Bone 48 (2011) 1087–1094

Contents lists available at ScienceDirect

Bone

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

Estimation of bone mineral density in children from diagnostic CT images: A comparison of methods with and without an internal calibration standard

Alexander H. Habashy [a], Xiaowei Yan [b,c], J. Keenan Brown [d], Xiaoping Xiong [b], Sue C. Kaste [a,e,f,*]

[a] Department of Radiological Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA
[b] Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, USA
[c] Biostatistics, University of Memphis, Memphis, TN, USA
[d] Mindways Software, Inc., Austin, TX, USA
[e] Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA
[f] Department of Radiology, University of Tennessee, Health Science Center, Memphis, TN, USA

ARTICLE INFO

Article history:
Received 5 August 2010
Revised 7 December 2010
Accepted 14 December 2010
Available online 23 December 2010

Edited by: Harry Genant

Keywords:
Bone mineral density
Diagnostic CT
Quantitative computed tomography
Calibration phantom
Lumbar spine
Contrast agent

ABSTRACT

We investigated the feasibility and potential limitations of estimating bone mineral density (BMD) from standard diagnostic computed tomography (dCT). We analyzed three sets of BMD measurements for L1 and L2, each performed by a novice and an expert, for intra- and interobserver variance ($n=43$ studies from 38 patients; median age, 13.2 years) using one BMD quantification system with (conventional quantitative computed tomography (QCT)) and two without (QCT and dCT) an external calibration phantom. Using ANOVA model, means of three sets of BMD measurements analyzed by the expert differed by 2.5 mg/cm^2; for the novice, by less than 1 mg/cm^2. Variation of measurement differences was less for the expert. Mean intra- and interobserver absolute standardized differences (ASD) were 1.77% and 1.8%, respectively. The mean ASD between phantom and phantom-less methods of QCT studies were 3.3%; mean ASD of phantom QCT versus phantom-less dCT was 14.3%. Regression modeling suggested compensation for sources of dCT BMD measurement bias can reduce the mean ASD of phantom QCT versus phantom-less dCT to 6.5%. Thus, phantom-less QCT of dCT adds clinically useful BMD information not typically attained from dCT, thereby augmenting patient care and presenting important possibilities for research without need for additional study.

Introduction

The assessment of bone mineral density (BMD) and fracture risk has become a common practice amongst patients suffering from diseases which may compromise the structural integrity of bone. Dual X-ray absorptiometry (DXA) is the most widely used means of determining BMD. However, the use of quantitative computed tomography (QCT) is growing due to its ability to measure volumetric BMD in g/cm^3 [1] instead of areal density as measured using DXA (g/cm^2). The direct volumetric assessment of BMD circumvents measuring errors that are known to occur in growing children as a result of changing bony morphology. QCT also has the ability to differentiate between cortical and trabecular bone, which allows monitoring of the response of the more metabolically active trabecular bone compartment to interventions designed to improve BMD [2]. Anatomy located outside of a region of interest (ROI) can be eliminated using the interactive capabilities of QCT. QCT also provides greater image resolution of tissue and can display images in multiple planes. Though QCT subjects patients to greater exposure to ionizing radiation than does DXA [3], contemporary low-dose techniques directed at focused regions of interest provide diagnostic information while minimizing such exposure.

The advancement in bone densitometry has brought forth new possibilities of investigating BMD. In 3D QCT images, a mid-vertebral trabecular volume is analyzed using an elliptical ROI with results presented as two measures, *Z*-score and *T*-score. A reference phantom is typically used to standardize measurements to a mineral reference and correct for potential analytical errors arising from beam hardening and radiation scattering. While the use of a phantom is standard among CT-acquired bone densitometry studies, the phantom is not utilized during routine diagnostic CT (dCT) examinations, namely pelvic and abdominal CT studies. Thus, measurements of BMD from a dCT image slice may not be a reliable measure compared with values derived using standard QCT performed with a phantom. Furthermore, diagnostic CTs are commonly performed after administration of an intravenous contrast agent, which has been shown to increase BMD measurements [4].

Our study is prompted by the need for a new and innovative way to measure BMD without the utilization of a phantom and minimizing patient exposure to ionizing radiation by retrospectively utilizing

* Corresponding author. Department of Radiological Sciences, 262 Danny Thomas Place, MSN #220, Memphis, TN 38105, USA. Fax: +1 901 595 3981.
E-mail address: skaste@stjude.org (S.C. Kaste).

8756-3282/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bone.2010.12.012

existing or prospective dCT data. In this study, we tested the accuracy and reliability of the Bone Investigational Toolkit®, (BIT; Mindways Software Inc.) which enables analysis of data sets acquired without the use of a CT calibration phantom. By focusing solely on statistical measurements of BMD, we sought to evaluate values generated from these two methods (standard QCT and BIT), seek out potential sources of error and assess the utility of BIT in a clinical setting. Accurate determination of BMD using this software would obviate the need for subjecting patients (particularly children) to additional radiation exposure, without compromising the accuracy of BMD measurements. Thus, patients with unanticipated BMD deficits who might benefit from further evaluation may be assessed retrospectively. If proven successful, the implications of this research may identify previously unrecognized patient cohorts with BMD deficits. For example, if BIT can be used to accurately analyze BMD from a dCT without use of a phantom, a more detailed analysis of bone mineral density over time could be obtained. Thus, such a determination of BMD could provide clinically valuable information not typically available with dCT. Most importantly, it may reduce the need for multiple examinations, consequently decreasing cumulative radiation exposure to patients.

Methods

We analyzed CT data collected between 2003 and 2008 of pediatric and young adult patients who had undergone both a standard lumbar spine QCT and a dCT study that included vertebral bodies L1 and L2, within 24 hours of each other. This retrospective study was performed with IRB approval and data were managed according to the Health Insurance Portability and Accountability Act (HIPAA) of 1996.

We based our study on analyses performed by two categories of observers: a novice [AHH] and an expert [JKB] user of the analysis software. The novice analyzed the BMD of L1 and L2 without any previous experience in processing QCT or dCT studies. The expert observer was the developer of the software and was highly experienced in the processing of QCT and dCT studies. To determine if a learning curve was associated with the software, we analyzed both inter- and intraobserver BMD estimates of the novice and the expert.

The QCT studies evaluated L1 and L2 BMD and incorporated an external calibration phantom (Mindways Software, Austin, TX) imaged simultaneously with the subject. The dCT studies were typically performed after administration of intravenous contrast (Visipaque/Omnipaque/Isovue, 1.5 cc/kg) and opacification of the gastrointestinal tract with oral contrast; no external calibration phantom was used.

The QCT data sets were analyzed with version 4.2 of QCT-PRO™ (Mindways Software, Austin, TX) using standard methods provided by the manufacturer [5]. The QCT and dCT data sets were both analyzed using QCT-PRO™ in conjunction with version 2.0 of the Bone Investigational Toolkit™ (BIT, Mindways Software, Austin, TX). With BIT, the CT calibration information necessary for the analysis was derived from a QCT quality assurance (QA) scan involving only phantoms. The QA phantom scans were acquired on the same CT scanner, using the same kVp and reconstruction methods as was used for patient imaging. QCT BMD studies as well as dCT studies were performed on a General Electric Lightspeed Ultra 8 Detector CT scanner until late January of 2008 at which time it was replaced with a General Electric VCT XT 64 Detector unit. The dCT scans for this study were acquired at 120 kVp and the QCT scans at 80 kVp. Both types of studies were reconstructed using a standard abdomen CT reconstruction method. Note the calibration data intrinsic to the QCT data sets was ignored when analyzing the QCT data sets with the BIT software. In the case of the dCT studies, the QA studies used for calibration purposes with BIT were acquired after all of the patient data were obtained and the creatinine levels of the patients were observed within no more than 1 month prior to the dCT.

In summary, three sets of BMD measurements were derived from applying two quantification systems (PRO and BIT) to two types of imaging data sets (QCT and dCT) for each subject in this study (see Fig. 1): (1) BMD results derived using conventional QCT data sets with external phantom calibration (this set of measurement is denoted as PRO/QCT), (2) BMD results derived from dCT data sets using BIT (denoted as BIT/dCT), and (3) BMD results derived from conventional QCT data sets using BIT ignoring phantom calibration information intrinsic to and available from the QCT data sets (denoted as BIT/QCT). The above analyses were each performed twice, 4 months apart, by both readers. Inter- and intra-observer variability was characterized using ANOVA methods. The relative measurement biases for the three BMD estimates for each study were compared using ANOVA methods. The results of the QCT method were considered the gold standard for bias comparisons. For assessment, we used L1 to compare results between users and methods used, while using L2 as validation.

Statistical analysis

For each subject, three BMD measurements were generated from two data sets: PRO quantification of QCT data, BIT quantification of QCT data, and BIT quantification of dCT data. The ultimate goal of this analysis was to obtain an estimation of BMD based on the dCT data, for which the PRO quantification of QCT data was taken as the gold standard. To reach this goal, we first compared PRO and BIT quantification of QCT data, from which the effect of phantom on the accuracy of BMD measurements could be estimated. We then compared BIT quantification of dCT data with PRO quantification of QCT data, from which other factors affecting the accuracy of BMD could be identified. Consequently, we obtained a good estimation of BMD for dCT data by removing biases caused by these factors. The BMD measurements were analyzed by two observers (novice and expert), and repeated twice for each observer. Thus, we could estimate systematic differences and variations within and between observers. We used ANOVA to assess the systematic difference (i.e., bias) between two measurements, performed by two observers, with two repetitions each, and effects of other factors such as weight, renal function, IV contrast and age of patients. The ANOVA model includes Measurement (two levels), and/or Observer (two levels), and/or Repetition (only one repetition; two levels for this factor), and Patient (35 levels) as factors. The ANOVA model is more powerful than the simple t-test, for example, to detect the effect of Measurement by removing the effect of Patient and Repetition from the errors of model. We used the standardized absolute difference of two observations to assess the absolute difference between two types of BMD measurements. For example,

$$\text{Absolute Standardized Difference} = 2 * |\text{Difference}| \,/\, (\text{Observation by PRO of QCT} + \text{Observation by BIT of dCT}),$$

where Difference = Observation by PRO of QCT − Observation by BIT of dCT. A good estimator based on dCT data should have a small bias as well as a small average absolute standardized difference between PRO of QCT and BIT of dCT. Using the regression model, we identified IV Contrast and patient age to be significant factors accounting for the difference in BMD results between PRO of QCT and BIT of dCT. As IV contrast and patient age are strongly correlated, we fitted a simple regression model,

$$\text{Difference} = \text{Intercept} + \text{Slope} * (\text{IV Contrast})$$

where Difference = PRO of QCT–BIT of dCT is for each patient, and IV Contrast is the volume (ml) of the intravenous contrast received by this patient. With this model, we will obtain a measure of BMD for the dCT data which improves the BIT measurement of dCT by removing a bias caused by IV Contrast.