European Radiology (2020) 30:3549–3557
https://doi.org/10.1007/s00330-020-06677-0

IMAGING INFORMATICS AND ARTIFICIAL INTELLIGENCE

Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network

Koichiro Yasaka[1] · Hiroyuki Akai[1] · Akira Kunimatsu[1] · Shigeru Kiryu[2] · Osamu Abe[3]

Received: 14 November 2019 / Revised: 1 January 2020 / Accepted: 27 January 2020 / Published online: 14 February 2020

Abstract

Objectives To investigate whether a deep learning model can predict the bone mineral density (BMD) of lumbar vertebrae from unenhanced abdominal computed tomography (CT) images.

Methods In this Institutional Review Board–approved retrospective study, patients who received both unenhanced CT examinations and dual-energy X-ray absorptiometry (DXA) of the lumbar vertebrae, in two institutions (1 and 2), were included. Supervised deep learning was employed to obtain a convolutional neural network (CNN) model using axial CT images, including the lumbar vertebrae as input data and BMD values obtained with DXA as reference data. For this purpose, 1665 CT images from 183 patients in institution 1, which were augmented to 99,900 (= 1665 × 60) images (noise adding, parallel shift and rotation were performed), were used. Internal (by using data of 45 other patients in institution 1) and external validations (by using data of 50 patients in institution 2) were performed to evaluate the performance of the trained CNN model. Correlations and diagnostic performances were evaluated with Pearson's correlation coefficient (r) and area under the receiver operating characteristic curve (AUC), respectively.

Results The estimated BMD values, according to the CNN model (BMD_{CNN}), were significantly correlated with the BMD values obtained with DXA ($r = 0.852$ ($p < 0.001$) and 0.840 ($p < 0.001$) for the internal and external validation datasets, respectively). Using BMD_{CNN}, osteoporosis was diagnosed with AUCs of 0.965 and 0.970 for the internal and external validation datasets, respectively.

Conclusions Using deep learning, the BMD of lumbar vertebrae could be predicted from unenhanced abdominal CT images.

Key Points

- *By applying a deep learning technique, the bone mineral density (BMD) of lumbar vertebrae can be estimated from unenhanced abdominal CT images.*
- *A strong correlation was observed between the estimated BMD and the BMD obtained with DXA.*
- *By using the estimated BMD, osteoporosis could be diagnosed with high performance.*

Keywords Artificial intelligence · Bone mineral density · Multidetector computed tomography · Deep learning · Osteoporosis

✉ Koichiro Yasaka
koyasaka@gmail.com

[1] Department of Radiology, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

[2] Department of Radiology, Graduate School of Medical Sciences, International University of Health and Welfare, 537-3 Iguchi, Nasushiobara, Tochigi 329-2763, Japan

[3] Department of Radiology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

Abbreviations

AUC	Area under the receiver operating characteristic curve
BMD	Bone mineral density
BMD_{CNN}	Bone mineral density obtained with a convolutional neural network
CNN	Convolutional neural network
CT	Computed tomography
DXA	Dual-energy X-ray absorptiometry
DICOM	Digital imaging and communications in medicine
ROC	Receiver operating characteristic
ROI	Region of interest

Introduction

Vertebral fractures and hip fractures are associated with hospitalisation and disability and are major causes of morbidity. Patients with these fractures have been demonstrated to be at increased risk of mortality [1]. In Europe, these fractures were reported to account for 1,180,000 quality-adjusted life years that were lost during 2010 [2]. Therefore, the prevention of these fractures is important. Osteoporosis, which is characterised by reduced bone mineral density (BMD), is one of the causes of these fractures. Therefore, knowing the BMD of patients could help physicians to prevent these fractures. Elderly patients tend to have lower BMD and, in this era of global population ageing, the prevalence of osteoporosis has been increasing. In 2010, 27.5 million Europeans were reported to have osteoporosis [3]. Dual-energy X-ray absorptiometry (DXA) enables estimation of BMD. Though measurements with this method is affected by atherosclerosis of aorta and sclerosis of vertebra, DXA is used most commonly to evaluate the BMD, and it is recommended by guidelines [3–5]. Patients diagnosed with osteoporosis are managed to prevent fragile fractures by recommending adequate dietary intake, fall prevention measures and/or pharmacological interventions. However, DXA is not necessarily readily available in every institution [3].

Computed tomography (CT) allows detailed evaluations of several diseases. As the number of elderly patients who tend to have multiple diseases is increasing throughout the world, CT examinations have been utilised increasingly. In 2015, 143.1 CT examinations per 1000 in the population were reported to be performed in the Organisation for Economic Co-operation and Development countries [6]. It is not rare that vertebral fractures, which can sometimes occur without causing serious symptoms, are incidentally detected at CT examinations performed for evaluations of other diseases [7]. If the BMD can be predicted from CT images, it will allow physicians to make appropriate strategies to prevent fragile fractures, which will benefit many patients.

Since Krizhevsky et al showed that deep learning with a convolutional neural network (CNN) can achieve high performance levels in visual recognition tasks [8], the application of deep learning to radiological imaging diagnosis has been gaining wide attention [9–11]. Successful applications of a deep learning technique to the evaluation of radiological images have been reported, such as image processing [12], the differential diagnosis of diseases [13, 14] and the detection of diseases [15–17]. This technique has also been applied to obtain parameters from radiological images (e.g. liver fibrosis stages from CT and magnetic resonance images [18, 19] and bone age estimation from plain radiographs of the hand [20]). We hypothesised that the BMD can be estimated from CT images with deep learning.

The purpose of this study was to investigate whether the BMD of the lumbar vertebrae can be estimated from unenhanced abdominal CT scans using a deep learning method.

Materials and methods

This retrospective study was approved by the Institutional Review Board and the requirement for obtaining written informed consent from patients was waived.

Overview of this study

In this study, supervised training of a CNN was performed using axial unenhanced abdominal CT images at the L1–4 level as input data and the BMD of the corresponding lumbar vertebrae measured with DXA as reference data. To manage the potential of an overfitting problem associated with deep learning, the CNN was trained with training datasets and the performance of the trained CNN was evaluated with internal and external validation datasets (hereafter, we denote these as validation datasets I and E, respectively) that were not included in the training datasets.

Patients

Patients who underwent unenhanced CT examinations (including any of L1, L2, L3 or L4 for training datasets; including all the L1, L2, L3 and L4 for validation datasets) within 500 days from the DXA examination were included in this study. CT examinations performed from April 2015 to September 2015 and those from October 2017 to April 2018 in institution 1 were used as the training dataset and validation dataset I, respectively. CT examinations performed from January 2018 to February 2018 in institution 2 were used as validation dataset E. The exclusion criteria were the following: patients with a history of lumbar surgery (including surgery for fracture), with severe scoliosis and with vertebrae that had a compression fracture or deformity (due to spondylosis) in L1–4 [5]. As for the validation group, patients whose CT images were included in the training datasets were also excluded.

CT scanning technique

For the training dataset and validation dataset I (institution 1), CT examinations were performed with Aquilion ONE (Canon Medical Systems). For validation dataset E, the following CT scanners were used: Aquilion PRIME, Aquilion Precision, Aquilion ONE (Canon Medical Systems) and Discovery CT 750 HD (GE Healthcare). The details of the scanning and image reconstruction parameters are shown in Table 1. Because the

Table 1 Scanning and reconstruction parameters of CT examinations

	Institution 1 (training dataset and validation dataset I)	Institution 2 (validation dataset E)	
Vendor	Canon Medical Systems ($n = 228$)	Canon Medical Systems ($n = 28$)	GE Healthcare ($n = 22$)
CT scanner	Aquilion ONE ($n = 228$)	Aquilion PRIME ($n = 16$) Aquilion Precision ($n = 8$) Aquilion ONE ($n = 3$) Aquilion ($n = 1$)	Discovery CT750 HD ($n = 22$)
Tube voltage (kVp)	120	120	120
Tube current (automatic exposure control)	SD = 13.0	SD = 13.0	Noise index = 11.36
Helical pitch	0.828 ($n = 228$)	0.8125 ($n = 27$) 0.828 ($n = 1$)	1.375 ($n = 16$) 0.984 ($n = 6$)
Gantry rotation time (s)	0.5	0.5	0.5
Reconstruction algorithm	Filtered back projection	Filtered back projection	Filtered back projection
Kernel for reconstruction	FC03	FC03	Standard
Slice thickness (mm)	5.0	5.0	5.0
Field of view (mm)	360–500 mm	340–500 mm	350–400 mm

aim of this study was to find reduced BMD status in opportunistic CT, the bone mineral equivalent phantom was not used.

Input image data

We used a cropped axial unenhanced CT image of the lumbar vertebrae as input data for the CNN (Fig. 1). From Picture Archiving and Communication System, CT images were extracted in the digital images and communications in medicine (DICOM) format. For the training datasets, up to 3 axial CT images were obtained from each lumbar vertebra at and near the mid-vertebral level, so that the model became robust against a selection of axial slice levels. For the validation datasets I and E, only 1 CT image was obtained from each lumbar vertebra at the mid-vertebral level. For example, if a vertebra was included in 5 axial CT images, the second to fourth images and the third image were included in the training dataset and the validation datasets, respectively.

Then, the DICOM format images were preprocessed with the Python 3.6.4 programming language (https://www.python.org/), pydicom 1.2.2 package (https://pydicom.github.io/pydicom/stable/index.html) and pillow 5.0.0 package (https://pillow.readthedocs.io/en/stable/#) on the command line. Regions around the vertebrae were cropped using a crop box that was 250 × 250 pixels (Fig. 1). These imaging data were resized to 96 × 96 pixels.

For the training dataset, augmentation was performed so that the models become robust against differences in the scanning condition (slight shift and rotation of the patients' positions and image noise). CT image data cropped with 15 parallel shifted crop boxes were generated. From these imaging data, rotated images (with 5 and 355 degrees; 15 × 2 images) and noise-added images (mean = 0, sigma = 15; 15 images) were also generated. Therefore, from one DICOM format CT image, 60 input image datasets (= 15 + 15 × 2 + 15 images) were generated by augmentation for the training dataset.

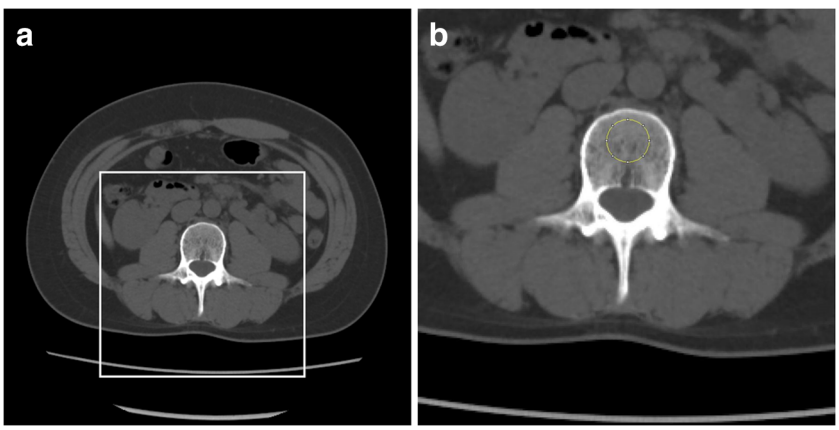

Fig. 1 **a** Pre-processing of input images. From the computed tomography images (512 × 512 pixels), a region including the vertebra (250 × 250 pixels; white line box) was cropped. The cropped images were resized to 96 × 96 pixels and used as input data to the convolutional neural network. **b** Region of interest (yellow circle) was put on vertebral body avoiding cortical bone and basivertebral vein

Reference standard

DXA examination was performed with Prodigy Primo (GE Healthcare). Patients underwent the examination in an antero-posterior position. The BMD of the lumbar vertebrae (L1, L2, L3, L4) and the % young adult mean, which was calculated using the BMD of L2–4 and the young adult mean value [5] in the DXA examination report, were recorded.

Deep learning with a CNN

Deep learning was performed with a computer equipped with a Core i9-7900X (Intel) central processing unit, a random access memory of 128 GB and a Quadro P6000 (NVIDIA) graphics processing unit. The programming language Python 3.6.4 and the deep learning framework of Chainer (https://chainer.org/) were used to perform deep learning with the CNN.

The preprocessed image data on the command line, as described in the previous section, were inputted to the CNN. The schema of the CNN structure is illustrated in Fig. 2. The CNN comprised four combinations of convolutional layers and max-pooling layers and three fully connected layers. Batch normalisation, which can accelerate the learning process and also reduce the risk of overfitting [21], was implemented within the CNN. Supervised learning was performed so that the mean squared error between the output data (BMD obtained with CNN [BMD_{CNN}]) and reference data (BMD measured with DXA) would become small. In the supervised training of the CNN, the following hyper-parameters were used: number of epochs, 20; optimiser, Adam [22]; and minibatch size, 15.

Validation with the trained CNN

After the completion of the supervised training, the performance of the trained CNN was evaluated using the validation datasets I and E. Any data in these datasets were not included in the training dataset.

CT value measurements

Previous studies aimed to estimate the BMD obtained with DXA from the CT values of the lumbar vertebrae [23, 24]. To compare the usefulness of the BMD_{CNN} and CT values in estimating the BMD, a radiologist (K.Y., with 9 years of imaging experience) placed a circular region of interest (ROI) with a diameter of 15 mm on the cancellous bone part of the vertebral body to measure the CT values of the vertebra for validation datasets I and E. In placing the ROIs, attention was paid so that the cortical bone and basivertebral vein were not included. The mean CT value (in Hounsfield unit [HU]) within the ROI was recorded.

Extrapolating the CNN to another vertebra

To assess whether the CNN could be used to estimate the BMD of other vertebrae that were not included in the training, a CNN model was also trained using the data of L2–4 in the training dataset, and the performance was tested using the data of L1 in the validation datasets I and E.

Statistics

The following statistical analyses were performed with R version 2.4.0 (https://www.r-project.org/). For continuous values, data were shown as the mean ± standard deviation.

All the BMD, BMD_{CNN} and CT value followed normal distribution according to Kolmogorov-Smirnov test (p values of 0.6273, 0.9988 and 0.9995, respectively). For each vertebra, Pearson's correlations between the BMD and BMD_{CNN} and between the BMD and CT values were calculated.

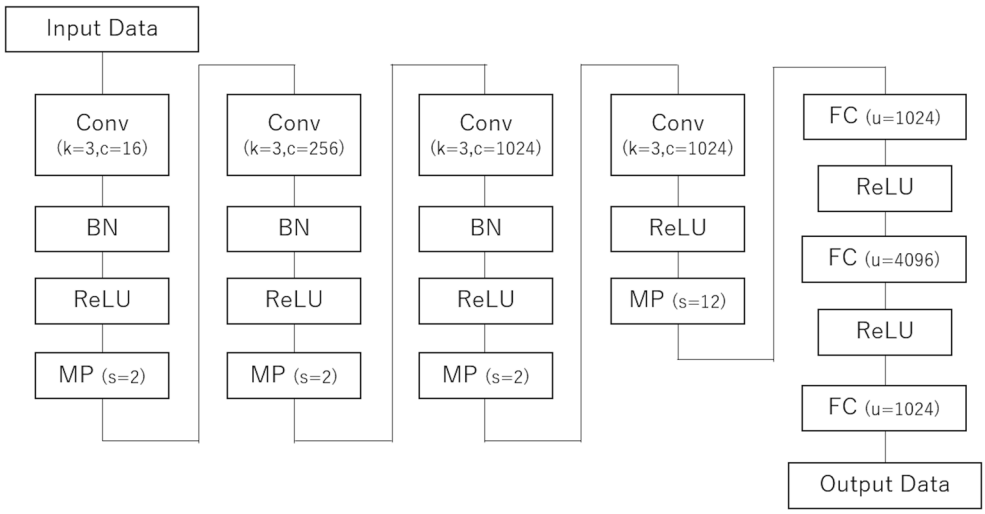

Fig. 2 Schema of the convolutional neural network. BN = batch normalisation, c = number of filters, Conv = convolutional layer, FC = fully connected layer, k = size of filters, MP = max-pooling layer, ReLU = rectified linear unit, s = size reduction rate, u = number of units

For each patient, the BMD_{CNN} and CT values were averaged for L2, L3 and L4. The performance of the averaged BMD_{CNN} and averaged CT values in diagnosing osteopenia and osteoporosis were evaluated with receiver operating characteristic (ROC) analysis and the area under the ROC curve (AUC) was calculated. Patients were diagnosed as having osteopenia and osteoporosis when the % young adult mean of the BMD evaluated with DXA was 80% or less and 70% or less, respectively [5]. A De Long test was performed to compare the AUCs of the averaged BMD_{CNN} and the averaged CT values in diagnosing osteopenia and osteoporosis [25]. Using cutoff values which achieve the Youden index, sensitivity, specificity and accuracy were calculated.

Results

Patients

For the training dataset, validation dataset I and validation dataset E, 183, 45 and 50 patients met the inclusion criteria and 1665, 180 and 200 vertebral CT images were obtained, respectively. For training dataset, a total of 99,900 images (= 1665 × 60) were generated with augmentation. Patient baseline characteristics are summarised in Table 2.

BMD measurement per vertebra analysis

The mean BMD values were 1.065 ± 0.234 g/cm^2 for the training dataset. The mean BMD values and CT values were 1.037 ± 0.220 g/cm^2 and 128.4 ± 56.8 HU, respectively, for the validation dataset I. The mean BMD values and CT values were 0.983 ± 0.235 g/cm^2 and CT 112.1 ± 59.3 HU, respectively, for the validation dataset E.

The correlation between the BMD and BMD_{CNN} of the lumbar vertebrae ($r = 0.852$ [95% confidence interval (CI), 0.806–0.887], $p < 0.001$) was higher than that between the BMD and CT values ($r = 0.425$ [95% CI, 0.297–0.538], $p < 0.001$) in validation dataset I (Fig. 3a and c). The correlation between the BMD and BMD_{CNN} of the lumbar vertebrae ($r = 0.840$ [95% CI, 0.794–0.877], $p < 0.001$) was also higher than that between the BMD and CT values ($r = 0.675$ [95% CI, 0.591–0.744], $p < 0.001$) in validation dataset E (Fig. 3b and d).

Reduced BMD per patient analysis

There were 12 and 21 patients with osteopenia in validation datasets I and E, respectively. The AUCs to diagnose osteopenia using the averaged BMD_{CNN} were 0.903–0.955 and they tended to be better than when the averaged CT value (AUC = 0.837–0.894; Table 3) was used.

There were two and nine patients with osteoporosis in validation datasets I and E, respectively. The AUCs to diagnose osteoporosis using the averaged BMD_{CNN} were 0.965–0.97 and they were better than when the averaged CT value (AUC = 0.829–0.953; Table 3) was used. A statistically significant difference between the averaged BMD_{CNN} and the averaged CT value to diagnose osteoporosis ($p = 0.013$) was observed in validation dataset E.

The sensitivity, specificity and accuracy in diagnosing osteopenia and osteoporosis are shown in Table 4.

Extrapolating the CNN to another vertebra

The correlation coefficients between the BMD_{CNN} values of L1 outputted by the CNN model trained with L2–4 data in the training dataset and the BMD of L1 derived from DXA were 0.839 (95% CI, 0.724–0.909; $p < 0.001$) and 0.810 (95% CI, 0.686–0.888; $p < 0.001$) for validation datasets I and E, respectively.

Discussion

Osteoporosis, which is characterised by a reduced BMD, is associated with fragile bone fractures. DXA is used as a reference standard for evaluations of the BMD. In this study, the estimated BMD values, by applying a deep learning technique to unenhanced abdominal CT images, were found to be highly correlated with the BMD values derived from DXA. Our method was found to be superior to simply measuring CT

Table 2 Patient baseline characteristics

	Training dataset	Validation dataset I	Validation dataset E
Men/women (n)	86/97	14/31	12/38
Mean age (years)	60.6 ± 15.5	63.5 ± 16.1	67.2 ± 12.3
Mean body height (cm)	161.2 ± 9.5	158.1 ± 9.3	157.3 ± 9.8
Mean body weight (kg)	55.8 ± 10.8	55.4 ± 14.0	54.2 ± 10.9
Mean body mass index (kg/m^2)	21.4 ± 3.4	22.1 ± 5.1	21.9 ± 3.9
Median (with interquartile range) time interval between the DXA and CT (days)	5.0 (1.0–55.5)	8.0 (1.0–113.0)	73.0 (32.2–203.8)

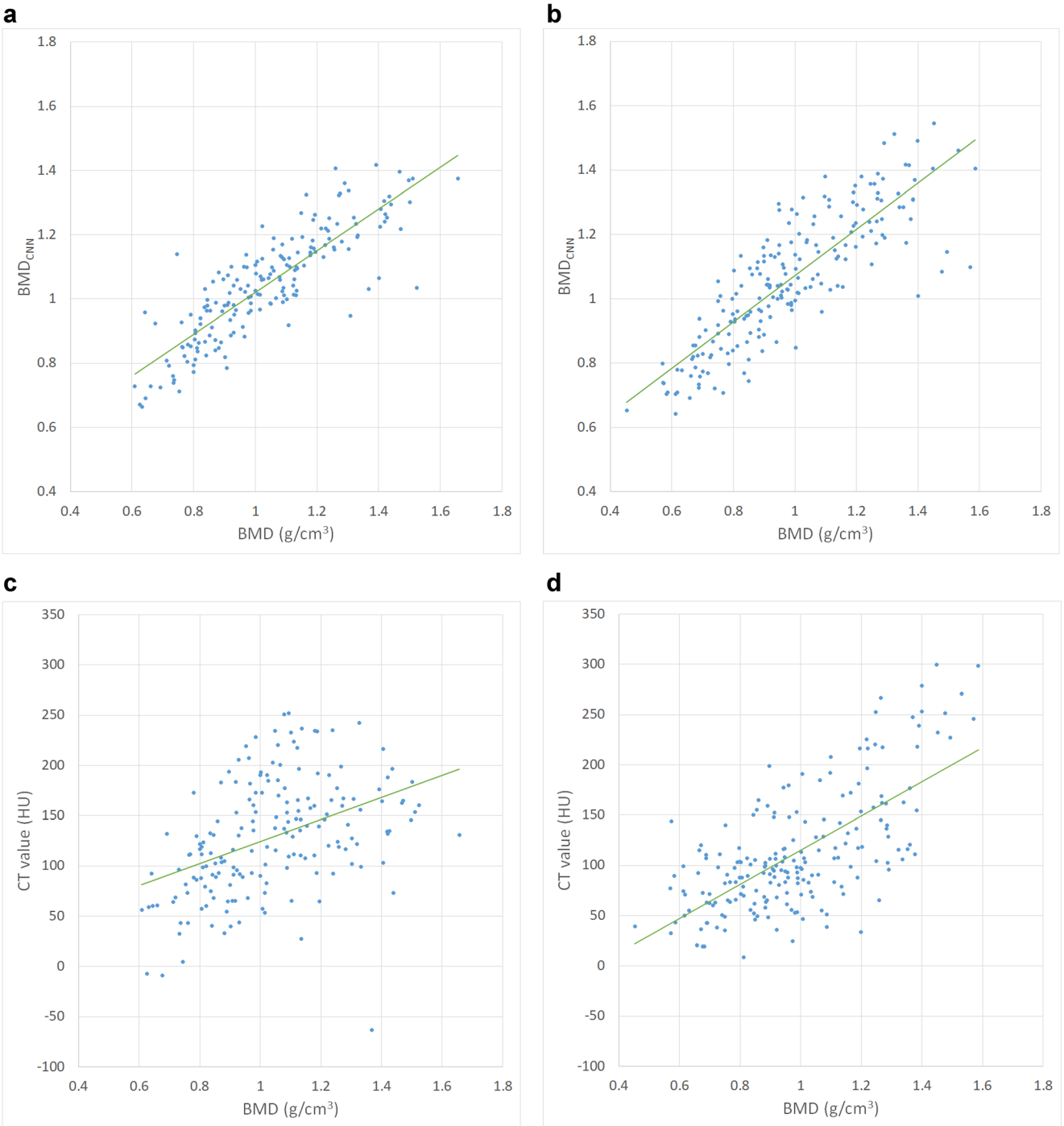

Fig. 3 Scatterplots of correlation between BMD_{CNN} and BMD (g/cm^2) in validation dataset I (**a**) and validation dataset E (**b**) and that between the CT values (Hounsfield unit) and BMD (g/cm^2) in validation dataset I (**c**) and validation dataset E (**d**)

values of the lumbar vertebrae in estimating the BMD. The CNN model trained with L2–4 data also allowed the estimation of the BMD of L1. By using the estimated BMD values, osteoporosis was diagnosed with higher performance than by using CT values of the lumbar vertebrae. These results were confirmed by external validation as well as internal validation.

Deep learning, which is one of the methods of artificial intelligence, has been applied to musculoskeletal radiological imaging diagnosis. However, to our knowledge, most of these studies have focused on detecting fractures on radiological images [15–17]. When fractures occur in patients, they may suffer from disability and/or pain. Resources to manage these conditions place a large burden on medical economics. Therefore, the prevention of fractures is rather important. As our model can be applied to CT images, which are used widely in daily clinical practice, patients with osteoporosis can be

Table 3 Diagnostic performance for diagnosing osteopenia and osteoporosis

	BMD_{CNN}	CT value	*p* value
Osteopenia			
Validation dataset I	0.955 (0.901–1.000)	0.894 (0.799–0.988)	0.167
Validation dataset E	0.903 (0.823–0.983)	0.837 (0.729–0.946)	0.242
Osteoporosis			
Validation dataset I	0.965 (0.887–1.000)	0.953 (0.882–1.000)	0.860
Validation dataset E	0.970 (0.915–1.000)	0.829 (0.693–0.965)	0.013

Area under the receiver operating characteristic curve (AUC) values with 95% confidence intervals are shown. A De long test was performed to compare AUC values between the BMD_{CNN} and CT values

diagnosed before the occurrence of fragile fractures on CT examinations performed for other purposes. In addition, these patients will have opportunities to be treated to prevent fragile fractures by being advised on adequate dietary intake, using fall prevention methods and receiving pharmacological interventions [3].

Some modalities are used to evaluate BMD. DXA is the method used most commonly for this purpose. However, DXA is not necessarily readily available at all institutions. Also, patients may not undergo this examination if they are not suspected of osteoporosis by specialised physicians. Quantitative ultrasonometry can also be used as a screening method to detect patients at risk for fractures. However, while the diagnostic DXA criteria established by the World Health Organization and recommended by the American Association of Clinical Endocrinologists apply only to the axial measurements, calcaneus bones of peripheral sites are measured with this modality. Therefore, quantitative ultrasonometry is not used to diagnose osteoporosis [26]. Quantitative CT can also be used to assess the BMD; however, a dedicated phantom should be placed beneath the patients when being scanned. While there are several modalities to evaluate the BMD, each has merits and demerits. Our proposed algorithm, which can be applied to routine clinical CT images, can become another strategy to estimate the BMD.

Our study revealed that the BMD_{CNN} was superior to the CT values of the lumbar vertebrae to estimate the BMD assessed with DXA. This means that when unenhanced abdominal CT images exist, our model allows the estimation of the past BMD of patients more precisely than simply measuring the CT values of the lumbar vertebrae. The time course of the BMD can also be more reliably estimated with our model. Also, our model will allow the estimation of the BMD in institutions where DXA is not available. While accurate reasons for the improved performance remain unclear due to the black-box nature of deep learning algorithms, we speculate that the texture of the trabecular bone in addition to CT attenuation might have been considered in estimating the BMD from CT images. The other reason might have come from the difference of modalities. According to Hendrickson et al, when mean and standard deviation for CT values of young adults (μref and σref, respectively) were used as reference, lumbar CT value T-scores calculated as (Lumbar CT value T-score = (CT value − μref) / σref) were reportedly lower than DXA T-scores [24].

The model trained with data of L2–4 allowed a precise estimation of the BMD of L1, with correlation coefficients between the BMD_{CNN} and BMD derived from DXA of 0.810–0.839. Vertebral compression fractures occur commonly in the thoracolumbar junction area

Table 4 Sensitivity, specificity and accuracy in diagnosing osteopenia and osteoporosis

		Cutoff	Sensitivity	Specificity	Accuracy
Osteopenia					
Validation dataset I	Averaged BMD_{CNN}	0.984	1.000 (12/12)	0.818 (27/33)	0.867 (39/45)
	Averaged CT value	107.29	0.917 (11/12)	0.818 (27/33)	0.844 (38/45)
Validation dataset E	Averaged BMD_{CNN}	1.010	0.810 (17/21)	0.862 (25/29)	0.840 (42/50)
	Averaged CT value	102.91	0.915 (19/21)	0.655 (19/29)	0.760 (38/50)
Osteoporosis					
Validation dataset I	Averaged BMD_{CNN}	0.849	1.000 (2/2)	0.930 (40/43)	0.933 (42/45)
	Averaged CT value	55.97	1.000 (2/2)	0.930 (40/43)	0.933 (42/45)
Validation dataset E	Averaged BMD_{CNN}	0.876	0.889 (8/9)	0.976 (40/41)	0.960 (48/50)
	Averaged CT value	72.02	0.778 (7/9)	0.829 (34/41)	0.820 (41/50)

Cutoff values were chosen to achieve the Youden index

(T10–L2) [27]. Our model may have the potential to estimate the BMD of these vertebrae that are at high risk of compression fractures.

Some limitations should be acknowledged in this study. First, the design of our study was retrospective. Future prospective studies are necessary to consolidate our results. Second, we did not assess the relationship between the BMD_{CNN} and fracture risk. Third, because contrast materials alter the CT number of the bone marrow as well as other abdominal tissues, our model, which was established by using unenhanced abdominal CT images as input data, cannot be applied to contrast-enhanced CT images. Fourth, although we used a single type of DXA machine, BMD measurements with DXA are not free from a reproducibility problem, which was reported to be instrument-related [28] or to be introduced by operators' and subjects' variability [29]. Fifth, our model cannot detect osteoporotic fracture on CT images. Whether osteoporotic fracture can be detected automatically with CNN on CT images or not needs to be assessed in future studies. Sixth, our result cannot be applied to patients with compression fracture because we excluded them from this study. However, it will not affect clinical management of them, because patients with fragility fracture can be diagnosed as osteoporosis without using BMD values [5]. Seventh, BMD values estimated by DXA is not necessarily equal to a true bone density measurement. Also, in identifying elderly who subsequently had a non-vertebral fracture, the sensitivity of DXA-determined osteoporosis was reportedly only 21–44% [30]. However, DXA is recommended by guidelines [3–5] and is most widely used to estimate BMD. Finally, our model was found to estimate the BMD of other vertebrae that were not included in the training. However, as we did not investigate the correlation of thoracic vertebrae, further investigation will be required to evaluate them.

In conclusion, by applying deep learning with a CNN, the BMD can be estimated from unenhanced abdominal CT images. There would be a possibility that osteoporosis is diagnosed with higher performance by using BMD_{CNN} values than CT values of the lumbar vertebrae; however, future prospective studies including a large number of patients would be required to consolidate this finding.

Funding information This study has received funding by JSPS KAKENHI Grant Number JP18K15542.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Koichiro Yasaka.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- Retrospective
- Diagnostic or prognostic study
- Multicentre study

References

1. Cooper C, Atkinson EJ, Jacobsen SJ, O'Fallon WM, Melton LJ 3rd (1993) Population-based study of survival after osteoporotic fractures. Am J Epidemiol 137:1001–1005
2. Hernlund E, Svedbom A, Ivergard M et al (2013) 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 8:136
3. Kanis JA, Cooper C, Rizzoli R et al (2019) European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int 30:3–44
4. Compston J, Cooper A, Cooper C et al (2017) UK clinical guideline for the prevention and treatment of osteoporosis. Arch Osteoporos 12:43
5. Orimo H, Nakamura T, Hosoi T et al (2012) Japanese 2011 guidelines for prevention and treatment of osteoporosis–executive summary. Arch Osteoporos 7:3–20
6. OECD Library (2017) Health at a glance 2017. Available via https://doi.org/10.1787/health_glance-2017-en. Accessed 20 Sept 2019
7. Bartalena T, Rinaldi MF, Modolon C et al (2010) Incidental vertebral compression fractures in imaging studies: lessons not learned by radiologists. World J Radiol 2:399–404
8. Krizhevsky A, Sutskever I, Hinton G (2012) ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing System 25 (NIPS 2012). Available via https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks. Accessed 20 Dec 2017
9. Yasaka K, Abe O (2018) Deep learning and artificial intelligence in radiology: current applications and future directions. PLoS Med 15: e1002707
10. Chartrand G, Cheng PM, Vorontsov E et al (2017) Deep learning: a primer for radiologists. Radiographics 37:2113–2131
11. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O (2018) Deep learning with convolutional neural network in radiology. Jpn J Radiol 36:257–272
12. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS (2018) Image reconstruction by domain-transform manifold learning. Nature 555: 487–492
13. Yasaka K, Akai H, Abe O, Kiryu S (2018) Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. Radiology 286:887–896
14. Kiryu S, Yasaka K, Akai H et al (2019) Deep learning to differentiate parkinsonian disorders separately using single midsagittal MR imaging: a proof of concept study. Eur Radiol 29:6891–6899

15. Urakawa T, Tanaka Y, Goto S, Matsuzawa H, Watanabe K, Endo N (2019) Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. Skeletal Radiol 48:239–244
16. Derkatch S, Kirby C, Kimelman D, Jozani MJ, Davidson JM, Leslie WD (2019) Identification of vertebral fractures by convolutional neural networks to predict nonvertebral and hip fractures: a registry-based cohort study of dual X-ray absorptiometry. Radiology 293:405–411
17. Cheng CT, Ho TY, Lee TY et al (2019) Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. Eur Radiol 29:5469–5477
18. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S (2018) Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid-enhanced hepatobiliary phase MR images. Radiology 287:146–155
19. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S (2018) Deep learning for staging liver fibrosis on CT: a pilot study. Eur Radiol 28:4578–4585
20. Larson DB, Chen MC, Lungren MP, Halabi SS, Stence NV, Langlotz CP (2018) Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. Radiology. 287:313–322
21. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. Cornell University Library. Available via http://arxiv.org/abs/1502.03167. Accessed 30 April 2017
22. Kingma DP, Ba JL (2014) Adam: a method for stochastic optimization. Cornell University Library. Available via http://arxiv.org/abs/1412.6980. Accessed 30 April 2017
23. Schreiber JJ, Anderson PA, Hsu WK (2014) Use of computed tomography for assessing bone mineral density. Neurosurg Focus 37:E4
24. Hendrickson NR, Pickhardt PJ, Del Rio AM, Rosas HG, Anderson PA (2018) Bone mineral density T-scores derived from CT attenuation numbers (Hounsfield units): clinical utility and correlation with dual-energy X-ray absorptiometry. Iowa Orthop J 38:25–31
25. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
26. Camacho PM, Petak SM, Binkley N et al (2016) American Association of Clinical Endocrinologists and American College of Endocrinology Clinical Practice Guidelines for the Diagnosis and Treatment of Postmenopausal Osteoporosis - 2016. Endocr Pract 22:1–42
27. Wood KB, Li W, Lebl DR, Ploumis A (2014) Management of thoracolumbar spine fractures. Spine J 14:145–164
28. Phillipov G, Seaborn CJ, Phillips PJ (2001) Reproducibility of DXA: potential impact on serial measurements and misclassification of osteoporosis. Osteoporos Int 12:49–54
29. Fuleihan GE, Testa MA, Angell JE, Porrino N, Leboff MS (1995) Reproducibility of DXA absorptiometry: a model for bone loss estimates. J Bone Miner Res 10:1004–1014
30. Schuit SC, van der Klift M, Weel AE et al (2004) Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam study. Bone 34:195–202