Original Article

The diagnostic cutoff value of vertebral bone quality score for osteoporosis is significantly influenced by the magnetic field in patients undergoing lumbar surgery

Fangke Hu[1#], Bing Chen[2,3#], Hanming Bian[1,3], Dong Zhao[1], Chao Chen[1], Gang Liu[1], Xinlong Ma[1], Zheng Wang[4], Qiang Yang[1]

[1]Department of Spine Surgery, Tianjin Hospital, Tianjin University, Tianjin, China; [2]Department of Spine Surgery, Shandong Provincial Third Hospital, Shandong University, Jinan, China; [3]Clinical College of Orthopedics, Tianjin Medical University, Tianjin, China; [4]Department of Orthopaedics, Chinese PLA General Hospital, Beijing, China

Contributions: (I) Conception and design: F Hu, B Chen, X Ma, Z Wang, Q Yang; (II) Administrative support: C Chen, X Ma, Z Wang, Q Yang; (III) Provision of study materials or patients: H Bian, D Zhao, Q Yang; (IV) Collection and assembly of data: F Hu, B Chen, G Liu; (V) Data analysis and interpretation: F Hu, B Chen, H Bian, C Chen, G Liu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Qiang Yang, MD, PhD. Department of Spine Surgery, Tianjin Hospital, Tianjin University, No. 406 Jiefangnan Road, Hexi District, Tianjin 300211, China. Email: yangqiang1980@126.com; Zheng Wang, MD, PhD. Department of Orthopaedics, Chinese PLA General Hospital, Haidian District, No. 28 Fuxing Road, Beijing 100853, China. Email: wangzheng301@163.com.

Background: The diagnostic efficacy of the magnetic resonance imaging (MRI)-based vertebral bone quality (VBQ) score for osteoporosis in patients undergoing lumbar surgery has been documented; however, whether the VBQ is influenced by different magnetic field strengths (1.5 T and 3.0 T) remains controversial. This study aimed to investigate the influence of magnetic fields on the diagnostic efficacy of the VBQ score via self-paired patients who underwent both 1.5 T and 3.0 T MRI scanning.

Methods: This retrospective study included consecutive patients aged >50 years who underwent ≥2 MRI examinations within a 6-month period at our department for degenerative lumbar surgery from June 2022 to June 2024. Patients were included only if they had both dual-energy X-ray absorptiometry (DXA) and quantitative computed tomography (QCT) results obtained within the same period. There were three MRI scanners in our institution: the GE 1.5 T (General Electric), the GE 3.0 T (General Electric), and the PH 3.0 T (Philips Healthcare, Amsterdam, the Netherlands). The self-paired VBQ scores were compared between the different magnetic fields and scan manufacturers using paired *t*-tests. The Steiger's Z test, correlation analysis, logistic regression analysis, and receiver operating characteristic (ROC) curve analysis were performed to explore the influence of the magnetic fields on the diagnostic value of the VBQ score.

Results: A total of 104 patients were included who underwent ≥2 MRI scans with available DXA and QCT results. The mean age was 65.3±8.8 years with 56.7% female participants. In self-paired analysis, a significantly higher VBQ score was found in 1.5 T compared with 3.0 T (3.2±0.59 *vs.* 2.71±0.66, P<0.001), whereas no difference was observed between different MRI scanners of 3.0 T. Correlation analysis and logistic regression analysis of the self-paired patients revealed that the VBQ 1.5 T was more accurately associated with osteoporosis than the VBQ 3.0 T, with the strongest relationship in VBQ 1.5 T with QCT (r=–0.529). ROC analysis revealed that the diagnostic efficacy for osteoporosis was comparable between the VBQ 1.5 T and the VBQ 3.0 T [QCT-defined osteoporosis: VBQ 1.5 T area under the curve (AUC) =0.709 *vs.* VBQ 3.0 T AUC =0.697; DeLong test P=0.904].

Conclusions: The VBQ 1.5 T has a significantly higher cutoff value for osteoporosis than the VBQ 3.0

T in patients undergoing lumbar surgery. It is essential to mention the magnetic resonance (MR) field when applying the VBQ score in the screening of osteoporosis.

Keywords: Vertebral bone quality (VBQ); osteoporosis; quantitative computed tomography (QCT); dual-energy X-ray absorptiometry (DXA); magnetic resonance imaging (MRI)

Submitted Jan 22, 2025. Accepted for publication Aug 07, 2025. Published online Sep 12, 2025.
doi: 10.21037/qims-2025-173
View this article at: https://dx.doi.org/10.21037/qims-2025-173

Introduction

It is essential to evaluate the preoperative bone mineral density (BMD) in patients undergoing lumbar surgery (1). The novel magnetic resonance imaging (MRI)-based vertebral bone quality (VBQ) score has emerged as an alternative for the screening of osteoporosis (2,3). The correlation of the VBQ with lumbar BMD as measured by dual-energy X-ray absorptiometry (DXA) and quantitative computed tomography (QCT) has been confirmed (4), and the diagnostic value for osteoporosis and osteopenia has also been summarized (2,3). Furthermore, the prognostic value of the VBQ score for postoperative complications such as cage subsidence and screw loosening has also been confirmed (5-8), making it an alternative tool for the radiation-free screening of osteoporosis (4).

Firstly proposed by Ehresman *et al.* in 2019 (9), the VBQ was adjusted to different MRI systems by the cerebral spinal fluid (CSF) signal adjustment. Initially, the VBQ score was proposed not to be influenced by the magnetic resonance (MR) magnetic field, and several authors from different regions also reported excellent intra- and inter-observer reliability for VBQ measurements (3,10-12). However, recent evidence has indicated that the magnetic field can significantly influence the VBQ score (4,13,14), with great variance in diagnostic cutoff values, and reported thresholds for osteoporosis as high as 3.705 for the VBQ 1.5 T compared with 2.605 for the VBQ 3.0 T (4).

By taking advantage of our clinical resources, this study aimed to investigate the consistency of the VBQ score under different MR fields and MR scanners, and summarize the screening efficacy of the VBQ between different magnetic fields. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-2025-173/rc) (15).

Methods

Patient inclusion and exclusion

This retrospective study enrolled all the consecutive patients aged >50 years who underwent degenerative lumbar surgery at the Spine Surgery Department of Tianjin Hospital between June 2022 and June 2024. The inclusion criteria were as follows: (I) availability of lumbar DXA, QCT, and ≥2 MRI scans within the period of 6 consecutive months; and (II) consent to join in the study. The exclusion criteria were as follows: (I) prior L1–L4 instrumentation; (II) spinal metastases or infection; (III) history of L1–L4 vertebral compression fractures with affecting ≥2 vertebral bodies; (IV) inadequate image quality or inability to measure the lumbar T-score, QCT, or VBQ; and (V) medication of teriparatide for anti-osteoporosis treatment. Collected data included including age, gender, body mass index (BMI), medical history, laboratory findings, DXA, QCT, and MRI results. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The study was approved by the institutional ethics board of Tianjin Hospital and the requirement for individual consent for this retrospective analysis was waived.

DXA examinations and QCT measurements

DXA examinations and BMD measurements were acquired with lumbar spine (L1–L4) using a GE Lunar iDXA scanner (enCORE software Version 16; GE Healthcare, Chicago, IL, USA) following manufacturer protocols. Lumbar computed tomography (CT) datasets were obtained using a GE 64-row CT scanner (CT660, GE Healthcare) and translated to a QCT workstation (Mindways QCT Pro, Version 6.1, Mindways Software, Austin, TX, USA) for the volumetric BMD (vBMD) measurements. Detailed

protocols were followed, as described elsewhere (16). According to the DXA diagnostic criteria for osteoporosis, we defined osteoporosis as T-score ≤–2.5 in the lumbar spine (17-19). According to the QCT criteria for the L1/L2 average vBMD according to the American College of Radiology classification, osteoporosis was defined as vBMD <80 mg/cm^3 (20,21).

MR examinations and VBQ measurements

There were three MR scanners available for lumbar imaging: a GE 1.5 T (HDXT ECHOSPEED 16 CHANNEL, GE Healthcare), a GE 3.0 T (GE Discovery MR750 3.0 T, GE Healthcare), and a PH 3.0 T (Philips Ingenia 3.0 T Cx, Philips Healthcare). Imaging parameters including repetition time (TR) and echo time (TE) are detailed in *Table 1*. All scans used 5 mm slice thickness. The VBQ scores were recorded as VBQ 1.5 T, VBQ GE 3.0 T, and VBQ PH 3.0 T, respectively.

Table 1 Imaging protocol of the three MR scanners

MR scanners	TR, ms	TE, ms
GE 1.5 T, median (range)	540 (480–840)	11.4 (8.4–12.6)
GE 3.0 T, median (range)	669 (606–872)	14.4 (9.38–14.6)
PH 3.0 T, median (range)	585 (583–771)	10.0 (10.0–16.5)

GE, General Electric; MR, magnetic resonance; PH, Philips; TE, echo time; TR, repetition time.

The VBQ score was calculated from the T1-weighted images (9). The MR images were measured by the Vue PACS software (Version 12.2.1.0104, Philips Healthcare). As shown in *Figure 1*, the midsagittal trabecular portions of the L1–L4 vertebral bodies were used to measure the signal intensities (SI). To normalize this median value, the SI of CSF adjacent to L3 was used as the denominator. The area of region of interest (ROI), mean SI, and standard deviation (SD) were recorded. An elliptical ROI of 140–160 mm^2 was used for the L1–L4 vertebral body and 20–40 mm^2 was used for the ventral CSF at L3. In cases of modic change, severe scoliosis, vertebral hemangiomas, and compression fracture, the remaining vertebral bodies or para-sagittal sections or CSF signals at L2/L4 could be used. Then, the VBQ score was calculated by taking the median SI of the L1–L4 vertebrae divided by the SI of the L3 CSF, and the formula was as follows (9,22):

$$VBQ\ score = \frac{SI\left(L1-L4\right)_{median}}{SI\left(L3_{CSF}\right)} \quad [1]$$

Two independent authors measured the VBQ score while blinded to DXA/QCT results, and the mean value was adopted. For patients undergoing multiple scans under the same MR scanner, the average value of the VBQ scores was adopted for the calculation.

Statistical analysis

Continuous variables were expressed as the mean ± SD,

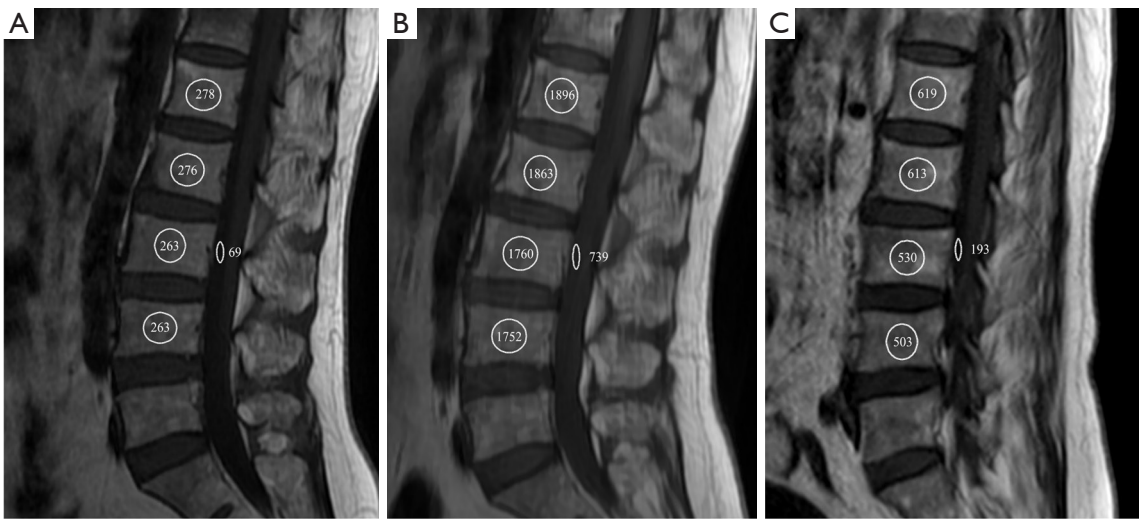

Figure 1 Representative VBQ measurements of a 67-year-old female patient across different MRI systems. (A) GE 1.5 T scanner; (B) GE 3.0 T scanner; (C) PH 3.0 T scanner. GE, General Electric; MRI, magnetic resonance imaging; PH, Philips; VBQ, vertebral bone quality.

and categorical variables are expressed as the frequency and percentage. Normality was assessed using Shapiro-Wilk tests. Continuous variables conforming to a normal distribution were analyzed with the Student's *t*-test. Paired continuous variables were analyzed with the paired *t*-test. One-way analysis of variance (ANOVA) was used for multiple-group analysis. The intraclass correlation coefficient (ICC, two-way mixed and absolute agreement), coefficient of variation (CV), and the Bland-Altman plot were performed to detect the consistency of the VBQ methods (SPSS AU online analysis, https://spssau.com/indexs.html). The correlations of the variables were analyzed by scatter plot and quantified by Pearson's correlation. Steiger's Z test was used to compare the significance of the self-paired correlations (https://www.psychometrica.de/correlation.html). Prognostic values were evaluated via binary logistic regression and expressed in odds ratio (OR) and 95% confidence interval (CI). Diagnostic performances were evaluated through receiver operating characteristic curve (ROC) analysis, and the area under the curve (AUC) was calculated. Youden's index was used to determine the threshold with the best sensitivity and specificity. DeLong's tests compared the significance of different AUCs (SPSS AU online analysis). Analyses were conducted using the software SPSS 26.0 (IBM Corp., Armonk, NY, USA). Figures were generated using the SigmaPlot software (Version 11.0, Systat Software, Inc., Erkrath, Germany). Statistical significance was assigned as a two-tailed P value <0.05.

Results

Patient characteristics

Figure 2 presents the patient selection flowchart. From an initial cohort of 1,505 patients who had lumbar MR and DXA examinations, 609 patients had available QCT results. After applying the inclusion criteria of at least two MR examinations performed within the period of 6 consecutive months of DXA and QCT, only 155 patients were included. According to the exclusion criteria, 51 patients were excluded, including: (I) lumbar instrumentation of L1–L4 (n=26); (II) spinal metastases or spinal infection (n=5); (III) multiple L1–L4 vertebral compression fractures (n=13); (IV) inadequate image inability to measure VBQ score (n=2); and (V) teriparatide treatment (n=5). The final analysis included 104 patients. Participants had mean age of 65.3±8.8 years, 56.7% were female (59/104), and the mean BMI was 25.9±3.7 kg/m^2. A total of 65 patients underwent two MR examinations, 27 patients underwent three scans, and 12 patients underwent ≥4 scans. The representative MR images and VBQ score calculation of a 67-year female patient are presented in *Figure 1*, showing VBQ values of 3.906 (GE 1.5 T), 2.451 (GE 3.0 T), and 2.961 (PH 3.0 T).

The consistency of VBQ score between different MRI fields and scanners

As presented in *Table 2* and *Figure 2*, among the 104 participants, 93 underwent GE 1.5 T, 55 received GE 3.0 T, and 42 were examined with PH 3.0 T. Baseline characteristics were comparable across groups. Significant higher VBQ was found in total 1.5 T compared with total 3.0 T (3.20±0.54 *vs.* 2.71±0.64, P<0.001 by *t*-test).

The total distribution of the VBQ scores of the 104 cases is presented in *Figure 3A*, and the VBQ was slightly higher at the 1.5 T than the 3.0 T. A total of 54 patients had repeat 1.5 T VBQ measurements (mean difference 0.38±0.34 between scans). ICC analysis revealed moderate reliability (ICC =0.694, 95% CI: 0.464–0.825), with CV =8.5% suggesting low variability. A total of 17 patients had repeat GE 3.0 T measurements (mean difference 0.41±0.33), demonstrating good reliability (ICC =0.820, 95% CI: 0.430–0.945) and CV =11.0%, indicating low variability for GE 3.0 T measurements.

As shown in *Table 2*, self-paired analysis of 12 patients undergoing both GE and PH 3.0 T MRI revealed comparable values (GE 3.0 T: 2.65±0.50 *vs.* PH 3.0 T: 2.70±0.63), showing no significant difference (*Figure 3B*, P=0.831, paired *t*-test). The Bland-Altman plots were performed to detect the consistency between VBQ GE 3.0 T and VBQ PH 3.0 T, and no difference was found (mean difference 0.052, P=0.389). As no difference was found, we combined the VBQ GE 3.0 T and VBQ PH 3.0 T into the group of VBQ 3.0 T.

Self-paired analysis of 62 patients undergoing both 1.5 T and 3.0 T MRI (*Table 2*) revealed significantly higher VBQ at 1.5 T (3.20±0.59 *vs.* 2.71±0.66; P<0.001, paired *t*-test). As presented in *Figure 3C*, in the Bland-Altman plot, the mean difference was 0.488 and significant difference was found (P<0.001). The histogram presented in *Figure 3D* shows a significantly higher VBQ in 1.5 T by the self-paired comparison.

In summary, although VBQ demonstrated acceptable consistency within the same scanner and negligible

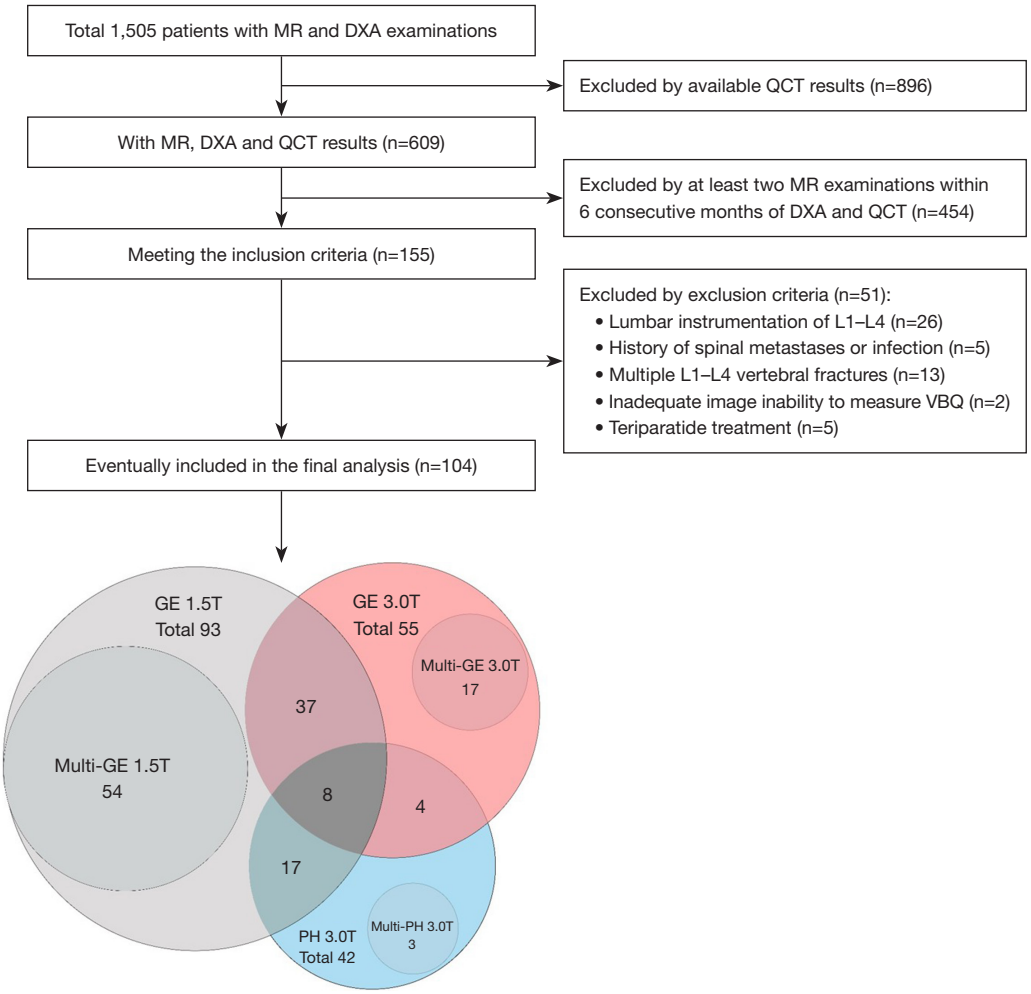

Figure 2 Flowchart of patient inclusion/exclusion process and distribution of the overlapped patients. DXA, dual-energy X-ray absorptiometry; MR, magnetic resonance; GE, General Electric; PH, Philips; QCT, quantitative computed tomography.

differences between 3.0 T scanner manufacturers, a significantly higher VBQ score could be detected in 1.5 T than 3.0 T.

VBQ 1.5 T was associated with osteoporosis more significantly than VBQ 3.0 T

Correlation analyses (*Table 3*) revealed that among 62 self-paired patients undergoing both 1.5 T and 3.0 T examinations, both the VBQ 1.5 T and VBQ 3.0 T were significantly correlated with QCT and age. Only the VBQ 1.5 T had significant correlations with the following variables of T-score, weight, and hemoglobin. For other variables, such as BMI, glucose, alkaline phosphatase, albumin, creatinine, glutamic-pyruvic transaminase, uric acid, triglyceride, total cholesterol, high- and low-density lipoprotein cholesterol, no correlation with VBQ 1.5 T and VBQ 3.0 T was found.

VBQ 1.5 T showed moderate correlation with 3.0 T values (r=0.582, *Table 3*). According to the QCT criteria for osteoporosis (*Figure 4A*), significantly higher VBQ 1.5 T compared with VBQ 3.0 T was found in both the self-paired patients of osteoporosis and non-osteoporosis (P<0.001, paired *t*-test). As shown in *Figure 4B*, VBQ 1.5 T was associated with QCT more significantly than VBQ 3.0 T (Steiger's Z test, P=0.046). The same result was found according to the T-score criteria for osteoporosis (Steiger's Z test, P=0.038).

Table 2 Characteristics of the overlapped patients according to different comparisons

Characteristics	Unpaired comparison				Self-paired comparison	
	GE 1.5 T group (N=93)	GE 3.0 T group (N=55)	PH 3.0 T group (N=42)	Total (N=104)	GE 3.0 T *vs.* PH 3.0 T (N=12)	1.5 T *vs.* total 3.0 T (N=62)
Gender: female/male	54/39	28/27	25/17	59/45	6/6	36/26
Age (years)	65.8±8.3	65.8±9.3	65±9.0	65.3±8.8	63.0±8.6	67.2±8.5
Weight (kg)	70.4±12.3	71.1±12	72.2±10.7	70.1±11.9	72.7±8.7	71.4±13.0
BMI (kg/m^2)	26.0±3.8	25.8±3.5	26.4±3.2	25.9±3.7	25.9±1.8	26.2±3.8
BMD (g/cm^2)	1.1±0.2	1.1±0.2	1.0±0.1	1.1±0.2	1.2±0.2	1.1±0.2
T-score	−0.8±1.9	−0.6±1.9	−0.6±1.9	−0.8±1.9	0.8±1.8	−0.9±1.8
QCT (mg/cm^3)	85.2±34.7	90.8±33.6	92.8±36.2	88.1±35.3	104.7±37.3	83.4±31.3
VBQ 1.5 T	3.2±0.5 (N=93)	3.1±0.5 (N=45)	3.3±0.7 (N=25)	3.2±0.5 (N=93)	3.2±0.7 (N=8)	3.2±0.6 (N=62)
VBQ 3.0 T total	2.7±0.7 (N=62)	2.6±0.6 (N=55)	2.9±0.6 (N=42)	2.7±0.6 (N=73)	2.7±0.5 (N=12)	2.7±0.7 (N=62)
GE 3.0 T	2.6±0.6 (N=45)	2.6±0.6 (N=55)	2.7±0.5 (N=12)	2.6±0.6 (N=55)	2.7±0.5 (N=12)	2.6±0.6 (N=45)
PH 3.0 T	3.0±0.7 (N=25)	2.7±0.6 (N=12)	2.9±0.6 (N=42)	2.9±0.6 (N=42)	2.7±0.6 (N=12)	3.0±0.7 (N=25)
Hemoglobin (g/L)	134.5±16.4	135.3±15.5	135.6±14.3	134.1±15.9	140.4±10.7	134.3±16.5
Triglyceride (mmol/L)	1.9±1.4	2.0±1.6	1.8±1.3	1.9±1.4	2.1±1.7	1.8±1.4
CHOL (mmol/L)	5.0±0.9	4.9±0.8	4.7±0.8	5.0±0.9	4.7±0.8	4.9±0.9
HDLC (mmol/L)	1.2±0.3	1.1±0.3	1.1±0.3	1.1±0.3	1.0±0.3	1.2±0.3
LDLC (mmol/L)	2.8±0.8	2.8±0.8	2.7±0.8	2.8±0.8	2.6±0.8	2.8±0.8

Continuous variables are expressed as the mean ± standard deviation. BMD, bone mineral density; BMI, body mass index; CHOL, total cholesterol; GE, General Electric; HDLC, high-density lipoprotein cholesterol; LDLC, low-density lipoprotein cholesterol; PH, Philips; QCT, quantitative computed tomography; VBQ, vertebral bone quality.

A high VBQ score was a risk factor for osteoporosis

Binary logistic regression (*Table 4*) revealed that using QCT diagnostic criteria, significantly higher ORs were observed in osteoporotic versus non-osteoporotic patients (VBQ total: OR =9.68; VBQ 1.5 T: OR =4.36; VBQ 3.0 T: OR =3.99). According to the T-score criteria for osteoporosis, only the VBQ 1.5 T (OR =3.21), not the VBQ total and VBQ 3.0 T, was detected as a risk factor for osteoporosis. Collectively, VBQ 1.5 T showed stronger associations with osteoporosis than VBQ 3.0 T.

Comparative diagnostic performance of VBQ across magnetic field strengths

ROC analyses evaluated the diagnostic efficacy for osteoporosis (*Table 5*). The diagnostic efficacy of T-score, BMD, and QCT for osteoporosis is presented in *Figure 5A,5B* as references. According to the QCT criteria for osteoporosis, reference standards showed the following: T-score (AUC =0.822, P<0.001) and BMD (AUC =0.810, P<0.001). For QCT-defined osteoporosis (*Figure 5C*), VBQ 1.5 T AUC =0.709 (P=0.005) and VBQ 3.0 T AUC=0.697 (P=0.008), with optimal cutoffs of: VBQ 1.5 T 3.018 (sensitivity =76.7%, specificity =62.5%) and VBQ 3.0 T 2.765 (sensitivity =60.0%, specificity =71.9%). DeLong tests revealed no significant AUC differences between T-score, BMD, and the VBQ scores (1.5 T *vs.* 3.0 T: P=0.904; 1.5 T *vs.* T-score: P=0.119; 3.0 T *vs.* T-score: P=0.164).

Using T-score criteria for osteoporosis, the diagnostic efficacy of QCT was AUC 0.896 (P<0.001). As presented in *Figure 5D*, a significant diagnostic effect of VBQ 1.5 T was detected (AUC =0.738, P=0.009; optimal cutoff =2.917), whereas no significance could be found for VBQ 3.0 T (AUC =0.586, P=0.346). DeLong tests indicated no

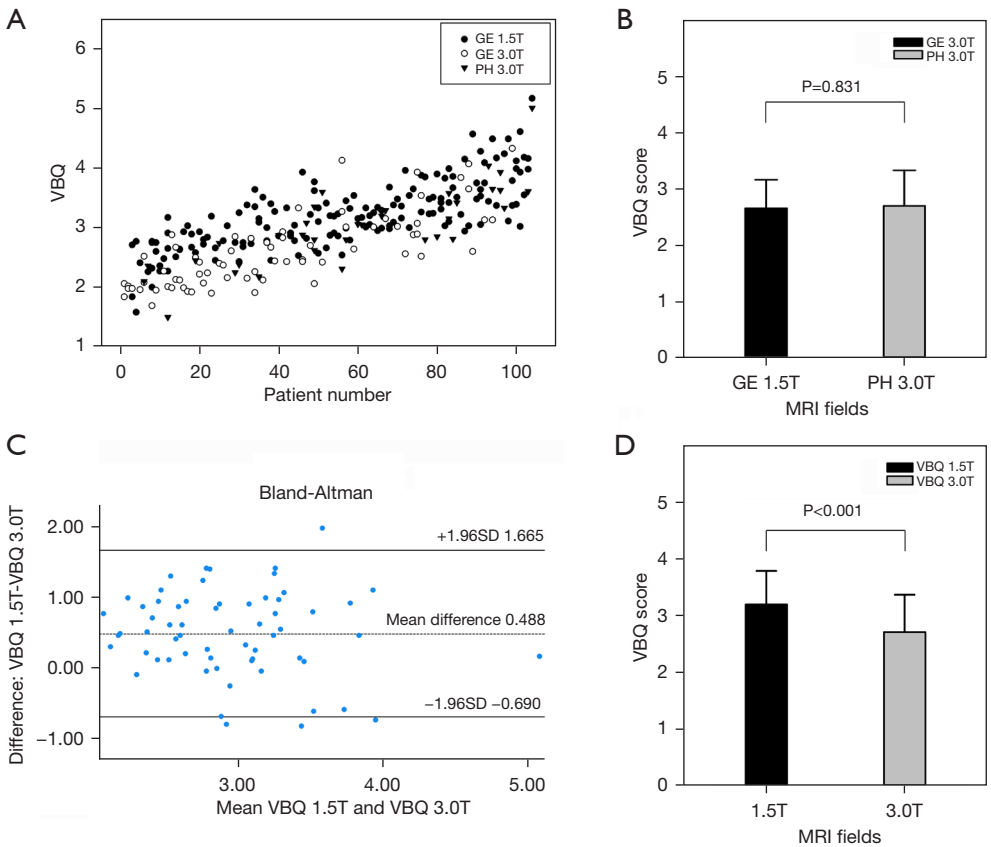

Figure 3 The consistency and difference of the VBQ score among different MRI fields and scanners. (A) VBQ score distribution of the 104 patients who underwent multi-MRI examinations; (B) no difference between GE 3.0 T and PH 3.0 T in 12 self-paired patients; (C) Bland-Altman plots between VBQ 1.5 T and VBQ 3.0 T; (D) significantly higher VBQ 1.5 T compared with VBQ 3.0 T in 62 self-paired patients. GE, General Electric; MRI, magnetic resonance imaging; PH, Philips; SD, standard deviation; VBQ, vertebral bone quality.

Table 3 Correlation analysis of the variables associated with osteoporosis by the self-paired patients

Variables	QCT (mg/cm^3)		VBQ 1.5 T		VBQ 3.0 T		Age		Weight (kg)		Hemoglobin (g/L)	
	r	P value	r	P value	r	P value	r	P value	r	P value	r	P value
T-score	0.679	<0.001*	−0.337	0.007*	−0.124	0.336	−0.139	0.282	0.436	<0.001*	0.359	<0.001*
QCT (mg/cm^3)	NA	NA	−0.529	<0.001*	−0.355	<0.001*	−0.490	<0.001*	0.285	0.030*	0.204	0.038*
VBQ 1.5 T	−0.529	<0.001*	NA	NA	0.582	<0.001*	0.375	<0.001*	−0.244	0.023*	−0.215	0.038*
VBQ 3.0 T	−0.355	<0.001*	0.582	<0.001*	NA	NA	0.315	0.007*	0.054	0.685	−0.146	0.256

Presented by the Pearson's correlation coefficient (r) and the Pearson's correlation test (P value). *, statistical significance of P<0.05. NA, not applicable; QCT, quantitative computed tomography; VBQ, vertebral bone quality.

difference between VBQ 1.5 T and VBQ 3.0 T, whereas significant superiority of QCT over the VBQ score could be found (1.5 T *vs.* 3.0 T: P=0.089; 1.5 T *vs.* QCT: P=0.020; 3.0 T *vs.* QCT: P=0.003). For the total VBQ score, the diagnostic effects were AUC =0.749 (QCT criteria) and AUC =0.708 (T-score criteria).

In summary, the diagnostic efficacy of the VBQ 1.5 T was comparable to that of VBQ 3.0 T. However, the AUCs of the VBQ scores were still not comparable to those of T-score, BMD, and QCT.

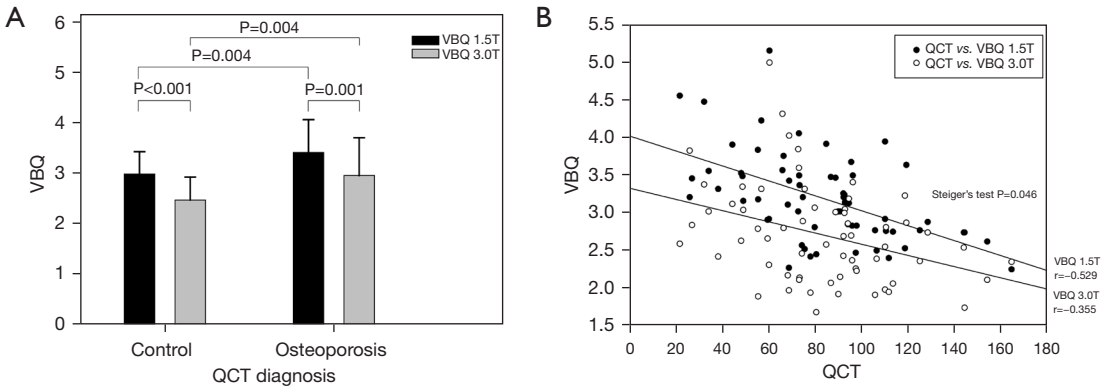

Figure 4 Magnetic field effects on the correlation of VBQ score with QCT-osteoporosis detected by the 62 self-paired patients. (A) Significantly higher VBQ 1.5 T compared with VBQ 3.0 T by QCT criteria; (B) correlation of VBQ 1.5 T/3.0 T with QCT. QCT, quantitative computed tomography; VBQ, vertebral bone quality.

Table 4 Binary logistic regression analysis of the variables for osteoporosis detected by the self-paired patients

Variables	QCT diagnosed osteoporosis					DXA diagnosed osteoporosis				
	Non-osteoporosis (n=32)	Osteoporosis (n=30)	P_1	OR (95% CI)	P_2	Non-osteoporosis (n=49)	Osteoporosis (n=13)	P_1	OR (95% CI)	P_2
T-score	0.03±1.71	−1.92±1.41	<0.001	0.43 (0.27–0.69)	<0.001	−0.32±1.58	−3.13±0.64	<0.001	NA	NA
BMD (g/cm^2)	1.18±0.21	0.94±0.17	<0.001	0.002 (0.00–0.05)	<0.001	1.13±0.19	0.79±0.08	<0.001	NA	NA
QCT (mg/cm^3)	107.1±21.1	57.4±17.2	<0.001	NA	NA	91.7±28.2	50.6±20.0	<0.001	0.93 (0.89–0.97)	<0.001
VBQ total	2.75±0.34	3.21±0.60	<0.001	9.68 (2.43–38.60)	0.001	2.91±0.54	3.23±0.44	0.053	2.95 (0.91–9.50)	0.070
VBQ 1.5 T	2.99±0.44	3.42±0.65	0.004	4.36 (1.48–12.88)	0.008	3.11±0.60	3.53±0.45	0.022	3.21 (1.10–9.40)	0.033
VBQ 3.0 T	2.47±0.46	2.96±0.75	0.004	3.99 (1.47–10.83)	0.007	2.68±0.68	2.82±0.59	0.515	1.35 (0.55–3.33)	0.509

Continuous variables are expressed as the mean ± standard deviation. P_1 value indicates the *t*-test between non-osteoporosis and osteoporosis; P_2 value indicates the binary logistic regression analysis of the risk factor for osteoporosis. BMD, bone mineral density; CI, confidence interval; DXA, dual-energy X-ray absorptiometry; NA, not available; OR, odds ratio; QCT, quantitative computed tomography; VBQ, vertebral bone quality.

Table 5 The diagnostic effect of the variables for osteoporosis detected by self-paired patients

Variables	QCT diagnosed osteoporosis						DXA diagnosed osteoporosis					
	AUC (95% CI)	P value	THRE	SENS	SPEC	YI	AUC (95% CI)	P value	THRE	SENS	SPEC	YI
T-score	0.822 (0.718–0.926)	<0.001	−1.650	0.700	0.875	0.575	NA	NA	NA	NA	NA	NA
BMD (g/cm^2)	0.810 (0.703–0.918)	<0.001	0.965	0.667	0.875	0.542	NA	NA	NA	NA	NA	NA
QCT (mg/cm^3)	NA	NA	NA	NA	NA	NA	0.896 (0.814–0.979)	<0.001	72.85	0.923	0.776	0.699
VBQ total	0.749 (0.625–0.873)	0.001	3.275	0.500	0.969	0.469	0.708 (0.552–0.864)	0.022	3.275	0.615	0.837	0.452
VBQ 1.5 T	0.709 (0.578–0.841)	0.005	3.018	0.767	0.625	0.392	0.738 (0.613–0.863)	0.009	2.917	1.000	0.490	0.490
VBQ 3.0 T	0.697 (0.565–0.829)	0.008	2.765	0.600	0.719	0.319	0.586 (0.409–0.762)	0.346	2.791	0.615	0.633	0.248

AUC, area under the curve; BMD, bone mineral density; CI, confidence interval; DXA, dual-energy X-ray absorptiometry; NA, not available; QCT, quantitative computed tomography; SENS, sensitivity; SPEC, specificity; THRE, threshold of VBQ score; VBQ, vertebral body quality; YI, Youden's index.

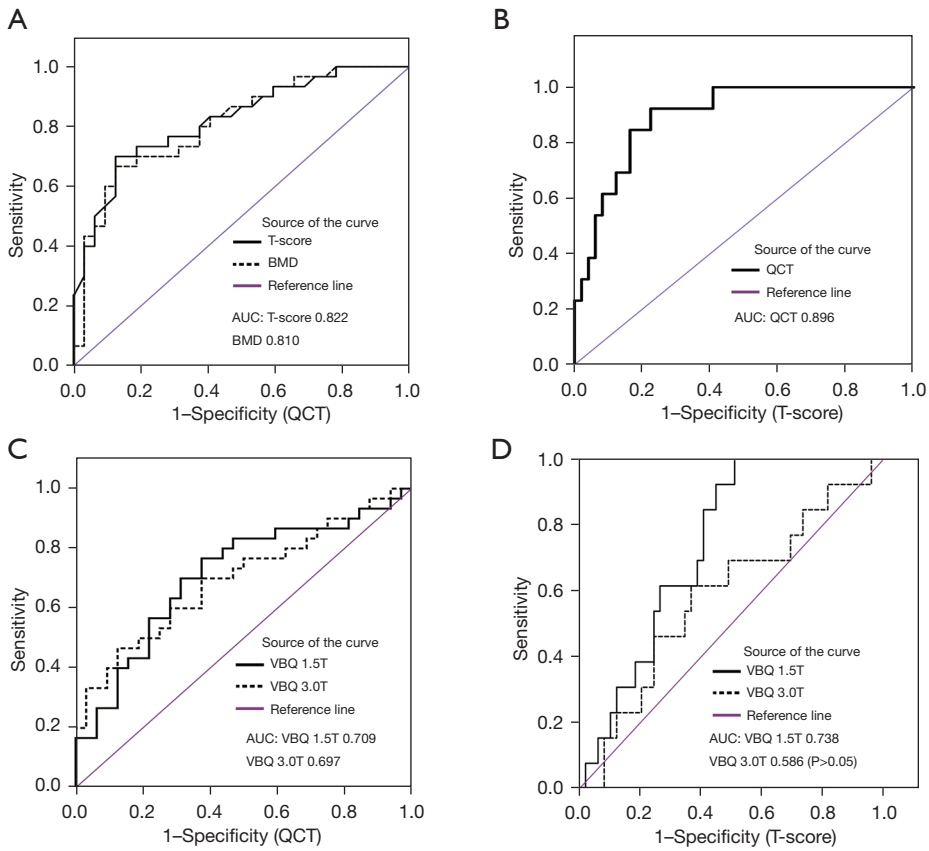

Figure 5 ROC curve analysis of the VBQ score for osteoporosis compared with T-score and QCT. (A) T-score/BMD performance using QCT reference; (B) QCT performance using T-score reference; (C) VBQ 1.5 T/3.0 T performance using QCT reference; (D) VBQ 1.5 T/3.0 T performance using T-score reference. AUC, area under the curve; BMD, bone mineral density; QCT, quantitative computed tomography; ROC, receiver operating characteristic; VBQ, vertebral bone quality.

Discussion

The purpose of this study was to explore the influence of magnetic fields on the diagnostic efficacy of the VBQ score via self-paired patients. We retrospectively analyzed 104 patients who had undergone at least two MR examinations, and the results demonstrated a significantly higher diagnostic threshold for VBQ 1.5 T than VBQ 3.0 T in patients undergoing lumbar surgery. The diagnostic efficacy for osteoporosis was comparable between the VBQ 1.5 T and VBQ 3.0 T.

The importance of osteoporosis assessment before lumbar surgery is well recognized (4,23). As MRI is performed as a routine examination before lumbar surgery, the novel MRI-based VBQ score has emerged as an opportunistic alternative tool for screening osteoporosis in the patients undergoing lumbar surgery (24). Early studies confirmed that osteoporosis is characterized by the localized replacement of adipocytes (25). The more severe the fat infiltration inside vertebral body, the more severe the degree of osteoporosis with impaired bone quality, which is reflected in an increased MR T1 SI (25,26). On this basis, Ehresman *et al.* overcame the baseline difference of the MR system by CSF signal adjustment and raised the VBQ score (1,27).

Initially, it was claimed that the VBQ score was generalizable across multiple MR systems from different manufacturers, such as Philips, GE, Toshiba, and Siemens; however, no details of the magnetic fields were provided (13,27,28). Recent studies have suggested to reconsider the influence of the magnetic field on the VBQ score (4,13,14). In 2023, Lin *et al.* (4) conducted a study to compare the VBQ 1.5 T and VBQ 3.0 T in patients undergoing spine surgery (n=452), which demonstrated that VBQ 1.5 T exhibited better discriminability for osteoporosis than VBQ

3.0 T, with the highest discriminative power of VBQ 1.5 T for QCT-osteoporosis (AUC =0.744); however, unpaired patients limited the statistical power of the study (4). Also in 2023, Liu *et al.* (13) conducted a study to assess the reproducibility of the VBQ score (n=136, underwent lumbar MRI twice within 3 months), and found that the MR technical factors, especially the magnetic field, had a significant impact on VBQ score; however, the exact influence was not explored (13). Liu *et al.* in 2023 (14) included 2,805 patients and found that the VBQ 1.5 T was significantly higher compared with VBQ 3.0 T (2.769±0.494 *vs.* 2.199±0.432); the authors also included 30 healthy adults who underwent MRI scans 4 times and found that the VBQ 1.5 T was significantly higher; however, the authors also had not explored the influence of magnetic fields on the diagnostic efficacy of the VBQ score.

Our research found that VBQ score was significantly influenced by MR field strength but not the MRI scanners. According to our self-paired comparison, a significantly higher VBQ score could be found in 1.5 T compared with 3.0 T (3.20±0.59 *vs.* 2.71±0.66). We considered the possible reason was that the VBQ score was only adjusted to the CSF as a lower limit, whereas the upper limit was not set. The upper limit of the T1-weighted SI was significantly influenced by the MR parameters, especially the magnetic fields. Without an upper limit, the cutoff value of the VBQ score varied greatly between MR 1.5 T and MR 3.0 T.

Prior VBQ studies have primarily used either DXA or QCT as the reference criteria for diagnosing osteoporosis, with limited investigation into how diagnostic criteria selection impacts VBQ performance (4,29). Recent evidence suggests that applying the conventional World Health Organization (WHO) Caucasian threshold (T-score ≤–2.5) to East Asian populations may overestimate osteoporosis prevalence, particularly for spine osteoporosis, and the cutoff value to East Asians could be as low as –3.7 (30). Additionally, ethnicity- and gender-specific BMD databases should be used for T-score calculations to ensure diagnostic accuracy (31,32). Similarly, for the QCT criteria, the traditional Caucasian cutoff (<80 mg/cm^3) appears to overestimate osteoporosis prevalence in Chinese women, with the study suggesting a more appropriate threshold of approximately 50 mg/cm^3 for East Asian populations (20). In our study, we retained the age- and gender-adjusted Caucasian reference values for spinal T-score calculations, along with the conventional QCT criterion (<80 mg/cm^3) for osteoporosis diagnosis. Notably, the diagnostic agreement between DXA and QCT was inconsistent, with 30/62 patients classified as osteoporotic by QCT compared to only 13/62 by DXA. Despite this discrepancy, both criteria consistently demonstrated a significantly higher VBQ cutoff value at 1.5 T compared to 3.0 T. Further research is needed to clarify the relationships among T-scores, QCT, and VBQ scores in East Asian populations.

Previous studies confirmed a weak-to-moderate correlation between the VBQ score and T-score/QCT. In a study by Lin *et al.* (4), the strongest correlation was observed between the VBQ 1.5 T and QCT (r=–0.511). In our study, the strongest relationship was also found in VBQ 1.5 T with QCT (r=–0.529), exceeding correlations reported in prior studies utilizing both 1.5 T and 3.0 T MR systems (r ranging from –0.27 to –0.358) (4).

Existing evidence indicates that the VBQ score has a moderate ability to predict osteoporosis (22,28,33). Our correlation and logistic regression analyses demonstrated that VBQ 1.5 T showed stronger associations with osteoporosis than VBQ 3.0 T. Meta-analysis showed that with the VBQ cutoff value of 3.02±0.38, the AUC could be 0.81 for the diagnosis of osteoporosis (2). In the present study, we found that the diagnostic efficacy for osteoporosis was AUC 0.709 for VBQ 1.5 T (optimum cutoff value of 3.018) and AUC 0.697 for VBQ 3.0 T (optimum cutoff value of 2.765). DeLong's test confirmed that VBQ diagnostic performance remained consistent across MR fields. However, the AUC of the VBQ score was still not comparable to that of T-score, BMD, and QCT. Clinically, we recommend that optimal thresholds should be institution-specific. Furthermore, the VBQ score should primarily serve as a preliminary osteoporosis screening tool due to its significant parameter-dependent variability (14).

Previous studies have reported that hyperlipidemia significantly increased the VBQ score which resembled the SI seen in osteoporosis (34); however, no significant influence was found in the present study of the hyperlipidemia-associated indexes such as triglyceride, total cholesterol, and high/low density lipoprotein cholesterol.

This is the first study using self-paired patients to comprehensively evaluate the influence of magnetic fields on the VBQ score; nevertheless, the study had several limitations. First, the study was performed in a lumbar surgery population, thus selection bias might have been present, and whether the findings can be generalized to the young healthy population still needs further evaluation. Second, because osteoporosis screening is more important than osteopenia screening in patients undergoing spine surgery, we did not explore the influence on osteopenia,

which will require further studies. Third, there were only three MR scanners in our institution, notably only one MR 1.5 T scanner, and whether our findings can be generalized to other MR scanners requires further research. Fourth, a clinical diagnosis of osteoporosis such as fragility fractures was not considered in the present study, and the rate of osteoporosis could have been underestimated. Fifth, given the retrospective design of the present study, we set the time span of MRI, DXA, and QCT examinations at 6 months. Although bone mass was not expected to change greatly within 6 months, future prospective cohort studies with shorter time span would be more persuasive. Sixth, the sample size from our single center was relatively small, limiting the statistical power of the analysis. Seventh, it was reported that extrinsic MRI parameters including spin-echo sequences and the TR/TE relaxation times might have influence on the VBQ score (13), and future studies with standardized MR parameters are needed to evaluate the issue. Eighth, it should be noted that the osteoporosis diagnostic thresholds (T-score and QCT) employed in this study were derived from Caucasian reference data, and their validity for East Asian populations requires additional investigation. Last, the effect of the potential confounding factors of the VBQ score, such as smoking status, hematologic abnormalities, treatments affecting bone marrow composition, as well as the relevance to the laboratory finding would need further research to clarify. All results of the manuscript should be interpreted within the above limitations.

Conclusions

The present study was performed to explore the influence of magnetic fields on the diagnostic efficacy of the VBQ score by self-paired patients. Through the retrospective inclusion of 104 patients who had received at least two MR examinations, significantly higher VBQ score was found in 1.5 T compared with 3.0 T. Correlation, logistic regression, and ROC analyses consistently demonstrated stronger osteoporosis associations for VBQ 1.5 T compared to VBQ 3.0 T, yet diagnostic accuracy was similar between field strengths. Given these field-strength-dependent differences, reporting magnetic field strength is crucial when interpreting VBQ scores for osteoporosis assessment.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-2025-173/rc

Data Sharing Statement: Available at https://qims.amegroups.com/article/view/10.21037/qims-2025-173/dss

Funding: This study was supported by Tianjin Science and Technology Plan Project "Unveiling and Directing" Major Project (No. 21ZXJBSY00130); Tianjin Education Commission Research Project (No. 2022YGYB11); National Natural Science Foundation of China (Nos. U24A20670, 82372419, 82072435); and National Key R&D Program of China (No. 2024YFC2400574).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-2025-173/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The study was approved by the institutional ethics board of Tianjin Hospital and individual consent for this retrospective analysis was waived.

References

1. Kadri A, Binkley N, Hernando D, Anderson PA. Opportunistic Use of Lumbar Magnetic Resonance Imaging for Osteoporosis Screening. Osteoporos Int 2022;33:861-9.
2. Hu F, Li X, Zhao D, Chen C, Liu G, Yang Q. The

diagnostic value of MRI-based vertebral bone quality score for osteoporosis or osteopenia in patients undergoing lumbar surgery: a meta-analysis. Osteoporos Int 2024;35:1881-95.
3. Chen A, Feng S, Lai L, Yan C. A meta-analysis of the value of MRI-based VBQ scores for evaluating osteoporosis. Bone Rep 2023;19:101711.
4. Lin W, He C, Xie F, Chen T, Zheng G, Yin H, Chen H, Wang Z. Assessment of bone density using the 1.5 T or 3.0 T MRI-based vertebral bone quality score in older patients undergoing spine surgery: does field strength matter? Spine J 2023;23:1172-81.
5. Ai Y, Chen Q, Huang Y, Ding H, Wang J, Zhu C, Song Y, Feng G, Liu L. MRI-based vertebral bone quality score for predicting cage subsidence by assessing bone mineral density following transforaminal lumbar interbody fusion: a retrospective analysis. Eur Spine J 2023;32:3167-75.
6. Li W, Zhu H, Hua Z, Miao D, Wang F, Tong T, Wang L. Vertebral Bone Quality Score as a Predictor of Pedicle Screw Loosening Following Surgery for Degenerative Lumbar Disease. Spine (Phila Pa 1976) 2023;48:1635-41.
7. Gao Y, Ye W, Ge X, Wang H, Xiong J, Zhu Y, Wang Z, Wang J, Tang P, Liu W, Cai W. Assessing the utility of MRI-based vertebral bone quality (VBQ) for predicting lumbar pedicle screw loosening. Eur Spine J 2024;33:289-97.
8. Li Q, Fu H, Yang H, Yang X, Wang L, Song Y. Novel MRI-Based Pedicle Bone Quality Score Independently Predicts Pedicle Screw Loosening after Degenerative Lumbar Fusion Surgery. Orthop Surg 2024;16:2372-9.
9. Ehresman J, Schilling A, Pennington Z, Gui C, Chen X, Lubelski D, Ahmed AK, Cottrill E, Khan M, Redmond KJ, Sciubba DM. A novel MRI-based score assessing trabecular bone quality to predict vertebral compression fractures in patients with spinal metastasis. J Neurosurg Spine 2020;32:499-506.
10. Roch PJ, Çelik B, Jäckle K, Reinhold M, Meier MP, Hawellek T, Kowallick JT, Klockner FS, Lehmann W, Weiser L. Combination of vertebral bone quality scores from different magnetic resonance imaging sequences improves prognostic value for the estimation of osteoporosis. Spine J 2023;23:305-11.
11. Özmen E, Biçer O, Meriç E, Circi E, Barış A, Yüksel S. Vertebral bone quality score for opportunistic osteoporosis screening: a correlation and optimal threshold analysis. Eur Spine J 2023;32:3906-11.
12. Pu M, Zhong W, Heng H, Yu J, Wu H, Jin Y, Zhang P, Shen Y. Vertebral bone quality score provides preoperative bone density assessment for patients undergoing lumbar spine surgery: a retrospective study. J Neurosurg Spine 2023. [Epub ahead of print]. doi: 10.3171/2023.1.SPINE221187.
13. Liu D, Kadri A, Hernando D, Binkley N, Anderson PA. MRI-based vertebral bone quality score: relationship with age and reproducibility. Osteoporos Int 2023;34:2077-86.
14. Liu XG, Chen X, Chen B, Liang PJ, Liu HH, Fu M. Vertebral bone quality different in magnetic resonance imaging parameters. J Orthop Surg Res 2023;18:772.
15. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF; STARD Group. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. Clin Chem 2015;61:1446-52.
16. Lin W, He C, Xie F, Chen T, Zheng G, Yin H, Chen H, Wang Z. Discordance in lumbar bone mineral density measurements by quantitative computed tomography and dual-energy X-ray absorptiometry in postmenopausal women: a prospective comparative study. Spine J 2023;23:295-304.
17. Kanis JA. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. WHO Study Group. Osteoporos Int 1994;4:368-81.
18. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R; National Osteoporosis Foundation. Clinician's Guide to Prevention and Treatment of Osteoporosis. Osteoporos Int 2014;25:2359-81.
19. Kanis JA, Melton LJ 3rd, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. J Bone Miner Res 1994;9:1137-41.
20. Wáng YXJ, Blake GM, Tang SN, Guermazi A, Griffith JF. Quantitative CT lumbar spine BMD cutpoint value for classifying osteoporosis among older East Asian women should be lower than the value for Caucasians. Skeletal Radiol 2024;53:1473-80.
21. Yu JS, Krishna NG, Fox MG, Blankenbaker DG, Frick MA, Jawetz ST, Li G, Reitman C, Said N, Stensby JD, Subhas N, Tulchinsky M, Walker EA, Beaman FD. ACR Appropriateness Criteria® Osteoporosis and Bone Mineral Density: 2022 Update. J Am Coll Radiol 2022;19:S417-32.
22. Ehresman J, Schilling A, Yang X, Pennington Z, Ahmed AK, Cottrill E, Lubelski D, Khan M, Moseley KF, Sciubba DM. Vertebral bone quality score predicts fragility fractures independently of bone mineral density. Spine J

2021;21:20-7.
23. Deshpande N, Hadi MS, Lillard JC, Passias PG, Linzey JR, Saadeh YS, LaBagnara M, Park P. Alternatives to DEXA for the assessment of bone density: a systematic review of the literature and future recommendations. J Neurosurg Spine 2023;38:436-45.
24. Chang HK, Hsu TW, Ku J, Ku J, Wu JC, Lirng JF, Hsu SM. Simple parameters of synthetic MRI for assessment of bone density in patients with spinal degenerative disease. J Neurosurg Spine 2022;36:414-21.
25. Rebuzzi M, Vinicola V, Taggi F, Sabatini U, Wehrli FW, Capuani S. Potential diagnostic role of the MRI-derived internal magnetic field gradient in calcaneus cancellous bone for evaluating postmenopausal osteoporosis at 3T. Bone 2013;57:155-63.
26. Cordes C, Baum T, Dieckmeyer M, Ruschke S, Diefenbach MN, Hauner H, Kirschke JS, Karampinos DC. MR-Based Assessment of Bone Marrow Fat in Osteoporosis, Diabetes, and Obesity. Front Endocrinol (Lausanne) 2016;7:74.
27. Ehresman J, Pennington Z, Schilling A, Lubelski D, Ahmed AK, Cottrill E, Khan M, Sciubba DM. Novel MRI-based score for assessment of bone density in operative spine patients. Spine J 2020;20:556-62.
28. Salzmann SN, Okano I, Jones C, Zhu J, Lu S, Onyekwere I, Balaji V, Reisener MJ, Chiapparelli E, Shue J, Carrino JA, Girardi FP, Cammisa FP, Sama AA, Hughes AP. Preoperative MRI-based vertebral bone quality (VBQ) score assessment in patients undergoing lumbar spinal fusion. Spine J 2022;22:1301-8.
29. Xu XM, Li N, Li K, Li XY, Zhang P, Xuan YJ, Cheng XG. Discordance in diagnosis of osteoporosis by quantitative computed tomography and dual-energy X-ray absorptiometry in Chinese elderly men. J Orthop Translat 2019;18:59-64.
30. Wáng YXJ, Griffith JF, Blake GM, Diacinti D, Xiao BH, Yu W, Su Y, Jiang Y, Guglielmi G, Guermazi A, Kwok TCY. Revision of the 1994 World Health Organization T-score definition of osteoporosis for use in older East Asian women and men to reconcile it with their lifetime risk of fragility fracture. Skeletal Radiol 2024;53:609-25.
31. Wáng YXJ, Blake GM, Xiao BH, Guglielmi G, Su Y, Jiang Y, Guermazi A, Kwok TCY, Griffith JF. East Asians' T-scores for the diagnosis of osteoporosis should be calculated using ethnicity- and gender-specific BMD reference ranges: justifications. Skeletal Radiol 2024;53:409-17.
32. Aparisi Gómez MP, Wáng YJ, Yu JS, Johnson R, Chang CY. Dual-Energy X-Ray Absorptiometry for Osteoporosis Screening: AJR Expert Panel Narrative Review. AJR Am J Roentgenol 2025. [Epub ahead of print]. doi: 10.2214/AJR.25.32802.
33. Soliman MAR, Aguirre AO, Kuo CC, Ruggiero N, Azmy S, Khan A, Ghannam MM, Almeida ND, Jowdy PK, Mullin JP, Pollina J. Vertebral bone quality score independently predicts cage subsidence following transforaminal lumbar interbody fusion. Spine J 2022;22:2017-23.
34. Aynaszyan S, Devia LG, Udoeyo IF, Badve SA, DelSole EM. Patient physiology influences the MRI-based vertebral bone quality score. Spine J 2022;22:1866-74.

Cite this article as: Hu F, Chen B, Bian H, Zhao D, Chen C, Liu G, Ma X, Wang Z, Yang Q. The diagnostic cutoff value of vertebral bone quality score for osteoporosis is significantly influenced by the magnetic field in patients undergoing lumbar surgery. Quant Imaging Med Surg 2025;15(10):10249-10261. doi: 10.21037/qims-2025-173

