Clin Orthop Relat Res (2019) 477:850-860
DOI 10.1097/CORR.0000000000000480

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

2017 Selected Proceedings of SOMOS

Proximal Femur Hounsfield Units on CT Colonoscopy Correlate With Dual-energy X-ray Absorptiometry

Daniel L. Christensen MD, Kyle E. Nappo MD, Jared A. Wolfe MD, Sean M. Wade MD, Daniel I. Brooks PhD, Benjamin K. Potter MD, Jonathan A. Forsberg MD, PhD, Scott M. Tintle MD

Received: 19 March 2018 / Accepted: 16 August 2018 / Published online: 4 September 2018

Abstract

Background Quantifying bone mineral density (BMD) on CT using commercial software demonstrates good-to-excellent correlations with dual-energy x-ray absorptiometry (DEXA) results. However, previous techniques to measure Hounsfield units (HUs) within the proximal femur demonstrate less successful correlation with DEXA results. An effective method of measuring HUs of the proximal femur from CT colonoscopy might allow for opportunistic osteoporosis screening.

Questions/purposes (1) Do proximal femur HU measurements from CT colonoscopy correlate with proximal femur DEXA results? (2) How effective is our single HU measurement technique in estimating the likelihood of overall low BMD? (3) Does the relationship between our comprehensive HU measurement and DEXA results change based on age, sex, or time between studies?

Methods This retrospective study investigated the measurement of HU of the femur obtained on CT colonoscopy studies compared with DEXA results. Between 2010 and 2017, five centers performed 9085 CT colonoscopy studies; of those, 277 (3%) also had available DEXA results and were included in this study, whereas 8809 (97%) were

One of the authors certifies that he (JAF) has received personal fees in an amount of less than USD 10,000 from The Solsidian Group LLC (Kensington, MD, USA), grants in an amount of USD 10,000 to USD 100,000 from Zimmer Biomet Inc (Warsaw, IN, USA), and personal fees in an amount of less than USD 10,000 from Clementia Pharmaceuticals Inc (Montreal, QC, Canada), all outside the submitted work. One of the authors certifies that he (SMT) has received grants in the amount of USD 750,000 from the Congressionally Directed Medical Research Program (CDMRP; Fort Detrick, MD, USA), outside the submitted work. One of the authors certifies that he (BKP) has received grants in the amount of more than USD 1,000,001 from both BUMED Wounded Ill and Injured (Falls Church, VA, USA) and CDMRP (Fort Detrick, MD, USA), outside the submitted work.

Clinical Orthopaedics and Related Research® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA approval status, of any drug or device before clinical use.

Each author certifies that his institution approved the human protocol for this investigation and that all investigations were conducted in conformity with ethical principles of research.

This work was performed at Walter Reed National Military Medical Center, Bethesda, MD, USA.

The views expressed in this article are those of the authors and do not reflect the official policy of the Department of the Army/Navy/Air Force, Department of Defense, or the US Government.

D. L. Christensen, K. E. Nappo, J. A. Wolfe, S. Wade, B. K. Potter, J. A. Forsberg, S. M. Tintle, Walter Reed National Military Medical Center, Uniformed Services University-Walter Reed Department of Surgery Orthopaedic Service, Bethesda, MD, USA

D. I. Brooks, Walter Reed National Military Medical Center, Department of Research Programs, Bethesda, MD, USA

S. M. Tintle (✉), Walter Reed National Military Medical Center, Uniformed Services University-Walter Reed Department of Surgery Orthopaedic Service 8901 Rockville Pike, Bethesda, MD 20889, USA, email: scott.tintle@gmail.com

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Wolters Kluwer

excluded for inadequate CT imaging, lack of DEXA screening, or lack of proximal femur DEXA results. The median number of days between CT colonoscopy and DEXA scan was 595 days; no patient was excluded based on time between scans because bone remodeling is a long-term process and this allowed subgroup analysis based on time between scans. Two reviewers performed HU measurements at four points within the proximal femur on the CT colonoscopy imaging and intraclass correlation coefficients were used to evaluate interrater reliability. We used Pearson correlation coefficients to compare the comprehensive (average of eight measurements) and a single HU measurement with each DEXA result—proximal femur BMD, proximal femur T-score, femoral neck BMD, and femoral neck T-score—to identify the best measurement technique within this study. Based on their lowest DEXA T-score, we stratified patients to a diagnosis of osteoporosis, osteopenia, or normal BMD. We then calculated the area under the receiver operator characteristic curves (AUCs) to evaluate the classification ability of a single HU value to identify possible threshold(s) for detecting low BMD. For each subgroup analysis, we calculated Pearson correlation coefficients between DEXA and HUs and evaluated each subgroup's contribution to the overall predictive model using an interaction test in a linear regression model.

Results The Pearson correlation coefficient between both the comprehensive and single HU measurements was highest compared with the proximal femur T-score at 0.75 (95% confidence interval [CI], 0.69–0.80) and 0.74 (95% CI, 0.68–0.79), respectively. Interobserver reliability, measured with intraclass correlation coefficients, for the comprehensive and single HU measurements was 0.97 (95% CI, 0.72–0.99) and 0.96 (95% CI, 0.89–0.98), respectively. Based on DEXA results, 20 patients were osteoporotic, 167 had osteopenia, and 90 patients had normal BMD. The mean comprehensive HU for patients with osteoporosis was 70 ± 30 HUs; for patients with osteopenia, it was 110 ± 36 HUs; and for patients with normal BMD, it was 158 ± 43 HUs ($p < 0.001$). The AUC of the single HU model was 0.82 (95% CI, 0.77–0.87). A threshold of 214 HUs is 100% sensitive and 59 HUs is 100% specific to identify low BMD; a threshold of 113 HUs provided 73% sensitivity and 76% specificity. When stratified by decade age groups, each decade age group demonstrated a positive correlation between the comprehensive HU and proximal femur T-score, ranging between 0.71 and 0.83 (95% CI, 0.59-0.91). Further subgroup analysis similarly demonstrated a positive correlation between the comprehensive HU and proximal femur T-score when stratified by > 6 months or < 6 months between CT and DEXA (0.75; 95% CI, 0.62-0.84) as well as when stratified by sex (0.70-0.76; 95% CI, 0.48-0.81). The linear regression model demonstrated that the overall positive correlation coefficient between HUs and the proximal femur T-score is not influenced by any subgroup.

Conclusions Our measurement technique provides a reproducible measurement of HUs within the proximal femur HUs on CT colonoscopy. Hounsfield units of the proximal femur based on this technique can predict low BMD. These CT scans are frequently performed before initial DEXA scans are done and therefore may lead to earlier recognition of low BMD. Future research is needed to validate these results in larger studies and to determine if these results can anticipate future fracture risk.

Level of Evidence Level III, diagnostic study.

Introduction

Low bone mineral density (BMD) is a silently growing epidemic, estimated to impact 71 million people by 2030 and affect more than half of the US population > 50 years old [42]. According to the World Health Organization, a diagnosis of osteoporosis and osteopenia, or low BMD, is made based on dual-energy x-ray absorptiometry (DEXA) scanning [17, 18]. The National Osteoporosis Foundation recommends that in otherwise healthy individuals, DEXA scanning should begin at age 65 years in women and age 70 years in men [9]. Consequently, as many as 25% of all osteoporotic fractures will occur before the first DEXA scan [6]. Given the increased rates of disability and mortality after osteoporotic hip fractures [1, 4, 7, 8, 19, 23], there is growing interest in using alternative screening techniques such as CT for opportunistic osteoporosis screening [12]. Opportunistic osteoporosis screening utilizes biomedical screening tests (radiology, laboratory, etc) obtained for other purposes to screen for patients at risk for osteoporosis and may indicate a patient is at risk, but it is not intended to replace the current gold standard, DEXA screening.

CT images can also be used to calculate quantitative BMD, which highly correlates ($R^2 > 0.9$) with DEXA results [11, 13, 28-30, 41, 44-46], but these calculations often require the use of additional proprietary software that is not available to all clinicians. In contrast, Hounsfield units (HUs) are a measure of attenuation on CT imaging that are available on most imaging platforms, and although they are less accurate than the software-produced BMD values, there is a strong relationship between HU of the lumbar spine and upper extremity when compared with DEXA results [5, 16, 21, 31, 32, 34, 35, 40]. However, HU measurements of the proximal femur are less common, use differing methods, and produce less impressive results [22, 26]. This may be secondary to ambiguity regarding the various techniques and locations to measure (such as the femoral head, femoral neck, trochanteric region, or a combination), accounting for the wide intra- and interrater

reliabilities (intraclass correlation coefficient [ICC], 0.5-0.9) and lower correlations (r = 0.4-0.5) between the femoral head, femoral neck, and greater trochanter HU and DEXA reported by Lee et al. [22].

To our knowledge, no study has described or evaluated a comprehensive measurement of proximal femur cancellous bone using HUs. Using HUs is practical and increases the generalizability of the measurements, incorporates the maximal amount of cancellous bone in the proximal femur at risk for fracture, and may be useful until proprietary quantitative CT software is readily available in all clinical settings. We created a technique to comprehensively measure proximal femur HUs and studied whether this technique would correlate with DEXA results and predict BMD status. We chose CT colonoscopy, an accepted screening method for colorectal cancer, over pelvis or hip CT because it is used in the same age demographic that is at risk for osteoporosis [24, 37], and images would be obtained with one imaging protocol, minimizing the variability associated with the technique used such as slice thickness, image interval, voltage used, etc. CT colonoscopy uses only a limited amount of oral contrast, limiting the effect of contrast on the proximal femur while also providing a standardized set of images for research purposes [29, 46]. A successful technique, especially if validated in other CT scans, would allow for a gross estimation of bone health for the millions of CT scans performed annually [3], providing patients who do not get DEXA screening a reasonable estimate of their bone health and stimulating potential osteoporosis treatment.

We therefore asked the following questions: (1) Do proximal femur HU measurements from CT colonoscopy correlate with proximal femur DEXA results? (2) How effective is our single HU measurement technique in estimating the likelihood of overall low BMD? (3) Does the relationship between our comprehensive HU measurement and DEXA change based on age, sex, or time between studies?

Patients and Methods

After institutional review board approval, we performed this retrospective study using the Military Health System Management Analysis and Reporting Tool (M2). We used Current Procedural Terminology codes within M2 and reviewed the electronic medical record to identify patients who underwent CT colonoscopy at military treatment facilities within the national capital region between 2010 and 2017 and then identified those who also underwent a DEXA scan. The electronic medical record was reviewed to obtain demographics including sex, age, body mass index (BMI), and DEXA results.

Between 2010 and 2017, five centers performed 9085 CT colonoscopy studies and of those, 277 (3%) patients had adequate CT colonoscopy imaging and DEXA screening. We excluded 8809 (97%) patients because they lacked adequate CT imaging available for review or did not have femoral neck or proximal femur (also referred to in some reports as "total femur," comprised of femoral neck, trochanteric and intertrochanteric regions) DEXA results. Additionally, if our results are to translate to all patients getting CTs, this highlights the potential usefulness as an opportunistic screening test in that 8809 patients had regular screening for colon cancer but no BMD screening. Twenty patients had a diagnosis of osteoporosis, 167 had osteopenia, and 90 patients had normal BMD (Table 1). We identified differences wherein patients with osteoporosis were older, had lower BMIs, and lower HUs on CT colonoscopy. The mean comprehensive HU measurement for patients with osteoporosis (70), osteopenia (110), and normal BMD (158) was statistically different. The median number of days between CT colonoscopy and DEXA scan was 595 days (interquartile range, 185–1076); 69 patients had < 6 months between scans. We included all patients regardless of time between CT colonoscopy and DEXA because bone remodeling is a long-term process and this allowed subgroup analysis based on time between scans. We stratified patients according to the National Osteoporosis Foundation guidelines based on the lowest T-score within the lumbar spine, proximal femur, or femoral neck as osteoporosis (T-score ≤ -2.5), osteopenia (-2.5 < T-score ≤ -1), or normal BMD (T-score > -1) [9]. Additionally, the categories of osteoporosis and osteopenia were combined into a low BMD category, defined as a T-score ≤ -1.0, for subsequent analysis. We did this because fracture risk is the main clinical outcome that providers are interested in and most fractures occur in patients with osteopenia [33, 36].

All CT colonoscopies were performed at one of five regional institutions obtaining both supine and prone axial imaging; approximately two-thirds of images included peak kilovoltage (kVp) values and slice thickness with all but one reporting 120 kVp and 1.5-mm slices reconstructed to 1-mm intervals. We developed a technique to measure HUs of the proximal femur to capture the maximal cancellous bone within the femoral neck and intertrochanteric region. Beginning superiorly and scrolling inferiorly, the femoral head is identified followed by the superior aspect of the femoral neck cortex. The first measurement is taken when the axial image displayed shows the femoral neck connecting to the greater trochanter, is no longer within the superior femoral neck cortex, and consists of predominantly cancellous bone; this is typically one or two image slices (1.5–3 mm) below the first appearance of the femoral neck cortex (Fig. 1A). The measurement consists of a rectangular area of cancellous bone immediately lateral to the femoral physeal scar and that includes the maximal amount of cancellous bone while avoiding cortical bone and the femoral head epiphyseal scar. After this, three

Table 1. Patient demographics

Patient information	Overall (n = 277)	Osteoporosis (n = 20)	Osteopenia (n = 167)	Normal BMD (n = 90)	Difference between osteoporosis versus osteopenia	Difference between osteoporosis versus normal BMD	Difference between osteopenia versus normal BMD
Age (years), mean (SD)	63 ± 8	66 ± 10	63 ± 8	61 ± 8	2 years older (95% CI, -2 to 7); p = 0.308	5 years older (95% CI, 0-10); p = 0.052	2 years older (95% CI, 0-4); p = 0.023
Sex (M/F)	41/236	3/17	19/148	19/71			
BMI (kg/m^2), mean (SD),	27 ± 6	24 ± 4	25 ± 5	30 ± 7	2 units lighter (95% CI, -4 to 1); p = 0.1214	6 units lighter (95% CI, 4-9); p < 0.001	5 units lighter (95% CI, 3-6); p < 0.001
Hounsfield units comprehensive measurement, mean (SD)	123 ± 46	70 ± 30	110 ± 36	158 ± 43	40 HU lower (95% CI, 25-55); p < 0.001	88 HU lower (95% CI, 72-105); p < 0.001	48 HU lower (95% CI, 37-58); p < 0.001
Hounsfield units single 12-mm measurement, mean (SD)	110 ± 46	56 ± 29	98 ± 37	144 ± 41	42 HU lower (95% CI, 27-56); p < 0.001	88 HU lower (95% CI, 73-104); p < 0.001	47 HU lower (95% CI, 36-57); p < 0.001

BMD = bone mineral density; CI = confidence interval; M/F = male/female; BMI = body mass index.

additional measurements are taken at 6-mm intervals corresponding to a total of four axial slices. Four measurements, labeled 0 mm, 6 mm, 12 mm, and 18 mm and termed "single measurements," are taken and averaged into a final measurement, which we defined as the "comprehensive HU measurement" (Fig. 1 A-D). The correlation of single HU measurements to DEXA results was compared with the comprehensive HU measurement correlations. We added this analysis because we believed that, although the comprehensive measurement might demonstrate a stronger correlation with DEXA results, a single HU measurement may increase the likelihood that this or similar techniques will be widely accepted and adopted by radiologists and ordering providers alike.

Hounsfield Unit Measurement Technique

We used Pearson correlation coefficients to evaluate the relationship between each of the four single measurements and the comprehensive HU measurement with each DEXA score associated with the hip: proximal femur BMD and T-score and femoral neck BMD and T-score. Pearson correlation coefficient was chosen over ICCs in this comparison because HU and T-scores are on two different scales. We chose to evaluate all of these because our HU measurement technique incorporates the femoral neck and the intertrochanteric region; thus, our measurement technique is similar to, but not strictly the same, as either femoral neck or proximal femur DEXA results. The DEXA result with the highest correlation with the comprehensive HU measurement was used to evaluate an acceptable single HU measurement and for the planned subgroup analysis. We used ICCs, using two-way, agreement of single measures, to evaluate the agreement between the individual HU measurements and our planned single HU measurement with the comprehensive HU measurement.

Identifying Overall Low Bone Mineral Density

We calculated the area under the receiver operating characteristic curves (AUCs) using the single best HU measurement alone identified on correlation analysis. The sensitivity, specificity, and positive and negative predictive values for identifying low BMD were calculated: we selected three thresholds to (1) maximize sensitivity; (2) maximize specificity; and (3) obtain a cutoff balanced between each [27, 38].

Statistical Analysis

Two reviewers (KEN, JAW) independently performed the HU measurement technique on each patient without

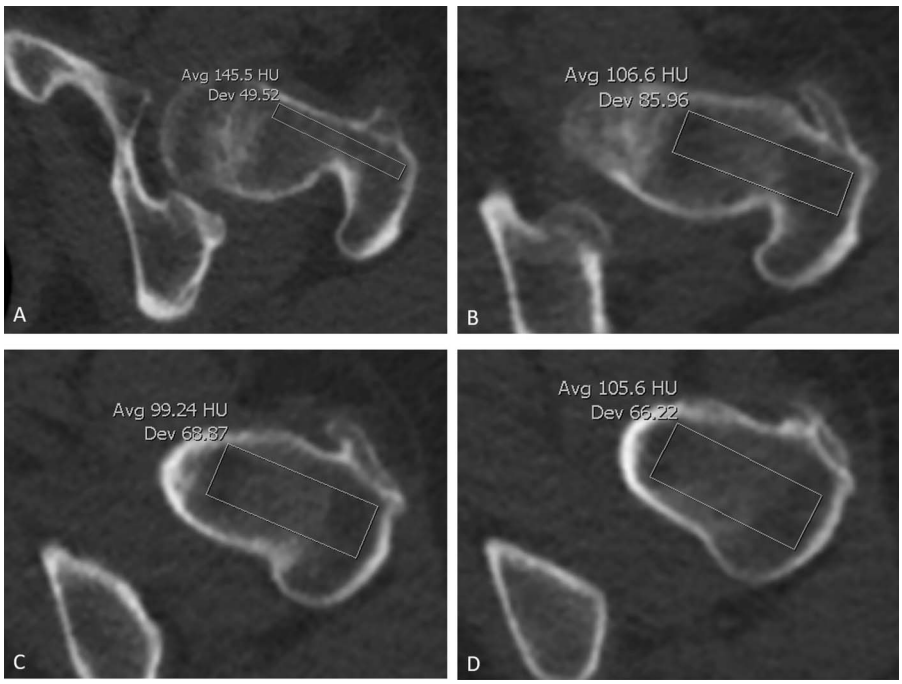

Fig. 1 A-D These are representative images of the four HU measurements: (**A**) The first and most superior measurement ("0-mm measurement"); (**B**) the second ("6-mm measurement"); (**C**) the third ("12-mm measurement"); and (**D**) the last and most distal measurement ("18-mm measurement"). Avg = average; Dev = deviation.

knowing the DEXA results; for two-way agreement of average measures, ICC was used to evaluate interrater reliability between each single measurement and the comprehensive measurement for each reviewer and results between reviewers were averaged together for all subsequent analyses [14]. We graded ICCs on a scale of excellent (ICC > 0.9), good (0.75 ≤ ICC < 0.9), moderate (0.5 ≤ ICC < 0.75), or poor (ICC < 0.5) [20]. Categorical data were compared using chi-square and ordinal data were compared with analysis of variance, presented as means with SDs and 95% confidence intervals (CIs). As HUs decrease with age, we sought to evaluate if our results remained consistent across age groups when stratified by decade of life [34]. In addition, we stratified our patients based on the time between CT colonoscopy and DEXA testing using a 6-month cutoff, selected as the point before which clinically relevant bone remodeling resulting from age or bisphosphonate treatment might alter DEXA or HUs results [2, 10, 25]. Finally, subgroup comparison by sex was performed. For each subgroup analysis, we calculated Pearson correlation coefficients between DEXA and HUs and evaluated each subgroup's contribution to the overall predictive model using an interaction test in a linear regression model. Finally, we performed decision curve analysis (DCA) to quantify the clinical utility of the HU measurement for two clinical end points: first, identifying patients who would benefit from DEXA screening and, second, identifying patients most likely to benefit from osteoporosis treatment [39]. All statistical analysis was performed using two-tailed testing with significance set at 0.05; we used the following packages in RStudio (Boston, MA, USA, version: 1.0.143): OptimalCutpoints, pROC, DecisionCurve, e1071, irr, and psych.

Results

The ideal HU measurement technique in our results is a region of interest consisting of cancellous bone within the proximal femur approximately 12 mm inferior to the superior femoral neck on axial imaging. All comprehensive and each single HU measurement (0 mm, 6 mm, 12 mm, and 18 mm) revealed a positive and linear correlation with each of the proximal femur DEXA results (Fig. 2A). The proximal femur T-score demonstrated the highest correlation with the comprehensive measurement and the 12-mm measurement displayed the highest correlation of all individual measurements (Table 2). The overall ICC of all four individual HU measurements was 0.764 (95% CI, 0.41–0.886) and the ICC between the individual 12-mm measurement and the comprehensive HU measurement was 0.942 (95% CI, 0.368–0.983) (Fig. 2B). Interobserver reliability was moderate to excellent between raters for both the comprehensive HU

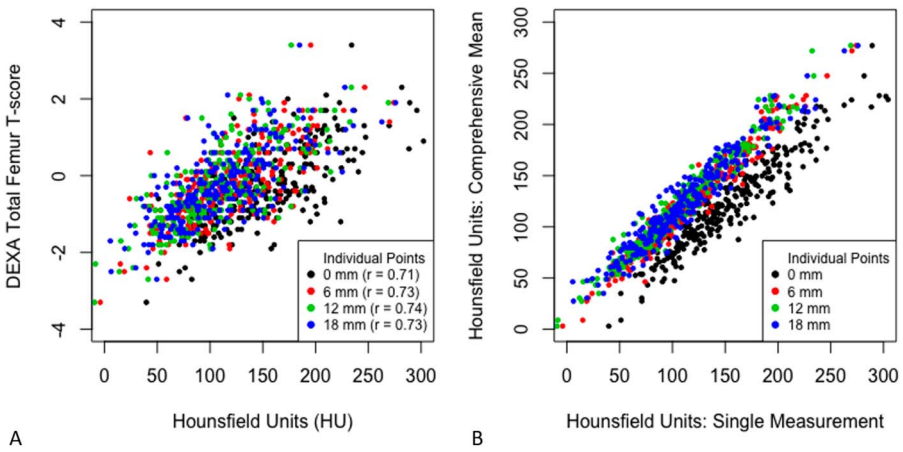

Fig. 2 A-B (**A**) This plot shows each individual HU measurement point versus DEXA proximal femur T-score with the Pearson correlation coefficient for each individual measurement in the legend. (**B**) This plot shows individual HU measurements versus the comprehensive (average of the four together) HU measurement.

measurement and the single 12-mm measurement with an ICC of 0.97 (95% CI, 0.72–0.99) and 0.96 (95% CI, 0.89–0.98), respectively.

The use of HU provides a moderately reliable classification of patients into low versus normal BMD. The AUC was 0.82 (95% CI, 0.77–0.87) when using the 12-mm HU measurement only (Fig. 3A). In addition, using the 12-mm measurement, a threshold of 214 HUs is 100% sensitive (95% CI, 98%-100% with specificity of 4%, 95% CI, 1%-11%), 59 HUs is 100% specific (95% CI, 96%-100% with sensitivity of 19%, 95% CI, 13%-25%), and 113 HUs provided 73% sensitivity (95% CI, 66%-79%) and 76% specificity (95% CI, 65%-84%) to identify low BMD (Table 3).

The overall relationships identified in this study, between HU and DEXA results, remained consistent when broken down by age, sex, or time between scans. The Pearson correlation coefficients between the comprehensive HU and DEXA proximal femur T-score, when stratified by any of the previously mentioned subgroups, are all consistent with our overall results (Table 4). When separate linear regression models were used to evaluate the relationship between DEXA proximal femur T-score and HU, while including subgroup as an additional factor, those models consistently identified the comprehensive HU (decade age group: $F[1,245] = 38.96$; sex: $F[1,251] = 297.50$; time between scans: $F[1,251] = 256.96$; $p < 0.001$ for all three models) as a strong predictor of DEXA proximal femur T-score. However, the subgroups as well as the interaction between HU and each subgroup were not major predictors of DEXA proximal femur T-score. Therefore, analysis indicates that the correlation coefficients between DEXA proximal femur T-score and HU reported do not differ based on subgroup (Table 4).

The DCA identified a positive net benefit to intervene for low BMD using the single 12-mm measurement at all risk thresholds above 0.23 (Fig. 3B), indicating that HU could be used clinically.

Table 2. Pearson correlation coefficients with 95% confidence intervals for all proximal femur DEXA results

Pearson correlation coefficients	Total femur BMD	Total femur T-score	Femoral neck BMD	Femoral neck T-score
Overall comprehensive measurement (95% CI)	0.73 (0.67-0.79)	**0.75 (0.69-0.80)**	0.727 (0.67-0.78)	0.741 (0.68-0.79)
Single 0-mm measurement (95% CI)	0.71 (0.64-0.77)	0.71 (0.64-0.76)	0.71 (0.65-0.76)	0.72 (0.66-0.77)
Single 6-mm measurement (95% CI)	0.712 (0.65-0.77)	0.73 (0.66-0.78)	0.72 (0.66-0.77)	0.73 (0.67-0.78)
Single 12-mm measurement (95% CI)	0.71 (0.65-0.77)	**0.74 (0.68-0.79)**	0.71 (0.64-0.76)	0.72 (0.66-0.78)
Single 18-mm measurement (95% CI)	0.69 (0.62-0.75)	0.73 (0.67-0.78)	0.67 (0.60-0.73)	0.70 (0.63-0.75)

Bolded values represent the highest Pearson correlation coefficients on which subsequent analysis in the paper was performed; the total femur T-score was used for comparison with Hounsfield unit results and the 12-mm measurement selected as the single measurement that most closely resembles the overall comprehensive measurement; DEXA = dual-energy x-ray absorptiometry; BMD = bone mineral density.

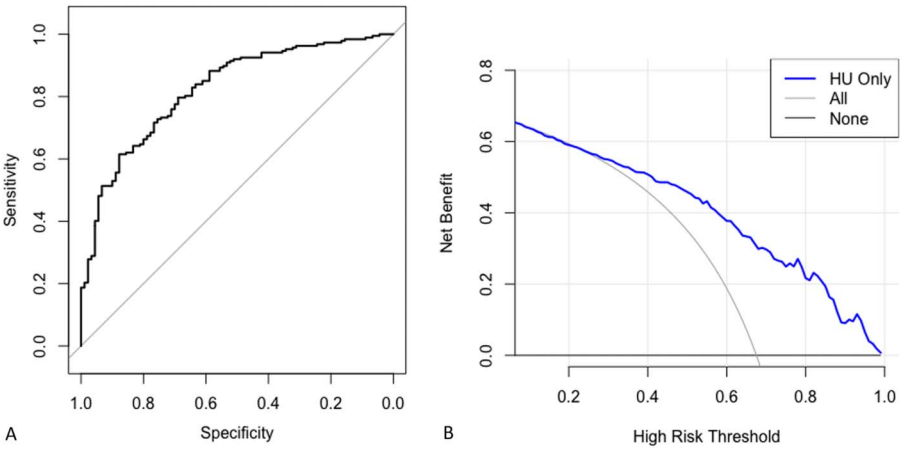

Fig. 3 A-B (**A**) This figure displays the receiver operating characteristic curve using the 12-mm HU measurement to evaluate for low BMD. (**B**) This figure displays the DCA plot for the 12-mm HU measurement model. The horizontal black line represents intervening on no patients with zero net benefit; the gray line represents intervening on all patients and the blue line represents intervening based on the HU model.

Discussion

As the number of Americans with low BMD increases, the relative ease of opportunistic screening (osteoporosis screening using imaging data obtained for other purposes) can serve to engage both patients and providers with earlier osteoporosis care with the goal of ultimately decreasing the number of fragility fractures and their associated morbidity and mortality [29, 44]. Because colorectal cancer screening likely occurs at higher rates than osteoporosis screening [15], our study of HUs to quantify BMD will enable nearly all clinicians with access to digital imaging software to use CT imaging for osteoporosis screening. Hounsfield units of the proximal femur on CT colonoscopy demonstrate a positive and linear correlation with DEXA results regardless of age or time between scans. Obtaining a single 12-mm measurement of cancellous bone within the proximal femur on CT colonoscopy takes 30 seconds, provides a very similar correlation to the comprehensive measurement (four measurements averaged together), and can be used to approximate DEXA results and thus clinical BMD classification.

There are several limitations of our study. Perhaps the most prominent is the retrospective nature and associated selection bias in the types of patients who underwent both a CT colonoscopy and DEXA scan. Although advantageous in that this may focus on the group at greatest risk (ie, patients in whom clinical concern prompts DEXA screening), the downside is potentially zeroing in too closely on a select subset of patients at higher risk for osteoporosis while limiting our results and knowledge of what the HU trends are in all patients who undergo CT colonoscopy. Furthermore, we identified differences in patient demographics, including BMI, when stratified by World Health Organization classification for BMD. Because BMI is known to influence DEXA results, it is unclear what influence this has on HU results, which may be a source of variability in our results if BMI affects HU at a differential rate relative to DEXA results. The limited number of patients, particularly those with osteoporosis, and relative short-term followup preclude any meaningful evaluation of long-term fracture risk. Use of an averaged HU measurement between two raters may decrease the variability seen by one clinician who obtains

Table 3. Hounsfield unit cutoffs based on a single 12-mm measurement

Method	HUs	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
Youden	101	62% (54-69)	88 % (79-94)	91% (85-93)	52% (45-70)
ROC (0,1)*	113	73% (66-79)	76% (65-84)	86% (79-90)	57% (49-69)
100% sensitive	214	100% (98-100)	4% (1-11)	69% (37-100)	100% (52-100)
100% specific	59	19% (13-25)	100% (96-100)	100% (90-100)	37% (28-100)

*ROC (0,1) represents the point on the ROC curve that comes closest to the Cartesian coordinate of (0,1); HUs = Hounsfield units; CI = confidence interval; ROC = receiver operating characteristic.

Table 4. Hounsfield unit and Pearson correlation subgroup analysis based on age and time between scans

Subgroup	Osteoporosis, mean ± SD	Osteopenia, mean ± SD	Normal BMD, mean ± SD	Comprehensive measurement versus total femur T-score correlation
Age younger than 50 years (n = 8)		126 ± 23	164 ± 51	0.87 (0.32-0.98)
Age 50-59 years (n = 102)	86 ± 29	123 ± 37	1612 ± 40	0.71 (0.59-0.80)
Age 60-69 years (n = 118)	78 ± 20	109 ± 32	146 ± 35	0.71 (0.61-0.79)
Age 70-79 years (n = 43)	42 ± 27	90 ± 35	184 ± 57	0.83 (0.70-0.91)
Age 80+ years (n = 6)	69 (n = 1)	72 ± 28	75 (n = 1)	0.69 (0.00-0.96)
Time < 6 months (n = 69)	80 ± 30	103 ± 27	150 ± 44	0.75 (0.62-0.84)
Time > 6 months (n = 208)	60 ± 29	112 ± 39	160 ± 42	0.75 (0.68-0.80)
Male	71 ± 10	103 ± 33	150 ± 41	0.70 (0.48-0.83)
Female	70 ± 33	111 ± 36	160 ± 43	0.76 (0.70-0.81)

BMD = bone mineral density.

one HU measurement in practice. Although averaging our measurements for statistical comparisons likely does decrease some variability, this effect is likely minimal because our ICC method used the stricter method of testing, evaluating not just for consistent results, but agreement between reviewers and with this, the results were still very high. The lack of a reference or phantom to calibrate HU measurements such as that used with calibrated CT-based BMD methods adds potential variability between scanners and settings as does the variability in both the CT and DEXA machines used. However, despite these limitations, our results remain consistent with similar studies of quantitative BMD and HU measurements in other body regions, which is reassuring that this technique reliably measures proximal femur BMD. We acknowledge that HU is not as accurate as calibrated, quantitative BMD calculations; however, for those without the necessary software, it provides a reasonable alternative. Although our results showed moderate-to-good correlations and a reasonable AUC, there is still a fair amount of unexplained variance in DEXA results, as predicted by HU, and this likely represents some of the differences seen between our results and those using quantitative BMD software. Although this study used CT colonoscopy, the use of any pelvic CT scan for opportunistic screening could likely provide similar HU results, although further research is needed to validate this claim. Finally, the preponderance of women in the study may limit the applicability to men; this should not affect our findings from a radiation physics perspective; however, we lacked an adequate number of male patients to definitively evaluate this.

Proximal Femur HUs Correlate With DEXA

Our recommended HU measurement technique from our results was a region of interest consisting of cancellous bone within the proximal femur approximately 12 mm inferior to the superior femoral neck on axial imaging. Previous studies of proximal femur HU do not provide detailed descriptions of their measurement technique nor have strong correlations between HU and DEXA results been reported [22, 26]. Our results demonstrate that both the comprehensive HU measurement technique and a single measurement provide moderate-to-good correlations with DEXA results. Identification of a single measurement was driven based on the desire to remove complexity and create a simple and effective screening technique. The high correlations between the single point and comprehensive measurement as well as high interobserver ICC are reassuring that this is a reproducible technique, perhaps even amenable to automation. We believe this to be a viable and clinically applicable option until the widespread use of potentially automated quantitative BMD calculations can be widely implemented on all CT imaging. In our technique, cortical bone was excluded to minimize artifact and facilitate reproducibility. Although this is consistent with many studies referenced, both cortical and cancellous bone mass contribute to bone strength and fracture prevention, and future research will dictate if there is a need to develop techniques to quantify cortical bone HU [43].

HU Measurements Effective in Estimating Overall Low BMD

The use of HU provides a moderately reliable classification of patients into low versus normal BMD. The trend in HU levels among patients with osteoporosis, osteopenia, and normal BMD in this study is consistent with published reports of the proximal femur and spine [5, 16, 21, 22, 26, 32, 34, 40]. Additionally, our correlation between HU and DEXA is consistent with those demonstrated between

DEXA and quantitative CT-based BMD calculations of the proximal femur, although there is an increased level of variance in our results [11, 13, 30, 41, 46]. The consistency of our results with published studies reinforces the belief that our HU measurement effectively estimates many patients with low BMD status. Although using a reference standard and calibration a priori of CT imaging enhances the validity of BMD calculations, a single HU measurement technique, in contrast, is readily available and provides results consistent with more complex CT-based BMD calculations. Although not ideal, this technique may potentially provide an additional data point for the millions of patients who may undergo CT scans but not DEXA screening. More importantly, however, is determining if this technique can be used for fracture prediction. DEXA is not a perfect predictor of fracture either nor is it perfect for dictating osteoporosis treatment. Fracture prevention treatments are best made, like many situations in medicine, by incorporating the clinical history, physical examination, laboratory data, and/or imaging data. This adds a valuable data point not previously available.

HU and DEXA Relationship Unchanged Based on Age, Sex, and Time Between Studies

The overall relationships between HU and DEXA results identified in this study remained consistent when categorized by age, sex, or time between scans. Although HUs are known to decrease with age, when compared among patients with osteoporosis, osteopenia, and normal BMD, the HU trends between groups as well as the correlation between HU and DEXA remained consistent with our overall results after age group stratification [34]. This is notable considering the statistical differences in age identified between BMD categories. The higher HU level identified in the patients in their 70s who had normal BMD is inconsistent with the remainder of our results and published evidence; future research is needed to understand this inconsistency. The time interval reported between imaging studies in previous reports varied or was not reported; we included all results regardless of time interval to improve our study generalizability [11, 13, 21, 30, 34, 41, 43]. However, when stratified by a 6-month time interval between scans, the results were consistent with our overall results. The 6-month time interval was chosen based on the idea that bone remodeling is a long-term process; this timeframe is described as a 200-day process for cancellous bone remodeling at the cellular level, whereas remodeling of all cancellous bone takes between 2 and 4 years [2, 10, 25]. Shorter time intervals may be ideal to reduce confounding between results; however, longer intervals are likely acceptable for comparisons based on our findings.

The use of DCA facilitates clinical interpretation of the model by quantifying potential benefits compared with two extremes, treating either all or none of the patients [39]. When evaluating results and deciding whether to intervene, clinicians must select their own individual "threshold" on the model for the intervention in question, considering the morbidity and mortality associated with treating all or none. Decision analysis provides guidance for what model is best for each clinician, especially when clinicians differ on what threshold is appropriate for an intervention. Clinical application of the DCA requires providers to identify the intervention (eg, DEXA screening, initiation of bisphosphonate treatment, etc) and the threshold or probability that a patient is at risk for low BMD, at which point the provider would initiate their chosen intervention. For example, an individual provider may choose to start a 70-year-old woman on bisphosphonate treatment at thresholds above 40%, whereas the provider may prefer a higher threshold such as 90% for the same woman at age 50 years. For all thresholds, use of the HU model in this study is as good or better than treating every patient. At the arbitrary threshold of 60%, for every 100 patients, intervening based on the HU model compared with intervening in all patients avoids 13 unnecessary interventions on those with normal BMD without missing any patients with low BMD. Compared with intervening on no patients at the same 60% threshold, use of the model results in intervening in 38 patients with low BMD per 100 without any unnecessary interventions.

Our measurement technique provides an accurate and reproducible measurement of proximal femur HU on CT colonoscopy that correlates with central BMD classification. These results may identify patients at risk of low BMD who may otherwise be unaware of this. Furthermore, these results provide an option to use CT-based BMD assessment for those institutions without access to the software necessary for quantitative BMD and may provide a stopgap technique until this software becomes widely clinically available. Our results, if confirmed in the multitude of pelvis, hip, or trauma CT scans performed annually, could increase the use of opportunistic screening for osteoporosis without additional software and may provide objective data that prompt providers and patients to obtain a DEXA scan when they might not otherwise [3]. Future research is needed to validate these results in larger studies and determine if HU measurements should be incorporated into routine radiology reporting, and finally to determine if these results are reflective of future hip and/or fragility fracture risk.

References

1. Abrahamsen B, van Staa T, Ariely R, Olson M, Cooper C. Excess mortality following hip fracture: a systematic epidemiological review. *Osteoporos Int*. 2009;20:1633–1650.
2. Allen MR, Burr DB. Bisphosphonate effects on bone turnover, microdamage, and mechanical properties: what we think we

know and what we know that we don't know. *Bone*. 2011;49: 56–65.
3. Berrington de González A, Mahesh M, Kim K-P, Bhargavan M, Lewis R, Mettler F, Land C. Projected cancer risks from computed tomographic scans performed in the United States in 2007. *Arch Intern Med*. 2009;169:2071–2077.
4. Bliuc D, Nguyen ND, Milch VE, Nguyen TV, Eisman JA, Center JR. Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. *JAMA*. 2009;301:513–521.
5. Buckens CF, Dijkhuis G, de Keizer B, Verhaar HJ, de Jong PA. Opportunistic screening for osteoporosis on routine computed tomography? An external validation study. *Eur Radiol*. 2015;25: 2074–2079.
6. Burge R, Dawson Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. *J Bone Miner Res.* 2007;22:465–475.
7. Cauley JA, Thompson DE, Ensrud KC, Scott JC, Black D. Risk of mortality following clinical fractures. *Osteoporosis Int*. 2000; 11:556–561.
8. Center JR, Nguyen TV, Schneider D, Sambrook PN, Eisman JA. Mortality after all major types of osteoporotic fracture in men and women: an observational study. *Lancet*. 1999;353:878–882.
9. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R; National Osteoporosis Foundation. Clinician's guide to prevention and treatment of osteoporosis. *Osteoporos Int*. 2014;25:2359–2381.
10. Eriksen EF. Cellular mechanisms of bone remodeling. *Rev Endocr Metab Disord*. 2010;11:219–227.
11. Fidler JL, Murthy NS, Khosla S, Clarke BL, Bruining DH, Kopperdahl DL, Lee DC, Keaveny TM. Comprehensive assessment of osteoporosis and bone fragility with CT colonography. *Radiology*. 2016;278:172–180.
12. Gausden EB, Nwachukwu BU, Schreiber JJ, Lorich DG, Lane JM. Opportunistic use of CT imaging for osteoporosis screening and bone density assessment: a qualitative systematic review. *J Bone Joint Surg Am.* 2017;99:1580–1590.
13. Gruber M, Bauer JS, Dobritz M, Beer AJ, Wolf P, Woertler K, Rummeny EJ, Baum T. Bone mineral density measurements of the proximal femur from routine contrast-enhanced MDCT data sets correlate with dual-energy x-ray absorptiometry. *Eur Radiol*. 2013;23:505–512.
14. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol*. 2012;8:23–34.
15. Halm EA, Beaber EF, McLerran D, Chubak J, Corley DA, Rutter CM, Doubeni CA, Haas JS, Balasubramanian BA. Association between primary care visits and colorectal cancer screening outcomes in the era of population health outreach. *J Gen Intern Med*. 2016;31:1190–1197.
16. Johnson CC, Gausden EB, Weiland AJ, Lane JM, Schreiber JJ. Using Hounsfield units to assess osteoporotic status on wrist computed tomography scans: comparison with dual energy x-ray absorptiometry. *J Hand Surg Am.* 2016;41:767–774.
17. Kanis JA, McCloskey EV, Johansson H, Oden A, Melton LJ, Khaltaev N. A reference standard for the description of osteoporosis. *Bone*. 2008;42:467–475.
18. Kanis JA, Melton LJ, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. *J Bone Miner Res.* 1994;9: 1137–1141.
19. Kern LM, Powe NR, Levine MA, Fitzpatrick AL, Harris TB, Robbins J, Fried LP. Association between screening for osteoporosis and the incidence of hip fracture. *Ann Intern Med*. 2005; 142:173–181.
20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163.
21. Lee S, Chung CK, Oh SH, Park SB. Correlation between bone mineral density measured by dual-energy x-ray absorptiometry and Hounsfield units measured by diagnostic CT in lumbar spine. *J Korean Neurosurg Soc.* 2013;54:384–389.
22. Lee SY, Kwon SS, Kim HS, Yoo JH, Kim J, Kim JY, Min BC, Moon SJ, Sung KH. Reliability and validity of lower extremity computed tomography as a screening tool for osteoporosis. *Osteoporos Int*. 2015;26:1387–1394.
23. Leibson CL, Tosteson ANA, Gabriel SE, Ransom JE, Melton LJ. Mortality, disability, and nursing home use for persons with and without hip fracture: a population-based study. *J Am Geriatr Soc.* 2002;50:1644–1650.
24. Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J, Dash C, Giardiello FM, Glick S, Johnson D, Johnson CD, Levin TR, Pickhardt PJ, Rex DK, Smith RA, Thorson A, Winawer SJ; American Cancer Society Colorectal Cancer Advisory Group, US Multi-Society Task Force, American College of Radiology Colon Cancer Committee. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology*. 2008;134:1570–1595.
25. Manolagas SC, Jilka RL. Bone marrow, cytokines, and bone remodeling. Emerging insights into the pathophysiology of osteoporosis. *N Engl J Med*. 1995;332:305–311.
26. Marinova M, Edon B, Wolter K, Katsimbari B, Schild HH, Strunk HM. Use of routine thoracic and abdominal computed tomography scans for assessing bone mineral density and detecting osteoporosis. *Curr Med Res Opin*. 2015;31:1871–1881.
27. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8:283–298.
28. Mueller DK, Kutscherenko A, Bartel H, Vlassenbroek A, Ourednicek P, Erckenbrecht J. Phantom-less QCT BMD system as screening tool for osteoporosis without additional radiation. *Eur J Radiol*. 2011;79:375–381.
29. Pickhardt PJ. Imaging and screening for colorectal cancer with CT colonography. *Radiol Clin North Am.* 2017;55:1183–1196.
30. Pickhardt PJ, Bodeen G, Brett A, Brown JK, Binkley N. Comparison of femoral neck BMD evaluation obtained using Lunar DXA and QCT with asynchronous calibration from CT colonography. *J Clin Densitom*. 2015;18:5–12.
31. Pickhardt PJ, Lee LJ, del Rio AM, Lauder T, Bruce RJ, Summers RM, Pooler BD, Binkley N. Simultaneous screening for osteoporosis at CT colonography: bone mineral density assessment using MDCT attenuation techniques compared with the DXA reference standard. *J Bone Miner Res.* 2011;26:2194–2203.
32. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Ann Intern Med*. 2013;158:588–595.
33. Sanders KM, Nicholson GC, Watts JJ, Pasco JA, Henry MJ, Kotowicz MA, Seeman E. Half the burden of fragility fractures in the community occur in women without osteoporosis. When is fracture prevention cost-effective? *Bone*. 2006;38:694–700.
34. Schreiber JJ, Anderson PA, Rosas HG, Buchholz AL, Au AG. Hounsfield units for assessing bone mineral density and strength: a tool for osteoporosis management. *J Bone Joint Surg Am.* 2011; 93:1057–1063.
35. Schreiber JJ, Gausden EB, Anderson PA, Carlson MG, Weiland AJ. Opportunistic osteoporosis screening–gleaning additional

information from diagnostic wrist CT scans. *J Bone Joint Surg Am*. 2015;97:1095–1100.
36. Siris ES, Chen Y-T, Abbott TA, Barrett-Connor E, Miller PD, Wehren LE, Berger ML. Bone mineral density thresholds for pharmacological intervention to prevent fractures. *Arch Intern Med*. 2004;164:1108.
37. US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW, García FAR, Gillman MW, Harper DM, Kemper AR, Krist AH, Kurth AE, Landefeld CS, Mangione CM, Owens DK, Phillips WR, Phipps MG, Pignone MP, Siu AL. Screening for colorectal cancer: US Preventive Services Task Force recommendation statement. *JAMA*. 2016;315:2564–2575.
38. Vermont J, Bosson JL, François P, Robert C, Rueff A, Demongeot J. Strategies for graphical threshold determination. *Comput Methods Programs Biomed*. 1991;35:141–150.
39. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565–574.
40. Wagner SC, Dworak TC, Grimm PD, Balazs GC, Tintle SM. Measurement of distal ulnar Hounsfield units accurately predicts bone mineral density of the forearm. *J Bone Joint Surg Am.* 2017;99:e38.
41. Weber NK, Fidler JL, Keaveny TM, Clarke BL, Khosla S, Fletcher JG, Lee DC, Pardi DS, Loftus EV, Kane SV, Barlow JM, Murthy NS, Becker BD, Bruining DH. Validation of a CT-derived method for osteoporosis screening in IBD patients undergoing contrast-enhanced CT enterography. *Am J Gastroenterol*. 2014;109:401–408.
42. Wright NC, Looker AC, Saag KG, Curtis JR, Delzell ES, Randall S, Dawson Hughes B. The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. *J Bone Miner Res.* 2014;29:2520–2526.
43. Yang L, Burton AC, Bradburn M. Distribution of bone density in the proximal femur and its association with hip fracture risk in older men: the MrOS study. *J Bone Miner Res.* 2012;27:2314–2324.
44. Ziemlewicz TJ, Binkley N, Pickhardt PJ. Opportunistic osteoporosis screening: addition of quantitative CT bone mineral density evaluation to CT colonography. *J Am Coll Radiol*. 2015;12:1036–1041.
45. Ziemlewicz TJ, Maciejewski A, Binkley N, Brett AD, Brown JK, Pickhardt PJ. Direct comparison of unenhanced and contrast-enhanced CT for opportunistic proximal femur bone mineral density measurement: implications for osteoporosis screening. *AJR Am J Roentgenol*. 2016;206:694–698.
46. Ziemlewicz TJ, Maciejewski A, Binkley N, Brett AD, Brown JK, Pickhardt PJ. Opportunistic quantitative CT bone mineral density measurement at the proximal femur using routine contrast-enhanced scans: direct comparison with DXA in 355 adults. *J Bone Miner Res.* 2016;31:1835–1840.