Utilizing Artificial Intelligence to Determine Bone Mineral Density Via Chest Computed Tomography

Rock H. Savage, BS,* Marly van Assen, PhD,*† Simon S. Martin, MD,*‡ Pooyan Sahbaee, PhD,*§ Lewis P. Griffith, BS*, Dante Giovagnoli, BS,* Jonathan I. Sperl, PhD,¶ Christian Hopfgartner, MSc,∥ Rainer Kärger, MSc,∥ and U. Joseph Schoepf, MD*

Purpose: The purpose of this study was to validate the accuracy of an artificial intelligence (AI) prototype application in determining bone mineral density (BMD) from chest computed tomography (CT), as compared with dual-energy x-ray absorptiometry (DEXA).

Materials and Methods: In this Institutional Review Board–approved study, we analyzed the data of 65 patients (57 female, mean age: 67.4 y) who underwent both DEXA and chest CT (mean time between scans: 1.31 y). From the DEXA studies, T-scores for L1-L4 (lumbar vertebrae 1 to 4) were recorded. Patients were then divided on the basis of their T-scores into normal control, osteopenic, or osteoporotic groups. An AI algorithm based on wavelet features, AdaBoost, and local geometry constraints independently localized thoracic vertebrae from chest CT studies and automatically computed average Hounsfield Unit (HU) values with kVp-dependent spectral correction. The Pearson correlation evaluated the correlation between the T-scores and HU values. Mann-Whitney U test was implemented to compare the HU values of normal control versus osteoporotic patients.

Results: Overall, the DEXA-determined T-scores and AI-derived HU values showed a moderate correlation ($r = 0.55; P < 0.001$). This 65-patient population was divided into 3 subgroups on the basis of their T-scores. The mean T-scores for the 3 subgroups (normal control, osteopenic, osteoporotic) were $0.77 ± 1.50$, $-1.51 ± 0.04$, and $-3.26 ± 0.59$, respectively. The mean DEXA-determined L1-L4 BMD measures were $1.13 ± 0.16$, $0.88 ± 0.06$, and $0.68 ± 0.06$ g/cm$^2$, respectively. The mean AI-derived attenuation values were $145 ± 42.5$, $136 ± 31.82$, and $103 ± 16.28$ HU, respectively. Using these AI-derived attenuation values, we compared the HU values of normal control versus osteoporotic patients.

Conclusion: Our results show that this AI prototype can successfully determine BMD in moderate correlation with DEXA. Combined with other AI algorithms directed at evaluating cardiac and lung diseases, this prototype may contribute to future comprehensive preventative care based on a single chest CT.

Key Words: artificial intelligence, bone mineral density, chest computed tomography

Bone mineral density (BMD) is an important marker of health, especially in the aged population. In the United States, an estimated 44 million adults suffer from low BMD, while ~10 million of those are currently living with osteoporosis. Osteoporosis can be defined as a disease of low BMD and altered bone microstructure that leads to bone fragility. It is well known that reduced BMD may be a sign of osteoporosis that leaves one at higher risk of fracture resulting in continued health complications and thus further medical cost and resource use. Weaver et al found that adults suffering a fracture concurrent with a diagnosis of osteoporosis have a greater hip fracture risk than those without prior osteoporosis.

MATERIALS AND METHODS

Study Population

This retrospective, single-center study was approved by the University Hospital’s Institutional Review Board, and

From the *Department of Radiology and Radiological Science, Division of Cardiovascular Imaging, Medical University of South Carolina, Charleston, SC; †Department of Radiology, University Medical Center Groningen, Groningen, The Netherlands; ‡Department of Diagnostic and Interventional Radiology, University Hospital Frankfurt, Frankfurt; §Siemens Healthineers, Forchheim, Germany; and ¶Siemens Medical Solutions, Malvern, PA.

U.I.S. is a consultant for and/or receives research support from Astellas, Bayer, Elucida BioImaging, General Electric, Guerbet, HeartFlow, and Siemens Healthineers. P.S. is a salaried employee of Siemens Medical Solutions. J.I.S., C.H., and R.K. are salaried employees/business partners of Siemens Healthineers. The remaining authors declare no conflicts of interest.

Correspondence to: U. Joseph Schoepf, MD, Heart & Vascular Center, Medical University of South Carolina, Ashley River Tower, 25 Courtenay Drive, Charleston, SC 29425-2260 (e-mail: schoepf@umc.edu).

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

DOI: 10.1097/RTI.0000000000004848
the need for informed consent was waived. Various institutional databases were accessed to identify patients who received both DEXA and chest CT studies within a 3-year period to ensure no significant change in BMD between scans. Other inclusion criteria included imaging studies of diagnostic quality, as determined by the Likert scale, and a minimum patient age of 30 years. As maximal BMD does not occur until approximately 30 years of age, T-scores are not applicable to patients younger than that threshold.

**Imaging Protocols**

The BMD measurements determined by DEXA images were produced by the Hologic Discovery A DEXA System (S/N81442) with software version 13.5.3.1. The Discovery model utilizes a linear x-ray fan-beam with switched-pulse dual energy (100 /140 kVp) and a multielement detector array. The scan time was 10 seconds with an exposure of 0.04 mGy for both lumbar spine and proximal femur scans.

Noncontrast chest CT studies were acquired during breath-hold at mid-inspiration using either single-source CT (SOMATOM Emotion, Siemens Healthineers, Forchheim, Germany) or second or third-generation dual-source CT (SOMATOM Definition Flash, SOMATOM Force, Siemens). Anatomic tube current modulation (CARE Dose 4D, Siemens) was applied per default on all systems. Single-source CT studies used the following parameters: collimation, 16×0.6 mm; tube voltage, 120 kV; tube current, 80 mA. For the second and third-generation dual-source CT scanners, the following parameters were used: collimation, 128×0.6 mm; tube voltage, 120 kV; tube current, 75 mA, and collimation, 192×0.6 mm; tube voltage, 120 kV; tube current, 80 mA, respectively. The average CT dose index was 0.49 ± 0.21 mGy, and the mean dose length product was 315.7 ± 216.4 mGy·cm, resulting in a mean effective dose of 4.74 ± 3.25 mSv.

**Image Analysis**

Two experienced observers, in consensus, rated image quality of the chest CT studies on a 5-point Likert scale. Only images with a rating of 4 or 5, by both observers, were included.

Once the DEXA scan had generated a measure of BMD, these values were then plotted as T-scores and Z-scores. The T-score is a comparison of the patient’s BMD with that of a sex-matched, young, and healthy individual. The Z-score is a comparison of the patient’s BMD with that of a sex and age-matched individual. These scores denote the difference between the measured densities and reference values in terms of SDs, which are then used to categorize patients into a normal, osteopenic, or osteoporotic level of BMD. A T-score below −2.5 indicates osteoporosis; between −2.5 and −1.0, osteopenia; and above −1.0, normal BMD. A Z-score below −2.0 indicates low bone mass, whereas above −2.0 indicates normal bone mass.

An AI algorithm based on wavelet features, AdaBoost, and local geometry constraints independently localized thoracic vertebrae within the chest CT studies using previously validated techniques. Adjacent to the center of each vertebral body, a vector approximately orthogonal to the top and bottom planes of the vertebrae was computed. In a subsequent step, regions of interest around each center were defined as cylinders with a height equaling 6 mm and radii equaling 4 mm for T1 to T3 and 6 mm for T4 to T12. In addition, the cylinder axes were all parallel to the normal vectors. Figure 1 illustrates the location of the individual cylinders.

For all voxels within a regions of interest, the corresponding HU were corrected by a kVp-dependent and filter-dependent correction factor obtained from independent phantom measurements. The correction factors varied between 0.52 and 1.20. The mean of the corrected HU values was subsequently reported.

**Statistical Analysis**

Statistical analysis used MedCalc (MedCalc Software bvba, version 16.8, Ostend, Belgium). All data were evaluated for normality using the Kolmogorov-Smirnov test. Parametric variables throughout the study were expressed as a mean ± SD, and nonparametric variables were expressed as a median with interquartile ranges (IQR). The Pearson correlation evaluated the correlation between the DEXA-determined T-scores and AI-derived attenuation values. The Mann-Whitney U test was implemented in order to compare the attenuation values of normal control patients versus the osteoporotic patients. Osteopenic patients were not included in the aforementioned analysis, as osteopenia is defined as a decrease in bone mass but is not the pathologic disease state of decreased bone mass; that designation is reserved for osteoporosis. A P-value <0.05 was considered to be statistically significant.

**RESULTS**

The inclusion criterion of a maximum of 3 years between DEXA and CT scans distilled the initial search database to 66 patients. One patient, below 20 years of age, was excluded from further analysis. In addition, all imaging studies received a score of 4 or 5 on the Likert scale and could therefore be included in the study.

The final study population consisted of 65 patients (mean age at DEXA, 66.4 ± 10.6 y; range, 36 to 99 y; mean age at CT, 67.0 ± 10.8 y; range, 38 to 100 y) including 57 female individuals (mean age at DEXA, 67.4 ± 9.7 y; range, 44 to 99 y; mean age at CT, 68.0 ± 9.8 y; range, 43 to 100 y) and 8 male individuals (mean age at DEXA, 59.3 ± 13.7 y; range, 36 to 79 y; mean age at CT, 59.5 ± 14.2 y; range, 38 to 80 y). The mean time between...
CT and DEXA studies was 1.31 ± 0.82 years. The mean body mass index of the study cohort was 28.7 ± 6.8 kg/m² (range, 16.5 to 52 kg/m²) (Table 1).

On the basis of their DEXA-determined L1-L4 T-scores, patients were first divided into normal control, osteopenic, or osteoporotic groups. These subpopulations included 38 (31 female individuals and 7 male individuals with a mean L1-L4 T-score = 0.77 ± 1.50), 24 (22 female individuals and 2 male individuals with a mean L1-L4 T-score = −1.51 ± 0.04), and 3 (3 female individuals with mean L1-L4 T-score = −3.26 ± 0.59) patients, respectively (Table 2).

The mean L1-L4 BMD for the normal control group was 1.13 ± 0.16 g/cm², the mean L1-L4 BMD for the osteopenic group was 0.88 ± 0.06 g/cm², and the mean L1-L4 BMD for the osteoporotic group was 0.68 ± 0.06 g/cm². The mean AI-derived attenuation values at T1-T12 for the normal control, osteopenic, and osteoporotic groups were as follows: 145 ± 2.5, 136 ± 3.18, and 103 ± 16.2 HU, respectively (Table 2). These mean AI-derived attenuation values were plotted against the BMD determined by DEXA and are visualized in Figure 2. In addition, the results shown in Figure 3 may demonstrate the ability of the AI to differentiate between diseased and nondiseased states. When comparing the AI-derived HU values between subgroups, a significant difference was found (P = 0.045), which suggests that this prototype may be efficacious in determining active disease states, or bone health status. Further work will be required in order to achieve this future goal of disease classification.

One of the first uses of predictive modeling toward bone health analysis can be found in the 1999 study by Gregory et al. In their work, they used Fourier transforms to create a spectral fingerprint to analyze an image’s principal components, chiefly trabecular bone, which were then classified via neural network. The authors claim an accuracy of ~80% in classifying histologic sections as normal, osteoarthritic, or osteoporotic by using this technique. This early focus on bone health with the assistance of computer modeling set the stage for the eventual expansion of such techniques into the clinical

### DISCUSSION

In this study, we evaluated an AI-based prototype algorithm designed to quantify BMD from noncontrast chest CTs compared with the traditional DEXA scan. Our results indicate that this AI-based prototype may be able to successfully determine BMD in a variety of patients regardless of age, sex, body mass index, or bone health status. The correlation of r = 0.55 between the DEXA-determined T-scores and AI-derived attenuation values indicates promising efficacy of this prototype in determining BMD from chest CT studies (Fig. 2). In addition, the results shown in Figure 3 may demonstrate the inability of the AI to differentiate between diseased and nondiseased states. When comparing the AI-derived HU values between subgroups, a significant difference was found (P = 0.045), which suggests that this prototype may be efficacious in determining active disease states, or bone health status, using noncontrast chest CT studies. However, further work will be required in order to achieve this future goal of disease classification.

Two representative cases of the AI’s display output are shown in Figures 4 and 5. Figure 4 is a case example of a patient with normal BMD, whereas Figure 5 is a case example of a patient with osteoporosis.

### TABLE 1. Basic Demographic Information of the Patients Included in This Study

<table>
<thead>
<tr>
<th>Demographic Markers</th>
<th>All Patients, n = 65</th>
<th>Normal Control, n = 38</th>
<th>Osteopenic, n = 24</th>
<th>Osteoporotic, n = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At DEXA</td>
<td>66.4 ± 10.6</td>
<td>66.6 ± 11.8</td>
<td>65.0 ± 8.7</td>
<td>72.0 ± 6.2</td>
</tr>
<tr>
<td>At chest CT</td>
<td>67.0 ± 10.8</td>
<td>67.3 ± 12.0</td>
<td>65.5 ± 8.8</td>
<td>72.7 ± 6.1</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female individuals</td>
<td>57 (88)</td>
<td>31 (82)</td>
<td>22 (92)</td>
<td>3 (100)</td>
</tr>
<tr>
<td>Male individuals</td>
<td>8 (12)</td>
<td>7 (18)</td>
<td>2 (8)</td>
<td></td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>48 (74)</td>
<td>31 (82)</td>
<td>15 (63)</td>
<td>2 (67)</td>
</tr>
<tr>
<td>Black</td>
<td>16 (25)</td>
<td>7 (18)</td>
<td>8 (33)</td>
<td>1 (33)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (1)</td>
<td>1 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.7 ± 6.8</td>
<td>29.8 ± 7.4</td>
<td>26.2 ± 4.8</td>
<td>23.4 ± 4.9</td>
</tr>
</tbody>
</table>

BMI indicates body mass index; CT, computed tomography; DEXA, dual-energy X-ray absorptiometry.

### TABLE 2. Comparison Between Subpopulations of DEXA-Determined Mean BMD, T-scores, and AI-derived Attenuation Values

<table>
<thead>
<tr>
<th>Scan Measures</th>
<th>Normal Control, n = 38</th>
<th>Osteopenic, n = 24</th>
<th>Osteoporotic, n = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEXA BMD (g/cm²)</td>
<td>1.13 ± 0.16</td>
<td>0.88 ± 0.06</td>
<td>0.68 ± 0.06</td>
</tr>
<tr>
<td>T-scores</td>
<td>0.77 ± 1.50</td>
<td>−1.51 ± 0.04</td>
<td>−3.26 ± 0.59</td>
</tr>
<tr>
<td>Chest CT Attenuation (HU)</td>
<td>145 ± 42.5</td>
<td>136 ± 31.8</td>
<td>103 ± 16.2</td>
</tr>
</tbody>
</table>

AI indicates artificial intelligence; CT, computed tomography; DEXA, dual-energy X-ray absorptiometry.

FIGURE 2. Scatter plot comparing the AI-derived HU values and DEXA-determined T-scores for each individual in the study population. r = 0.55, P < 0.001.
setting. Consequently, while the methods used in Gregory and colleagues were based on histologic sections, our present study transfers this paradigm into the realm of preventative identification of pathology, in patients at risk, by harnessing AI to provide a measure of bone density without the need for additional scans or procedures.

Outside of imaging, the use of AI in the context of bone health has been explored by Chiu et al. who implemented an artificial neural network attuned to 7 predictive demographic and lifestyle variables to predict osteoporosis with a sensitivity of 78.3% and a specificity of 73.3%. Their findings contributed to the work of several others, all of which help to demonstrate the efficacy of AI in predicting osteoporotic disease. Nevertheless, the computational algorithms in these studies required the careful collection of potentially erroneous self-reported data. The prototype algorithm in the present study is not attuned to qualitative data of medical records but rather to quantitative imaging with a higher degree of reproducibility and objectivity.

In the following years, several AI-based approaches have been described for the diagnosis of osteoporotic disease involving anatomic locations that differ from the site of original diagnostic determination. In the study by Kavitha et al., a radial basis function kernel (RBF kernel)-based support vector machine measured mandibular cortical width via dental panoramic radiographs in order to identify low BMD in postmenopausal women. The use of support vector machine in the context of bone density quantification had been previously found to be efficacious in a prior study by Lee et al. Although the work by Kavitha and colleagues does support that novel algorithms can be used to determine BMD using imaging from a location other than that of the initial site of diagnosis, their study was limited to women above the age of 50 and so may not be able to be extrapolated to the population at large. The present study aims to be more inclusive with a broader patient population that is more indicative of the general population requiring measures of BMD.

Although lumbar vertebrae and proximal femur scans are the most typically used sites for determining BMD via DEXA, prior studies have shown that thoracic vertebrae provide reliable bone density measures, as compared with lumbar vertebrae. The study by Wong et al. found a correlation of $r = 0.86$ in thoracic bone density versus lumbar bone density, and the study by Lenchik et al. found a correlation of $r = 0.87$. Utilizing thoracic vertebrae for BMD determination is crucial in the application of the AI prototype implemented in this study, and the aforementioned works prove the reliability of this method.
The limitations of this study are in large part due to the pathophysiology and epidemiology of osteoporosis. A maximum of 3 years between DEXA and chest CT studies was set to ensure there would be no significant change in BMD during the time between the 2 image acquisitions.13 Because of this narrow window of time, a sample size of only 65 patients was able to be procured. Larger patient populations should confirm our findings going forward. Similarly, due to the pathophysiology of osteoporotic disease, a significant proportion of the study participants were female individuals. Table 1 shows that 88% of the study participants overall and 92% and 100% of the osteopenic and osteoporotic patients, respectively, were women. With the increasing importance of osteoporosis in an aging general population, more data on AI performance in men would be desirable. Furthermore, only patients included in this study had a diagnosis of osteoporosis. Following works should attempt to include more osteoporotic patients in order to corroborate the accuracy of the prototype in differentiating between diseased and nondiseased states. In addition, these future works could consider a dichotomic analysis that would include osteopenic patients in either the diseased, or nondiseased states, as this was not carried out in the present study. Conversely, the Kruskal-Wallis test could be utilized to determine differences between individual groups for disease classification. Our convenience sample included patients who had undergone regular radiation dose diagnostic noncontrast CT of the chest. Further studies should focus on AI performance in low radiation-dose CT studies, as they are used, for example, for lung cancer screening purposes.

The prevalence and cost associated with osteoporotic disease is significant, particularly in developed nations, and there have long been measures by which to determine BMD.2–5,8 However, the current gold standard of DEXA is limited in its diagnostic scope to measures of BMD and therefore unable to contribute to other aspects of health assessment.9,10 This limited efficacy can be ameliorated by the implementation of the AI prototype discussed in this work, which functions on the basis of chest CT. The implementation of AI in analyzing chest CT studies for the evaluation of various disease processes including lung cancer, cardiovascular disease, and chronic obstructive pulmonary disease has been found efficacious in numerous works.10–21 Our results indicate that this AI prototype provides a reliable measure of BMD in moderate correlation with DEXA. Combined with other AI algorithms directed at evaluating cardiovascular and lung diseases, among others, this prototype may aid in future comprehensive preventative care to the population at large based on a single, possibly low-dose, chest CT.

REFERENCES