



OPEN Radiomics models based on thoracic and upper lumbar spine in chest LDCT to predict low bone mineral density

Yaling Pan^{1,4}, Yidong Wan^{2,3,4}, Yinbo Wu¹, Chunmiao Lin¹, Qin Ye¹, Jing Liu¹, Hongyang Jiang¹, Huogen Wang^{2,3}✉ & Yajie Wang¹✉

This study aims to develop and validate different radiomics models based on thoracic and upper lumbar spine in chest low-dose computed tomography (LDCT) to predict low bone mineral density (BMD) using quantitative computed tomography (QCT) as standard of reference. A total of 905 participants underwent chest LDCT and paired QCT BMD examination were retrospectively included from August 2018 and June 2019. The patients with low BMD ($n = 388$) and the normal ($n = 517$) were randomly divided into a training set ($n = 622$) and a validation set ($n = 283$). Radiomics features (RFs) were extracted from the single and consecutive vertebrae in chest LDCT images to construct the single vertebra RFs models, mixed RFs models and Radscore models, respectively. The performance of these models was evaluated by the area under the curve (AUC) of receiver operator characteristic curve, using QCT as standard of reference. The Radscore models, mixed RFs models, and single vertebra RFs models yielded the AUC values ranging from 0.809 to 0.906, 0.792 to 0.883, and 0.731 to 0.884 for predicting low BMD in the validation set, respectively. For predicting low BMD, the Radscore model of L1-L2 vertebrae yielded the highest AUC of 0.906, and of T1-T3 yielded the lowest AUC of 0.809 ($P < 0.05$), respectively. However, there was no significant difference among the AUC values of three Radscore models constructed on the vertebrae of T4-T6 (AUC = 0.855), T7-T9 (AUC = 0.845), and T10-T12 (AUC = 0.871) for predicting low BMD in the validation set ($P > 0.1$). The Radscore model of L1-L2 have potential to serve as an important tool for predicting and screening low BMD from normal in chest LDCT images.

Keywords Bone mineral density, Radiomics, Osteoporosis, Quantitative computed tomography

With the aging population of society, osteoporosis has become a serious public health problem worldwide. It is a silent osteopathy characterized by reduced bone mass and increased bone fragility with aging^{1,2}. After 50 years of age, the risk of osteoporotic fractures is 50% for women and 20% for men³. The majority of the causal fragility fracture of vertebrae occurs in osteopenic patients⁴. If not detected and interfered in time, these fractures would lead to high morbidity and mortality⁵. Hence, the early identification of patients at high risk of low bone mineral density (BMD) including osteopenia and osteoporosis in an earlier stage plays a significant role in the progress of disease prevention.

The diagnosis of osteopenia and osteoporosis mainly relies on the BMD testing. Dual energy X-ray absorptiometry (DXA) is the most widely used method for BMD measurement and has been adopted as a standard⁶. Quantitative computed tomography (QCT) has been used to quantify BMD in CT images. In contrast to DXA, QCT can avoid BMD overestimation due to extra-osseous calcification, spinal degeneration and other sclerotic lesions⁷. Despite these advantages, QCT is not yet widely available because of specialized phantom⁸. Therefore, it is beneficial to investigate the role of alternative techniques for accurate prediction of low BMD, such as radiomics.

Radiomics introduces a new quantitative approach for disease assessment based on medical images, it can exploit high throughput imaging features that fail to be visually recognized and further quantitatively analyze the

¹Center for Rehabilitation Medicine, Department of Radiology, Zhejiang Provincial People's Hospital (Affiliated People's Hospital), Hangzhou Medical College, Hangzhou 310014, Zhejiang, China. ²HiThink Research, Hangzhou 310023, Zhejiang, China. ³Zhejiang Herymed Technology Co., Ltd., Hangzhou 310023, Zhejiang, China. ⁴Yaling Pan and Yidong Wan contributed equally to this work. ✉email: 0621318@zju.edu.cn; wangyajie3964@163.com

lesion heterogeneity by applying advanced data model algorithm to aid in clinical decision making⁹. Currently, radiomics analysis is commonly used in clinical oncology applications and contributes to the tumor diagnosis, prognosis prediction and efficacy evaluation^{10,11}, only a limited number of studies investigated this approach on osteoporosis and bone mineral loss^{12,13}. Radiomics analysis may be useful for the assessment of microstructure changes of trabecular bone, which may be suited for abnormal bone mass screening^{13,14}. There are some patients without undergoing DXA because osteoporosis is asymptomatic until major incidental fragile fractures occur¹⁵, but they frequently received chest low-dose CT (LDCT) examinations during annual health check-ups. Even if a small number of chest LDCT images was used to opportunistically screen for low BMD, the impact could be substantial. But, reports on the use of radiomics analysis derived from chest LDCT for low BMD prediction is still lacking.

Therefore, the purpose of our study was to develop and validate different radiomics models based on thoracic and upper lumbar spine in chest LDCT images for low BMD prediction, with QCT as standard of reference. In addition, the performance of radiomics modeling on the single and consecutive vertebrae were investigated, respectively.

Method and materials

This study was approved by institutional ethics committee of our hospital, and informed consent was waived because of a retrospective study. The identifying information were desensitized before using in order to protect the participant's privacy. We stated that all methods were performed in accordance with the relevant guidelines and regulations.

Study population

Between August 2018 and June 2019, the healthy check-up population aged 50 years or older who had undergone chest LDCT and paired QCT BMD examination at our hospital were retrospectively included. Exclusion criteria: (I) obvious lesions in vertebrae, such as bone islands, cystic lesion, tumors; (II) vertebral fracture; (III) spinal surgery and severe imaging artifacts. A total of 905 participants were included in the final analysis (590 men and 315 women). The age ranged from 50 to 88 years old (mean, 59.7 ± 8.2 years). We randomly split the participants at a ratio of 7:3 into training set ($n = 622$) and validation set ($n = 283$).

LDCT image acquisition and QCT examination

Chest LDCT images were acquired using the CT scanner (Somatom Definition AS+, Siemens Healthcare, Forchheim, Germany). The following CT protocol was used: 120 kVp, automatic tube current, average 30mAs, 1.0-mm reconstruction slice thickness, 512×512 matrices, and 500-mm scan field of view. An asynchronously phantom (Mindways, Austin, TX, USA) was scanned to calibrate the BMD of lumbar vertebra with the same CT parameters. A computer workstation (Mindways QCT Pro 6.1, Austin, TX, USA) was used to delineate the separate volume of interest (VOI) on the center of L1 and L2 vertebral bodies; these VOIs were then utilized to calculate the trabecular BMD. The mean BMD of L1 and L2 was used to determine final reference diagnosis. The thresholds to classify mean BMD into reference diagnosis were in line with the American College of Radiology in 2013¹⁶, characterizing the groups as: normal $> 120 \text{ mg/cm}^3$; low BMD including osteoporosis and osteopenia $\leq 120 \text{ mg/cm}^3$.

Workflow

The workflow of this study is summarized in Fig. 1, and it can be divided into five parts including: imaging acquisition, VOI segmentation, feature extraction, feature selection and radiomics model construction and evaluation. Chest LDCT images and paired BMD data were collected by experienced radiologists. VOIs were fully automatically segmented using open source segmentation codes. Quantitative radiomics features were extracted from target VOIs. Finally, the bone mass classifiers were constructed by the machine learning algorithm.

Image segmentation and preprocessing

The deep learning-based open-source segmentation codes were employed to complete the VOI segmentation. Each vertebra was individually labeled, and corresponding segmentation masks were generated. To ensure the accuracy of VOIs, the generated labels and segmentation masks were checked by a musculoskeletal radiologist with three years of experience. ITK-SNAP software (version 3.8c; <http://itksnap.org>) was used to manually modify the VOIs if necessary. Finally, the VOI of each thoracic and lumbar vertebra from T1 to L2 (T1-L2) was obtained in all chest LDCT images.

To normalize different image specification, image resampling and gray-level normalization were conducted before the process of radiomics feature extraction based on machine learning, ensuring consistency in both image intensity and resolution. To minimize the influence of the original intensity distribution, the gray level of the image was normalized to a scale of 64. Moreover, to account for variations in spatial resolution during image acquisition, the voxel sizes were resampled to a uniform size of $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ by bilinear interpolation method.

Radiomics feature extraction

Firstly, the radiomics features were extracted in PyCharm platform (version:2022.1.3, community edition) by PyRadiomics (version:3.0.1) software package. The radiomics features included six types: the first-order features, gray level co-occurrence matrix (GLCM) features, gray level dependence matrix (GLDM) features, gray level run length matrix (GLRLM) features, gray level size zone matrix (GLSZM) features, neighboring gray tone difference matrix (NGTDM) features. For each VOI of vertebra, we extracted radiomics features separately. After feature extraction, the training set was used to select appropriate features and develop the classifier. We normalize the

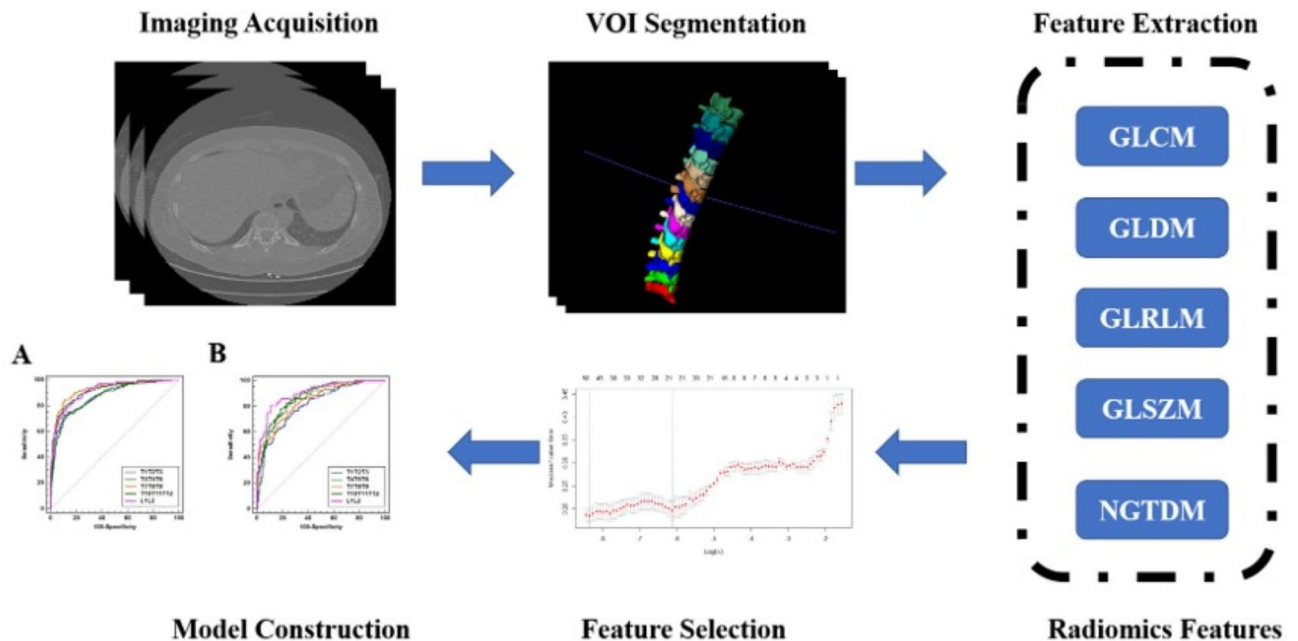


Fig. 1. The workflow of this study.

different scales among variables. In the training set, radiomics features were individually subtracted by the mean value of each feature and divide by their respective standard deviation value. Then, the mean value and standard deviation were also applied to normalize the independent validation set.

Feature selection and model construction

A two-step method was employed to feature selection, including univariate analysis and least absolute shrinkage and selection operator (LASSO) method. First, we applied Wilcoxon test to selected the significant features between normal and low BMD groups, and features were retained with $P < 0.05$. Second, LASSO regression was employed to obtain the most robust and non-redundant radiomics features. The complexity of LASSO algorithm is impacted by a tuning parameter λ with the rule that as the value of λ increases, the penalty for each variable coefficient also increases. In this study, we select the largest value of λ such that error is within one standard error of the cross-validated errors for minimum λ which is represent the minimum mean cross-validated error. Only variables with non-zero coefficients were selected in LASSO regression. The best λ value were selected by five cross-fold-validation. A radiomics signature was constructed from linear combinations of the selected features weighted by their respective LASSO coefficients. After that, the selected features were employed to construct the normal and low BMD classifier by a linear regression. The LASSO-based linear regression algorithm offers superior interpretability compared to other machine learning algorithms such as support vector machine (SVM) or random forest. The coefficients derived from linear regression provide clear insights into the contribution of each feature to the model, making it easier to understand and interpret the results. A radiomics score (Radscore) was constructed from linear combinations of the selected features weighted by their corresponding LASSO coefficients. The Radscore was calculated using the following formula:

$$\text{Radscore} = \sum_{i=0}^n c_i \times X_i + b \quad (1)$$

where X_i represents the i th selected feature, C_i is the corresponding feature coefficient, and b is the intercept. The Radscore was used as a predictor of low BMD in the subsequent analysis.

Construction of the radiomics models

The linear regression algorithm was used to construct three types of radiomics models at vertebral level of T1-L2 in the training set, namely, the single vertebra RFs model based on the radiomics features (RFs) of each vertebra of T1-L2, respectively. The mixed RFs model based on the mixed radiomics features, and the Radscore model based on the average Radscores, of consecutive vertebrae in term of T1-T3, T4-T6, T7-T9, T10-T12, L1-L2, respectively. The performance of these models for the prediction of low BMD was evaluated in both the training and validation sets.

Statistical analysis

The radiomics models were established and verified using Rstudio through the analyses of radiomics features. All statistical analyses were performed in Rstudio (version: 2022.02.3) and MedCalc 15.2.2 and SPSS 22.0. The LASSO regression was performed using the “glmnet” package in Rstudio. Independent-samples T test was used to analyze continuous variables. Chi-square test was used to analyze the differences in categorical variables.

	Total (n = 905)	Training set (n = 622)	Validation set (n = 283)
Sex, male/female	590/315	405/217	185/98
Age, years old	59.7 ± 8.2	58.8 ± 7.9 [#]	61.6 ± 8.5 [#]
BMD			
L1, mg/cm ³	130.94 ± 36.95	131.59 ± 37.98	129.52 ± 34.61
L2, mg/cm ³	127.05 ± 37.74	127.32 ± 38.43	126.45 ± 36.26
Normal, n	517	355	162
Osteopenia, n	309	212	97
Osteoporosis, n	79	55	24
Low BMD, n	388	267	121

Table 1. Demographic characteristics of participants. [#]represents as $P < 0.05$ between training and validation sets; BMD, bone mineral density. Low BMD (include osteopenia and osteoporosis).

Vertebra	Feature number	Training set			Validation set		
	(Significant/Lasso)	AUC (95%CI)	Sensitivity (%)	Specificity (%)	AUC (95%CI)	Sensitivity (%)	Specificity (%)
T1	73/35	0.850 (0.820–0.880)	67.8	86.5	0.731 (0.671–0.790)	57.9	74.1
T2	79/29	0.830 (0.798–0.863)	70.4	83.4	0.800 (0.747–0.852)	61.2	83.3
T3	80/21	0.858 (0.828–0.887)	82.4	72.1	0.809 (0.760–0.859)	75.2	69.1
T4	75/25	0.862 (0.833–0.891)	74.9	83.9	0.826 (0.779–0.874)	67.8	75.9
T5	80/12	0.808 (0.775–0.842)	71.5	77.2	0.768 (0.712–0.824)	64.5	73.5
T6	80/35	0.862 (0.833–0.891)	74.9	86.2	0.862 (0.818–0.905)	71.1	82.7
T7	78/27	0.878 (0.851–0.904)	79.4	77.8	0.804 (0.753–0.855)	78.5	67.3
T8	77/56	0.927 (0.906–0.947)	87.6	85.6	0.821 (0.773–0.869)	76.0	71.0
T9	76/28	0.879 (0.852–0.905)	78.6	82.0	0.818 (0.769–0.867)	66.9	81.5
T10	65/33	0.874 (0.846–0.901)	73.4	86.8	0.853 (0.808–0.898)	73.6	86.4
T11	68/34	0.887 (0.862–0.913)	82.8	78.6	0.812 (0.761–0.863)	75.2	74.1
T12	75/24	0.859 (0.831–0.888)	78.7	77.8	0.814 (0.765–0.862)	76.0	71.0
L1	79/21	0.897 (0.873–0.921)	79.8	85.6	0.884 (0.845–0.923)	76.9	85.2
L2	80/32	0.900 (0.877–0.924)	86.1	79.2	0.877 (0.838–0.916)	84.3	74.1

Table 2. The feature number and performance of radiomics models constructed on single vertebra for predicting low BMD. AUC, area under the curve of the receiver operating characteristic (ROC) curve. BMD, bone mineral density. CI: confidence interval. LASSO, least absolute shrinkage and selection operator.

Sensitivity, specificity, and the area under curve (AUC) of the receiver operating characteristic (ROC) curve and 95% confidence interval (CI) in the training and validation sets were calculated to evaluate the accuracy of the radiomics models, respectively. Delong’s test was used to assess the difference of ROC curve. $P < 0.05$ was considered statistically significant.

Results
Demographic characteristics

The characteristics of participants were summarized in Table 1. A total of 905 participants were included in the study. 517 participants had normal bone mass, and 388 patients had low BMD. There was significant difference in age ($P < 0.05$), but no significant difference in the distribution of bone mass between training and validation sets ($P > 0.05$). This indicates that the rationality of distribution in this study.

The performance of radiomics models based on single vertebra

The feature number and performance of fourteen single vertebra RFs models for predicting low BMD were shown in Table 2. The AUC values of the single vertebra RFs models for predicting low BMD ranged from 0.808 (95%CI: 0.775–0.842) to 0.927 (95%CI: 0.906–0.947) in the training set, and 0.731 (95%CI: 0.671–0.790) to

0.884 (95%CI: 0.845–0.923) in the validation set, respectively. Among fourteen models, the single vertebra RFs model of L1 yielded the highest AUC of 0.884 (95%CI: 0.845–0.923), subsequently of L2 (AUC=0.877, 95%CI: 0.838–0.916), and of T1 yielded the lowest AUC of 0.731 (95%CI: 0.671–0.790) for low BMD prediction in validation set.

Number of occurrences and coefficients of radiomics features

The number of occurrences of frequently occurred twenty-two features which contributed to develop the single vertebra RFs models was shown in Fig. 2. The frequent occurrence of first-order features, GLCM features, and GLSZM features suggests that these features contribute significantly to develop the single vertebra RFs models. The number of occurrences of GLSZM_HighGrayLevelZoneEmphasis is thirteen, which means it contributed to develop thirteen single vertebra RFs models of T1–L2 except T5. Twenty-one and thirty-two features of nonzero coefficients were selected to establish the L1 and L2 models with LASSO regression, respectively, and their corresponding coefficients were shown in Fig. 3.

The performance of radiomics models based on consecutive vertebrae

The feature number and performance of mixed RFs models and Radscore models for predicting low BMD were summarized in Table 3. The mixed RFs models and Radscore models yielded the AUC values ranging from 0.865 (95%CI: 0.837–0.894) to 0.913 (95%CI: 0.890–0.935), 0.870 (95%CI: 0.842–0.898) to 0.930 (95%CI: 0.910–0.950) in the training set, 0.792 (95%CI: 0.740–0.844) to 0.883 (95%CI: 0.843–0.922), and 0.809 (95%CI: 0.759–0.859) to 0.906 (95%CI: 0.871–0.940) in the validation set for predicting low BMD, respectively. For the prediction of low BMD, Radscore model of L1–L2 vertebrae yielded the highest AUC of 0.906, and of T1–T3 yielded the lowest AUC of 0.809 in the validation set ($P < 0.05$), respectively. There is no significant difference among the AUC values of three Radscore models constructed on the vertebrae of T4–T6 (AUC=0.855), T7–T9 (AUC=0.845), and T10–T12 (AUC=0.871) for predicting low BMD in the validation set ($P > 0.1$), shown in Fig. 4.

Discussion

Osteoporosis and osteopenia are diseases in which screening can have a great effect on patient outcome¹⁷. Patients often do not recognize the seriousness of decreased bone mass, therefore, do not participate in the DXA screening program voluntarily until the fracture occurred¹⁸. There is a growing consensus regarding the need for alternative screening methods to overcome the limitations and underuse of DXA. Chest LDCT is one of the most widely conducted medical screening during annual health check-ups and expected to apply for opportunistic

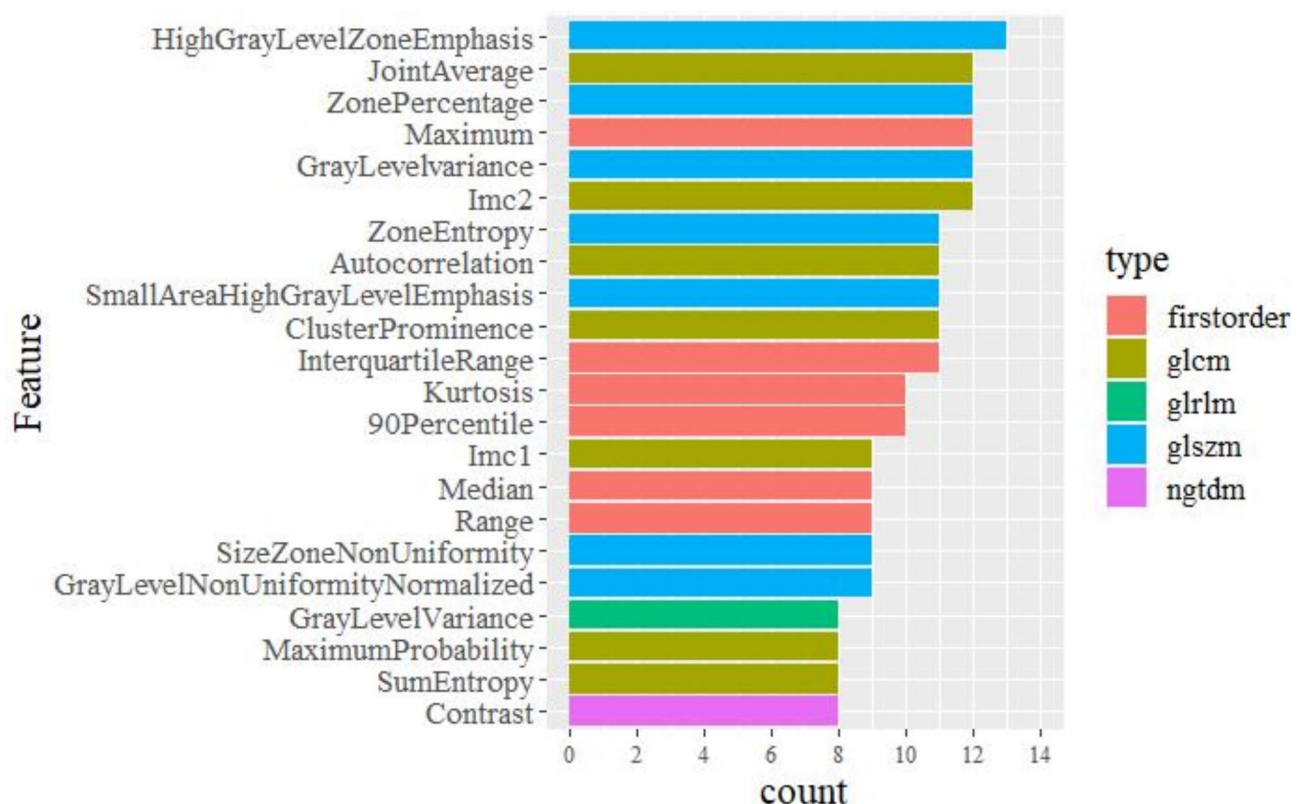


Fig. 2. The number of occurrences of frequently occurred 22 significant features which contributed to develop the single vertebra radiomics features models. The top feature was most frequently occurred in the fourteen single vertebra radiomics features models. The first-order features and second-order features including glcm, glrlm, glszm, and ngtdm features were highlighted in different color.

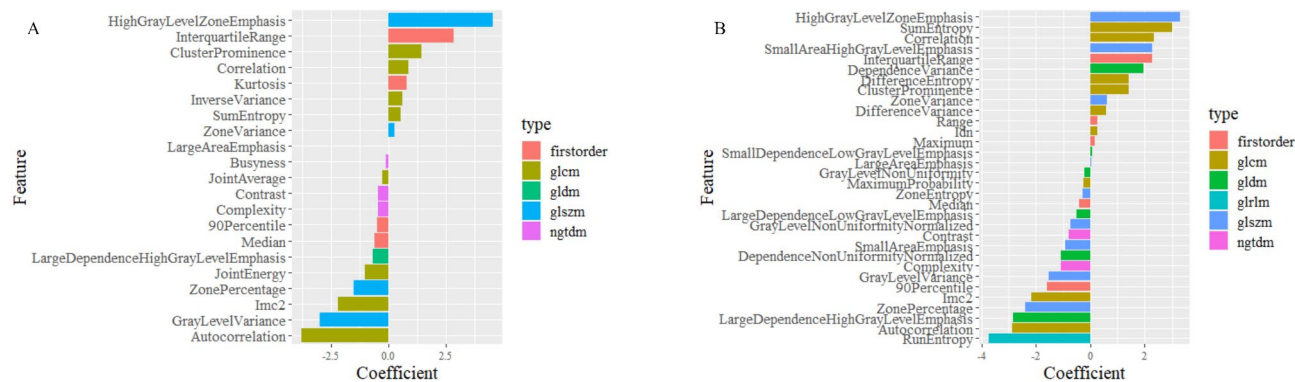


Fig. 3. The histograms of the radiomics features of L1 (A) and L2 (B) in corresponding single vertebra radiomics features model and their coefficients after LASSO: the y-axis indicates the selected radiomics features and the x-axis represents the coefficient of radiomics features. The type of features was highlighted in different color. LASSO: the least absolute shrinkage and selection operator.

Consecutive vertebrae	Feature number	Training set			Validation set		
	(Significant/Lasso)	AUC (95%CI)	Sensitivity (%)	Specificity (%)	AUC (95%CI)	Sensitivity (%)	Specificity (%)
Mixed radiomics features models							
T1-T3	232/31	0.865 (0.837–0.894)	0.779	0.792	0.795 (0.743–0.846)	0.669	0.784
T4-T6	235/37	0.871 (0.843–0.898)	0.764	0.828	0.820 (0.771–0.869)	0.719	0.759
T7-T9	231/41	0.881 (0.854–0.907)	0.783	0.839	0.792 (0.740–0.844)	0.653	0.735
T10-T12	208/45	0.904 (0.880–0.927)	0.843	0.809	0.831 (0.784–0.878)	0.793	0.728
L1-L2	159/36	0.913 (0.890–0.935)	0.775	0.893	0.883 (0.843–0.922)	0.736	0.864
Radscore models							
T1-T3	232/31	0.876 (0.849–0.903)	0.730	0.893	0.809 (0.759–0.859)	0.620	0.833
T4-T6	235/37	0.870 (0.842–0.898)	0.719	0.879	0.855 (0.812–0.899)	0.645	0.870
T7-T9	231/41	0.930 (0.910–0.950)	0.843	0.893	0.845 (0.802–0.890)	0.711	0.765
T10-T12	208/45	0.915 (0.892–0.937)	0.802	0.887	0.871 (0.831–0.912)	0.760	0.815
L1-L2	159/36	0.917 (0.893–0.938)	0.813	0.851	0.906 (0.871–0.940)	0.818	0.840

Table 3. The feature number and performance of mixed radiomics features models and Radscore models base on consecutive vertebrae for predicting low BMD. AUC, area under the curve of the receiver operating characteristic (ROC) curve. BMD, bone mineral density. CI: confidence interval. LASSO, least absolute shrinkage and selection operator.

screening. In this study, we developed and validated different radiomics models based on thoracic and upper lumbar spine in chest LDCT images to predict low BMD, using QCT as standard of reference, which can help to identify low BMD from normal.

Radiomics is a promising technique using computerized quantitative imaging analysis to extract a large number of image-related features to assist in diagnosing diseases^{19,20}. First-order features describe the statistic parameters of voxel gray-level HU intensities within the VOI²¹. Radiomics texture features extracted from medical bone images contain meaningful information, which could be used to predict bone mineral disorders more accurately²². As proven in previous studies, these features might describe an indirect surrogate of BMD^{21,23}. In this study, we have extracted six classes of quantitative radiomics features based on the chest LDCT, describing the distribution and spatial arrangement of voxel intensities within the VOI. Firstly, we developed the fourteen radiomics models based on the radiomics features from single vertebra of T1 to L2, respectively. Among fourteen models, our single vertebra RFs model of L1 yielded the highest AUC of 0.884 (95%CI: 0.845–0.923), subsequently of L2 (AUC=0.877, 95%CI: 0.838–0.916) for low BMD prediction in the validation set. Our results showed that the radiomics analysis based on single vertebra in chest LDCT images can discriminate low BMD from normal, and the single vertebra RFs model of L1 showed acceptable predictive performance.

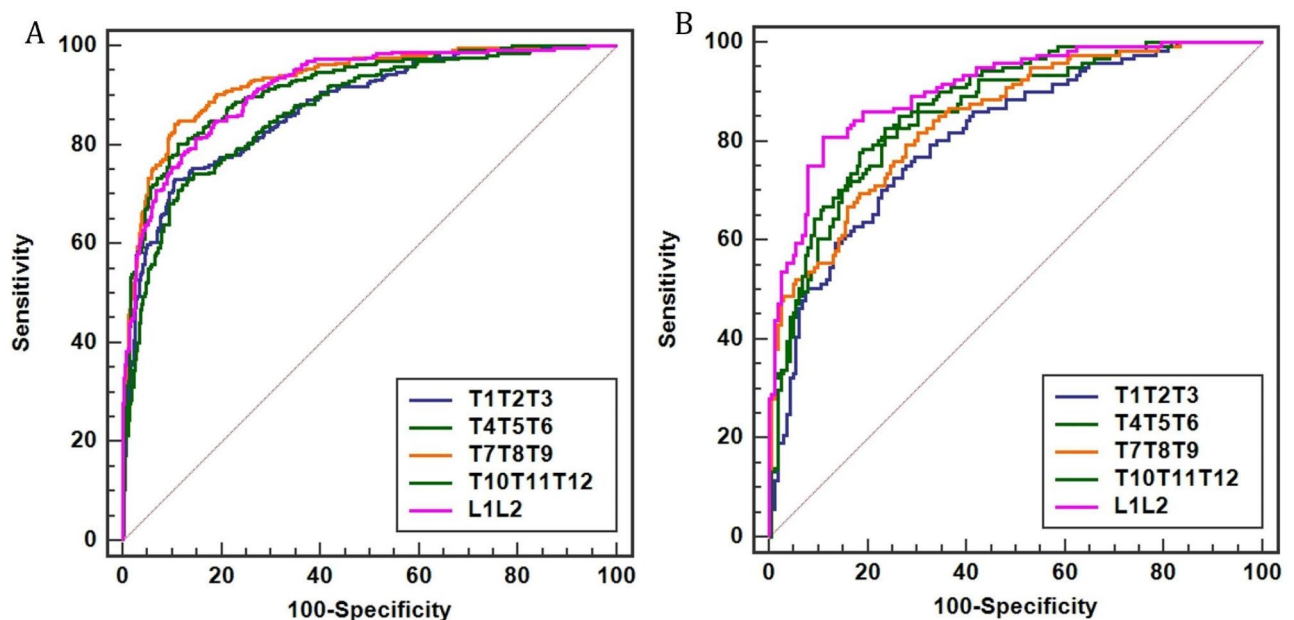


Fig. 4. ROC curves of low BMD prediction in the training (A) and validation (B) sets based on the Radscore models. ROC, receiver operating characteristic; BMD, bone mineral density.

Both first-order features and texture features significantly contributed to the radiomics models²¹. In our study, the frequent occurrence of first-order features, GLCM features, and GLSZM features in the different single vertebra RFs models showed that these features contribute significantly for the development of these models based on chest LDCT. Previous studies have shown a significant association between BMD-decrease and the first order statistics change²⁴. It is reasonable to speculate that the first-order parameters could describe a regional BMD representative of the vertebrae cancellous bone²¹. Some studies have shown that there is a certain correlation between the texture parameters of cancellous bone extracted from radiological images and bone microstructure^{25,26}. Moreover, GLCM features have been reported to be linked to known pathological changes in the osteoporotic bone^{23,27–29}. Yao et al.³⁰ have used three first-order features and two GLSZM features to predict low BMD based the fat-water imaging of dual-energy spectral CT (T11–L2), and showed that original_firstorder_90Percentile and wavelet_HLH_firstorder_Minimum contribute more to the radiomics model, and GLSZM zone% and Size Zone Nonuniformity also had contributions to the radiomics model³⁰. Wang et al.³¹ have reported that original_firstorder_Mean was the most predictive feature, and osteoporosis had a lower mean gray level intensity than non-osteoporosis. In our study, the GLSZM_HighGrayLevelZoneEmphasis contributed to develop thirteen out of fourteen single vertebra RFs models. Our results also meant that the grey value distribution within the VOI between the normal and low BMD were different³⁰.

Besides, we further developed the radiomics models based on mixed RFs and average Radscores of consecutive thoracic and upper lumbar vertebrae in chest LDCT images. The Radscore models and mixed RFs models yielded the AUC values ranging from 0.809 to 0.906 and 0.792 to 0.883 for predicting low BMD in the validation set, respectively. Several studies^{12,32,33} have confirmed that the lumbar spine is the good observation site for bone loss in all skeletal structures. In this study, the Radscore model of L1–L2 was superior to other Radscore models of consecutive thoracic vertebrae ($P < 0.05$) in the validation set. However, there is no significant difference among the AUC values of three Radscore models constructed on the consecutive thoracic vertebrae of T4–T6 (AUC = 0.855), T7–T9 (AUC = 0.845), and T10–T12 (AUC = 0.871) for predicting low BMD in the validation set ($P > 0.1$). The AUC values of our Radscore model of L1–L2 in the training and validation sets were 0.917 and 0.906 respectively, indicating that it could effectively distinguish low BMD from normal.

Recently, several radiomics studies have focused on assessment of osteoporosis and osteopenia based on DXA, magnetic resonance (MR) or CT. Rastegar et al.¹³ investigated the diagnostic accuracy of bone mass classification based on DXA images using machine learning algorithms, showing a small sample capacity and a lower AUC that ranging between 0.50 and 0.78. He et al.³⁴ proposed the classification method of normal vs. osteopenia, normal vs. osteoporosis, and osteopenia vs. osteoporosis using radiomics based on T1-weighted and T2-weighted sagittal lumbar spine MR images, and the AUC values were 0.810, 0.797, and 0.769, respectively. There are several studies indicating radiomics approach based on CT images has good potential for osteoporosis prediction. Lim et al.¹⁴ evaluated the prediction performance of femoral osteoporosis using random forest algorithm analysis with radiomics features in 500 patients underwent both DXA and abdomen-pelvic CT, and their AUC values to predict osteoporosis were 0.959 in the training set and 0.96 in the validation set, respectively. Jiang et al.³⁵ included 386 vertebral bodies from lumbar spine CT scans to construct and validate a radiomics signature model to screen for lumbar spine osteoporosis using DXA as standard, and the AUC values of the radiomics signature models were 0.96 in the training set and 0.92 in the testing set. Most previous studies^{14,34,35} used the areal BMD measured by DXA as the reference standard, which may have a potential effect on the accuracy of these studies³¹.

It is well known that QCT can measure true volumetric BMD and provide more accurate BMD values than DXA⁷. Compared to two-dimensional image analysis, three-dimensional image analysis may be more useful for evaluating osteoporosis because the vertebrae have a three-dimensional complex structure. Moreover, Both the VOIs of radiomics features extraction and BMD measurement with QCT were three-dimensional. Hence, we used QCT as standard of reference to develop and validate radiomics models for low BMD prediction in chest LDCT images. In this study, our radiomics models were constructed on thoracic and upper lumbar spine in chest LDCT images, and the performance of our Radscore model of L1-L2 was comparable to the results of the aforementioned studies^{13,14,31,34,35}. Our Radscore model of L1-L2 had potential to serve as an important tool to screen low BMD based on chest LDCT. However, the procedure for radiomics analysis may seem complex. In this study, vertebrae segmentation was performed using an auto-segmentation algorithm. The progress of automated segmentation might allow the integration of the feature extraction and calculation into a software program³⁵. Low BMD screening by radiomics analysis may be easier and more effective.

There are several limitations in this study. First, this study was a retrospective study performed in a single center, which may have a possibility of participant selection bias. Second, we did not include any clinical features such as age, sex, BMI, only focusing on the chest LDCT images analysis. Finally, our study population was divided into two groups: normal and low BMD groups, which consist of osteopenia and osteoporosis patients. Add this was to ensure the balance of the number of participants between the groups because the patients with osteoporosis was significantly less than the normal and osteopenic patients in our included healthy check-up population. Moreover, the low BMD prediction can prompt early clinical attention. Thus, this study was focused on the prediction of low BMD rather than osteoporosis.

In conclusion, we developed and validated chest LDCT-based different radiomics models (i.e., single vertebra RFs model, mixed RFs models, and Radscore models) on thoracic and upper lumbar spine, with QCT as the standard of reference. Our Radscore model of L1-L2 had potential to serve as an important tool to predict low BMD in chest LDCT images. This radiomics method may provide valuable information in the prediction and screening of low BMD based on chest LDCT without additional radiation exposure.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 15 July 2024; Accepted: 6 December 2024

Published online: 28 December 2024

References

1. Consensus development conference: Diagnosis, prophylaxis, and treatment of osteoporosis. *Am J Med* **94**, 646–650 (1993). [https://doi.org/10.1016/0002-9343\(93\)90218-e](https://doi.org/10.1016/0002-9343(93)90218-e)
2. He, J., Fang, H. & Li, N. Vertebral bone marrow diffusivity in normal adults with varying bone densities at 3T diffusion-weighted imaging. *Acta Radiol.* **59**, 89–96. <https://doi.org/10.1177/0284185117704235> (2018).
3. Liu, J., Curtis, E. M., Cooper, C. & Harvey, N. C. State of the art in osteoporosis risk assessment and treatment. *J. Endocrinol. Invest.* **42**, 1149–1164. <https://doi.org/10.1007/s40618-019-01041-6> (2019).
4. Pasco, J. A. et al. The population burden of fractures originates in women with osteopenia, not osteoporosis. *Osteoporos. Int.* **17**, 1404–1409. <https://doi.org/10.1007/s00198-006-0135-9> (2006).
5. Pisani, P. et al. Major osteoporotic fragility fractures: Risk factor updates and societal impact. *World J. Orthop.* **7**, 171–181. <https://doi.org/10.5312/wjo.v7.i3.171> (2016).
6. Kanis, J. A. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: Synopsis of a WHO report. WHO Study Group. *Osteoporos. Int.* **4**, 368–381. <https://doi.org/10.1007/bf01622200> (1994).
7. Löffler, M. T. et al. Improved prediction of incident vertebral fractures using opportunistic QCT compared to DXA. *Eur. Radiol.* **29**, 4980–4989. <https://doi.org/10.1007/s00330-019-06018-w> (2019).
8. Zhao, Y. et al. Fully automated radiomic screening pipeline for osteoporosis and abnormal bone density with a deep learning-based segmentation using a short lumbar mDixon sequence. *Quant. Imaging Med. Surg.* **12**, 1198–1213. <https://doi.org/10.21037/qims-21-587> (2022).
9. Kumar, V. et al. Radiomics: The process and the challenges. *Magn. Reson. Imaging.* **30**, 1234–1248. <https://doi.org/10.1016/j.mri.2012.06.010> (2012).
10. Raman, S. P., Chen, Y., Schroeder, J. L., Huang, P. & Fishman, E. K. CT texture analysis of renal masses: Pilot study using random forest classification for prediction of pathology. *Acad. Radiol.* **21**, 1587–1596. <https://doi.org/10.1016/j.acra.2014.07.023> (2014).
11. Wang, H. et al. Semiquantitative computed tomography characteristics for lung adenocarcinoma and their association with lung cancer survival. *Clin. Lung Cancer.* **16**, e141–e163. <https://doi.org/10.1016/j.clcc.2015.05.007> (2015).
12. Xie, Q. et al. Development and validation of a machine learning-derived radiomics model for diagnosis of osteoporosis and osteopenia using quantitative computed tomography. *BMC Med. Imaging.* **22**, 140. <https://doi.org/10.1186/s12880-022-00868-5> (2022).
13. Rastegar, S. et al. Radiomics for classification of bone mineral loss: A machine learning study. *Diagn. Interv. Imaging.* **101**, 599–610. <https://doi.org/10.1016/j.diii.2020.01.008> (2020).
14. Lim, H. K., Ha, H. I., Park, S. Y. & Han, J. Prediction of femoral osteoporosis using machine-learning analysis with radiomics features and abdomen-pelvic CT: A retrospective single center preliminary study. *PloS One.* **16**, e0247330. <https://doi.org/10.1371/journal.pone.0247330> (2021).
15. Curtis, E. M., Moon, R. J., Harvey, N. C. & Cooper, C. The impact of fragility fracture and approaches to osteoporosis risk assessment worldwide. *Bone* **104**, 29–38. <https://doi.org/10.1016/j.bone.2017.01.024> (2017).
16. Radiology, A. C. o. ACR–SPR–SSR practice parameter for the performance of musculoskeletal quantitative computed tomography (QCT). *Am. Coll. Radiol. Res.* **1076**, 6 (2018).
17. Grimes, D. A. & Schulz, K. F. Uses and abuses of screening tests. *Lancet* **359**, 881–884. [https://doi.org/10.1016/s0140-6736\(02\)09748-5](https://doi.org/10.1016/s0140-6736(02)09748-5) (2002).
18. Curtis, J. R. et al. Longitudinal trends in use of bone mass measurement among older americans, 1999–2005. *J. Bone Min. Res.* **23**, 1061–1067. <https://doi.org/10.1359/jbmr.080232> (2008).
19. Kang, B. et al. T1 stage clear cell renal cell carcinoma: A CT-based radiomics nomogram to estimate the risk of recurrence and metastasis. *Front. Oncol.* **10**, 579619. <https://doi.org/10.3389/fonc.2020.579619> (2020).

20. Mao, N. et al. Radiomics nomogram of DCE-MRI for the prediction of axillary lymph node metastasis in breast cancer. *Front. Oncol.* **10**, 541849. <https://doi.org/10.3389/fonc.2020.541849> (2020).
21. Biamonte, E. et al. Artificial intelligence-based radiomics on computed tomography of lumbar spine in subjects with fragility vertebral fractures. *J. Endocrinol. Invest.* **45**, 2007–2017. <https://doi.org/10.1007/s40618-022-01837-z> (2022).
22. Hirvasniemi, J. et al. Bone texture analysis for prediction of incident radiographic hip osteoarthritis using machine learning: Data from the Cohort hip and cohort knee (CHECK) study. *Osteoarthr. Cartil.* **27**, 906–914. <https://doi.org/10.1016/j.joca.2019.02.796> (2019).
23. Valentini, A. et al. Opportunistic osteoporosis screening in multi-detector CT images via local classification of textures. *Osteoporos. Int.* **30**, 1275–1285. <https://doi.org/10.1007/s00198-019-04910-1> (2019).
24. Zou, D., Li, W., Deng, C., Du, G. & Xu, N. The use of CT Hounsfield unit values to identify the undiagnosed spinal osteoporosis in patients with lumbar degenerative diseases. *Eur. Spine J.* **28**, 1758–1766. <https://doi.org/10.1007/s00586-018-5776-9> (2019).
25. Jeong, H., Kim, J., Ishida, T., Akiyama, M. & Kim, Y. Computerised analysis of osteoporotic bone patterns using texture parameters characterising bone architecture. *Br. J. Radiol.* **86**, 20101115. <https://doi.org/10.1259/bjr.20101115> (2013).
26. Baum, T. et al. Trabecular bone structure analysis of the spine using clinical MDCT: Can it predict vertebral bone strength? *J. Bone Min. Metab.* **32**, 56–64. <https://doi.org/10.1007/s00774-013-0465-6> (2014).
27. Kawashima, Y. et al. Using texture analysis of head CT images to differentiate osteoporosis from normal bone density. *Eur. J. Radiol.* **116**, 212–218. <https://doi.org/10.1016/j.ejrad.2019.05.009> (2019).
28. S, R. & R, M. An automatic bone disorder classification using hybrid texture feature extraction with bone mineral density. *Asian Pac. J. Cancer Prev.* **19**, 3517–3523. <https://doi.org/10.31557/apjcp.2018.19.12.3517> (2018).
29. Muehlethaler, U. J. et al. Vertebral body insufficiency fractures: Detection of vertebrae at risk on standard CT images using texture analysis and machine learning. *Eur. Radiol.* **29**, 2207–2217. <https://doi.org/10.1007/s00330-018-5846-8> (2019).
30. Yao, Q. et al. Radiomics Nomogram based on dual-energy spectral CT imaging to diagnose low bone mineral density. *BMC Musculoskelet. Disord.* **23**, 424. <https://doi.org/10.1186/s12891-022-05389-4> (2022).
31. Wang, J. et al. Prediction of osteoporosis using radiomics analysis derived from single source dual energy CT. *BMC Musculoskelet. Disord.* **24**, 100. <https://doi.org/10.1186/s12891-022-06096-w> (2023).
32. Samelson, E. J. et al. QCT measures of bone strength at the thoracic and lumbar spine: The Framingham study. *J. Bone Min. Res.* **27**, 654–663. <https://doi.org/10.1002/jbmr.1482> (2012).
33. Mao, S. S. et al. Thoracic quantitative computed tomography (QCT) can sensitively monitor bone mineral metabolism: Comparison of thoracic QCT vs lumbar QCT and dual-energy X-ray absorptiometry in detection of age-related change in bone Mineral Density. *Acad. Radiol.* **24**, 1582–1587. <https://doi.org/10.1016/j.acra.2017.06.013> (2017).
34. He, L. et al. Radiomics based on lumbar spine magnetic resonance imaging to detect osteoporosis. *Acad. Radiol.* **28**, e165–e171. <https://doi.org/10.1016/j.acra.2020.03.046> (2021).
35. Jiang, Y. W., Xu, X. J., Wang, R. & Chen, C. M. Radiomics analysis based on lumbar spine CT to detect osteoporosis. *Eur. Radiol.* **32**, 8019–8026. <https://doi.org/10.1007/s00330-022-08805-4> (2022).

Acknowledgements

Y.L.P. discloses support for the research of this work from the National Natural Science Foundation of China [grant number 82302303], Medical Science and Technology Project of Zhejiang Province [grant number 2023KY493]. Y.J.W. discloses support for publication of this work from Medical Science and Technology Project of Zhejiang Province [grant number 2023KY452, 2024KY625]. Y.B.W. discloses support for publication of this work from Medical Science and Technology Project of Zhejiang Province [grant number 2025KY551].

Author contributions

YJ.W., HG.W. contributed to study conception and design. YB.W., CM.L., Q.Y., J.L., H.Y. J contributed to data collection. YL.P., YD.W., YJ.W., HG.W. was responsible for statistical analysis, tables and charts. YL.P., YD.W. was responsible for drafting manuscript. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-82642-x>.

Correspondence and requests for materials should be addressed to H.W. or Y.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024