Automated Opportunistic Osteoporosis Screening in Routine Computed Tomography of the Spine: Comparison With Dedicated Quantitative CT

Nico Sollmann,1,2,3,4 Maximilian T. Löf,5 Malek El Husseini,1 Anjany Sekuboyina,1 Michael Dieckmeyer,1 Sebastian Rühling,1 Claus Zimmer,1,2 Bjoern Menze,6,7 Gabby B. Joseph,4 Thomas Baum,1 and Jan S. Kirschke1,2

1Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany
2TUM-Neuroimaging Center, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany
3Department of Diagnostic and Interventional Radiology, University Hospital Ulm, Ulm, Germany
4Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, CA
5Department of Diagnostic and Interventional Radiology, University Medical Center Freiburg, Freiburg im Breisgau, Germany
6Image-Based Biomedical Modeling, Department of Informatics, Technical University of Munich, Munich, Germany
7Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland

ABSTRACT
Opportunistic osteoporosis screening in nondedicated routine computed tomography (CT) is of increasing importance. The purpose of this study was to compare lumbar volumetric bone mineral density (vBMD) assessed by a convolutional neural network (CNN)-based framework in routine CT to vBMD from dedicated quantitative CT (QCT), and to evaluate the ability of vBMD and surrogate measurements of Hounsfield units (HU) to distinguish between patients with and without osteoporotic vertebral fractures (VFs). A total of 144 patients (median age: 70.7 years, 93 females) with clinical routine CT (eight different CT scanners, 120 kVp or 140 kVp, with and without intravenous contrast medium) and dedicated QCT acquired within ≤ 30 days were included. Vertebral measurements included (i) vBMD from the CNN-based approach including automated vertebral body labeling, segmentation, and correction of the contrast media phase for routine CT data (vBMD_OPP), (ii) vBMD from dedicated QCT (vBMD_QCT), and (iii) noncalibrated HU from vertebral bodies of routine CT data as previously proposed for immanent opportunistic osteoporosis screening based on CT attenuation. The intraclass correlation coefficient (ICC) for vBMD_QCT versus vBMD_OPP indicated better agreement (ICC = 0.913) than the ICC for vBMD_QCT versus noncalibrated HU (ICC = 0.704). Bland-Altman analysis showed data points from 137 patients (95.1%) within the limits of agreement (LOA) of −23.2 to 25.0 mg/cm³ for vBMD_QCT versus vBMD_OPP. Osteoporosis (vBMD < 80 mg/cm³) was detected in 89 patients (vBMD_QCT) and 88 patients (vBMD_OPP), whereas no patient crossed the diagnostic thresholds from normal vBMD to osteoporosis or vice versa. In a subcohort of 88 patients (thoracolumbar spine covered by imaging for VF reading), 69 patients showed one or more prevalent VFs, and the performance for discrimination between patients with and without VFs was best for vBMD_OPP (area under the curve [AUC] = 0.862; 95% confidence interval [CI], 0.771–0.953). In conclusion, automated opportunistic osteoporosis screening in routine CT of various scanner setups is feasible and may demonstrate high diagnostic accuracy for prevalent VFs. © 2022 The Authors. Journal of Bone and Mineral Research published by Wiley Periodicals LLC on behalf of American Society for Bone and Mineral Research (ASBMR).

KEY WORDS: bone QCT; fracture prevention; fracture risk assessment; osteoporosis; screening

Introduction
Osteoporosis is a highly prevalent skeletal disease, and a major complication is fragility fractures after low-energy trauma, which primarily manifest as vertebral fractures (VFs).1–4 Dual-energy X-ray absorptiometry (DXA) represents the widely used standard to diagnose osteoporosis and to estimate the individual fracture risk by measuring the areal bone

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Received in original form September 28, 2021; revised form May 12, 2022; accepted May 18, 2022.
Address correspondence to: Nico Sollmann, MD, PhD, Department of Diagnostic and Interventional Neuroradiology, Klinikum rechts der Isar, Technical University of Munich, Ismaninger Str. 22, 81675 Munich, Germany. E-mail: nico.sollmann@tum.de

Journal of Bone and Mineral Research, Vol. 37, No. 7, July 2022, pp 1287–1296.
DOI: 10.1002/jbmr.4575
© 2022 The Authors. Journal of Bone and Mineral Research published by Wiley Periodicals LLC on behalf of American Society for Bone and Mineral Research (ASBMR).
mineral density (aBMD). Due to many important drawbacks of the DXA technique, a frequently used alternative is quantitative computed tomography (QCT), which allows identification of osteopenic (80–120 mg/cm²) or osteoporotic subjects (<80 mg/cm²) based on the volumetric BMD (vBMD) as proposed by the American College of Radiology (ACR) for scanning at the lumbar spine. 

Osteoporosis is underdiagnosed in clinical practice worldwide. One main reason is that quantitative imaging by DXA or QCT is often only performed when fragility fractures have already occurred. An appealing option increasingly considered is opportunistic vBMD assessment in nondedicated routine CT data that have been acquired for other purposes than osteoporosis screening (eg, staging on oncologic patients). This provides the potential of identifying the undiagnosed patients early, and commencing fracture prevention therapies that otherwise would not be possible. For instance, a cutoff vBMD value of 145 mg/cm² was 100% sensitive and 57% specific for DXA-defined osteopenia.  

In another study, opportunistic CT significantly improved discrimination between patients with and without incidental VFs compared to DXA. However, extraction of vBMD in opportunistically used CT is commonly based on manual definition of only planar or small volumetric regions of interest (ROIs) not covering the entire structural information of a vertebral body, and might be user-dependent as well as time-consuming in case of measurements for multiple vertebrae or in large patient populations.

With advances in deep learning, (semi-)automated extraction of vBMD from vertebral bodies becomes possible, as demonstrated by recent retrospective investigations. A freely available framework (https://anduin.bonescreen.de) using a convolutional neural network (CNN)-based architecture has been introduced, allowing for retrospective opportunistic extraction of vBMD after fully automated vertebral body labeling and segmentation in routine noncontrast or contrast-enhanced CT. Compared to DXA-based aBMD, vBMD extracted from routine CT with this framework improved the discrimination of osteoporotic VFs (area under the curve [AUC] = 0.885 versus 0.668). However, the correlation of vBMD measurements from this CNN-based approach with results from conventional QCT has not been elucidated to date.

The hypothesis of this study was that a CNN-based framework for opportunistic osteoporosis assessment, including automated vertebral labeling, segmentation, and correction for the contrast media phase, can deliver vBMD measurements that correlate well with those from dedicated QCT scans at the lumbar spine, and that it can discriminate well between patients with and without osteoporotic VFs. Furthermore, we hypothesized that it outperforms another method of opportunistic osteoporosis screening based on noncalibrated measurements of CT attenuation values in Hounsfield units (HU), with the commonly used method of enclosing the anterior trabecular region of the vertebral bodies.

Materials and Methods

Study design and patient population

This retrospective monocentric study was approved by the local institutional review board (reference number 27/195/5R) and conducted in accordance with the Declaration of Helsinki. The requirement of written informed consent was waived due to the retrospective design of this study.

A search of the Picture Archiving and Communication System (PACS) of our tertiary care center (University Hospital rechts der Isar, Technical University of Munich, Munich, Germany; ~60,000 patients/year with hospitalized care and ~250,000 patients/year receiving outpatient treatment) was performed to identify eligible patients, which covered the interval of September 2004 to December 2019 regarding the time point of image acquisition. The inclusion criteria were as follows: (i) age >18 years; (ii) availability of conventional QCT and routine CT imaging data including the lumbar spine (acquired for other purposes than osteoporosis assessment and used herein as opportunistic imaging). Patients were excluded if they met the following exclusion criteria: (i) no spatial overlap between the vertebral bodies covered by the field of view (FOV) of QCT and routine CT imaging data; (ii) no vertebral body without implanted hardware (eg, spinal instrumentation) captured by the routine CT imaging data; (iii) QCT imaging not evaluable (eg, abortion of the QCT exam after the survey scan due to multiple lumbar vertebral fractures or degenerative disease with extensive sclerosis, rendering representative vBMD measurements impossible); (iv) no sagittal reformulation of the spine with a bone kernel (resolution of ≤3 mm) or no image data with a craniocaudal resolution of at least 1 mm available for routine CT data (as a prerequisite for the CNN-based framework used in this study); (v) time interval between QCT and routine CT imaging acquisitions of >30 days.

QCT

Dedicated QCT exams for osteoporosis assessment were performed in the supine position using four different systems (iCT 256 and iQon, Philips Healthcare; Somatom Definition AS and Somatom Sensation Cardiac 64, Siemens Healthineers) combined with standard software packages (QCT Pro, Mindways Software or Syngo Osteo CT, Siemens Healthineers). A dedicated phantom was placed below the lumbar spine. A lateral projection survey scan was first obtained, followed by scanning with the FOV including at least three vertebral bodies. ROIs to include vertebral trabecular bone while sparing cortical bone or the entry zone of the basivertebral vein were then defined, followed by extraction of level-wise vBMD values. Fractured vertebrae and vertebral bodies with foreign material were spared, aiming at vBMD measurements from at least three vertebral bodies (between T12 and L4) in each patient, if possible under the described premises.

Routine CT

Image acquisition

Routine CT scanning was performed in the supine position with eight different systems (Brilliance 64, ICT 256, Ingenuity Core 128, Ingenuity CT, and iQon, Philips Healthcare; Somatom Definition AS, Somatom Definition AS +, and Somatom Sensation Cardiac 64, Siemens Healthineers). Depending on the clinical indication for imaging, some scans were performed subsequent to administration of either both oral (Barilux Scan, Sanochemia Diagnostics) and intravenous (lomeron 400, Bracco) contrast medium or only intravenous contrast medium (n = 38 patients), with imaging acquired in the portalvenous phase (triggered by a threshold of CT attenuation surpassed in the ROI placed in the aorta or after a delay of 70 seconds). The remaining patients were scanned without administration of any contrast media (n = 106 patients). Acquisitions were performed in helical mode...
with a peak tube voltage of 120 kVp \((n = 134\) patients\) or 140 kVp \((n = 10\) patients\) and adaptive tube load.

**Assessment of VFs**

Sagittal reformations of the spine with a bone kernel were used to screen for VFs by a radiologist (5 years of experience in spine imaging), using the classification system described by Genant and colleagues.\(^2\) Furthermore, to form subgroups of patients with at least one VF (Genant grades 1, 2, or 3) or no VFs, only patients were considered who had imaging available that covered the thoracolumbar spine (at least from T6 to L5 in order to exclude missing a VF due to only minor coverage of the spine). Specifically, to be included as a patient without VFs, one of the two criteria had to be fulfilled: (i) the routine CT imaging scan used for opportunistic extraction of vBMD covered at least the region from T6 to L5 and no VF was identified; (ii) cross-sectional imaging by CT or magnetic resonance imaging (MRI) at a time point later than QCT was available that covered at least the region from T6 to L5 and clearly did not show any VFs.

**Data processing and calibration**

Labeling of vertebral bodies and segmentations derived from the routine CT scans were achieved in fully automated fashion using an in-house-developed CNN-based framework (https://anduin.bonescreen.de; Figs. 1A-D and 2A-D).\(^{17-21}\) Regarding segmentations, two masks were generated: (i) masks of the entire trabecular compartment of the vertebral bodies;\(^{17-21}\) (ii) masks of central axial slices of the anterior trabecular region of the vertebral bodies.\(^{25,26}\)

For the segmentations enclosing the entire trabecular compartment of vertebrae, asynchronous calibration was applied by converting HU to vBMD, using kVp and scanner-specific equations (vBMD = calibration factor \times\ HU mg/mL). All calibration factors were obtained by asynchronous phantom measurements with a QSA-717 phantom (QRM Quality Assurance in Radiology and Medicine GmbH) with four different hydroxyapatite inserts and ranged from 0.68 to 0.73 for 120 kVp and from 0.77 to 0.83 for 140 kVp. Further, automated detection and correction of the contrast media phase was achieved to minimize the contrast medium-induced error, which is also implemented in the CNN-based framework using a two-dimensional anatomy-guided DenseNet model (https://anduin.bonescreen.de).\(^{27}\)

The ROIs placed in the anterior trabecular region of the vertebral bodies were used to extract level-wise vertebral HU, as previously described by Pickhardt and colleagues\(^{25,26}\) as another method for imminent opportunistic osteoporosis assessment based on noncalibrated measurements of CT attenuation values using HU.

**Quality assessment and visualization**

Both QCT scans with respective selection of ROIs for vBMD extraction as well as all generated labels and segmentation masks for opportunistic use of routine CT data were checked visually by a radiologist (5 years of experience in spine imaging). Values from vertebrae with severe degenerative changes (eg, Modic type 3 endplate changes), implanted hardware (eg, due to spinal instrumentation), or from fractured vertebrae were not considered for further analyses regarding extraction of vBMD or HU, which were restricted to the vertebral bodies that were captured by the FOV in both the QCT and routine CT (one to three vertebral bodies per patient).

For data derived from the CNN-based framework, reconstructions passing through the center of mass of the vertebral bodies.

---

**Fig. 1.** Automated vertebral body labeling and segmentation – I. CT scan of a 69-year-old man visualized as virtual radiograph in lateral projection (A) and as planar reconstructions with segmentation masks enclosing the vertebral bodies in lateral views (B-D). Labeling and automated segmentation of vertebral bodies \(T_7-L_5\) was achieved using a CNN-based framework. The segmentation of \(L_1\) is shown in D for axial, coronal, and lateral views, with the bottom right image showing only the trabecular compartment being enclosed by a segmentation mask. CNN = convolutional neural network; CT = computed tomography.
were generated from routine CT data and overlaid with segmentation masks, supplemented by generation of virtual radiographs in lateral projection (Figs. 1A-D and 2A-D). Any patient cases with segmentation masks generated by the CNN-based approach that did not entirely and correctly enclose the trabecular component of vertebral bodies (eg, due to Schmorl nodes, Fig. 2).

**Fig. 2.** Automated vertebral body labeling and segmentation – II. CT scan of a 72-year-old woman visualized as virtual radiograph in lateral projection (A) and as planar reconstructions with segmentation masks enclosing the vertebral bodies in lateral views (B,C). Labeling and automated segmentation of vertebral bodies (T7–L3) was achieved using a CNN-based framework. The segmentation of L3 is shown in D for axial, coronal, and lateral views, with the bottom right image showing only the trabecular compartment being enclosed by a segmentation mask. The lateral images show osteoporotic VFs of T7 and T8. CNN = convolutional neural network; CT = computed tomography; VF = vertebral fracture.

**Fig. 3.** Flowchart for study inclusion. This figure depicts the inclusion and exclusion criteria applied to arrive at the cohort size of this study (n = 144 patients). For differentiation between patients with and without osteoporotic VFs, the subgroup of patients who had the thoracolumbar spine covered by imaging was considered (n = 88 patients). VF = vertebral fracture.
fused vertebral bodies, or thoracolumbar or lumbosacral transitional anatomy) were excluded from extraction of vBMD or HU.

Statistics

Statistical analysis was performed using STATA (version 16; StataCorp LP) and Prism (version 6; GraphPad Software Inc.). Descriptive statistics were calculated for patient characteristics and vBMD from QCT (vBMD_QCT), vBMD from the opportunistic CNN-based approach using routine CT data (vBMD_OPP), and for noncalibrated measurements of CT attenuation in HU. In each patient, values from different spinal levels were averaged to provide one mean value for each parameter per patient. Furthermore, vBMD_QCT was plotted against vBMD_OPP, followed by calculation of the line of best fit with 95% confidence intervals (CIs) based on linear regression. In addition, vBMD_QCT was plotted against vBMD_OPP also in relation to scanner type and manufacturer (eight different scanners from two manufacturers), software package used for evaluation of QCT data (two different software packages), application of an intravenous contrast medium (noncontrast or contrast-enhanced imaging in portalvenous phase), and peak tube voltage (120 kVp or 140 kVp).

The intraclass correlation coefficient (ICC) was calculated between vBMD_QCT as the reference standard of this study and vBMD_OPP as well as noncalibrated HU values. Additionally, the root mean square coefficient of variation (RMSCV) was calculated based on the vBMD_QCT and vBMD_OPP data as well as HU values. A Bland-Altman analysis was performed to assess the agreement between vBMD_QCT and vBMD_OPP. The 95% CI was used to determine the limits of agreement (LOA; mean difference ± 1.96 standard deviation [SD] of the difference). Furthermore, using the diagnostic criteria of the ACR based on vBMD ranges, the vBMD_QCT and vBMD_OPP data were used for categorization into patients with osteoporosis (<80 mg/cm³), osteopenia (80–120 mg/cm³), and normal vBMD (>120 mg/cm³). (6,8)

Receiver operating characteristic (ROC) analysis was performed, using grading into patients without and patients with VFs (Genant grades 1, 2, and 3) combined with the vBMD_QCT, vBMD_OPP, and HU values. The ROC analysis was performed in the subcohort of patients with imaging available that covered the thoracolumbar spine (at least from T₆ to L₅). In this subcohort, the vBMD_QCT, vBMD_OPP, and noncalibrated HU values were further compared between patients with and without VFs. (Genant grades 1, 2, and 3) combined with the vBMD_QCT, vBMD_OPP, and HU values. The ROC analysis was performed in the subcohort of patients with imaging available that covered the thoracolumbar spine (at least from T₆ to L₅). In this subcohort, the vBMD_QCT, vBMD_OPP, and noncalibrated HU values were further compared between patients with and without

Fig. 4. This figure shows the number of detected VFs as box plots (median / interquartile ranges) with minimum and maximum whiskers (individual data point of each patient shown as a dot) for all grades according to the Genant classification together (Genant grades 1–3), as well as for Genant grades 1, 2, and 3 separately (A). Furthermore, the absolute numbers of VFs per vertebral level (T₆–L₅) and in relation to the Genant classification (Genant grades 1–3) are plotted (B). VF = vertebral fracture.

Fig. 5. vBMD (in mg/cm³) from QCT versus opportunistic CT – I. Analysis of dedicated QCT (vBMD_QCT; x-axis) versus opportunistic routine CT (vBMD_OPP, y-axis) showed a line of best fit $y = 0.95x + 2.80$, with a 95% CI for the slope of 0.878 to 1.017 ($R^2 = 0.834$). CI = confidence interval; QCT = quantitative computed tomography; vBMD = volumetric bone mineral density.
osteoporotic VFs using Mann-Whitney tests. A p value <0.05 was considered statistically significant.

**Results**

Overall, 144 patients were analyzed in this study (93 females and 51 males, median age: 70.7 years, range: 20.5–89.3 years; Fig. 3). The median time interval between QCT and routine CT imaging was 3 days (range: 0–30 days).

A total number of 318 vertebral bodies were analyzed for vBMD and HU extraction, with each patient contributing with one to three vertebral bodies. Included vertebral bodies were T12 (n = 22 patients), L1 (n = 75 patients), L2 (n = 84 patients), L3 (n = 92 patients), and L4 (n = 45 patients). According to VF reading, 69 patients showed at least one VF, whereas 19 patients did not show any VFs. The remaining 56 patients had no imaging available that captured the thoracolumbar spine (at least from T6 to L3). Details regarding the average number and categorization of VFs are shown in Figs. 4A and 4B.

The mean vBMD for QCT was slightly higher than the mean vBMD opportunistically derived from routine CT (mean vBMD_QCT ± SD: 70.6 ± 28.8 mg/cm³, range: 16.9–141.0 mg/cm³; mean vBMD_OPP ± SD: 69.6 ± 29.9 mg/cm³, range: 18.7–160.7 mg/cm³). The ICC for vBMD_QCT and vBMD_OPP was higher (ICC = 0.913; 95% CI, 0.881–0.936) than

---

**Fig. 6.** vBMD (in mg/cm³) from QCT versus opportunistic CT – II. Analysis of dedicated QCT (vBMD_QCT; x-axis) versus opportunistic routine CT (vBMD_OPP, y-axis) depending on the scanner type and manufacturer (A), software package used for evaluation of QCT data (B), application of contrast media (C), and the peak tube voltage used during scanning (D). Depending on the scanner type and manufacturer, the lines of best fit were

\[ y = 1.01x - 3.39 \] (Brilliance 64), \[ y = 0.59x + 36.97 \] (ICT 256), \[ y = 0.73x + 17.21 \] (Ingenity Core 128), \[ y = 0.91x + 223.20 \] (Ingenity CT), \[ y = 0.86x + 7.64 \] (iCT 256), \[ y = 0.91x + 4.51 \] (Sensation Cardiac 64), \[ y = 0.95x + 10.06 \] (Somatom Definition AS), and \[ y = 1.36x - 12.96 \] (Somatom Definition AS+). For vBMD_OPP, values from scanners by Philips Healthcare are shown in green, and values from scanners by Siemens Healthineers are shown in orange (A). For the two softwares used for QCT evaluation, the lines of best fit were \[ y = 0.98x + 1.00 \] (Syngo Osteo CT, green) and \[ y = 0.85x + 8.87 \] (QCT Pro Mindways, orange) (B). Depending on the administration of contrast media, the lines of best fit were \[ y = 0.91x + 5.86 \] (noncontrast imaging, green) and \[ y = 1.07x - 6.68 \] (contrast-enhanced imaging in portalvenous phase, orange) (C). Depending on the peak tube voltage, the lines of best fit were \[ y = 0.94x + 3.63 \] (120 kVp, green) and \[ y = 1.10x - 13.05 \] (140 kVp, orange) (D). QCT = quantitative computed tomography; vBMD = volumetric bone mineral density.
the ICC for vBMD_QCT and noncalibrated HU values (ICC = 0.704; 95% CI, 0.050–0.882). Furthermore, the RMSCV amounted to 14.4% (vBMD_QCT versus vBMD_OPP) and to 24.3% (vBMD_QCT versus noncalibrated HU values). The line of best fit for vBMD_QCT against vBMD_OPP was given by \( y = 0.95x + 2.80 \) (95% CI of slope, 0.878–1.017; \( R^2 = 0.834 \); Fig. 5). Figure 6 plots vBMD_QCT against vBMD_OPP in relation to scanner type (Fig. 6A), software used for QCT evaluation (Fig. 6B), administration of intravenous contrast media (Fig. 6C), and peak tube voltage (Fig. 6D).

Regarding the Bland-Altman analysis, data points from 137 patients (95.1%) were within the LOA of −23.2 to 25.0 mg/cm³, with the data points of the remaining seven patients (4.9%) being located outside of the LOA (Fig. 7). In six of these patients, vBMD_OPP values were clearly higher than the respective vBMD_QCT values. In retrospective case review, these patients showed spondylophytes included in the segmentation masks.

Using the diagnostic criteria of the ACR for categorization into normal, osteopenic, or osteoporotic vBMD values, osteoporosis was detected in 89 patients for vBMD_QCT and in 88 patients for vBMD_OPP, respectively. Regarding vBMD_QCT, a normal vBMD was detected in nine patients and values in the range of osteopenia were detected in 46 patients. Based on vBMD_OPP, six patients were considered normal and 50 patients were categorized as osteopenic. None of the included patients crossed the diagnostic thresholds from normal vBMD to osteoporosis or vice versa depending on opportunistic or dedicated QCT measurements.

Patients with VFs showed significantly lower vBMD_QCT (mean vBMD_QCT ± SD: 67.4 ± 27.3 mg/cm³ versus 98.2 ± 20.6 mg/cm³, \( p < 0.001 \)), vBMD_OPP (mean vBMD_OPP ± SD: 66.4 ± 25.9 mg/cm³ versus 105.0 ± 25.7 mg/cm³, \( p < 0.001 \)), and noncalibrated HU values (mean HU ± SD: 89.5 ± 37.7 versus 133.6 ± 29.1, \( p < 0.001 \)). For discrimination between patients with and without VFs, vBMD_OPP showed the best performance (AUC = 0.862; 95% CI, 0.771–0.953; Fig. 8B). The performance of vBMD_QCT and noncalibrated HU values was similar (AUC for vBMD_QCT = 0.815; 95% CI, 0.721–0.908; AUC for noncalibrated HU = 0.824; 95% CI, 0.728–0.920; Figs. 8A and C).

Discussion

We used a CNN-based approach on routine CT scans to enable opportunistic vBMD measurements at the spine, implementing fully automated vertebral body labeling, segmentation, and correction for the contrast media phase with subsequent manual filtering steps (for removal of vertebrae with severe degenerative

Fig. 7. Bland-Altman plot. The difference and average for vBMD (in mg/cm³) based on Bland-Altman analysis using the opportunistic measurements from routine CT as well as dedicated QCT are plotted in this graph, together with the neutral line and two lines defining the LOA based on a 95% CI, −23.2 to 25.0 mg/cm³. CI = confidence interval; LOA = limits of agreement; QCT = quantitative computed tomography; vBMD = volumetric bone mineral density.

Fig. 8. ROC curve. Analysis with a ROC curve was performed to discriminate between patients with and without osteoporotic VFs, using a subcohort of patients who had the thoracolumbar spine covered by imaging. The AUC was calculated for vBMD from dedicated quantitative computed tomography (vBMD_QCT) (A), opportunistic CT (vBMD_OPP) (B), and for noncalibrated HU (C). The vBMD_OPP showed the best performance (AUC = 0.862; 95% CI, 0.771–0.953). The performance of vBMD_QCT and noncalibrated HU values was similar (AUC for vBMD_QCT = 0.815; 95% CI, 0.721–0.908; AUC for noncalibrated HU = 0.824; 95% CI, 0.728–0.920). AUC = area under the curve; HU = Hounsfield units; ROC = receiver operating characteristic; QCT = quantitative computed tomography; vBMD = volumetric bone mineral density; VF = vertebral fracture.
changes, fractures, or implanted hardware). This approach was compared to dedicated QCT as well as noncalibrated measurements of CT attenuation values in HU. The main findings were as follows: (i) vBMD_QCT showed a better correlation with vBMD_OPP than with noncalibrated HU measurements; (ii) discrimination between patients with and without osteoporotic VFs was slightly better when using vBMD_OPP compared to vBMD_QCT or noncalibrated HU values.

Opportunistic assessment of vBMD is of high interest for the clinical setting, given that it can save additional exams by DXA or dedicated QCT for the purpose of osteoporosis screening and fracture risk assessment, thus reducing imaging caseloads, costs, and radiation exposure for the patient. Therefore, sparing dedicated DXA or QCT would require that opportunistic vBMD assessments are reliable and accurate, particularly given the common clinical environment with multiple scanners and different scanning protocols in the light of imaging indications other than related to osteoporosis. Previous work using an entirely different CNN-based network reported on significant correlations between aBMD values obtained with the used network and DXA for internal and external validation datasets ($r = 0.852$ and $0.840$), yet only used unenhanced CT acquisitions. Further, it trained CNNs on DXA and CT data, which may, however, propagate the inaccuracies of DXA to CT measures. Another CNN-based approach indicated high correlations ($r > 0.98$) of vBMD values with those obtained with QCT, yet with rather homogeneous scanning protocols and without explicitly incorporating dedicated corrections for the contrast media phase. Further, for low-dose chest CT scans obtained for lung cancer screening to extract vBMD of thoracic and upper lumbar vertebrae using deep learning, mean errors of 2.2 to 4.0 mg/cm$^3$ compared to dedicated QCT were observed. Although these results are promising, they stem from a highly selective cohort with imaging being performed only on one CT system and with a similar scanning protocol without application of a contrast agent. Hence, these findings may not be generalizable to other scanning environments with different CT systems and imaging protocols. In the present study, a CNN-based framework was used on routine CT data gathered from multiple scanners (eight CT systems from two different vendors) and variable scanning protocols (peak tube voltage of 120 kVp or 140 kVp) with and without application of intravenous contrast media, thus providing an approach for the immanent routine clinical scenario. In this regard, the better agreement between vBMD_QCT as the reference method and vBMD_OPP compared to noncalibrated measurements of HU values is a necessary step towards more accurate opportunistic osteoporosis assessment. Particularly the presence of contrast media has been considered a relevant issue for opportunistic use of CT data; yet, the herein used algorithm accounts for this by automated detection of the contrast media phase with high accuracy, thus minimizing the contrast medium–induced error in vBMD assessments.

Most previous work using deep learning–based opportunistic osteoporosis screening has not investigated the performance of derived measurements for differentiating between patients with and without VFs. According to the results of this study, discrimination between patients with and without VFs based on vBMD_OPP was non-inferior in relation to vBMD_QCT. Specifically, it showed a similar AUC as compared to a recent study that used a previous version of the same CNN-based algorithm (AUC = 0.885 for trabecular vBMD), yet in the primary setting of comparing results with DXA in a different cohort. In this regard, the cohort of the present study was constituted of elderly patients that were suspected to suffer from osteoporosis, given that clinical indication for QCT was made. This also implicated that a large proportion of included patients showed at least one VF. Yet, despite this complex cohort with a high prevalence of VFs and related impairments of vertebral body shape that also affected the alignment and curvature of the spine, only 11 patients were excluded from the analyses due to incorrect segmentations with respect to the vertebral bodies of interest (between T12 and L4). For such cases, the used CNN-based approach has the option for manual corrections when considered necessary.

In addition to opportunistic vBMD extractions using our CNN-based framework, we performed noncalibrated measurements of HU derived from ROI definition in the anterior trabecular region of the vertebral bodies, which was previously implemented by Pickhardt and colleagues as a method for opportunistic screening using CT attenuation values. The approach is straightforward to use and can be applied directly within the PACS, yet does inherently and intentionally not use any compensation for bias due to tilting of vertebrae (eg, due to scoliosis), contrast media, or variations in the scanning protocol. We speculate that due to such factors, the agreement between noncalibrated measurements of HU and vBMD_QCT was inferior to that of vBMD_OPP and vBMD_QCT. Importantly, is that HU values are scanner-specific, thus attenuation thresholds that are used for osteoporosis screening cannot be obtained from one and then applied to another scanner. In this regard, it has to be stated that the HU cutoff value (cutoff for osteoporosis of 145 HU) has not been published as an official guideline to diagnose osteoporosis. Further, it has been derived as an average for T12 to L5 (threshold for 100% sensitivity for osteoporosis) and from scanning using 120 kVp; yet, thresholds may change due to the impact of peak tube voltage. Based on the results of the present study, noncalibrated measurements of HU values should be avoided for osteoporosis assessment, and automated pipelines as used in this study may facilitate swift and more accurate vBMD extraction from CT even during clinical routine in the near future.

Although agreement between vBMD_QCT and vBMD_OPP was better than between vBMD_QCT and noncalibrated measurements of HU and no patient crossed the diagnostic thresholds from normal vBMD to osteoporosis or vice versa (depending on opportunistic or dedicated QCT measurements), the LOA nevertheless spanned a rather broad range of values. A major reason for this may relate to the definition of the ROIs for extraction of vBMD_QCT and vBMD_OPP, which were not the same for opportunistic use of routine CT data with our CNN-based approach and the dedicated QCT exams. If the ROI for vBMD_OPP extraction is chosen so that it closely resembles that of vBMD_QCT extraction, the LOA might be narrowed; however, a part of the ROI definition procedure for QCT is done manually for both QCT software packages considered in this study, which is not fully standardized and, thus, we were not able to exactly reproduce the ROIs of the dedicated QCT exams. Furthermore, we refrained from manual corrections of segmentation masks produced by the CNN-based approach, hence these masks may have included also degenerative changes (eg, spondylophytes) that could have raised vBMD_OPP values in some patients, resulting in widened LOA. This approach has been chosen intentionally to reduce manual interference with the algorithm as far as possible for the purpose of fully automated vertebral body labeling and segmentation in routine CT data.
However, it needs to be emphasized that manual correction of labels and segmentation masks in particular is seamlessly possible within the CNN-based application, thus could principally be performed for some patients to optimize ROI definitions and the LOA.

There are limitations to this retrospective study. First, there was a patient selection bias toward elderly patients, which is related to the requirement of availability of both routine CT and dedicated QCT. On the one hand, osteoporosis has thus already been suspected in the majority of included patients, which resembles the common clinical scenario in most centers where dedicated QCT is only requested when osteoporosis has already been suspected or VFs have already occurred. Yet, on the other hand, this patient population may be the primary use case for our CNN-based approach because alternative approaches such as extracting noncalibrated measurements of HU may become particularly prone to inaccuracies related to spinal degeneration. Second, the subcohort of patients without VFs was considerably small, which is related to the study design where dedicated QCT had to be available, which mostly existed for patients with high suspicion of osteoporosis or sustained previous VFs. Third, we have not been able to homogeneously extract vBMD from L1 to L3 in all patients, which would have been the most common reference levels. This is related to exclusion of vertebrae for vBMD and HU extraction on an individual basis due to various causes such as lumbar spinal hardware, VFs, or severe degenerative changes. However, this resembles the common scenario in clinical routine. Fourth, this study and previous work have in common that existing CT data were used to retrospectively extract vBMD or noncalibrated HU \cite{12,21,25,26}; yet, the transition to prospective study designs and ultimate use in clinical settings needs to be achieved to fully exploit the advantages of opportunistic osteoporosis screening.

In conclusion, this study used an automated CNN-based framework for labeling and segmentation of vertebral bodies for opportunistic extraction of vBMD, implementing correction for the presence of contrast media and the distinct CT system and scanning protocol used. Better agreement between vBMD values from dedicated QCT and those derived from the CNN-based approach using routine CT data was observed when compared to noncalibrated measurements of HU, whereas our CNN-based approach performed slightly better in discriminating between patients with and without VFs compared to dedicated QCT as well as noncalibrated HU measurements. Thus, opportunistic osteoporosis screening and fracture risk assessment for routine CT of the spine could become feasible using vBMD with fully automated vertebral body labeling, segmentation, and correction for the contrast medium phase, even for the common clinical scenario with multiple scanners and varying scanning protocol environments. In contrast, noncalibrated measurements of HU values should be avoided for these purposes.

Acknowledgments

The present work was supported by the European Research Council (grant agreement No. 963904 – Bonescreen: to JSK), the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG; project 432290010: to TB and JSK), the B. Braun Foundation (project BBST-D-19-00106: to NS), and the Joachim Herz Foundation (to NS). Open Access funding enabled and organized by Projekt DEAL.

Author Contributions

Nico Sollmann: Conceptualization; data curation; formal analysis; funding acquisition; investigation; resources; validation; visualization; writing – original draft; writing – review and editing. Maximilian T. Löffler: Data curation; investigation; methodology; resources; software; writing – review and editing. Malek El Husseini: Conceptualization; investigation; methodology; resources; software; writing – review and editing. Anjany Sekuboyina: Investigation; methodology; resources; software; writing – review and editing. Sebastian Rühling: Data curation; investigation; methodology; validation; writing – review and editing. Claus Zimmer: Funding acquisition; investigation; project administration; resources; supervision; writing – review and editing. Bjoern Menze: Conceptualization; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; writing – review and editing. Gabby B. Joseph: Data curation; formal analysis; software; validation; writing – review and editing. Thomas Baum: Conceptualization; data curation; formal analysis; funding acquisition; investigation; project administration; resources; writing – review and editing. Jan S. Kirschke: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; writing – review and editing.

Conflicts of Interest

Jan S. Kirschke and Anjany Sekuboyina are co-founders of Bonescreen GmbH. The other authors have no conflicts of interest to report.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Sanchez-Riera L, Wilson N, Kamalaraj N, et al. Osteoporosis and fragility fractures. Best Pract Res Clin Rheumatol. 2010;24(6):793-810.
2. Schousboe JT. Epidemiology of vertebral fractures. J Clin Densitom. 2016;19(1):8-22.
3. Compston JE, McClung MR, Leslie WD. Osteoporosis. Lancet. 2019; 393(10169):364-376.
4. Hernlund E, Svedbom A, Ivergard M, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos. 2013;8:136.
5. Jain RK, Vokes T. Dual-energy X-ray absorptiometry. J Clin Densitom. 2017;20(3):291-303.
6. Löffler MT, Sollmann N, Mei K, et al. X-ray-based quantitative osteoporosis imaging at the spine. Osteoporos Int. 2019;31(2):233-250.
7. Link TM, Kazakia G. Update on imaging-based measurement of bone mineral density and quality. Curr Rheumatol Rep. 2020;22(5):13.
8. American College of Radiology. ACR-SPR-SSR practice parameter for the performance of musculoskeletal quantitative computed tomography (QCT). American College of Radiology, Reston. 2018. https://www.acr.org/-/media/ACR/Files/Practice-Parameters/QCT.pdf?la. Accessed May 30, 2022.
9. Chesnut CH 3rd. Osteoporosis, an underdiagnosed disease. JAMA. 2001;286(22):2865-2866.
10. Loffler MT, Kallweit M, Niederreiter E, et al. Epidemiology and reporting of osteoporotic vertebral fractures in patients with long-term hospital records based on routine clinical CT imaging. Osteoporos Int. 2022;33(3):685-694.
11. Boutin RD, Lenchik L. Value-added opportunistic CT: insights into osteoporosis and sarcopenia. AJR Am J Roentgenol. 2020;215(3):582-594.
12. Weaver AA, Beavers KM, Hightower RC, Lynch SK, Miller AN, Stitzel JD. Lumbar bone mineral density phantomless computed tomography measurements and correlation with age and fracture incidence. Traffic Inj Prev. 2015;16(Suppl 2):S153-S160.
13. Loffler MT, Jacob A, Valentinitsch A, et al. Improved prediction of incident vertebral fractures using opportunistic QCT compared to DXA. Eur Radiol. 2019;29(9):4980-4989.
14. Fang Y, Li W, Chen X, et al. Opportunistic osteoporosis screening in multi-detector CT images using deep convolutional neural networks. Eur Radiol. 2020;30(7):4107-4116.
15. Pan Y, Shi D, Wang H, et al. Automatic opportunistic osteoporosis screening using low-dose chest computed tomography scans obtained for lung cancer screening. Eur Radiol. 2020;30(6):3549-3557.
16. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Ake O. Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network. Eur Radiol. 2020;30(6):3549-3557.
17. Sekuboyina A, Husseini ME, Bayat A, et al. VerSe: a vertebrae labelling and segmentation benchmark for multi-detector CT images. Med Image Anal. 2021;73:102166.
18. Loffler MT, Sekuboyina A, Jacob A, et al. A vertebral segmentation dataset with fracture grading. Radiol Artif Intell. 2020;2(4):e190138.
19. Loffler MT, Jacob A, Schar A, et al. Automatic opportunistic osteoporosis screening in routine CT: improved prediction of patients with prevalent vertebral fractures compared to DXA. Eur Radiol. 2021;31(8):6069-6077.
20. Loffler MT, Sollmann N, Burian E, et al. Opportunistic osteoporosis screening reveals low bone density in patients with screw loosening after lumbar semi-rigid instrumentation: a case-control study. Front Endocrinol (Lausanne). 2020;11:552719.
21. Ruhling S, Scharr A, Sollmann N, et al. Proposed diagnostic volumetric bone mineral density thresholds for osteoporosis and osteopenia at the cervicothoracic spine in correlation to the lumbar spine. Eur Radiol. Published online ahead of print April 6, 2022. doi: 10.1007/s00330-022-08721-7.
22. Adams JE. Quantitative computed tomography. Eur J Radiol. 2009;71(3):415-424.
23. Engelke K, Adams JE, Armbrecht G, et al. Clinical use of quantitative computed tomography and peripheral quantitative computed tomography in the management of osteoporosis in adults: the 2007 ISCD Official Positions. J Clin Densitom. 2008;11(1):123-162.
24. Genant HK, Wu CY, van Kuijk C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res. 1993;8(9):1137-1148.
25. Pickhardt PJ, Lee LJ, del Rio AM, et al. 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(9):2194-2203.
26. 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(8):588-595.
27. Ruhling S, Navarro F, Sekuboyina A, et al. Automated detection of the contrast phase in MDCT by an artificial neural network improves the accuracy of opportunistic bone mineral density measurements. Eur Radiol. 2022;32(3):1465-1474.
28. Garner HW, Paturzo MM, Gaudier G, Pickhardt PJ, Wessell DE. Variation in attenuation in L1 trabecular bone at different tube voltages: caution is warranted when screening for osteoporosis with the use of opportunistic CT. AJR Am J Roentgenol. 2017;208(1):165-170.