Title

HR-pQCT cross-calibration using standard vs. Laplace-Hamming binarization approach

Journal

JBMR Plus, 8(10)

ISSN

2473-4039

Authors

Sadoughi, Saghi
Subramanian, Aditya
Ramil, Gabriella
et al.

Publication Date

2024-09-01

DOI

10.1093/jbmrpl/ziae116

HR-pQCT Cross-Calibration Using Standard vs. Laplace-Hamming Binarization Approach

Saghi Sadoughi PhD, Aditya Subramanian, Gabby Ramil MS, Minhao Zhou PhD, Andrew J. Burghardt, Galateia J. Kazakia PhD

Bone Quality Research Lab, Department of Radiology and Biomedical Imaging, University of California San Francisco, CA, USA.

Running title: Standard vs. Laplace-Hamming Cross-Calibration for XCTII

Author contribution:

Saghi Sadoughi: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing-original draft preparation, Writing-review & editing

Aditya Subramanian: Investigation, Methodology, Software, Writing-review & editing

Gabby Ramil: Investigation, Methodology, Software, Writing-review & editing

Minhao Zhou: Investigation, Visualization, Writing-review & editing

Andrew J. Burghardt: Conceptualization, Writing-review & editing

Galateia J. Kazakia: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Visualization, Writing-review & editing

Corresponding author:

Galateia J. Kazakia
Department of Radiology and Biomedical Imaging
185 Berry St., Lobby 6, Suite 350
San Francisco, CA 94107
Galateia.Kazakia@ucsf.edu
(415) 353-4534

Abstract

High-Resolution Peripheral Quantitative Computed Tomography (HR-pQCT) has emerged as a powerful imaging technique for characterizing bone microarchitecture in the human peripheral skeleton. The second-generation HR-pQCT scanner provides improved spatial resolution and a shorter scan time. However, the transition from the first-generation (XCTI) to second-generation HR-pQCT scanners (XCTII) poses challenges for longitudinal studies, multi-center trials, and comparison to historical data. Cross-calibration, an established approach for determining relationships between measurements obtained from different devices, can bridge this gap and enable the utilization and comparison of legacy data. The goal of this study was to establish cross-calibration equations to estimate XCTII measurements from XCTI data, using both the standard and Laplace-Hamming (LH) binarization approaches. Thirty-six volunteers (26–85 years) were recruited and their radii and tibiae were scanned on both XCTI and XCTII scanners. XCTI images were analyzed using the manufacturer's standard protocol. XCTII images were analyzed twice: using the manufacturer's standard protocol and the LH segmentation approach previously developed and validated by our team. Linear regression analysis was used to establish cross-calibration equations. Results demonstrated strong correlations between XCTI and XCTII density and geometry outcomes. For most microstructural outcomes, although there were considerable differences in absolute values, correlations between measurements obtained from different scanners were strong, allowing for accurate cross-calibration estimations. For some microstructural outcomes with a higher sensitivity to spatial resolution (e.g., trabecular thickness, cortical pore diameter), XCTII standard protocol resulted in poor correlations between the scanners, while our LH approach improved these correlations and decreased the difference in absolute values and the proportional bias for other measurements. For these reasons and due to the improved accuracy of our LH approach compared to the standard approach, as established in our previous study, we propose that investigators should use the LH approach for analyzing XCTII scans, particularly when comparing to XCTI data.

Keywords: HR-pQCT, Laplace-Hamming Binarization, Gaussian Binarization, cross-calibration, estimation error

1. Introduction

High-Resolution Peripheral Quantitative Computed Tomography (HR-pQCT) has emerged as a powerful non-invasive imaging technique for in vivo characterization of bone microarchitecture in the human peripheral skeleton. HR-pQCT quantifies geometric, densiometric, microstructural, and biomechanical properties of trabecular and cortical bone [1–3], contributing to fracture risk assessment [4,5]. Further, HR-pQCT data are used in the study of bone biology, disease progression, and treatment outcomes [6–10].

The first-generation HR-pQCT scanner (XCTI) has an isotropic voxel size of 82 μm. At this resolution, XCTI image analysis relies on tissue density and histomorphometric model assumptions [11,12]. The second-generation HR-pQCT scanner (XCTII) provides an improved voxel size of 61 μm. This higher resolution allows for direct measurement of all trabecular parameters using distance transform methods that do not rely on density and model assumptions [13]. Higher resolution also provides improved visualization and quantification of fine features within both the trabecular and cortical compartments. However, the transition from first-generation to second-generation scanners poses challenges for longitudinal studies, multi-center trials, and the comparison of XCTII-derived data to historical XCTI-derived results, as these activities require ensuring compatibility and comparability of measurements across different scanners.

Cross-calibration is an established approach for determining relationships between measurements obtained from different devices, allowing for the estimation of measurements on newer devices using data collected from older or different devices [14–16]. The objective of this study is to estimate second-generation HR-pQCT measurements from first-generation scanner data using cross-calibration. By establishing reliable cross-calibration equations, we aim to bridge the gap between different generations of HR-pQCT scanners and enable the utilization of the rich body of historical XCTI data in conjunction with the advanced capabilities of second-generation scanners. This approach holds promise for facilitating longitudinal studies, improving comparability across different research centers, and enhancing the clinical application of HR-pQCT in bone health assessment. In the following sections, we will describe the methodology employed for cross-calibration, present the validation results, and discuss the implications of our findings. We will perform cross-calibration using both the manufacturer's

standard XCTII Gaussian-based binarization approach, as well as the Laplace-Hamming segmentation approach we previously developed and validated for XCTII [17].

2. Materials and Methods

2.1. Study Participants

Based on the sample size in our previously published study that highlighted the effect of Laplace-Hamming segmentation approach on XCTII density and structural outcome metrics, 36 volunteers (16 women, 20 men; 56 ± 13 years; age range 26–85; **Table 1**) were recruited to ensure sufficient eligible tibia and radius scans for cross-calibration [17]. Women known to be pregnant or breast-feeding and men and women with metal implants at both scan sites (radius and tibia) were excluded. There were no other health- or bone-related inclusion or exclusion criteria for enrollment. The institutional review board at the University of California, San Francisco approved the study protocol, and all participants provided written informed consent prior to their involvement in the study.

2.2. Image Acquisition

Volunteers were scanned on both first-generation (XCTI) and second-generation (XCTII) HR-pQCT scanners (Scanco Medical, Brüttisellen, Switzerland) using the manufacturer's standard in vivo protocols. For XCTI, the scan settings were: source potential 60 kVp, tube current 900 mA, isotropic 82 μm nominal resolution. Scans were acquired 9.5 mm and 22.5 mm proximal from the joint for distal radius and tibia, respectively, covering a length of 9.0 mm (110 slices). For XCTII, the scan settings were: source potential 68 kVp, tube current 1460 mA, isotropic 61 μm nominal resolution. Scans were acquired 9 mm and 22 mm proximal from the joint for distal radius and tibia, respectively, covering a length of 10.2 mm (168 slices); compared to XCTI, XCTII captured approximately 0.5 mm more bone at either end of the scan volume. The non-dominant forearm and lower leg were scanned. In the case where a history of fracture or surgery was reported on the non-dominant side, the contralateral side was scanned. Prior to scanning and throughout the study, the scanners were routinely calibrated using scanner-specific quality control phantoms. All in vivo scans were visually inspected for motion artifacts using the manufacturer's grading scheme to ensure adequate quality for all the images included for further analyses (grades 1 to 3) [18].

2.3. Image Analysis

XCTI images were analyzed using the manufacturer's standard XCTI patient evaluation protocol, which involved semi-automated contouring of the periosteal and endosteal cortical boundaries, a Laplace-Hamming filter and a fixed global threshold for binarization [1,11]. The standard microstructural quantification scripts were applied, which use a combination of direct (e.g., trabecular number) and indirect measurement techniques (e.g., cortical thickness).

XCTII images were analyzed twice – first with the manufacturer's standard Gaussian segmentation approach and second with the Laplace-Hamming (LH) segmentation approach we developed and validated in-house [17]. In both XCTII analysis protocols, an auto-contouring process was first used to identify the periosteal and endosteal cortical boundaries; these auto-contours were checked and corrected manually if necessary. In the standard protocol (i.e., the Gaussian segmentation approach), a Gaussian filter and fixed BMD thresholds were then used to extract the trabecular and cortical bone. In our developed LH segmentation approach, a Laplace-Hamming filter followed by a fixed global threshold was used to extract the trabecular and cortical bone. In both analysis protocols, once the trabecular and cortical binary masks were created, direct measurement of all parameters was performed using distance transform methods [13]. The LH approach affected only the microstructural parameters derived from the binary masks; volumetric density and geometry measurements remain identical for both the standard and LH analysis protocols [17]. For both XCTI and XCTII, complete stacks were analyzed; no registration between XCTI and XCTII volumes was performed.

2.4. Cross-Calibration

To compare the agreement between the measurements obtained from the two scanners, a combination of cross-validation and bootstrapping approaches was utilized. For radius or tibia scans, the 36 data points were first randomly divided into two sets: one set with 12 data points and another with 24 data points. The 12-data points set was set aside for validation purposes (to determine the estimation error after cross-calibration), while the remaining 24 data points were used for cross-calibration (using linear regression analysis). During this step, 1000 sets of 12 data points were randomly selected (with replacement) for establishing cross-calibration equations, resulting in 1000 sets of slopes and intercepts

for each output parameter (**Table 1**). The average values from these 1000 sets were recorded as the slope and intercept for that specific outcome measure.

Next for the validation step, the cross-calibration equations were applied to the first-generation output parameters to estimate the second-generation outputs, defined as XCTII*. To evaluate the accuracy of these estimates, the cross-calibration error was determined by comparing XCTII* with the respective actual measured values on XCTII (i.e., XCTII). This procedure was performed twice, once for the XCTII standard analysis approach and a second time for the XCTII LH approach.

2.5. Statistical Analysis

Participant demographics were compared between the training and test sets using unpaired student's t-test and Pearson's chi-squared test. Linear regression analysis was used to establish cross-calibration equations. Correlation strength was defined as strong ($R^2 > 0.9$), moderate ($0.7 < R^2 < 0.9$) and weak ($R^2 < 0.7$). Bland-Altman plots were used to explore the differences between scanners. Statistical analyses were performed using JMP 16 software (SAS Institute Inc., Cary, NC, USA) with significance set at $p < 0.05$.

3. Results

3.1. Cross-Calibration

Of the total 72 scans (36 radius, 36 tibia), 2 radius scans were excluded from analyses due to motion.

There were strong correlations between XCTI and XCTII density outcomes (Tt.BMD, Tb.BMD, Ct.BMD) at both radius and tibia ($R^2 > 0.88$; **Table 2**). Bland-Altman plots showed good agreement for these outcomes between scanners at both sites although radius Ct.BMD was slightly underestimated by XCTII relative to XCTI (**Supplemental Figure S1**; Note: The R^2 values in **Table 2** and the **Supplementary Figures** are different because **Table 2** presents the average results from all 1000 bootstrapping iterations, while the **Supplementary Figures** present representative results from 1/1000 bootstrapping iterations).

Geometry outcomes Tb.Ar and Ct.Ar were strongly correlated between XCTI and XCTII at both the radius and tibia ($R^2 > 0.91$; **Table 2**). Bland-Altman plots showed good agreement between scanners and proportional biases were small (**Supplemental Figure S2**).

For trabecular microstructure outcomes, there were strong correlations between scanners for BV/TV at both the radius and tibia ($R^2 = 0.97$ and 0.99, respectively; **Table 2**) and Tb.1/N.SD at radius ($R^2 = 0.92$; **Table 2**). There were moderate correlations between scanners at both the radius and tibia for Tb.N ($R^2 = 0.76$ and 0.82, respectively; **Table 2**) and Tb.Sp ($R^2 = 0.88$ and 0.79, respectively; **Table 2**). Among all the outcomes, Tb.Th showed the weakest correlation between scanners at both the radius and tibia ($R^2 = 0.69$ and 0.31, respectively; **Table 2**). Relative to XCTI, BV/TV, Tb.Th, Tb.Sp, and Tb.1/N.SD were overestimated by XCTII at both the radius and tibia, while Tb.N was underestimated by XCTII at both sites. Bland-Altman plots showed a strong proportional bias for BV/TV at both the radius and tibia, for Tb.Sp at the radius, and for Tb.N at the tibia (**Supplemental Figures S3** and **S4**). The analysis of XCTII data using our developed LH approach greatly reduced the observed proportionality bias (**Supplemental Figures S3** and **S4**).

For cortical microstructure outcomes, there was a strong correlation for Ct.Th at both radius and tibia between scanners ($R^2 > 0.92$; **Table 2**). There was a moderate correlation for Ct.Po at both the radius and tibia between scanners ($R^2 = 0.83$ at both sites; **Table 2**). Ct.Po.Dm showed the weakest correlations at both the radius and tibia ($R^2 = 0.21$ and 0.38, respectively; **Table 2**). Relative to XCTI, Ct.Po was underestimated by XCTII at both the radius and tibia, while Ct.Po.Dm was overestimated by XCTII at both sites (**Supplemental Figures S5** and **S6**). There was a strong proportional bias in Ct.Po at both the radius and tibia, which was eliminated using our LH approach (**Supplemental Figures S5** and **S6**).

3.2. Validation

Applying the established cross-calibration equations to our test set XCTI data, the estimated values for XCTII outcomes were calculated and compared against the measured values to compute percent error. Linear regression and Bland-Altman plots that compare measured XCTII vs estimated XCTII* outcomes using both standard and LH approaches are presented in **Figure 1** to **Figure 3**.

For trabecular outcomes, the mean percent error was between 0.80% and 4.65% for density and geometry outcomes (Tt.BMD, Tb.BMD, Tb.Ar), and between 1.68% and 7.58% for microstructural

outcomes (BV/TV, Tb.N, Tb.Th, Tb.Sp, Tb.1/N.SD). For cortical outcomes, the mean percent error was between 0.61% and 10.71% for density and geometry outcomes (Ct.BMD, Ct.Ar, Ct.Pm, Ct.Th), and between 6.81% and 21.98% for microstructural outcomes (Ct.Po, Ct.Po.Dm). Using the LH approach for cross-calibration resulted in smaller percent errors for BV/TV, Tb.Th, and Ct.Po.Dm, but larger percent errors for Tb.N, Tb.Sp, and Ct.Po (**Table 3**). More specifically, percent errors for Tb.Th from LH were significantly smaller at both the radius and tibia ($p = 0.04$ and < 0.001, respectively; **Table 3**); percent error for Ct.Po.Dm was significantly smaller at the radius ($p = 0.006$; **Table 3**).

4. Discussion

The goal of this study was to explore the feasibility and accuracy of estimating second-generation HR-pQCT measurements from the first-generation scanner with cross-calibration using linear regression equations. For most density, geometry, and microstructural outcomes, cross-calibration accurately estimated and eliminated the differences between scanners. However, there were large estimation errors for some microstructural outcomes, notably cortical porosity.

As expected, and consistent with previous studies, cross-scanner correlations for density and geometry outcomes were stronger compared to those for microstructural outcomes [14–16]. Density and geometry metrics are less sensitive to resolution and image artifacts (including motion artifacts) due to their quantification over larger scales, lending them greater stability across scanners. This is also in line with previous multi-center studies reporting a smaller variability in density and geometry measures across different XCTI scanners compared to microstructural measures [19,20]. Conversion of density and geometry metrics from first-generation to second-generation scanners is therefore feasible.

We found large differences in the absolute values of some microstructural outcomes between the two scanners. Consistent with the findings of Agarwal et. al. and Manske et. al. [14,15], we found that XCTII overestimated BV/TV, Tb.Th, and Tb.Sp and underestimated Tb.N and Ct.Po relative to XCTI. This can be explained by the inherent differences in the acquisition and analysis approach of the two scanners, including differences in resolution, image analysis (filtering and thresholding), and quantitative analysis approach (XCTII direct vs XCTI indirect). The standard XCTII analysis approach also obscures some fine features in both the trabecular and cortical compartments in the segmentation step and thickens larger trabeculae, therefore resulting in overestimated BV/TV, Tb.Th and Tb.Sp and underestimated

Tb.N and Ct.Po, as established in our previous publication [17]. Despite these differences in absolute values, the correlations between the two scanners were strong for these parameters, explaining the relatively small percent error in estimated values. However, for some other microstructural outputs such as Tb.Th and Ct.Po.Dm, in addition to the large differences in absolute values, the correlations were weak between the two scanners, resulting in relatively larger percent errors in the estimated values. This could be due to the higher sensitivity of these outcomes to spatial resolution, suggesting that estimating XCTII data from XCTI data for such measures may not be recommended using XCTII standard analysis approach. Using our developed XCTII LH approach, which is more consistent with the analysis approach of XCTI, improved the correlations and, therefore, resulted in smaller estimate errors for these metrics. In addition, we found stronger cross-scanner correlations and better agreement between the measured and estimated values at the tibia compared to the radius due to fewer motion artifacts in the tibia scans compared to the radius scans.

Our study has a number of strengths. First, including both the standard and LH binarization approaches provides direct comparisons and insight into the effects of different segmentation approaches on the cross-calibration process and the consequent estimations. Next, combining cross-validation and bootstrapping takes advantage of the benefits of both two methods – bootstrapping provides estimates for the parameters, and cross-validation provides estimates of the test error.

Our study also has a number of important limitations. First, despite the proven improved accuracy and reproducibility of the LH binarization approach compared to the standard approach for XCTII for the cohort studied here, the effect of LH on HR-pQCT outcome metrics has not been examined for disease cohorts with pathology-specific bone microarchitecture (e.g., great cortical bone loss with chronic kidney disease). Thus, the effect of LH should be evaluated on a cohort-by-cohort basis before any generalization can be made upon further validation. Additionally, although bootstrapping was deployed to minimize the cross-calibration bias, sex was not matched between the tibia training and test sets, potentially contributing to the proportional biases and limiting the robustness of the cross-calibration results when extrapolated to larger and more diverse cohorts. However, this bootstrapping approach can be readily adapted to produce additional cross-calibration equations in diverse cohorts, which can potentially be integrated with our cross-calibration equations with improved generalizability. Last, it's worth mentioning that since we chose to perform our analysis on full scan regions for both XCTI and

XCTII, reference line placement inconsistency may be a potential contributor to our reported error values. However, all our scans were acquired by one skilled operator to minimize such errors.

In summary, we found good agreement between density and geometry outcomes measured by XCTI and XCTII; there were large differences in the absolute values for some of the microstructural outcomes, suggesting that they should be directly compared between the two scanners. However, we found moderate to strong correlations for these microstructural outcomes between scanners despite the large differences in absolute values, showing that XCTII outcomes can be estimated from XCTI measurements using the established cross-calibration equations. For microstructural outputs such as Tb.Th and Ct.Po.Dm that have a higher sensitivity to spatial resolution, the standard segmentation approach on XCTII resulted in weak correlations between the scanners, while the LH approach we developed and validated in-house resulted in stronger correlations between the two scanners for these outcomes. These stronger correlations were accompanied by reduced percent errors in these outcomes, as well decreased differences in absolute values and the proportional bias in Bland-Altman plots for other measurements. For these reasons, and considering the improved accuracy of our LH approach compared to the standard approach established in our previous publication [17], we propose that investigators should use the LH approach for analyzing scans on XCTII, particularly when comparing to XCTI data.

5. Disclosures

Saghi Sadoughi – None

Aditya Subramanian – None

Gabby Ramil – None

Minhao Zhou – None

Andrew J. Burghardt – None

Galateia J. Kazakia – None

6. Acknowledgements

UCSF Department of Radiology and Biomedical Imaging Pilot Grant (MRBPLG2004)

UCSF Core Center for Musculoskeletal Biology in Medicine, National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (P30 AR075055)

National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (R01 AR060700)

Office of Research Infrastructure Programs, NIH (S10 OD021803)

7. Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

8. References

MacNeil JA, Boyd SK. Accuracy of high-resolution peripheral quantitative computed tomography for measurement of bone quality. Med Eng Phys. 2007;29(10):1096–105.

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):1678–86.

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):1203–13.

Mikolajewicz N, Bishop N, Burghardt AJ, Folkestad L, Hall A, Kozloff KM, Lukey PT, Molloy-Bland M, Morin SN, Offiah AC, Shapiro J, van Rietbergen B, Wager K, Willie BM, Komarova S V., Glorieux FH. HR-pQCT Measures of Bone Microarchitecture Predict Fracture: Systematic Review and Meta-Analysis. Journal of Bone and Mineral Research. 2020;35(3):446–59.

Cheung WH, Hung VWY, Cheuk KY, Chau WW, Tsoi KKF, Wong RMY, Chow SKH, Lam TP, Yung PSH, Law SW, Qin L. Best Performance Parameters of HR-pQCT to Predict Fragility Fracture: Systematic Review and Meta-Analysis. Journal of Bone and Mineral Research. 2021;36(12):2381–98.

Dalzell N, Kaptoge S, Morris N, Berthier A, Koller B, Braak L, Van Rietbergen B, Reeve J. Bone micro-architecture and determinants of strength in the radius and tibia: Age-related changes in a population-based study of normal adults measured with high-resolution pQCT. Osteoporosis International. 2009;20(10):1683–94.

Kawalilak CE, Johnston JD, Olszynski WP, Kontulainen SA. Characterizing microarchitectural changes at the distal radius and tibia in postmenopausal women using HR-pQCT. Osteoporosis International. 2014;25(8):2057–66.

306 Seeman E, Delmas PD, Hanley DA, Sellmeyer D, Cheung AM, Shane E, Kearns A, Thomas T, Boyd
307 SK, Boutroy S, Bogado C, Majumdar S, Fan M, Libanati C, Zanchetta J. Microarchitectural
308 deterioration of cortical and trabecular bone: differing effects of denosumab and alendronate. Journal of
309 Bone and Mineral Research. 2010;25(8):1886–94.

310 Nishiyama KK, Cohen A, Young P, Wang J, Lappe JM, Guo XE, Dempster DW, Recker RR, Shane E.
311 Teriparatide increases strength of the peripheral skeleton in premenopausal women with idiopathic
312 osteoporosis: a pilot HR-pQCT study. J Clin Endocrinol Metab. 2014 Jul;99(7):2418–25.

313 Bacchetta J, Boutroy S, Vilayphiou N, Juillard L, Guebre-Egziabher F, Rognant N, Sornay-Rendu E,
314 Szulc P, Laville M, Delmas PD, Fouque D, Chapurlat R. Early impairment of trabecular
315 microarchitecture assessed with HR-pQCT in patients with stage II-IV chronic kidney disease. Journal
316 of Bone and Mineral Research. 2010;25(4):849–57.

317 Laib A, Rüegsegger P. Calibration of trabecular bone structure measurements of in vivo three-
318 dimensional peripheral quantitative computed tomography with 28-µm-resolution microcomputed
319 tomography. Bone. 1999;24(1):35–9.

320 Krause M, Museyko O, Breer S, Wulff B, Duckstein C, Vettorazzi E, Glueer C, Püschel K, Engelke K,
321 Amling M. Accuracy of trabecular structure by HR-pQCT compared to gold standard µcT in the radius
322 and tibia of patients with osteoporosis and long-term bisphosphonate therapy. Osteoporosis
323 International. 2014;25(5):1595–606.

324 Manske SL, Zhu Y, Sandino C, Boyd SK. Human trabecular bone microarchitecture can be assessed
325 independently of density with second generation HR-pQCT. Bone. Elsevier Inc.; 2015;79:213–21.

326 Manske SL, Davison EM, Burt LA, Raymond DA, Boyd SK. The Estimation of Second-Generation HR-
327 pQCT From First-Generation HR-pQCT Using In Vivo Cross-Calibration. Journal of Bone and Mineral
328 Research. John Wiley and Sons Inc.; 2017 Jul 1;32(7):1514–24.

329 Agarwal S, Rosete F, Zhang C, McMahon DJ, Guo XE, Shane E, Nishiyama KK. In vivo assessment of
330 bone structure and estimated bone strength by first- and second-generation HR-pQCT. Osteoporosis
331 International. Springer London; 2016 Oct 1;27(10):2955–66.

332 Mikolajewicz N, Zimmermann EA, Rummler M, Hosseinitabatabaei S, Julien C, Glorieux FH, Rauch F,
333 Willie BM. Multisite longitudinal calibration of HR-pQCT scanners and precision in osteogenesis
334 imperfecta. Bone. Elsevier Inc.; 2021 Jun 1;147.

335 Sadoughi S, Subramanian A, Ramil G, Burghardt AJ, Kazakia GJ. A Laplace–Hamming Binarization
336 Approach for Second-Generation HR-pQCT Rescues Fine Feature Segmentation. Journal of Bone and
337 Mineral Research. John Wiley and Sons Inc; 2023 Jul 1;38(7):1006–14.

338 Pialat JB, Burghardt AJ, Sode M, Link TM, Majumdar S. Visual grading of motion induced image
339 degradation in high resolution peripheral computed tomography: Impact of image quality on measures
340 of bone density and micro-architecture. Bone. Elsevier Inc.; 2012;50(1):111–8.

Burghardt AJ, Pialat JB, Kazakia GJ, Boutroy S, Engelke K, Patsch JM, Valentinitsch A, Liu D, Szabo E, Bogado CE, Zanchetta MB, McKay HA, Shane E, Boyd SK, Bouxsein ML, Chapurlat R, Khosla S, Majumdar S. Multicenter precision of cortical and trabecular bone quality measures assessed by high-resolution peripheral quantitative computed tomography. Journal of Bone and Mineral Research. 2013;28(3):524–36.

Boutroy S, Khosla S, Sornay-Rendu E, Zanchetta MB, McMahon DJ, Zhang CA, Chapurlat RD, Zanchetta J, Stein EM, Bogado C, Majumdar S, Burghardt AJ, Shane E. Microarchitecture and Peripheral BMD are Impaired in Postmenopausal White Women with Fracture Independently of Total Hip T-Score: An International Multicenter Study. Journal of Bone and Mineral Research. 2016;31(6):1158–66.

9. Figures and Tables

Figure 1: Regression and Bland-Altman plots for Tt.vBMD **(A-D)**, Tb.vBMD **(E-H)**, Ct.vBMD **(I-L)** Tb.Ar **(M-P)**, and Ct.Ar **(Q-T)** assessed using measured XCTII values and the estimated XCTII* values from XCTI at the radius (left) and tibia (right). On regression plots, the dashed line indicates the line of unity. On Bland-Altman plots, the solid black line indicates the mean difference, the dashed lines indicate the 95% limits of agreement, and the gray line indicates zero.

Figure 2. Regression and Bland-Altman plots for BV/TV **(A-D)**, Tb.N **(E-H)**, Tb.Th **(I-L)**, Tb.Sp **(M-P)**, Tb.1/N.SD **(Q-T)**, Ct.Po **(U-X)**, and Ct.Po.Dm **(Y-AB)** assessed using the measured XCTII values and the estimated XCTII* values from XCTI at the radius using the standard (left) and LH (right) approach. On regression plots, the dashed line indicates the line of unity. On Bland-Altman plots, the solid black line indicates the mean difference, the dashed lines indicate the 95% limits of agreement, and the gray line indicates zero.

Figure 3. Regression and Bland-Altman plots for BV/TV **(A-D)**, Tb.N **(E-H)**, Tb.Th **(I-L)**, Tb.Sp **(M-P)**, Tb.1/N.SD **(Q-T)**, Ct.Po **(U-X)**, and Ct.Po.Dm **(Y-AB)** assessed using measured the XCTII values and the estimated XCTII* values from XCTI at the tibia using the standard (left) and LH (right) approach. On regression plots, the dashed line indicates the line of unity. On Bland-Altman plots, the solid black line indicates the mean difference, the dashed lines indicate the 95% limits of agreement, and the gray line indicates zero.

Table 1. Participant characteristics categorized by Training and Test sets for bootstrapping. Values are shown as mean ± SD, or count (percentage). Unpaired student's t-test and Pearson's chi-squared test were performed between the Training and Test sets when applicable.

Table 2. Regression analysis results for trabecular and cortical outcomes reported for both the standard and LH approach. All values represent averages derived from the cross-calibration procedure.

Table 3. Mean absolute percent error between measured XCTII and estimated XCTII* outcomes by cross-calibration for trabecular and cortical outcomes, reported for both the standard and developed LH approach. The *p*-values are obtained from the paired student's t-test performed between the percent errors from the standard and LH approach.