OSTEOPOROSIS is the most prevalent metabolic bone disease,\(^1\) resulting in bone mineral loss, altered bone microarchitecture, and increased patient fragility. Reduced bone mineral density (BMD) not only increases the incidence of fragility fractures, but also poses an equal challenge to subsequent surgical treatment. In patients undergoing internal spinal fixation or fusion surgery, bone mass plays an important role in the surgical approach, disease prognosis, and need for postoperative antiosteoporotic treatment. Low BMD can lead to an increased risk of

**ABBREVIATIONS**

AUC = area under the curve; BMD = bone mineral density; DEXA = dual-energy x-ray absorptiometry; FA = fractional anisotropy; ICC = intraclass correlation coefficient; NPV = negative predictive value; PPV = positive predictive value; ROC = receiver operating characteristic; ROI = region of interest; SI = signal intensity; TBS = trabecular bone score; VBQ = vertebral bone quality.
screw loosening, fusion failure, vertebral body collapse, and pseudarthrosis formation. Therefore, it is particularly important to have timely and adequate information about the patient’s bone mineral condition before surgery.

At present, dual-energy x-ray absorptiometry (DEXA) is recommended by the WHO as the gold standard for assessing BMD and diagnosing osteoporosis. DEXA can obtain BMD information based on the relative absorption of x-rays in specific areas such as the lumbar spine, femoral neck, forearm, etc., and then compare it with a standard reference population to obtain a T-score for diagnostic evaluation. However, DEXA is not fully utilized in clinical practice. Although most orthopedic surgeons accept that DEXA should be routinely performed in older women (> 65 years) or patients with a previous history of fragility fractures, selective use remains in other patients for bone densitometry, such as those with herniated discs, spinal stenosis, nonsenior patients, and other groups. Studies have shown that fewer than 30% of patients at high risk for fracture have been screened using DEXA. There is growing evidence, however, indicating that DEXA has its limitations. The 2D imaging-based feature of DEXA makes it unable to distinguish between cortical and cancellous bone but assesses the entire vertebral body, incorporating not only the vertebral body and posterior structures but also the intervertebral disc tissue between the vertebral bodies. It has been proven that DEXA may incorrectly estimate bone fragmentation as high bone mass or include too much soft tissue in the measurement to affect the final reading, leaving the possibility of inadequate BMD measurement in patients for the interpretation of DEXA results.

For these reasons, scholars have focused on other imaging modalities and developed tools such as Hounsfield units, trabecular bone score (TBS), etc., as supplements to the traditional DEXA method. Recently, another novel MRI-based score, the vertebral bone quality (VBQ) score, has gained our attention. Ehresman et al. developed this novel BMD opportunistic assessment tool by developing the previous “M-score.” Compared with the healthy population, patients with osteoporosis have significantly increased bone marrow adipose content. The shorter relaxation time of adipose tissue in the T1 phase of MRI causes it to show a higher signal intensity (SI). Using this feature, the VBQ score expresses a patient’s BMD profile by measuring the SI of L1–4 in the noncontrast T1 phase and dividing it by the L3 CSF SI for normalization, which enables comparison between patients. This novel VBQ score has been evaluated in a small number of studies at this stage. Kadri et al. performed opportunistic screening for osteoporosis using the overall VBQ score and an L1 single-level VBQ score in patients undergoing thoracolumbar spine surgery and obtained good diagnostic ability (areas under the curve [AUCs] 0.806 and 0.779). Optimal diagnostic thresholds were found at an overall VBQ score of 2.95 and L1 single-level VBQ score of 3.0. Schilling et al. evaluated the interobserver agreement of VBQ scores and found good agreement between vertebral and CSF SI measurements, while there was wide variability in the agreement of the final VBQ scores, with intraclass correlation coefficients (ICCs) ranging from 0.667 (good) to 0.998 (excellent). External validation of the novel VBQ score continues to be insufficient, however, and doubts remain as to whether there are differences in CSF SI in different measurement regions and whether VBQ scores are consistent between MRI machines.

To resolve these queries, we designed this study with the purpose of examining the diagnostic value of the novel VBQ score by assessing BMD in patients undergoing lumbar spine surgery, evaluating whether there are differences in CSF measurements at different levels of L1–4, and evaluating the consistency of VBQ scores between MRI machines with different field strengths.

**Methods**

IRB approval was obtained for this study from the Second Affiliated Hospital of Soochow University and the informed consent process was exempted due to the retrospective nature of the study.

**Patient Cohort**

We reviewed 1801 female patients who underwent surgery for lumbar disc herniation at our center between September 2018 and September 2021. The study inclusion criteria were patients aged > 50 years with available DEXA scans and lumbar spine MRI. The exclusion criteria were 1) lack of preoperative DEXA scans or lumbar spine MRI (noncontrast T1-weighted images); 2) internal fixation in the lumbar spine; 3) scoliosis deformity in the lumbar spine; 4) narrow CSF circulation space preventing placement of the region of interest (ROI) by the software; 5) L1–4 vertebral fracture, according to Genant’s semiquantitative technique (≥ grade II); 6) lesions in the L1–4 vertebral body such as hemangioma, bone destruction, and tumor; and 7) patients with a combination of other diseases affecting bone metabolism, such as Paget’s disease, any bone tumor, osteogenesis imperfecta, etc.

Two hundred thirty-seven patients met these criteria; 137 were excluded because they were examined using other MRI machines with different parameters and the sample sizes used were small. Therefore, 100 patients examined on the Achieva 1.5-T MRI machine (Philips Healthcare) were included in the study (Fig. 1). In addition, 24 of these patients were found to have undergone lumbar spine scans on another Achieva 3-T MRI machine at our center during the retrospective process. Using these 24 patients as controls, we evaluated the consistency of VBQ scores between different MRI machines.

We collected detailed demographic characteristics of the enrolled patients including age, sex, BMI, comorbidities (hypertension, diabetes, hyperlipidemia), medication history (glucocorticoids), previous history of fragility fractures of the spine, and some potentially relevant laboratory indices.

**DEXA**

DEXA scans were completed on the same bone densitometer (Lunar Prodigy, GE Healthcare) at our center using standard methods. The BMDs (g/cm²) of the anterior-posterior lumbar spine (L1–4) and the left femoral neck were measured, and the corresponding T-scores were obtained from the database provided by the manufacturer.
Because lumbar spine T-scores are likely to be affected by degeneration, we classified patients into different BMD subgroups according to their lowest T-scores in this study: T-score ≥ −1.0 for normal BMD, between −1.0 and −2.5 for osteopenia, and ≤ −2.5 for osteoporosis.²

MRI
The lumbar spine scan was performed using a standard method on the same Achieva 1.5-T MRI machine at our center. Scanning parameters were fractional anisotropy (FA) 90.00, TR 582.04 msec, TE 12.00 msec, slice thickness 4.00 mm, and slice gap 0.40 mm. Twenty-four of the included patients had undergone a lumbar spine examination using an Achieva 3-T MRI machine (Philips Healthcare). Scanning parameters were FA 80.00, TR 520 msec, TE 6.28 msec, slice thickness 4.00, and slice gap 0.40 mm. All sequences used in this study were noncontrast T1-weighted turbo spin echo sequences, excluding other sequences such as contrast enhancement or inversion recovery. All radiology images were uploaded to a standard radiology PACS workstation (Neusoft PACS/RIS) for subsequent analysis.

VBQ Score
The measurement and calculation of VBQ scores were performed using the method described by Ehresman et al.⁸ First, the midsagittal image of the patient’s lumbar spine was acquired on the MRI noncontrast T1 phase. By placing an elliptical ROI in the trabecular region of the L1–4 vertebral body (VB), the software then automatically generated an averaged MRI SI for the corresponding region. Similarly, ROIs were also placed in the CSF circulation space posterior to the L1–4 vertebrae to obtain the SI of the CSF (SI\textsubscript{CSF}). The measurement schematic is shown in Fig. 2. The VBQ score was calculated according to the following formula: median SI\textsubscript{VB} (L1–4)/SI\textsubscript{CSF} (L3). In addition, we obtained single-level VBQ scores for each vertebra of L1–4, which were calculated as SI\textsubscript{VB}/SI\textsubscript{CSF}.

Requirements for placement of vertebral ROIs should be noted: 1) avoid the cortical bone and the posterior vertebral venous plexus; 2) the ROI covers as much of the trabecular region as possible while not requiring a consistent ROI size, because there are differences in vertebral body shape and size; and 3) because of the influence of the patient’s position during the examination, or the occurrence of mild lumbar scoliosis, slippage, and rotation (which are quite common and difficult to avoid in clinical practice), the midsagittal section selected for measurement should be as close as possible to the majority of the vertebrae. The placement of CSF ROIs should also be noted; i.e., the ROI should be placed in the CSF directly posterior to the

FIG. 1. Flowchart of the included and excluded patients used in the study.

Unauthenticated | Downloaded 05/05/24 11:31 AM UTC
vertebral body while avoiding the inclusion of descending nerve roots and other soft tissues (the T2 phase can be used to assist in identifying areas of CSF).

All measurements were performed by a trained observer (M.P.). Repeated measurements of the 100 patients were performed 3 months later by M.P. and another clinician (H.H.) to test for inter- and intraobserver agreement. The measurements were blinded and conducted independently by the observers, who did not obtain any other information about the patients.

Statistical Analysis

All data in this study were collected through Microsoft Excel, with continuous variables expressed as means ± standard deviations and categorical variables expressed as frequencies and percentages. Normality and chi-square tests were performed for each variable using the Shapiro-Wilk and Levene tests. Subsequent statistical methods were selected based on the results. Consistency tests for continuous variables were performed using ANOVA or the Kruskal-Wallis test. Consistency tests for categorical variables were performed using the chi-square or Fisher’s exact test. The Pearson correlation coefficient was used to describe the degree of linear correlation between the variables. Receiver operating characteristic (ROC) analysis and AUC were used to evaluate the diagnostic efficacy of the indices. The Friedman test was used to investigate the consistency of SI in different regions of the CSF, and the Wilcoxon signed-rank test was used as a post hoc paired test. The Wilcoxon signed-rank test was also used to evaluate the consistency of VBQ scores between different MRI machines. Data in this study were considered statistically significant for p values < 0.05. All statistical analyses and visualizations were performed using R (version 4.2.1, R-Foundation for Statistical Computing; r-project.org).

Results

Patient Characteristics

The 100 patients included in this study were all Asian females, with a mean age of 66.1 ± 9.4 (range 50–87) years. With the exception of DEXA T-scores and VBQ scores, there were no significant differences between the different BMD subgroups for the remaining variables (p > 0.05). There were no patients with a history of smoking or alcohol abuse identified during the review. The detailed demographic characteristics and radiological data are shown in Table 1.
Radiological Data Analysis

The MRI lumbar scans and DEXA examinations used in this study were completed preoperatively with a mean interval of 3.4 ± 6.8 (range 0–28) days.


dbq
dbq

VCF Score

The mean VBF scores of patients among BMD subgroups were 2.81 ± 0.28 (normal), 3.06 ± 0.36 (osteopenia), and 3.43 ± 0.37 (osteoporosis), respectively. The single-level VBF scores of L1–4 and overall VBF scores were significantly different between groups (p < 0.001). Scatterplots and Pearson correlation coefficients revealed a moderate negative linear correlation between VBF scores and lowest DEXA T-scores (r = −0.524), while the correlation coefficients between VBF scores and lumbar spine T-scores or femoral neck T-scores alone were −0.482 and −0.469, respectively (Fig. 3). The results of the ROC analysis showed that the VBF score had a good ability to identify people with reduced BMD and osteoporosis according to the maximum Youden index were 3.06 (sensitivity 0.636, specificity 0.870, positive predictive value [PPV] 0.942, negative predictive value [NPV] 0.417) and 3.05 (sensitivity 0.875, specificity 0.618, PPV 0.459, NPV 0.891), respectively (Table 2). The single-level VBF score for L1–4 was not superior to the overall VBF score in diagnostic ability (p > 0.1). Single-level and overall VBF scores had good inter- and intraobserver agreement between the two observers (ICC ≥ 0.9; Table 3).

CSF SI

The CSF SIs at different levels of L1–4 in 100 patients were 122.8 ± 17.5 (L1), 119.0 ± 17.1 (L2), 118.0 ± 16.5 (L3), and 116.2 ± 17.7 (L4), respectively. The Friedman test revealed a significant difference between them (p < 0.001), and the VBF scores calculated from each SI CSF of L1–4 were also inconsistent (p < 0.001). Wilcoxon signed-rank tests were performed for each level of L1–4, and the results showed that the measured values of CSF signals and calculated VBF scores at the L2, L3, and L4 levels were generally consistent (p > 0.05), while at the L1 level they were significantly different (p < 0.001; Supplementary Fig. 1).

| TABLE 1. Demographic and radiological data of 100 patients |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Demographic     | Normal, n = 23  | Osteopenia, n = 45 | Osteoporosis, n = 32 | p Value |
| Mean age ± SD, yrs | 63.5 ± 9.8     | 65.8 ± 9.5       | 68.4 ± 8.6       | 0.152*          |
| Mean BMI ± SD    | 25.8 ± 4.4      | 25.0 ± 3.7       | 23.4 ± 3.6       | 0.065*          |
| History of VCFs, n (%)† | 3 (13)         | 5 (11.1)         | 7 (21.9)         | 0.391‡          |
| Comorbidities, n (%) |               |                 |                 |                |
| Hypertension     | 7 (30.4)        | 17 (37.8)        | 12 (37.5)        | 0.818§          |
| Hyperlipidemia   | 3 (13)          | 6 (13.3)         | 2 (6.2)          | 0.653‡          |
| Diabetes         | 5 (21.7)        | 9 (20.0)         | 7 (21.9)         | >0.99‡           |
| Mean laboratory values ± SD |
| Calcium, mmol/L   | 2.35 ± 0.17     | 2.28 ± 0.17      | 2.31 ± 0.19      | 0.596¶          |
| Phosphate, mmol/L | 1.16 ± 0.19     | 1.16 ± 0.13      | 1.15 ± 0.21      | 0.982¶          |
| Albumin, g/L      | 40.4 ± 3.3      | 40.5 ± 4.2       | 40.2 ± 3.7       | 0.975¶          |
| Hemoglobin, g/L   | 127.4 ± 13.0    | 128.0 ± 12.6     | 125.0 ± 11.2     | 0.392¶          |
| Glucose, mmol/L   | 5.66 ± 1.00     | 5.88 ± 1.55      | 6.04 ± 1.62      | 0.872¶          |
| Mean DEXA T-scores ± SD |
| Lumbar spine     | 0.61 ± 1.16     | −1.32 ± 0.74     | −3.13 ± 0.68     | <0.001¶         |
| Femoral neck     | −0.02 ± 0.70    | −1.34 ± 0.70     | −2.22 ± 0.71     | <0.001¶         |
| Lowest overall   | −0.22 ± 0.69    | −1.76 ± 0.39     | −3.18 ± 0.65     | <0.001¶         |
| Mean single-level VBF scores ± SD |
| L1               | 2.89 ± 0.33     | 3.18 ± 0.42      | 3.53 ± 0.50      | <0.001*         |
| L2               | 2.87 ± 0.26     | 3.07 ± 0.36      | 3.45 ± 0.42      | <0.001¶         |
| L3               | 2.74 ± 0.27     | 2.95 ± 0.36      | 3.32 ± 0.38      | <0.001*         |
| L4               | 2.65 ± 0.35     | 3.04 ± 0.42      | 3.26 ± 0.38      | <0.001¶         |
| Mean VBF scores ± SD | 2.81 ± 0.28    | 3.06 ± 0.36      | 3.43 ± 0.37      | <0.001*         |

VCF = vertebral compression fracture.

No patients with a history of smoking or alcohol abuse were identified during the review. Boldface type indicates statistical significance.

* ANOVA.
† History of VCFs represents patients with a previous history of spinal fracture (excluding L1–4).
‡ Fisher’s exact test.
§ Chi-square test.
¶ Kruskal-Wallis test.
Interdevices

Twenty-four of the 100 patients included had received a lumbar spine scan for other purposes on an Achieva 3-T MRI machine with a different field strength from the same device manufacturer (Philips Healthcare). The mean age of these 24 patients was 61.4 ± 8.3 (range 50–77) years. The mean interval between the two MRI sessions was 129.1 ± 92.8 (range 13–341) days. VBQ scores were 2.91 ± 0.43 for the 1.5-T machine versus 3.08 ± 0.62 for the 3-T machine, and the Wilcoxon signed-rank test showed no significant difference between them (p = 0.107; Fig. 5).

Discussion

Because of the inadequate utilization of traditional DEXA in clinical practice, some novel BMD assessment tools have been developed, with benefits as well as some limitations. For example, the CT-based Hounsfield units have an excellent correlation with conventional DEXA results and are less susceptible to vertebral degeneration, but investigations have shown that different scan parameters (especially tube voltage) affect the measurement of Hounsfield units. The International Society for Clinical Densitometry also mentioned in its official statement that rigorous routine maintenance calibration of CT equipment is required to ensure equipment stability and obtain accurate, comparable readings using this method. The DEXA image-based TBS can provide bone microstructure information different from the conventional DEXA T-score, but it does not escape the limitations of 2D imaging.

Recently, Ehresman et al. developed a novel MRI-based VBQ score as an opportunistic screening tool for osteoporosis by previously developing the M-score. Because MRI is free of radiation exposure and is readily available as a routine preoperative examination tool for patients undergoing spine surgery, we believe that opportunistic use of this novel VBQ score is uniquely valuable as a preoperative BMD assessment for patients undergoing lumbar spine surgery.

Compared with previous studies, our study has the following advantages. First, the 100 patients included in this study were imaged on the same MRI machine with fixed parameters, minimizing the bias caused by external differences. Second, all radiological data in this study were obtained as part of the preoperative examination of patients undergoing spine surgery, and the very short interval between DXA and MRI provided reliable obser-
vations. Third, we performed interdevice agreement tests using 24 patients who underwent examinations on MRI machines with two different field strengths; thus, these patients served as their own controls.

In our study, we found differences in VBQ scores between all BMD subgroups (p < 0.05). The Pearson correlation coefficient revealed a moderate negative linear correlation of VBQ scores with the lowest T-scores (r = −0.524), with lumbar spine T-scores (r = −0.482), and with formal neck T-scores (r = −0.469). The results of the ROC analysis showed that the novel VBQ score had a good ability to identify this segment in both people with reduced BMD and patients with osteoporosis (AUCs 0.79 and 0.81, respectively), and a cutoff point of 3.06 was obtained for identifying reduced BMD and 3.05 for identifying osteoporosis, based on the maximum Youden index. These two cutoff points have a high PPV and NPV, respectively, while the other side is more average in each age group (Table 2).

Based on the above results, we believe that a VBQ score < 3.05 can be used to exclude patients with osteoporosis (NPV = 0.913), while further examination (such as using DEXA or quantitative CT) is recommended for those with a VBQ score ≥ 3.05 (especially ≥ 3.06) to clarify their BMD status, as they are very likely to have reduced BMD or osteoporosis. Additionally, the diagnostic efficacy of the single-level VBQ scores was not superior to that of the

![FIG. 4. ROC curve analysis showing the ability of VBQ scores to identify patients with reduced BMD and osteoporosis. Figure is available in color online only.](image)

| TABLE 2. Prevalence, sensitivity, specificity, PPV, and NPV according to age subgroup |
|---|---|---|---|---|---|---|---|---|---|
| Age Range (yrs) | Prevalence (%) | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
| | Reduced BMD* | Osteo† | Reduced BMD | Osteo | Reduced BMD | Osteo | Reduced BMD | Osteo | Reduced BMD | Osteo |
| 50–59 | 66.7 | 14.8 | 44.4 | 75.0 | 88.9 | 73.9 | 88.9 | 33.3 | 44.4 | 94.4 |
| 60–69 | 80.6 | 36.1 | 62.1 | 84.6 | 85.7 | 52.2 | 94.7 | 50.0 | 35.3 | 85.7 |
| 70–79 | 79.3 | 41.4 | 78.3 | 91.7 | 83.3 | 52.9 | 94.7 | 57.9 | 50.0 | 90.0 |
| 80–89 | 87.5 | 37.5 | 71.4 | 1 | NA | 40.0 | 83.3 | 50.0 | NA | 1 |
| Total | 77.0 | 32.0 | 63.6 | 87.5 | 87.0 | 61.8 | 94.2 | 51.9 | 41.7 | 91.3 |

NA = not applicable for the presence of 0; Osteo = osteoporosis.
* VBQ score ≥ 3.06.
† VBQ score ≥ 3.05.
overall VBQ score. Given the potential limitations of measuring only one single vertebra, using single-level VBQ scores was not a primary recommendation. The results above are generally consistent with previous studies.\(^{8,13}\) Therefore, it can be considered that the new VBQ score can serve as a proxy for BMD to some extent and can help us to identify those patients who need further examination.

It is worth noting, however, that in this study we obtained very close diagnostic thresholds when using the VBQ score to identify people with reduced BMD as well as patients with osteoporosis, and they did not have both high PPV and NPV, which we speculate may be related to the measurement stability of the VBQ score. During our measurements, we found that although the midsagittal section of the lumbar spine was chosen for the measurement of VBQ scores, the fact that the selected plane is not the standard midsagittal section for every vertebral body of L1–4 due to incorrect positioning of the patient during the examination or the presence of lumbar scoliosis, vertebral rotation, or vertebral slippage may affect the accuracy of the measurement. This is a limitation of VBQ scores, and some of these factors are difficult to avoid. The limitations of vertebral measurements were counterbalanced by taking the median (which can better avoid the effect of extreme values compared with the mean) and using it as the numerator in the VBQ score calculation formula, which has a relatively small impact on the final results. What is of greater concern to us is using the CSF measurement as the denominator, which has a greater impact on the final calculation. In addition, the flow space of CSF is narrow, and measuring it with a small ROI casts doubt on its accuracy and stability. We measured CSF signals at different levels of L1–4 in all patients, and the results of the Friedman test confirmed the differences (p < 0.001). Multiple paired tests showed that CSF measurements at the L2, L3, and L4 levels were generally consistent, while those at the L1 levels showed significant differences and influenced the final VBQ scores. We consider the following as possible reasons for the variance.

During MRI scans, the flow of CSF causes changes in magnetization intensity, which results in the inhomogeneity of its signal. Second, because of the existence of the physiological curvature of the spine, the CSF flow space is not completely horizontal when the human body is lying.

![Graph](image.png)

**FIG. 5.** The Bland-Altman plot shows that the VBQ scores obtained from the same group of patients on the two MR machines are essentially identical. The _dotted dark blue line_ represents the mean difference (~0.16). The _dashed light blue lines_ represent the 95% CI (~1.03 to 0.70). All measurements were within the 95% CI, which implies a good level of agreement for these data. Figure is available in color online only.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>0.946</td>
<td>0.985</td>
<td>0.922</td>
<td>0.945</td>
<td>0.896</td>
<td>0.926</td>
<td>0.914</td>
<td>0.931</td>
</tr>
<tr>
<td>L2</td>
<td>0.954</td>
<td>0.992</td>
<td>0.933</td>
<td>0.943</td>
<td>0.902</td>
<td>0.932</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>0.946</td>
<td>0.990</td>
<td>0.935</td>
<td>0.958</td>
<td>0.917</td>
<td>0.934</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L4</td>
<td>0.938</td>
<td>0.986</td>
<td>0.932</td>
<td>0.946</td>
<td>0.912</td>
<td>0.927</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ITE = interobserver; ITA = intraobserver; VB = vertebral body.
in a horizontal position. Compared with the L1–4 levels, the L1 level is usually more posterior, that is, closer to the coil, and the signal strength here is a bit higher. Finally, the spinal cord usually bundles at the L1 level to form the conus medullaris and extends downward to the cauda equina. The CSF space at the L1 level tends to be narrower, making it inconvenient to place an ROI for measurement, while at the L3 level the space is usually wider. For these reasons, we consider the L3 level as a reasonable choice for the measurement area of the CSF signal in VBQ score calculation, but one needs to be careful to avoid free nerves in the spinal canal. Additionally, it is acceptable to use the L2 or L4 level measurements as surrogates for L3 in cases of obstructed CSF flow.

At the end of the study, we compared the VBQ scores of the same 24 patients examined with different magnetic field strengths. The results showed that their VBQ scores obtained in different MRI machines were generally consistent (p = 0.107). This finding suggests that the VBQ scores did eliminate part of the interdevice variation, allowing for comparison between patients.

Limitations of the Study

We acknowledge the limitations of this study. First, because aging females are a population at high risk of osteoporosis and fractures, we highly selected female patients older than 50 years of age who underwent spinal surgery as our study observation subjects, which may lead to the occurrence of selective bias, and some of the results of this study may not be generalizable to a broader population. Second, as many patients were excluded during the retrospective process because of the lack of available materials or because they were examined using other machines with different parameters, we used a cohort with a small sample size for the analysis. Therefore, a multicenter large sample size for the analysis and a prospective study population are still needed. Lastly, the VBQ score is still a relatively simple tool for the measurement area of the CSF signal in VBQ score calculation, but one needs to be careful to avoid free nerves in the spinal canal. Additionally, it is acceptable to use the L2 or L4 level measurements as surrogates for L3 in cases of obstructed CSF flow.

Conclusions

Our study shows that the novel VBQ score can reflect patients’ BMD to a considerable extent. A VBQ score < 3.05 can tentatively exclude osteoporosis, while a VBQ score ≥ 3.05 (especially ≥ 3.06) is highly suspicious of the presence of reduced BMD, and further examination (such as using DEXA) is recommended to clarify the diagnosis. The VBQ score appeared to have no significant differences between MRI machines with varied field strengths and scanning parameters, offering the possibility of comparison between patients. It should be noted that the process of CSF signal measurement is a critical stage affecting the stability of this score. CSF signal measurements at the L3 level are used as the usual choice, while those at L2 or L4 levels can be used as an approximate substitute. Using the novel VBQ score as a preoperative BMD assessment tool for patients undergoing lumbar spine surgery provides an additional opportunity for osteoporosis screening, which not only helps to improve the diagnosis of osteoporosis but also provides guidance for the successful performance of surgery.

References

Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Shen, Pu, Zhong. Acquisition of data: Pu, Yu, Wu, Jin. Analysis and interpretation of data: Pu, Zhong. Drafting the article: Pu. Critically revising the article: Shen, Zhong, Heng, Zhang. Reviewed submitted version of manuscript: Shen, Pu, Zhong, Heng, Yu, Jin, Zhang. Approved the final version of the manuscript on behalf of all authors: Shen. Statistical analysis: Pu. Administrative/technical/material support: Heng, Yu. Study supervision: Shen, Zhang.

Supplemental Information
Online-Only Content
Supplemental material is available with the online version of the article.
Supplementary Fig. 1. https://thejns.org/doi/suppl/10.3171/2023.1.SPINE221187.

Correspondence
Yixin Shen: Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu Province, China. 972679925@qq.com.