Bone 103 (2017) 325–333

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

Phantomless calibration of CT scans for measurement of BMD and bone strength—Inter-operator reanalysis precision

David C. Lee [a,*], Paul F. Hoffmann [a], David L. Kopperdahl [a], Tony M. Keaveny [b,c]

[a] O.N. Diagnostics LLC, Berkeley, CA, USA
[b] Department of Mechanical Engineering, University of California, Berkeley, CA, USA
[c] Department of Bioengineering, University of California, Berkeley, CA, USA

ARTICLE INFO

Article history:
Received 21 October 2016
Revised 1 June 2017
Accepted 21 July 2017
Available online 1 August 2017

Keywords:
BMD
Bone strength
Precision
Repeatability
Osteoporosis
Finite element analysis

ABSTRACT

Patient-specific phantomless calibration of computed tomography (CT) scans has the potential to simplify and expand the use of pre-existing clinical CT for quantitative bone densitometry and bone strength analysis for diagnostic and monitoring purposes. In this study, we quantified the inter-operator reanalysis precision errors for a novel implementation of patient-specific phantomless calibration, using air and either aortic blood or hip adipose tissue as internal calibrating reference materials, and sought to confirm the equivalence between phantomless and (traditional) phantom-based measurements. CT scans of the spine and hip for 25 women and 15 men (mean ± SD age of 67 ± 9 years, range 41–86 years), one scan per anatomic site per patient, were analyzed independently by two analysts using the VirtuOst software (O.N. Diagnostics, Berkeley, CA). The scans were acquired at 120 kVp, with a slice thickness/increment of 3 mm or less, on nine different CT scanner models across 24 different scanners. The main parameters assessed were areal bone mineral density (BMD) at the hip (total hip and femoral neck), trabecular volumetric BMD at the spine, and vertebral and femoral strength by finite element analysis; other volumetric BMD measures were also assessed. We found that the reanalysis precision errors for all phantomless measurements were ≤0.5%, which was as good as for phantom calibration. Regression analysis indicated equivalence of the phantom- versus phantomless-calibrated measurements (slope not different than unity, $R^2 \geq 0.98$). Of the main parameters assessed, non-significant paired mean differences ($n = 40$) between the two measurements ranged from 0.6% for hip areal BMD to 1.1% for mid-vertebral trabecular BMD. These results indicate that phantom-equivalent measurements of both BMD and finite element-derived bone strength can be reliably obtained from CT scans using patient-specific phantomless calibration.

1. Introduction

Quantitative analysis of computed tomography (CT) scans can be performed clinically to identify patients at high risk of fracture [1–3] based on calibrated measurements of bone mineral density (BMD) at the spine and hip, as well as measurements of bone strength when combined with finite element analysis [4–7]. In any quantitative CT analysis, proper calibration of the scan is required to correct for variations in scanner settings and attenuation characteristics [8], any related beam-hardening, and patient-specific characteristics such as body size, all of which can alter the attenuation characteristics [9]. Without such corrections, BMD and bone strength measurements can vary across different CT scanners or with different scan protocols or over time, confounding interpretation and clinical utility.

The most widely used method for calibrating CT scans utilizes an external calibration phantom [10]. However, the need for a phantom, which must be placed under the patient during scanning, adds expense and increases the logistical burden of clinical imaging. Various approaches have been proposed to calibrate without an external calibration phantom. One approach is to not calibrate the scan [11], but this amounts to not performing quantitative densitometry and is therefore questionable for diagnostic or monitoring purposes because attenuation values can vary widely depending on the specific scanner and scan protocol [9,12]. Another approach is to pre-calibrate a particular CT scanner using a calibration phantom, or to pre-calibrate CT-based BMD measurements via DXA, and then use that pre-calibration for scans of future patients on that or similar scanners [13–16]. While this approach is preferable to not calibrating, both approaches are not specific to the individual patient and thus ignore potentially important calibration issues associated with variations in patient body size and habitus; nor can they be applied retrospectively over any appreciable period of time. A third

* Corresponding author at: 2150 Shattuck Ave. STE 610, Berkeley, CA 94704, USA.
E-mail addresses: david.lee@ondiagnostics.com (D.C. Lee), paul.hoffmann@ondiagnostics.com (P.F. Hoffmann), david.kopperdahl@ondiagnostics.com (D.L. Kopperdahl), tonykeaveny@berkeley.edu (T.M. Keaveny).

http://dx.doi.org/10.1016/j.bone.2017.07.029
8756-3282/

approach, which represents a patient-specific phantomless calibration, is to utilize the patient's own internal tissues as the calibrating reference materials [17,18]. One such approach involves sampling the attenuation of a region of muscle, then further processing those attenuation data into components assumed to be associated with pure muscle tissue and pure adipose tissue, thereby providing attenuation data for two known reference materials [19]. However, one limitation of that approach is its poor repeatability [17], presumably due to the challenges of choosing the region of muscle in a repeatable fashion as well as consistently separating out the pure tissue components.

Overcoming these limitations with patient-specific phantomless calibration, we report here on an alternative implementation having improved precision. Already validated clinically against DXA for accurately measuring areal BMD at the hip [2,3], this technique utilizes as calibrating reference materials the external air that is visible on the patient's CT scan and one of either the patient's blood or adipose tissue adjacent to the bone being assessed. Applying this technique to a diverse cohort of patients and CT scanners, we sought to quantify its inter-operator reanalysis precision for measuring both BMD and finite element-derived bone strength, at both the spine and hip; we also sought to confirm the equivalence of the phantomless versus traditional phantom-based measurements.

2. Methods

2.1. CT scans for the study sample

The study comprised of a reanalysis of pre-existing research-quality clinical-resolution CT scans that had been analyzed in prior clinical drug trials at O.N. Diagnostics. Permission was obtained from the original sources for reanalysis of the CT scans, and additional internal review board approval for this reanalysis was not necessary due to the retrospective, de-identified nature of the dataset.

We randomly selected a sample of CT scans from a larger pool of scans in order to minimize sources of measurement bias in the selected scans. Eight prior multi-center studies at O.N. Diagnostics had scans that were available for reanalysis. Of those studies, eligible CT scans for inclusion were those that were: 1) acquired on a multi-detector CT scanner at 120 kVp; 2) available for both the spine and hip for an individual patient; 3) imaged with a mineral phantom pair that included an external phantom for calibration and a quality-assurance phantom for beam hardening correction; 4) reconstructed utilizing a slice thickness of 3 mm or less and a standard kernel (GE: standard; Siemens: B30; Philips: B, C; Toshiba: FC12, FC13); 5) without imaging artifacts that would preclude analysis; and 6) were not previously utilized in the development of the phantomless calibration method. Typical exposure values were set according to patient height and weight (ranging from 25 to 195 mAs for spine scans, and 50–390 mAs for hip scans) or were determined by the scanner's automatic exposure control function (e.g. noise index = 25 HU or quality reference mAs = 160). For any of the studies that assessed longitudinal effects, only baseline scans were included, so no scans reflected any treatment effects. The CT scans for 1032 subjects thus identified were acquired on 58 scanners. From those, we randomly selected 40 subjects—providing sufficient statistical power for a precision study [20,21]—while limiting the number of scans from any single CT scanner to no more than eight. The resulting scans were derived from 24 unique CT scanners, representing nine different CT scanner models (Table 1). Fourteen of these 24 scanners were represented in development of the phantomless calibration method. The cohort consisted of 25 women and 15 men, spanning a wide range of age, weight, height, and body mass index (Table 2).

2.2. Phantom and phantomless calibrations

The VirtuOst software (version 2.1, O.N. Diagnostics, Berkeley, CA), written in the Python programming language and utilizing NumPy and SciPy software libraries, was used for both the phantom and phantomless calibrated analyses. Both methods of calibration were performed separately at the spine and hip since attenuation characteristics can differ at each site due to site-specific differences in body habitus. For the phantom calibration, the same type of external calibration phantom (Model 3 Phantom, Mindways Software, Inc., Austin, TX) was utilized for each patient (Fig. 1). Following the manufacturer-supplied specifications [22], the attenuation values (Hounsfield Units) of the chambers in the calibration phantom were sampled and the images were calibrated into equivalent-BMD units (mg/cm^3) of a K_2HPO_4-water mixture. The user specified the range along the length of the phantom from which to sample data (avoiding any shading or other artifacts) and then over that range cylinders were automatically registered to each chamber. The final volume of interest (VOI) was determined by removing an outer layer from each chamber to eliminate volume averaging with the surrounding phantom substrate. To help account for potential beam hardening between the locations of the external calibration phantom and the bone of interest, a quality-assurance torso phantom (Mindways Software, Inc., Austin, TX) was also scanned on top of the external calibration phantom (Fig. 1), typically within one day of each subject's CT exam. The resulting quality-assurance phantom scan was calibrated as described above and the measured BMD of its central chamber was used to generate a ratio versus the reference value of BMD for that chamber [23]; this field-uniformity correction ratio was then used to scale the corresponding subject's BMD measurement to correct for any field-uniformity effects, the same ratio used both for the spine and hip scans. Across the 40 scans, values of this ratio varied from 0.92–1.05. To assess the sensitivity of the overall reanalysis precision error to the measurement of this ratio (which itself contributes some degree of measurement error), we also calculated reanalysis precision with this ratio set to 1.0 for all scans.

For the phantomless calibration, the attenuation values were sampled for external air and either abdominal aortic blood tissue for assessment of the spine or pelvic visceral adipose tissue from the ischioanal fossa for assessment of the hip (Fig. 1). Since both reference tissues

Table 1
Forty scans each for the spine and hip were used in this analysis, acquired from nine different CT scanner models, from 24 different scanners.

CT scanner model[a]	Number of scanners	Number of scans[b]
GE BrightSpeed	3	4
GE LightSpeed Ultra	1	1
GE LightSpeed VCT	4	8
GE LightSpeed 16	5	8
Philips Brilliance 16	2	2
Philips Brilliance 64	3	7
Siemens Sensation 40	1	1
Siemens Sensation 64	1	2
Toshiba Aquilion	4	7
Total	24	40

[a] All scans were acquired at 120 kVp and reconstructed with a standard reconstruction kernel with a slice thickness of up to 3 mm.
[b] Number of paired spine and hip scans (n = 40 total for each type of scan).

Table 2
Descriptive statistics for the analyzed cohort.

	Women	Men	Pooled
Number of subjects	25	15	40
Age (yrs)	64 ± 9 (41–80)	72 ± 6 (65–86)	67 ± 9 (41–86)
Height (cm)	157 ± 6 (141–166)	177 ± 6 (167–185)	164 ± 12 (141–185)
Weight (kg)	58 ± 9 (39–73)	100 ± 18 (66–124)	74 ± 24 (39–124)
BMI (kg/m^2)	23.7 ± 3.9 (18.2–31.5)	31.7 ± 4.0 (23.3–36.7)	26.7 ± 5.5 (18.2–36.7)

BMI—Body Mass Index.
Values are mean ± standard deviation (range in parentheses).