a,b,c}$	$3.69^{a,b,c}$	$2.56^{a,b,c}$	$1.46^{a,b}$	$8.40^{a,b,c}$	$5.45^{a,b,c}$	$2.12^{a,b}$	$3.11^{a,b}$	$1.17^{a,b,c}$	$2.04^{a,b,c}$	$6.37^{a,b,c}$
New (o = 6)	$0.34^{b,c}$	$2.09^{a,b,c}$	$1.50^{b,c}$	0.87^{b}	$4.90^{a,b,c}$	$3.24^{b,c}$	1.32^{b}	$1.94^{a,b,c}$	$0.72^{a,b,c}$	$0.96^{b,c}$	$3.73^{a,b,c}$

(a)

TIBIA	**Positioning Precision** SD_{RMS} **[mm]**	**Bone Parameter Precision** CV_{RMS} **[%]**							**Biomechanics Precision** CV_{RMS} **[%]**		
		Tt.BMD	**Ct.BMD**	**Tb.BMD**	**Ct.Th**	**Ct.Po**	**Tb.N**	**Tb.Sp.SD**	**$L_{failure}$**	**Ct.LF_{prox}**	**Ct.LF_{dist}**
Short-term intra-op (n = 16)*											
Experienced (o = 8)	0.13 ± 0.07^{a}	$0.26 \pm 0.15^{a,c}$	$0.19 \pm 0.10^{a,c}$	0.26 ± 0.21^{a}	$0.94 \pm 0.50^{a,c}$	$0.69 \pm 0.33^{a,c}$	0.31 ± 0.18^{a}	0.48 ± 0.26^{a}	$0.07 \pm 0.03^{a,c}$	0.42 ± 0.16^{a}	$0.92 \pm 0.45^{a,c}$

TIBIA	**Positioning Precision** SD_{RMS} *[mm]*	**Bone Parameter Precision** CV_{RMS} *[%]*							**Biomechanics Precision** CV_{RMS} *[%]*		
		Tt.BMD	**Ct.BMD**	**Tb.BMD**	**Ct.Th**	**Ct.Po**	**Tb.N**	**Tb.Sp.SD**	**$L_{failure}$**	**Ct.LF_{prox}**	**Ct.LF_{dist}**
New (o = 6)	$0.11 \pm 0.03^{a,c}$	0.31 ± 0.06^{c}	0.30 ± 0.07	0.28 ± 0.08	$0.52 \pm 0.29^{a,c}$	0.66 ± 0.22^{c}	0.31 ± 0.05	0.34 ± 0.02^{c}	$0.06 \pm 0.01^{a,c}$	0.39 ± 0.07	$0.75 \pm 0.24^{a,c}$
Long-term intra-op (n = 55; o = 2)	0.36	0.72^{c}	0.58^{c}	0.75	3.01	0.73^{c}	0.93	1.59	0.18^{c}	0.90	2.45
Inter-operator (n = 55)											
Experienced (o = 8)	$0.30^{a,b,c}$	$0.61^{a,b,c}$	$0.42^{a,b,c}$	$0.65^{a,b}$	$1.97^{a,b,c}$	$1.72^{a,b,c}$	$0.85^{a,b}$	$1.43^{a,b}$	$0.20^{a,b,c}$	$0.84^{a,b,c}$	$2.02^{a,b,c}$
New (o = 6)	$0.16^{a,b,c}$	$0.30^{b,c}$	$0.21^{b,c}$	0.32^{b}	$1.02^{a,b,c}$	$0.94^{b,c}$	0.41^{b}	$0.69^{a,b,c}$	$0.08^{a,b,c}$	$0.44^{b,c}$	$1.32^{a,b,c}$

(b)

[a] Short-term intra-operator vs. inter-operator comparison; p<0.001.

[b] Experienced vs. new operator comparison; p<0.001.

[c] Radius vs. tibia; p<0.001.

n = number of images.

o = number of operators.

[*] For short-term intra-operator precision and corresponding bone parameters, values refer to mean ± standard deviation.