

A new diagnostic score to detect osteoporosis in patients undergoing lumbar spine MRI

Michele Bandirali¹ · Giovanni Di Leo² · Giacomo Davide Edoardo Papini² · Carmelo Messina¹ · Luca Maria Sconfienza^{2,3} · Fabio Massimo Olivieri⁴ · Francesco Sardanelli^{2,3}

Received: 21 June 2014 / Revised: 15 January 2015 / Accepted: 2 March 2015 / Published online: 22 April 2015
© European Society of Radiology 2015

Abstract

Objectives Signal intensity of lumbar-spine at magnetic resonance imaging (MRI) correlates to bone mineral density (BMD). Our aim was to define a quantitative MRI-based score to detect osteoporosis on lumbar-spine MRI.

Methods After Ethics Committee approval, we selected female patients who underwent both lumbar-spine MRI and dual-energy X-ray absorptiometry (DXA) and a reference group of 131 healthy females (20–29 years) who underwent lumbar-spine MRI. We measured the intra-vertebral signal-to-noise ratio in L1–L4. We introduced an MRI-based score (M-score), on the model of T-score. M-score diagnostic performance in diagnosing osteoporosis was estimated against DXA using receiver operator characteristic (ROC) analysis.

Results We included 226 patients (median age 65 years), 70 (31 %) being osteoporotic at DXA. MRI signal-to-noise ratio correlated to BMD ($r=-0.677$, $P<0.001$). M-score negatively correlated to T-score ($r=-0.682$, $P<0.001$). Setting a 90 %-specificity, an M-score thresh-

old of 5.5 was found, distinguishing osteoporosis from non-osteoporosis (sensitivity 54 %; ROC AUC 0.844). Thirty-one (14 %) patients had a fragility fracture, with osteoporosis detected in 15 (48 %) according to M-score and eight (26 %) according to T-score ($P=0.016$).

Conclusions M-score obtained on lumbar spine MRI is a quantitative method correlating with osteoporosis. Its diagnostic value remains to be demonstrated on a large prospective cohort of patients.

Key Points

- M-score is a quantitative score potentially screening osteoporosis on lumbar-spine MRI;
- This method showed good intra- and inter-reader reproducibility;
- M-score may be used for identifying patients who should undergo DXA.

Keywords Osteoporosis · Magnetic resonance imaging · Lumbar spine · Dual-energy X-ray absorptiometry · Bone mineral density

✉ Carmelo Messina
carmelomessina.md@gmail.com

Michele Bandirali
michele.bandirali@hotmail.it

Giovanni Di Leo
gianni.dileo77@gmail.com

Giacomo Davide Edoardo Papini
docgde@gmail.com

Luca Maria Sconfienza
io@lucasconfienza.it

Fabio Massimo Olivieri
olivieri@gmail.com

Francesco Sardanelli
francesco.sardanelli@unimi.it

¹ Scuola di Specializzazione in Radiodiagnostica, Università degli Studi di Milano, via Festa del Perdono 7, 20122 Milano, Italy

² Radiology Unit, IRCCS Policlinico San Donato, via Morandi 30, 20097 San Donato Milanese, Italy

³ Dipartimento di Scienze Biomediche per la Salute, Università degli Studi di Milano, via Morandi 30, 20097 San Donato Milanese, Italy

⁴ Mineralometria Ossea Computerizzata e Ambulatorio Malattie Metabolismo Minerale e Osseo, Servizio di Medicina Nucleare, IRCCS Fondazione Ca' Granda Ospedale Maggiore Policlinico, via Francesco Sforza 9, 20122 Milano, Italy

Introduction

Osteoporosis is characterized by compromised bone strength and increased risk of fracture due to reduced bone mineral density (BMD) and quality [1, 2]. Osteoporosis typically occurs in postmenopausal women and the elderly population and in patients with prolonged glucocorticoids substitution [3, 4]. This problem has tremendous implications, as the elderly are the fastest-growing age group worldwide [5]. About 50 % of Caucasian women over 50 will have at least a fragility fracture in their lifetime and hip fractures will rise worldwide from 1.7 million in 1990 to 6.3 million in 2050 [5]. Twenty percent of patients with hip fractures die within one year and another 20 % will require permanent nursing home care [6]. Vertebral fractures usually imply minor complications but they are associated with substantial disability and a 10-fold risk of further vertebral fractures [7].

Dual energy X-ray absorptiometry (DXA) is a well-standardized and easy-to-use quantitative imaging modality that has high precision (maximum acceptable precision error 2 %–2.5 %) and uses a nearly negligible radiation dose [8]. In 1994, BMD testing with DXA has been adopted as a standard of reference in managing osteoporosis using the T-score to classify BMD [9]. Also, risk fractures of osteoporotic patients can be estimated using the FRAX[®] tool, as suggested by the World Health Organization (WHO). Although early BMD assessment is necessary for prevention of fractures and complications [10], osteoporosis is still an underdiagnosed condition, and in the United States nearly half of all female Medicare beneficiaries have never undergone DXA [11] and more than 80 % of all patients with a major fragility fracture do not undergo DXA nor receive any relevant drug therapies [12, 13].

Apart from DXA, osteoporosis can be detected using other methods, such as quantitative computed tomography, quantitative ultrasound, and others [2]. However, just like DXA, these represent diagnostic tools patients should undergo without any perceived reason, as osteoporosis is definitely an asymptomatic disease. A different approach could be that of extracting data regarding BMD while patients undergo imaging examinations for other clinical indications. This would represent an opportunistic screening, allowing for rapid identification of high-risk cohorts of patients in whom further evaluation (DXA) or treatment is warranted. A recent study by Pickhardt et al. [14] demonstrated how routine abdominal computed tomography obtained for other clinical indications could be used for quantitative diagnosis of osteoporosis. As there are more than 80 million computed tomography examinations performed each year in the US, the idea of extracting more information from imaging data collected for other purposes holds merit. Indeed, the possibility of accurately detecting osteoporosis-related fragility fractures has already been demonstrated [15].

Lumbar spine magnetic resonance imaging (MRI) is another imaging examination that is frequently requested worldwide in the elderly population, mainly due to low back pain [16, 17]. Previous studies demonstrated that the bone marrow signal intensity on the standard T1-weighted images is negatively correlated with BMD and osteoporosis [18–20]. This correlation may open a window of possibility to detect osteoporosis using MRI. Moreover, osteoporosis-related fragility fractures may be accurately detected by MRI [21]. Thus, the aim of our study was to define a new quantitative lumbar spine MRI-based method on standard T1-weighted images for diagnosing osteoporosis, having lumbar spine DXA as a reference standard.

Materials and methods

Study population

This preliminary single-centre retrospective study received Ethics Committee approval. We searched our database from June 2010 to December 2013 for all female patients who underwent both a lumbar spine MRI for low back pain and a lumbar and hip DXA. Exclusion criteria were: 1) time elapsed between the two examinations longer than six months; 2) ethnicity other than Caucasian; 3) injection of MRI contrast material; 4) known oncologic patients; 5) traumatic vertebral injuries; 6) known or suspected demyelinating disease; and 7) presence of relevant image artefacts.

To calculate an MRI-based score equivalent to the T-score used in DXA, we used a reference group consisting of 131 healthy Caucasian females aged 20–29 years [22], with a normal body mass index (BMI), i.e., 19–25 kg/m² [23], who had undergone a lumbar MRI for low back pain. The same exclusion criteria used for the patient group were applied.

Magnetic resonance methods and image analysis

All MRI examinations were performed at 1.5 T (Magnetom Sonata Maestro Class, Siemens Medical Solutions, Erlangen, Germany). The imaging protocol always included a sagittal T1-weighted spin-echo sequence (TR=600 ms; TE=11 ms; slice thickness=4 mm; squared field of view=280 mm; matrix 320×320), optimal for evaluating vertebral fatty marrow [18]. A reader with a 3-year experience in spinal MRI and a 4-year experience in DXA manually segmented vertebral bodies from L1 to L4 (similarly to DXA). A region of interest (ROI) was placed in the vertebral body excluding cortical bone, subchondral abnormalities, focal lesions (e.g., hemangiomas), and posterior venous plexus. Three ROIs were used for each vertebra, each of them acquired on a different slice, with their mean used for analysis. A ROI was also positioned in an artefact-free site outside the patient to measure the noise. Signal-to-noise ratio (SNR) was obtained

by dividing the intra-vertebral signal intensity by the standard deviation of the noise. Image analysis is shown on Panel A of Figs. 1 and 2. To assess intra-reader reproducibility, measurements were repeated by the same reader after 30 days on a subset of 40 randomly chosen patients. To assess inter-reader reproducibility, a second independent reader with a 1-year experience in spinal MRI and a 2-year experience in DXA repeated measurements on the same subset.

We also evaluated the presence of vertebral fractures using the Genant visual semi-quantitative method [24]. In particular, we counted only moderate (grade 2, 25 % to 40 % loss of height) or severe (grade 3, >40 % loss of height) compression deformities to avoid ambiguity related to subjective borderline or mild compression deformities [14].

Dual-energy X-ray absorptiometry methods and image analysis

All DXA examinations were performed using a Hologic QDR-Discovery A unit (Hologic Inc., Bedford, MA, USA). Data were acquired and processed according to the International Society for Clinical Densitometry (ISCD) guidelines [25]. Both the lumbar spine and a non-dominant femur were examined. The presence of spinal degenerative joint disease was recorded as stated in the DXA report. Patients were considered as obese when BMI was $\geq 30 \text{ kg/m}^2$, according to the World Health Organization (WHO) guideline [23].

Statistical analysis

Data were reported as mean and standard deviation (SD) or as median and interquartile (IQ) range.

Intra- and inter-reader reproducibility was evaluated using the Bland-Altman method [26], obtaining a percent reproducibility, as already reported [27]. Differences in SNR among the four vertebrae were evaluated using the Friedman test for paired data. However, as consistent with the DXA practice [25], the correlation analysis (see below) was performed using the median value of the SNRs obtained in vertebrae from L1 to L4 ($\text{SNR}_{\text{L1-L4}}$). The correlation between $\text{SNR}_{\text{L1-L4}}$ and lumbar BMD was initially estimated on the whole sample of patients using the Spearman correlation coefficient. We excluded the BMD of the femur from the analysis because the inter-site correlation is considered inadequate for prediction of fracture risk [28] and ISCD guidelines emphasize that the BMD on a site is predictive of fracture risk for that site [25].

As lumbar DXA has limitations in obese patients [23] and in patients with spinal degenerative joint disease [29, 30], subsequent correlation analysis was performed by excluding those patients.

The diagnostic performance of $\text{SNR}_{\text{L1-L4}}$ was estimated by means of a receiver operator characteristic (ROC) analysis using lumbar-spine DXA as a reference standard. We established

thresholds that would yield an arbitrarily chosen near-90 % sensitivity or near-90 % specificity and also calculated positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+) and negative likelihood ratio (LR-).

We introduced an MRI-based score for diagnosing osteoporosis, named the M-score. In DXA, the dependence of BMD on the densitometer is removed normalizing the BMD obtained for an individual patient to the mean BMD of a reference population studied using the same densitometer, obtaining the so-called T-score [28]. Similarly, we measured the $\text{SNR}_{\text{L1-L4}}$ in all subjects of the reference group and calculated the mean (SNR_{ref}) and the standard deviation (SD_{ref}). The M-score was defined according to a formula as follows:

$$\text{M-score} = \frac{\text{SNR}_{\text{L1-L4}} - \text{SNR}_{\text{ref}}}{\text{SD}_{\text{ref}}}$$

The correlation between M-score and T-score was estimated using the Spearman correlation coefficient. The diagnostic performance of the M-score was estimated using ROC analysis, keeping the T-score as a reference standard. We defined thresholds that would yield high near-90 % sensitivity or near-90 % specificity and calculated PPV, NPV, LR+, and LR-.

The subset of patients with $\text{BMI} \geq 30 \text{ kg/m}^2$ or affected by spinal degenerative joint disease was separately analyzed. Patients with $\text{T-score} > -1$ were considered as normal, the remaining ones as non-normal. The same classification was used for the M-score according to the diagnostic thresholds distinguishing normal from non-normal patients and also osteoporotic from non-osteoporotic patients (McNemar test). The comparison of median M-score or T-score between patients with and those without at least a vertebral fracture was performed on the whole population of patients using the Mann-Whitney *U* test. For patients with at least one vertebral fracture, the osteoporosis rate was evaluated according to a T-score or an M-score and results were compared (McNemar test). Calculations were performed using SPSS v20 (SPSS Inc., Chicago, IL, USA).

Results

Patient characteristics

A total of 226 female patients met the inclusion criteria (median age 65 years; IQ range 58–73 years), with a median BMI of 26 kg/m^2 (IQ range $23\text{--}29 \text{ kg/m}^2$) and a median FRAX[®] of 8 % (IQ range 5–13 %) for major fractures and 2 % (IQ range 1–5 %) for hip fractures. The median time elapsed between the two examinations was 56 days (IQ range 26–113 days).

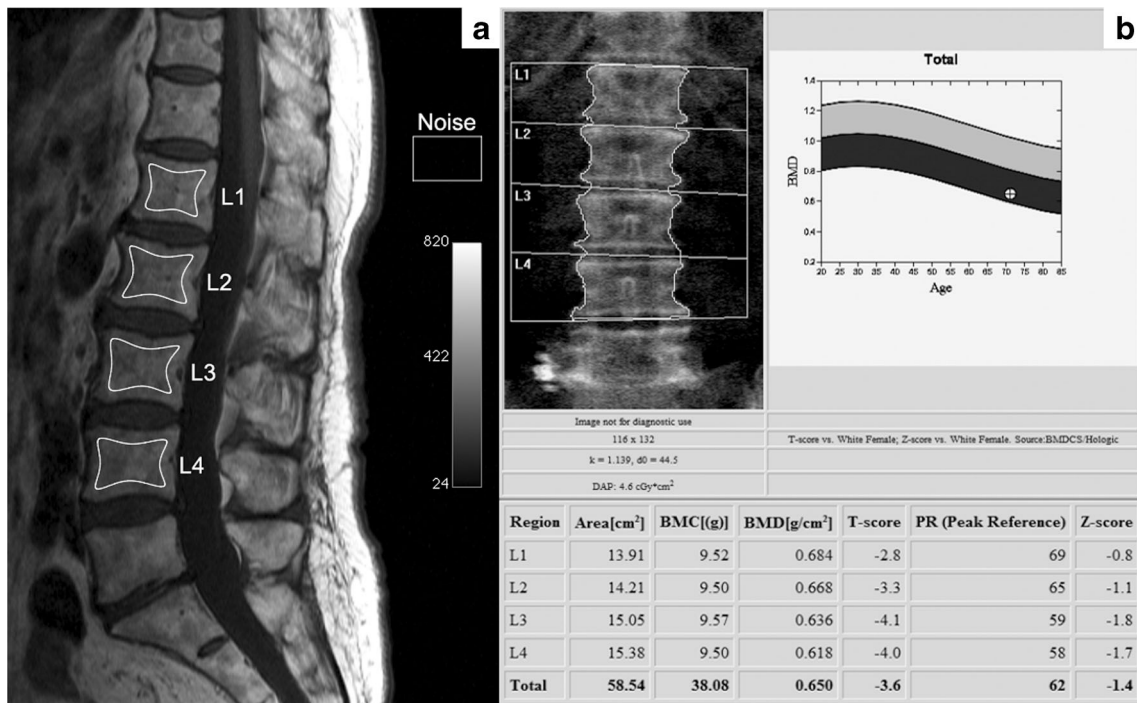


Fig. 1 Seventy-one-year-old woman with a body mass index of 23 kg/m². Panel A shows a sagittal T1-weighted magnetic resonance image of the lumbar spine. Signal intensity is measured on four regions of interest manually segmented from L1 to L4, plus a region of interest placed

outside the patient for the measurement of the noise. No fractures are shown. Panel B shows the report from dual-energy X-ray absorptiometry of the same patient. T-score=-3.6 and M-score=8, both in the range of osteoporosis

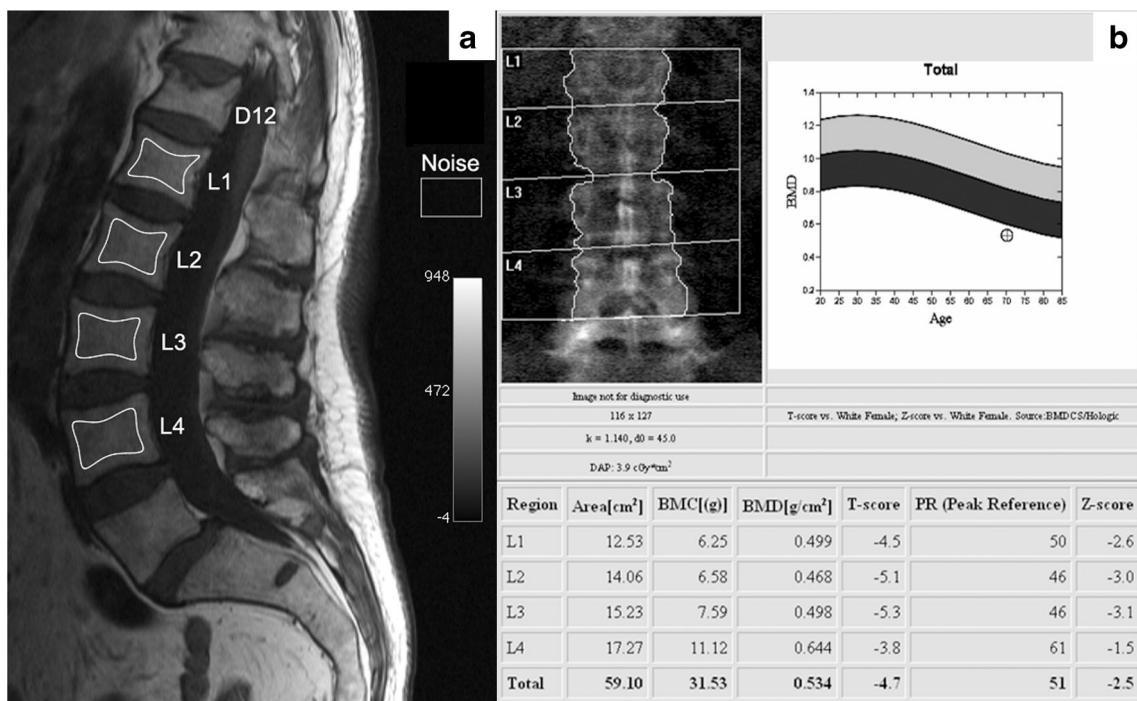


Fig. 2 Seventy-year-old woman with a body mass index of 24 kg/m². Panel A shows a sagittal T1-weighted magnetic resonance image of the lumbar spine. Signal intensity is measured on four regions of interest manually segmented from L1 to L4, plus a region of interest placed

outside the patient for the measurement of the noise. A moderate fragility fracture on D12 is visible. Panel B shows the report from dual-energy X-ray absorptiometry of the same patient. T-score=-4.7 and M-score=9, both in the range of osteoporosis

According to WHO criteria, 37 patients (16 %) had a normal T-score, 119 patients (53 %) had osteopenia, and 70 patients (31 %) had osteoporosis. Twenty-eight patients (12 %) had a spinal degenerative joint disease, 31 (14 %) were obese, and 3 (1 %) had both conditions. Of 226 patients, 31 (14 %) had a moderate ($n=24$) or severe ($n=7$) vertebral fracture.

The reference group consisted of 131 subjects, without any abnormalities on MRI.

Reproducibility of measurements and comparison among vertebrae

Intra-reader reproducibility ranged between 89 % and 93 % across vertebrae, while inter-reader reproducibility ranged between 78 % and 84 %. L1 showed the lowest reproducibility (89 % and 78 %, respectively). Intra- and inter-reader reproducibility of SNR_{L1-L4} resulted 93 % and 83 %, respectively.

The median SNR_{L1} was 43 (IQ range 32–66), the median SNR_{L2} was 39 (IQ range 29–62), the median SNR_{L3} was 36 (IQ range 27–54), and the median SNR_{L4} was 36 (IQ range 26–52), with a statistically significant difference ($P<0.001$). SNR_{L1-L4} was 38 (IQ range 29–58).

Correlation analysis and diagnostic performance of SNR_{L1-L4}

SNR_{L1-L4} negatively correlated to lumbar BMD on the entire sample of 226 patients, ($r=-0.344$; $P<0.001$), and it also negatively correlated when excluding 62 patients with BMI ≥ 30 kg/m² or affected by spinal degenerative joint disease ($r=-0.677$; $P<0.001$). For distinguishing osteoporosis from non-osteoporosis, we found a threshold of 36 when setting near-90 % sensitivity (with 61 % specificity) and a threshold of 51 when setting near-90 % specificity (with 60 % sensitivity); at ROC analysis, the area under the curve (AUC) was 0.852. For distinguishing normal from low-BMD patients, we found a threshold of 41.5 when setting near-90 % sensitivity (with 46 % specificity) and a threshold of 24 when setting near-90 % specificity (with 49 % sensitivity), with AUC=0.834 (Table 1).

Correlation analysis and diagnostic performance of M-score

For the 164 patients with BMI <30 kg/m² and not affected by spinal degenerative joint disease, M-score negatively

Table 1 Discriminating thresholds and diagnostic performance of SNR_{L1-L4} using dual-energy X-ray absorptiometry of the total spine as a reference standard in non-obese patients and in patients without spinal degenerative joint disease

Variable	Diagnostic Performance of SNR _{L1-L4} for Distinguishing Osteoporosis From Non-osteoporosis*		Diagnostic Performance of SNR _{L1-L4} for Distinguishing Normal From Low BMD [§]	
	Threshold for Achieving About 90 % Sensitivity	Threshold for Achieving About 90 % Specificity	Threshold for Achieving About 90 % Sensitivity	Threshold for Achieving About 90 % Specificity
AUC	0.852		0.834	
SNR _{L1-L4}	≥ 36	≥ 51	≤ 41.5	≤ 24
Sensitivity				
TP/(TP+FN)	92 % (46/50)	60 % (30/50)	90 % (35/39)	49 % (19/39)
95%CI	81–98	45–74	76–97	32–65
Specificity				
TN/(TN+FP)	61 % (69/114)	90 % (103/114)	46 % (57/125)	90 % (113/125)
95%CI	51–70	83–95	37–55	84–95
PPV				
TP/(TP+FP)	51 % (46/91)	73 % (30/41)	34 % (35/103)	61 % (19/31)
95%CI	40–61	57–86	25–44	42–78
NPV				
TN/(TN+FN)	95 % (69/73)	84 % (103/123)	93 % (57/61)	85 % (113/133)
95%CI	87–99	76–90	84–98	78–91
LR+	2.33	6.22	1.65	5.07
LR-	0.13	0.44	0.22	0.57

AUC=area under the curve; SNR=signal-to-noise ratio; TP=true positive; TN=true negative; FP=false positive; FN=false negative; BMD=bone mineral density; PPV=positive predictive value; NPV=negative predictive value; LR+=positive likelihood ratio; LR-=negative likelihood ratio; 95%CI=95 % confidence interval

*Non-osteoporosis was defined as having osteopenia or normal BMD of the total spine. Here non-osteoporosis is considered a negative finding

§ Osteopenia was combined with osteoporosis and normal BMD was considered a positive result

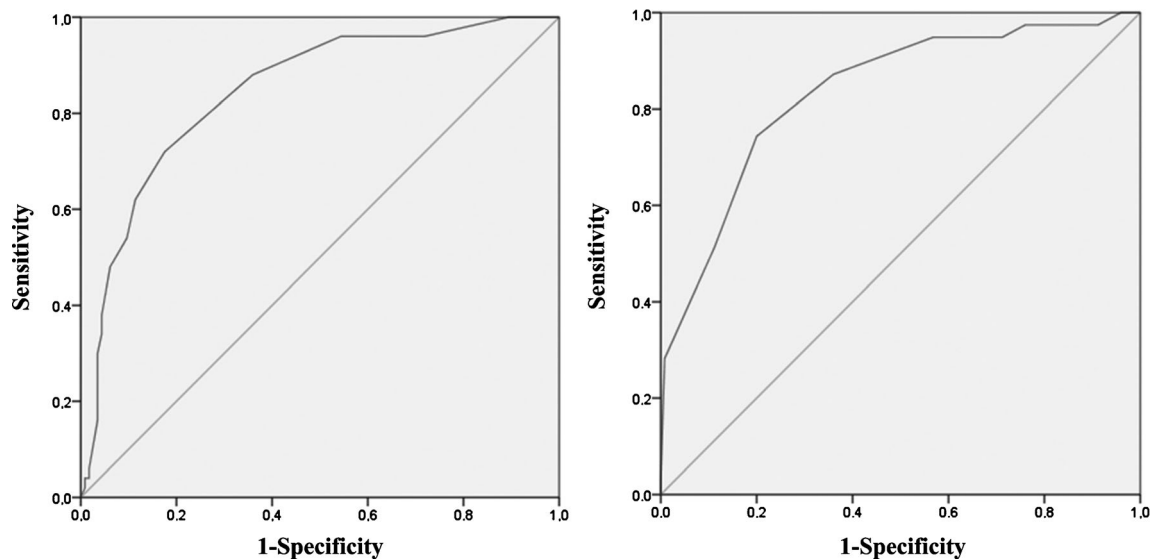


Fig. 3 Receiver operator characteristic curve of M-score for the discrimination of osteoporosis from non-osteoporosis patients (*left*) and for the discrimination of normal from low-bone mineral density patients (*right*): area under the curve were 0.844 and 0.835, respectively

correlated to T-score ($r=-0.682$; $P<0.001$). For distinguishing osteoporosis from non-osteoporosis, we found a threshold of 2.5 when setting near-90 % sensitivity (with 64 % specificity) and a threshold of 5.5 when setting near-

90 % specificity (with 54 % sensitivity), with AUC=0.844 (Fig. 3). For distinguishing normal from low-BMD patients, we found a threshold of 2.5 when setting near-90 % sensitivity (with 64 % specificity) and a threshold of 0.5 when setting

Table 2 Discriminating thresholds and diagnostic performance of M-score using dual-energy X-ray absorptiometry of the total spine as a reference standard in non-obese patients and in patients without spinal degenerative joint disease

	Diagnostic Performance of M-score for Distinguishing Osteoporosis From Non-osteoporosis*		Diagnostic Performance of M-score for Distinguishing Normal From Low BMD [§]	
	Threshold for Achieving About 90 % Sensitivity	Threshold for Achieving About 90 % Specificity	Threshold for Achieving About 90 % Sensitivity	Threshold for Achieving About 90 % Specificity
AUC	0.844		0.835	
M-score	≥ 2.5	≥ 5.5	≤ 2.5	≤ 0.5
Sensitivity				
TP/(TP+FN)	88 % (44/50)	54 % (27/50)	87 % (34/39)	51 % (20/39)
95%CI	76–96	39–68	73–96	35–68
Specificity				
TN/(TN+FP)	64 % (73/114)	90 % (103/114)	64 % (80/125)	89 % (111/125)
95%CI	55–73	83–95	55–72	82–94
PPV				
TP/(TP+FP)	52 % (44/85)	71 % (27/38)	43 % (34/79)	59 % (20/34)
95%CI	41–63	54–85	32–55	41–75
NPV				
TN/(TN+FN)	92 % (73/79)	82 % (103/126)	94 % (80/85)	85 % (111/130)
95%CI	84–97	74–88	87–98	78–91
LR+	2.45	5.60	2.42	4.58
LR-	0.19	0.51	0.20	0.55

For the definition of M-score see the text. AUC=area under the curve; TP=true positive; TN=true negative; FP=false positive; FN=false negative; BMD=bone mineral density; PPV=positive predictive value; NPV=negative predictive value; LR+=positive likelihood ratio; LR-=negative likelihood ratio; 95%CI=95 % confidence interval

*Non-osteoporosis was defined as having osteopenia or normal BMD of the total spine. Here non-osteoporosis is considered a negative finding

[§] Osteopenia was combined with osteoporosis and normal BMD was considered a positive result

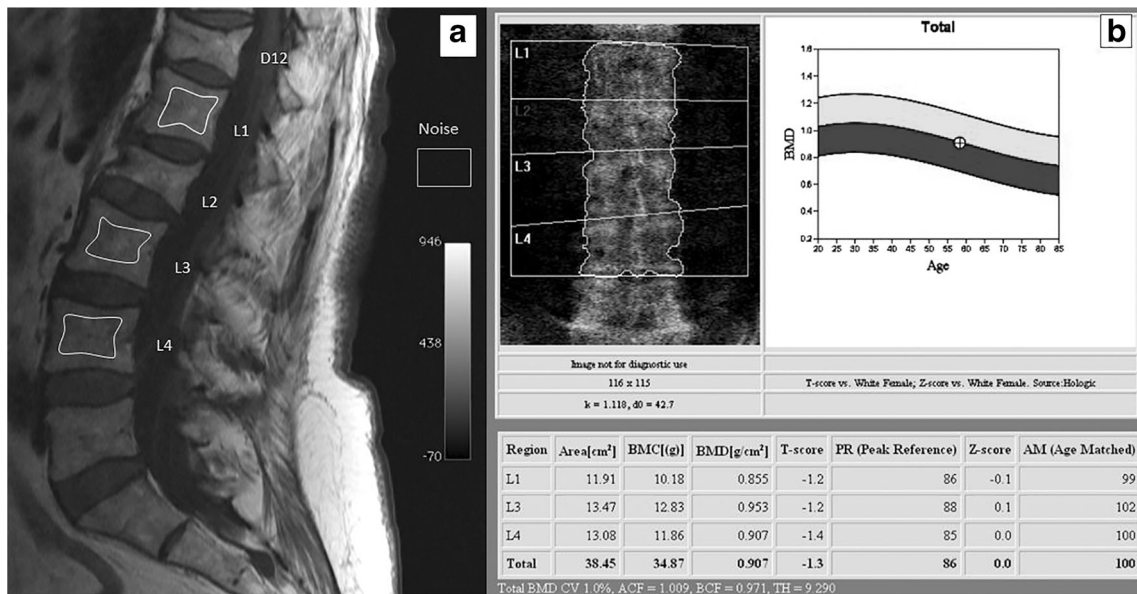


Fig. 4 Seventy-four-year-old woman with a body mass index of 27 kg/m². Panel A shows a sagittal T1-weighted magnetic resonance image of the lumbar spine. Signal intensity is measured on three regions of interest manually segmented (L1, L3 and L4), plus a region of interest placed outside the patient for the measurement of the noise. Two moderate

fractures at D12 and L2 vertebrae are visible. Panel B shows the report from dual-energy X-ray absorptiometry of the same patient. T-score=-1.3 is in the range of osteopenia while M-score=11 is in the range of osteoporosis

near-90 % specificity (with 51 % sensitivity), with AUC=0.835 (Table 2 and Fig. 3).

Analysis of patients with spinal degenerative joint disease or obese patients

Of the 62 patients with BMI≥30 kg/m² or affected by spinal degenerative joint disease, 35 (56 %) had a normal T-score, while the remaining 27 (44 %) had osteopenia or osteoporosis. According to the M-score, they were six (10 %) and 56 (90 %), respectively (P<0.001). Conversely, according to DXA, six (10 %) patients were osteoporotic, while the remaining 56 (90 %) were non-osteoporotic; according to the M-score, they were 24 (39 %) and 38 (61 %), respectively (P<0.001). Of the six osteoporotic patients at DXA, none (0 %) had a vertebral osteoporotic fracture; while of the 24 osteoporotic patients at MR, 12 (50 %) had a vertebral osteoporotic fracture.

Analysis of patients with respect to vertebral fractures

We observed 31 patients with at least one vertebral fracture, showing a median M-score of 4.1, significantly higher (P=0.001) than the M-score (2.5) of the remaining 195 patients without any fractures. In these two groups, the T-score was -2.1 and -1.6, respectively, with a not statistically significant difference (P=0.277). Regarding the 31 patients with vertebral fractures, 15 (48 %) had osteoporosis according to the M-score, while only 8 (26 %) had osteoporosis according to the T-score (P=0.016). An example of a patient with vertebral

fractures characterized by a non-osteoporotic T-score but with an osteoporotic M-score is shown in Fig. 4.

Discussion

Even though DXA represents a well-standardized quantitative imaging technique in the management of osteoporosis [31], few DXA examinations are performed in subjects at high risk of fragility fractures and few patients with fragility fractures are properly evaluated and treated for osteoporosis [32, 33]. This is probably due to the lack of a screening policy. DXA has an acceptable specificity but a low detection rate [28], so a case finding strategy is applied and patients identified for a fragility fracture or by the presence of other strong risk factors are thereafter assessed with DXA [28].

T1-weighted spin-echo images are best done to evaluate the cellular content in bone marrow because of the high fat content interspersed with hematopoietic elements [18]. The hydrophobic carbon-hydrogen groups in fat result in a short T1 relaxation time because of very efficient spin-lattice relaxation. Diffusely increased T1-weighted hyperintensity also indicates decreased cellularity of bone marrow and increased fat content. This pattern can be associated with osteoporosis [18]. In fact, osteoporotic individuals, whether of primary or secondary origin, have more fat in the bone marrow than age-matched controls [34]. It has been proposed that such an increase in bone marrow fat may be associated with a compensation mechanism for osteoporosis-related changes in trabecular microarchitecture. Due to trabecular thinning and loss of

bone volume, adipocytes may infiltrate spaces previously occupied by trabecular bone [35].

In this study we introduced a new quantitative lumbar spine MRI-based method on standard T1-weighted images using lumbar DXA as a reference standard for identifying patients at increased risk of osteoporosis. We demonstrated that lumbar spine MRI routinely performed for low back pain can be used as an opportunistic screening tool for osteoporosis, with accuracy values around 83 % having DXA as a reference standard. Obviously, the aim of this new index is not to replace DXA with MRI. It is rather to provide a reliable, easy-to-use, and highly reproducible tool that – among patients undergoing lumbar spine MRI for different reasons – may allow for an opportunistic “screening” and identify those subjects who need to undergo DXA.

We found that lumbar spine vertebrae are not equivalent to each other in terms of SNR, thus justifying the use of their mean, as in lumbar DXA, with excellent intra- and inter-reader reproducibility. Based on the whole sample of 226 patients, SNR_{L1-L4} slightly negatively correlated with BMD. However, this correlation was twice as high when excluding obese patients and patients affected with spinal degenerative joint disease, which may have resulted in a spurious BMD. In these patients lumbar BMD may be increased, resulting in false assurance of a low fracture risk [28, 29]. Furthermore, we have to consider that in DXA practice, osteoporosis is diagnosed on a T-score instead of BMD. This reduces the difficulties associated with differences in calibration between instruments and is referred to as the T-score [28]. Thus, we introduced the M-score as a predictor of osteoporosis, built on the model of the T-score, with high diagnostic performance. When setting specificity at 90 %, we obtained an M-score threshold of 5.5 with a PPV of 71 %. This means that for every 100 patients with $M\text{-score} \geq 5.5$, 71 will have a lumbar spine T-score ≤ -2.5 (osteoporosis). Interestingly, when applying the M-score threshold ≥ 5.5 to 62 patients obese or affected with spinal degenerative joint disease (excluded from correlation analysis), we found that 24/62 (39 %) could be classified as osteoporotic, compared to only 6/62 (10 %) applying the T-score threshold of ≤ -2.5 (osteoporosis). We may argue that MRI is not influenced by spinal degenerative joint disease nor by a $BMI \geq 30 \text{ kg/m}^2$, with sensitivity potentially higher than that of DXA. In fact, considering the presence of vertebral fragility fractures for clinical diagnosis of osteoporosis [9], we showed that 12 of 24 (50 %) patients classified as osteoporotic with the use of MRI had a vertebral fracture (true positives), versus none of 6 (0%) patients that were osteoporotic at DXA. The hypothesis of a higher MRI sensitivity for osteoporosis is supported also by our findings of fractures in the whole sample. We obtained a median M-score in patients with a vertebral fracture (4.1) significantly higher ($P=0.001$) than the M-score in patients without fractures (2.5), while the T-score was not significantly different, being -2.1 versus

-1.6 , respectively ($P=0.277$). Moreover, among 31 patients with a vertebral fracture, 15 (48 %) had osteoporosis according to the M-score, while 8 (26 %) had osteoporosis according to the T-score ($P=0.016$).

This study has limitations. First, it has a retrospective design, with a non-simultaneous MRI and DXA. However, the median time elapsed between the two examinations was relatively low and it is unlikely it had influenced the correlation between DXA and MRI. Second, DXA is an imperfect reference standard: most patients having low-trauma fractures have non-osteoporotic BMD and BMD evaluations have limitations in obese patients and in patients affected with spinal degenerative joint disease. Thus, data regarding these patients may deserve further analysis. Third, regarding the reference group of young woman undergoing MRI, bone densitometry data were lacking because they did not undergo DXA. Moreover, we assumed that they were healthy from considering the bone status. Fourth, the use of M-score values is valid for the single MR system they are built on. Thus, to use in clinical practice, each single centre should build its own thresholds based on different SNRs values provided by each MR system. Finally, the signal evaluation by the placement of the ROI within the vertebra and outside of the patient are simplified methods of measurement with respect to the more accurate spectroscopic measurement.

Despite the limitations described, our preliminary study showed that routine lumbar spine MRI T1-weighted sequence can be used to predict osteoporosis, opening a new opportunity to extend the screening for osteoporosis without any additional imaging, radiation exposure, cost, or patient time to a large cohort of patients daily undergoing lumbar spine MRI for low back pain. Its diagnostic value remains to be demonstrated on a large prospective cohort of patients.

Acknowledgments The scientific guarantor of this publication is Michele Bandirali. The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article. The authors state that this work has not received any funding. One of the authors, Giovanni Di Leo, has significant statistical expertise. Institutional Review Board approval was obtained. Written informed consent was waived by the Institutional Review Board. Methodology: retrospective, case-control study, performed at one institution.

References

1. NIH consensus development panel on osteoporosis prevention, diagnosis and therapy (2001) Osteoporosis prevention, diagnosis and therapy. *JAMA* 285:785–795
2. Link TM (2012) Osteoporosis imaging: state of the art and advanced imaging. *Radiology* 263:3–17
3. Glaser DL, Kaplan FS (1997) Osteoporosis. Definition and clinical presentation. *Spine* 22:12S–16S

4. McDonnell JM, Lane JM, Zimmerman PA (1987) Osteoporosis: definition, risk factors, etiology, and diagnosis. *AAOHN* 35:527–530
5. Sambrook P, Cooper C (2006) Osteoporosis. *Lancet* 367:2010–2018
6. Office of the Surgeon General (US) (2004) Bone Health and Osteoporosis: A Report of the Surgeon General. Rockville (MD): Office of the Surgeon General (US). 5, The Burden of Bone Disease. Available via <http://www.ncbi.nlm.nih.gov/books/NBK45502/>. Accessed 20 Jun 2014
7. Melton LJ 3rd, Atkinson EJ, Cooper C, O'Fallon WM, Riggs BL (1999) Vertebral fractures predict subsequent fractures. *Osteoporos Int* 10:214–221
8. Bandirali M, Lanza E, Messina C et al (2013) Dose absorption in lumbar and femoral dual energy X-ray absorptiometry examinations using three different scan modalities: an anthropomorphic phantom study. *J Clin Densitom* 16:279–282
9. World Health Organization (1994) Technical report: Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: a report of WHO study group. Switzerland, Geneva
10. Kühn JP, Hernando D, Meffert PJ et al (2013) Proton-density fat fraction and simultaneous R2* estimation as an MRI tool for assessment of osteoporosis. *Europ Radiol* 23:3432–3439
11. King AB, Fiorentino DM (2011) Medicare payment cuts for osteoporosis testing reduced use despite tests' benefit in reducing fractures. *Health Aff* 30:2362–2370
12. Leslie WD, Giangregorio LM, Yogendran M et al (2012) A population-based analysis of the post-fracture care gap 1996–2008: the situation is not improving. *Osteoporos Int* 23:1623–1629
13. Solomon DH, Finkelstein JS, Katz JN, Mogun H, Avorn J (2003) Underuse of osteoporosis medications in elderly patients with fractures. *Am J Med* 115:398–400
14. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N (2013) Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Ann Intern Med* 158:588–595
15. Bazzocchi A, Spinnato P, Albisinni U, Battista G, Rossi C, Guglielmi G (2012) A careful evaluation of scout CT lateral radiograph may prevent unreported vertebral fractures. *Eur J Radiol Sep* 81:2353–2357
16. Hart LG, Deyo RA, Cherkin DC (1995) Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine* 20:11–19
17. Deyo RA, Weinstein JN (2001) Low back pain. *N Engl J Med* 344:363–370
18. Shah LM, Hanrahan CJ (2011) MRI of spinal bone marrow: part I, techniques and normal age-related appearances. *AJR Am J Roentgenol* 197:1298–1308
19. Shen W, Chen J, Punyanitya M, Shapses S, Heshka S, Heymsfield SB (2007) MRI-measured bone marrow adipose tissue is inversely related to DXA-measured bone mineral in Caucasian women. *Osteoporos Int* 18:641–647
20. Shen W, Scherzer R, Gantz M et al (2012) Relationship between MRI-measured bone marrow adipose tissue and hip and spine bone mineral density in African-American and Caucasian participants: the CARDIA study. *J Clin Endocrinol Metab* 97:1337–1346
21. Bazzocchi A, Garzillo G, Fuzzi F et al (2014) Localizer sequences of magnetic resonance imaging accurately identify osteoporotic vertebral fractures. *Bone* 61:158–163
22. Kanis JA, Glüer CC (2000) An update on the diagnosis and assessment of osteoporosis with densitometry. Committee of Scientific Advisors, International Osteoporosis Foundation. *Osteoporos Int* 11:192–202
23. Nielson CM, Srikanth P, Orwoll ES (2012) Obesity and fracture in men and women: an epidemiologic perspective. *J Bone Miner Res* 27:1–10
24. Genant HK, Wu CY, Van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res* 8:1137–1148
25. Lewiecki EM, Gordon CM, Baim S et al (2007) International Society for Clinical Densitometry Adult and Pediatric Official Positions. *Bone* 43:1115–1121
26. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
27. Bandirali M, Sconfienza LM, Aliprandi A et al (2014) In vivo differences among scan modes in bone mineral density measurement at dual-energy X-ray absorptiometry. *Radiol Med* 119:257–260
28. Kanis JA (2002) Diagnosis of osteoporosis and assessment of fracture risk. *Lancet* 359:1929–1936
29. Liu G, Peacock M, Eilam O, Dorulla G, Braunstein E, Johnston CC (1997) Effect of osteoarthritis in the lumbar spine and hip on bone mineral density and diagnosis of osteoporosis in elderly men and women. *Osteoporos Int* 7:564–569
30. Schneider DL, Bettencourt R, Barrett-Connor E (2006) Clinical utility of spine bone density in elderly women. *J Clin Densitom* 9:255–260
31. Osteoporosis prevention, diagnosis, and therapy (2000) NIH Consensus Statement (2000 March 27–29) 17(1): 1–36. Available via: <http://consensus.nih.gov/2000/2000Osteoporosis111PDF.pdf>. Accessed 20 Jun 2014
32. Elliot-Gibson V, Bogoch ER, Jamal SA (2004) Practice patterns in the diagnosis and treatment of osteoporosis after a fragility fracture: a systematic review. *Osteoporos Int* 15:767–778
33. Kanis JA, Borgstrom F, De Laet C et al (2005) Assessment of fracture risk. *Osteoporos Int* 16:581–589
34. Meunier P, Aaron J, Edouard C, Vignon G (1971) Osteoporosis and the replacement of cell populations of the marrow by adipose tissue. A quantitative study of 84 iliac bone biopsies. *Clin Orthop* 80:147–154
35. Rosen CJ, Ackert-Bicknell C, Rodriguez JP, Pino AM (2009) Marrow fat and the bone microenvironment: developmental, functional, and pathological implications. *Crit Rev Eukaryot Gene Expr* 19:109–124