Opportunistic Quantitative CT Bone Mineral Density Measurement at the Proximal Femur Using Routine Contrast-Enhanced Scans: Direct Comparison With DXA in 355 Adults

Timothy J Ziemlewicz,1 Alyssa Maciejewski,1 Neil Binkley,2 Alan D Brett,3 J Keenan Brown,3 and Perry J Pickhardt1

1Department of Radiology, University of Wisconsin, Madison, WI, USA
2Osteoporosis Clinical Research Program, University of Wisconsin, Madison, WI, USA
3Mindways Software, Inc., Austin, TX, USA

ABSTRACT
For patients undergoing routine contrast-enhanced CT examinations, an opportunity exists for concurrent osteoporosis screening without additional radiation exposure or patient time using proximal femur CT X-ray absorptiometry (CTXA). We investigated the effect of i.v. contrast enhancement on femoral neck CTXA T-score measurement compared with DXA. This cohort included 355 adults (277 female; mean age, 59.7 ± 13.3 years; range, 21 to 90 years) who underwent standard contrast-enhanced CT assessment at 120 kVp over an 8-year interval, as well as DXA BMD assessment within 100 days of the CT study (mean 46 ± 30 days). Linear regression and a Bland-Altman plot were performed to compare DXA and CTXA results. CTXA diagnostic sensitivity and specificity was evaluated with DXA as the reference standard. There was good correlation between DXA and CTXA (r² = 0.824 for both areal BMD and T-scores) and the SD of the distribution of residuals was 0.063 g/cm² or 0.45 T-score units. There was no trend in differences between the two measurements and a small bias was noted with DXA T-score +0.18 units higher than CTXA. CTXA had a sensitivity for discriminating normal from low bone mineral density of 94.9% (95% CI, 90.6% to 97.4%). For opportunistic osteoporosis screening at routine post-contrast abdominopelvic CT scans, CTXA produces T-scores similar to DXA. Because femoral neck CTXA BMD measurement is now included in the WHO Fracture Risk Assessment Tool (FRAX) tool, this opportunistic method could help to increase osteoporosis screening because it can be applied regardless of the clinical indication for CT scanning. © 2016 American Society for Bone and Mineral Research.

KEY WORDS: SCREENING; OSTEOPOROSIS; BONE MINERAL DENSITY; CONTRAST-ENHANCED; QCT

Introduction

Osteoporosis conveys a significant lifetime fracture risk, but can be detected with screening and subsequently treated, decreasing the number of fractures in patients at risk.1–3 Despite this, osteoporosis screening with dual-energy X-ray absorptiometry (DXA) remains underused, with nearly one-half of female Medicare beneficiaries never having undergone screening4 and 80% of those with a major osteoporosis-related fracture not having prior bone mineral density (BMD) testing.5 Meanwhile, abdominopelvic CT is a common procedure with 130.6 studies being performed per every 1000 Medicare beneficiaries in the United States annually.6 Previous work has shown that CT scans obtained for other indications can be used to assess BMD at the spine and may provide an acceptable initial screen.7–10 Although these CT vertebral measures can be used to identify those at highest risk for fracture, none provide 10-year fracture risk, which can truly guide when further workup and treatment is indicated.

The World Health Organization Fracture Risk Assessment Tool (WHO-FRAX) allows estimation of 10-year fracture risk, which guides treatment initiation decisions.11,12 This tool uses a femoral neck BMD or T-score along with patient characteristics and risk factors to estimate the 10-year fracture risk. Currently, the WHO-FRAX tool includes femoral neck QCT measurement using the Mindways QCT system (Mindways Software, Austin, TX, USA). A previous study has shown a significant correlation between QCT and DXA at the femoral neck utilizing unenhanced CT.13 Prior studies including CT of the femoral neck show minimal differences between precontrast and postcontrast values, suggesting contrast-enhanced CT (CECT) may be useful for opportunistic screening.14,15 Comparison between post-contrast CT and DXA is lacking; therefore, the purpose of this study was to investigate the correlation between and diagnostic...
Patients and Methods

Patient Cohort

This study was performed under a waiver of informed consent from our institutional review board. A medical records search identified 727 patients who had a CECT scan of the abdomen and pelvis within 100 days prior to or following a DXA examination (mean 46 ± 30 days) over an 8-year period. A total of 372 patients were excluded due to at least one of the following reasons:

1. Imaging performed at 140 kVp;
2. Inferior scan extent not including the lesser trochanters;
3. Slice thickness >5 mm;
4. Failure of 3D software to process the images;
5. Only noncontrast CT images being available.

Our final cohort included 355 adults (78 male, 277 female; mean age, 59.7 ± 13.3 years; range, 21 to 90 years) who underwent standard CECT at 120 kVp within 100 days of a DXA.

DXA and CT acquisition

DXA of the proximal femur was performed using standard techniques according to International Society for Clinical Densitometry (ISCD) guidelines using GE Healthcare LunarProdigy densitometers (Madison, WI, USA). At least one valid femoral neck T-score was required for inclusion. Femoral neck T-scores were calculated using the U.S. National Health and Nutrition Examination Survey (NHANES) III reference database. The T-score was used to categorize each patient as normal, osteopenic, or osteoporotic according to the WHO classification.(12,16) Osteopenia is defined as a DXA T-score of between −1.0 and −2.4 and osteoporosis is defined as a T-score less than or equal to −2.5, with low BMD encompassing both osteoporosis and osteopenia. For the purposes of this study, individual patients were categorized according to their femoral neck T-score.

Multidetector CT (MDCT) scanning of the abdomen and pelvis was performed on multiple scanners (GE Healthcare, Waukesha, WI, USA) using a 120 kVp setting and variable mA, pelvis was performed on multiple scanners (GE Healthcare, similar to those used in Hologic DXA devices (Hologic Inc., Bedford, MA, USA) for proximal femur analysis (total hip, femoral neck [FN], intertrochanter, and trochanter) were identified automatically on the projected image by the software. The left hip was chosen for analysis in 351 cases, with the right hip used in 4 cases because of metallic hardware at the left hip. The automatically identified ROIs were visually checked to verify that the lower extent of the intetrochanter ROI was set at the lower junction of the lesser trochanter and the femoral shaft and that the FN axis and FN ROI position and size were appropriate. Results for aBMD (g/cm²) were reported in terms of equivalent calibrated aqueous potassium phosphate density and were stored in the QCT Pro database for export as text files. FN T-scores were calculated using the manufacturer’s CTXA reference database.

Data analyses

Linear regression analysis was used to compare CTXA and DXA BMD measurements and T-score results. Bland-Altman plots were used to test for trends in differences between BMD results once a linear mapping between CTXA and DXA had been applied. Sensitivity and specificity with 95% confidence intervals (CIs) was calculated for CTXA utilizing DXA as the reference standard. All statistics were performed with Microsoft Excel (Microsoft, Redmond, WA, USA).

Results

Mean aBMD on DXA was 0.888 ± 0.150 g/cm² and on CTXA it was 0.657 ± 0.119 g/cm². This correlated with a mean T-score of −1.07 ± 1.08 and −1.25 ± 1.08, respectively. Linear regression analysis showed good correlation between DXA and CTXA (R² = 0.825 for both BMD and T-scores) and the SD of the distribution of residuals (SEE) was 0.063 g/cm² or 0.45 T-score
units (Figs. 1, 2, and 3). A Bland-Altman plot of $T$-scores indicated no trend in differences between the two measurements and a small bias with DXA $T$-score +0.18 units higher than CTXA (Fig. 4).

There were 35 patients (9.9%) in the osteoporotic range on DXA and 162 patients (45.6%) in the osteopenic range. There were 48 patients (13.5%) in the osteoporotic range on CTXA and 173 patients (48.7%) in the osteopenic range. $T$-score categorization of normal, osteopenic, or osteoporotic differed between DXA and QCT in 63 (17.7%) patients with 77.8% (49/63) of these being patients who had a lower $T$-score on CTXA. Overall 10 patients with low BMD on DXA were categorized as normal BMD at CTXA, with the average DXA $T$-score in this group being $-1.23 \pm 0.11$ (range, $-1.0$ to $-1.4$).

Applying a correction of +0.18 $T$-score units to the CTXA $T$-score based on the bias noted on the Bland-Altman plot leads to a difference between DXA and QCT in 61 (17.2%) patients. Of these, 23 patients with low BMD on DXA would be categorized as normal BMD at CTXA, with the average $T$-score in this group being $-1.33 \pm 0.03$ (range, $-1.0$ to $-1.9$).

When using DXA as the gold standard, CTXA has a sensitivity of 94.9% (95% CI, 90.6% to 97.4%) and a specificity of 79.1% (95% CI, 71.8% to 85.0%) for discriminating low BMD from normal. Positive and negative predictive values are 85.6% (95% CI, 79.4% to 92.5%) and 92.5% (95% CI, 86.4% to 96.2%), respectively. For discriminating osteoporotic from nonosteoporotic patients CTXA has a sensitivity of 88.6% (95% CI, 72.3% to 96.3%) and specificity of 95.0% (95% CI, 91.8% to 97.0%) using DXA as the gold standard. Positive and negative predictive values are 65.9% (95% CI, 50.6% to 78.7%) and 98.7% (95% CI, 96.4% to 99.6%), respectively.

**Discussion**

$T$-score results from CTXA studies performed at the hip on CECT examinations showed high correlation with DXA $T$-scores obtained within 100 days, which could allow for expansion of the additive technique of CTXA for opportunistic screening, increasing overall screening rates. The SEE of 0.5 $T$-score units requires further consideration as to whether a correction factor should be used when categorizing risk. As these CTXA measurements were performed retrospectively, BMD measurement could be performed either prospectively or retrospectively.
when requested by a referring physician from nearly any CT scan that includes the pelvis.

Prior studies comparing CT aBMD with DXA noted good correlation with relatively little bias between the two modalities. Pickhardt and colleagues\textsuperscript{(13)} compared noncontrast CT colonoigraphy studies with DXA noting a 0.3 T-score bias, with DXA scores being higher, which is in line with the 0.18 T-score bias in the same direction on this larger study. That study, which only included females, showed an $r^2$ of 0.91 and SEE of 0.38 T-score units as compared with $r^2$ of 0.83 and SEE of 0.45 T-score units in this study. Weber and colleagues,\textsuperscript{(19)} when studying a younger patient population with CECT, showed a similar correlation with $r^2$ of 0.84 using a different software vendor. That study did not report a bias from a Bland-Altman plot, but did report average DXA T-score 0.18 higher than that calculated with CT, which is equivalent to the bias noted in this study. Our study is the largest series comparing DXA with CECT including the pelvis, which is a relatively commonly performed diagnostic examination in older adults for a wide variety of indications. Taking advantage of the BMD data available on these scans, which hitherto has largely been ignored, would have a positive impact on osteoporosis screening.

Based on the results of this study, adding 0.18 T-score units to CTXA values when reporting for low BMD screening could be considered, although in our cohort using this additive measure actually resulted in an increased number of patients with low BMD on DXA who would be classified in the normal range on CT. In addition, previous studies evaluating the precision of DXA have demonstrated a 0.4 T-score unit variability at the FN,\textsuperscript{(20)} similar to the SEE of 0.45 T-score units in this study. With the purpose of opportunistic screening being the identification of those with low BMD rather than exact quantification, we would suggest using uncorrected T-score values from these routine CT scans, whether performed with or without i.v. contrast. This is particularly justified because patients who are identified to have low BMD and increased fracture risk by CTXA and who might warrant therapy should have a DXA examination as baseline for future follow-up and treatment monitoring. This approach to opportunistic screening is perhaps best achieved by increasing sensitivity for low BMD at the price of specificity to capture patients who need the next step of workup. The sensitivity for low BMD of 94.9% in this study supports the use of CECT for opportunistic screening in this fashion. In addition, the negative predictive value for low BMD of 92.5% and for osteoporosis of 98.2% are particularly valuable in identifying patients who have not been screened and will require no further workup. The cost-effectiveness of this two-step screening process, in which patients with increased fracture risk on CTXA receive a follow-up DXA, is beyond the scope of this work, but would likely be very favorable assuming only a small additional charge would apply to the add-on BMD evaluation at CT. Regardless, there is a clear need for increased screening because of the number of eligible patients who are not currently being screened.\textsuperscript{(4,21)}

The implementation of an opportunistic screening program should lead to minimal change in the CT workflow.\textsuperscript{(22)} The service could be offered as either a prospective order, which could potentially be prompted by electronic medical records based on prior BMD evaluation and risk factors, or via an order following reporting of the CT scan. In the prospective paradigm the technologist could obtain the clinical risk factor information that allows for calculation of a 10-year fracture risk via the WHO-FRAX tool.\textsuperscript{(11)} The postprocessing can be performed with high reliability by dedicated CT technologists, who are adept at image postprocessing from other volume-rendering protocols, especially with an integrated quality control program.\textsuperscript{(23)} An additional benefit of this opportunistic approach is that it allows an initial screen with no additional patient time or radiation. As we look toward an era of accountable care organizations, this value-added approach also could lead to a more efficient utilization of existing resources.

An additional consideration when approaching the idea of opportunistic screening via QCT is the use of spine data. Multiple studies have reported using QCT measurements obtained from the spine for osteoporosis screening.\textsuperscript{(8,9,24)} Unfortunately, the variability with these measurements can be significant following contrast administration, reported as high as 30.3% in one study,\textsuperscript{(14)} favoring the use of hip data in the postcontrast setting. In addition, because the WHO-FRAX tool utilizes only FN data, the appropriate T-score and aBMD to calculate a 10-year fracture risk is obtained only with hip CTXA QCT.\textsuperscript{(11)} Meanwhile, spine data is reported in mg/cm$^3$ according to reporting standards for QCT without an equivalent fracture risk calculation available.\textsuperscript{(25)}
The asynchronous calibration method utilized in this study is unique because rather than use of internal reference standards, a phantom is scanned at a point in time remote from the acquisition of patient images. The ISCD position paper on QCT documents variability in HU values between scanners when evaluating synthetic bone samples. Given this variability it is imperative to perform calibration of each scanner at varying densities and varying table heights, a calibration that was performed in our study. Although the daily calibration of the CT scanner to a HU value of zero for water does not guarantee a correct BMD value over the time between phantom calibrations, which may limit applicability of this data, good correlation with acceptable noise limits have been reported for the asynchronous method when applied to the retrospective extraction of BMD estimates from CT images acquired approximately 5 years prior to obtaining the asynchronous calibration data. The observed SEE for this study similarly supports the notion that the unknown variability in CT scanner performance and drift between the acquisition of CT data and asynchronous calibration data was not sufficient to degrade BMD measurement quality to the extent that it severely compromises the utility of this method for opportunistic screening.

The limitations of the study include the use of asynchronous scanner calibration, although daily CT calibration scans were performed. More routine phantom calibration could be performed in a future, prospective study to confirm the reproducibility of these results. There is currently limited data to validate asynchronous, contrast-enhanced CTA and the results of this study are not robust enough to advocate replacement of DXA screening with CTA performed on postcontrast CT. There was no collection of clinical fracture risk factors, which when combined with BMD to calculate a fracture risk is a potential area for future study. Only 120 kVp scans were evaluated and there is a need for data from scans performed at 140 kVp as well. This study also does not evaluate the clinical or cost-effectiveness of reporting BMD during CECT and this would be a useful area for future study.

In conclusion, utilizing CECT to calculate CTA BMD values results in a small T-score discrepancy that may allow identification of patients needing further evaluation with dedicated DXA prior to management decisions. Prospective evaluation of this technique is warranted to evaluate correlation of obtained values with fracture risk. This postprocessing step from routine CT obtained for any indication could improve osteoporosis screening because it can be applied in a prospective or retrospective fashion.

Disclosures

ADB is an employee of Mindways Software, Inc. JKB is a stockholder and employee of Mindways Software, Inc. All other authors state that they have no conflicts of interest.

Acknowledgments

Authors’ roles: Study design: TJZ, AB, JKB, and PJP. Study conduct: TJZ, AM, and PJP. Data collection: TJZ, AM, and PJP. Data analysis: TJZ, AM, ADB, and PJP. Data interpretation: TJZ, NB, ADB, and PJP. Drafting manuscript: TJZ and PJP. Revising manuscript content: AM, NB, ADB, JKB. Approving final version of manuscript: TJZ, AM, NB, ADB, JKB, and PJP. TJZ takes responsibility for the integrity of the data analysis.

References


