Calcified Tissue International (2026) 117:151
https://doi.org/10.1007/s00223-026-01610-x

ORIGINAL RESEARCH

Fracture Risk Stratification by REMS: A 5 Years Prospective Longitudinal Multicentre Study

Carla Caffarelli[1] · Maria Luisa Brandi[2] · Giuseppe Girasole[3] · Andrea Giusti[3] · Maurizio Muratore[4] · Stefano Gonnelli[1]

Received: 9 July 2026 / Accepted: 17 September 2026
© The Author(s) 2026

Abstract

Osteoporosis is a skeletal disorder characterized by low bone mass and deterioration of bone microarchitecture. Current predictive tools, mainly based on DXA-BMD, do not full capture all the factors contributing to bone deterioration. By combining information on bone quantity (T-score) and skeletal fragility (Fragility Score, FS), the Radiofrequency Echographic Multi-Spectrometry provides an integrated measure of fracture susceptibility known as fracture risk stratification. We aimed To evaluate the ability of REMS fracture risk stratification through a 5 years prospective longitudinal multicentre study. A non-parametric Mann–Whitney test was used to assess differences of REMS T-score and FS between fractured and non-fractured patients. Additionally, Chi-squared tests, Cochran-Armitage trend tests, class-specific unadjusted logistic regressions and standardized residual analyses were performed to evaluate the proposed stratification model. A total of 4443 female patients was recruited and $n=4017$ underwent REMS scans. The Chi-squared and the Cochran-Armitage trend tests showed that the distribution of fractured patients differed across the 7 risk classes with a strong positive monotonic trend at lumbar spine ($Z=47.83$; $p<0.0001$) and femoral neck ($Z=75.41$; $p<0.0001$). From a clinical perspective, both Class-specific binary logistic regression and standardized residuals clearly identified two opposite risk profiles: patients belonging to the low-risk classes exhibited a significantly reduced probability of fracture, whereas those in the high-risk classes showed a markedly increased fracture risk. This 5-year longitudinal study supports the clinical consistency of a REMS composite approach for fracture risk stratification based on the integration of REMS BMD T-score and Fragility Score.

Keywords REMS · Fracture risk · T-score · Fragility score

Introduction

Osteoporosis is a skeletal disorder that is chronic and progressive in nature. It is characterized by low mineral density and microarchitectural deterioration of bone tissue, leading to an increased bone fragility and susceptibility to fractures [1]. Globally, it represents a major public health issue, particularly in elderly population [2–5]. Due to its silent progression accompanied by an absent screening program and a poor awareness of the problem, people generally receive a diagnosis of osteoporosis when a fracture has already occurred. Thus, the high social and economic impact of osteoporosis makes the early identification of individuals at high-risk of utmost importance.

Currently, the standard technique used for the diagnosis of osteoporosis is the dual-energy X-ray absorptiometry (DXA). According to the World Health Organization (WHO) diagnostic criteria, osteoporosis is defined by a T-score ≤ -2.5 at lumbar spine, femoral neck or total hip [6]. Despite its widespread use and clinical utility, DXA has several well-recognized limitations. First, it provides a 2D areal measurement of BMD and not a true 3D volumetric assessment, which may not fully reflect bone strength:

Maria Luisa Brandi, Giuseppe Girasole, Andrea Giusti, and Maurizio Muratore have equally contributed to this work.

✉ Carla Caffarelli
carlacaffarelli04@gmail.com

[1] Division of Internal Medicine, Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy

[2] Observatory on Fragility Fractures Italia (OFF Italia), Florence, Italy

[3] SC Rheumatology, ASL 3 Genovese, Genoa, Italy

[4] ASL-Le, Ospedale Vito Fazzi, Lecce, Italy

Published online: 01 October 2026

areal BMD, in fact, is susceptible to bone size and, thus, will overestimate fracture risk in individuals with a small skeletal size, who will have lower areal BMD than normal-sized individuals, and vice versa for individuals with a large skeletal size [7]. Second, DXA does not capture information about bone quality, including factors such as microarchitecture, turnover, mineralization and micro-damage, all of which significantly contributing to overall bone strength and fracture risk [8]. Finally, DXA involves exposure to ionizing radiation. Although DXA involves low radiation exposure, it may be less suitable for frequent monitoring or for the examinations of fragile population, in addition to high-costs and non-portable equipment that limit the accessibility in some healthcare settings [9].

In response to these limitations, recent advances in medical imaging and signal processing have led to the development of innovative diagnostic techniques able to provide a more detailed and clinically relevant assessment about bone status. Among these, Radiofrequency Echographic Multi Spectrometry (REMS) is a non-invasive, radiation-free and portable technology based on the analysis of native radiofrequency ultrasound signals. The technology provides an estimate of BMD and derived parameters (T-score and Z-score) reflecting the structural and compositional properties of the bone. Several studies have demonstrated that REMS shows a high diagnostic agreement with DXA [10, 11] and a strong ability to avoid clinically relevant artefacts that may lead to BMD overestimation in DXA acquisitions [12, 13] and its efficacy for short-term monitoring of bone health [14, 15]. In addition, REMS has introduced also the Fragility Score (FS), a proprietary parameter designed to enhance the prediction of fracture risk by covering microstructural information not captured by traditional BMD measures. Previous published studies have demonstrated that the FS is more strongly associated with a history of fractures than the DXA and REMS T-score alone, suggesting its superior discriminative capability in identifying individuals at high-risk of fracture [16–18]. All these characteristics support the use of the REMS technology in osteoporosis management as a reliable and accessible alternative to DXA [9, 19].

In this context, an accurate and accessible method for the assessment of fracture risk is essential as many fragility fractures occur in individuals with T-score values above the osteoporotic threshold, thus highlighting the importance of adopting more comprehensive methods for assessing fracture risk [4, 20, 21]. Currently, several risk assessment tools combining different risk factors (including DXA-BMD or not) have been developed, with different levels of complexity, but only few of them (QFracture, Fracture Risk Assessment (FRAX) tool, Garvan, Derived Fracture Risk Assessment (DeFRA) among others) have been validated in a population-based settings according to a proper methodological approach [22]. However, none of these perform better than the others [23] and none of these is able to meet different specific needs [24]. Considering that both BMD and bone quality contribute to define bone strength and fracture risk, the combination of these two sources of information could potentially enhance diagnostic accuracy.

To this end, a novel composite parameter that integrates data from both REMS T-score and FS could generate an integrated measure of fracture susceptibility. The underlying hypothesis is that this approach could provide a more comprehensive assessment of skeletal fragility, thereby improving the stratification of patients based on fracture risk and supporting more effective clinical decision-making procedures.

Therefore, the primary aim of the present 5-year longitudinal multicentre study was to evaluate the ability of the REMS fracture risk stratification approach, based on the combined use of REMS BMD T-score and FS, to classify women into predefined classes of increasing fracture risk. Specifically, lumbar spine classes were evaluated against major osteoporotic fractures (MOF), whereas femoral neck classes were evaluated against incident hip fractures. By integrating information related to bone quantity and bone quality with a single non-ionizing assessment, this approach was intended to explore whether REMS-derived parameters may support fracture risk stratification beyond conventional densitometric assessment alone. In this context, the present study was designed to assess the clinical behaviour and internal consistency of predefined REMS risk classes over a 5-year follow-up period.

Materials and Methods

This multicentre prospective observational study was conducted across the participating clinical centres. The patient recruitment started on January 2016 and lasted until July 2019. Patient follow-up lasted until July 2024. The study protocol was approved by the Ethics Review Board of the participating institutions.

Patients and Measurements

The clinical data were collected from a study involving patients who fulfilled the following inclusion criteria: Caucasian ethnicity, female gender, age between 30 and 90 years, medical prescription for spinal and/or femoral densitometric examinations and absence of significant ambulation difficulties. All the patients voluntarily entered the study after giving written informed consent. According to the study protocol, patients underwent an ultrasound REMS scan of lumbar spine and femoral neck. REMS acquisitions were

performed at the anatomical site(s) specified in the medical prescription (femoral neck and/or lumbar spine). During the follow-up period, patients were regularly (every 6 months) seen in-person or contacted by phone to monitor for any occurrence of incident fractures, which were confirmed through medical reports based on imaging investigations (such as radiographic imaging, vertebral morphometry, etc., when available) acquired during the follow-up period lasting up to 5 years. Investigators responsible for fracture ascertainment were blinded to the participants' baseline REMS results and assigned risk classes. For the purposes of the study, fractures resulting from trauma were excluded in order to focus exclusively on fractures attributed to bone fragility.

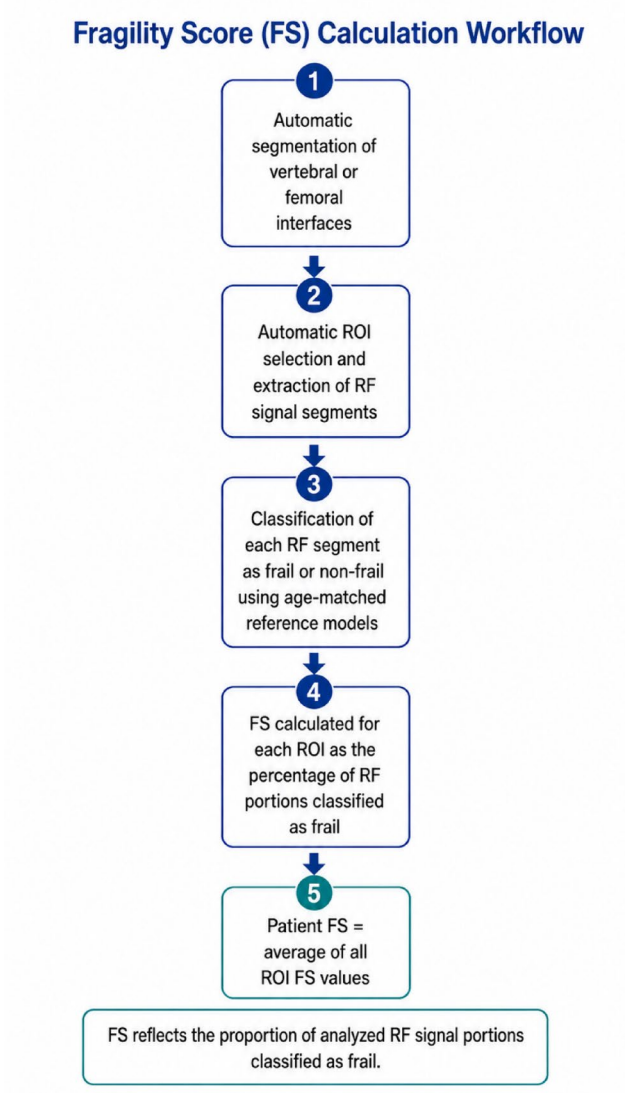

Fig. 1 Diagram representing the workflow of Fragility Score calculation

Examinations

REMS scans were carried out using an EchoStation device (Echolight S.p.a., Lecce, Italy) equipped with a convex transducer operating at a nominal frequency of 3.5 MHz, following the manufacturer's instructions. The scanning procedure has been thoroughly described in previously published works [25, 26]. In brief, lumbar scans were conducted by positioning the echographic transducer trans-abdominally under the sternum to initially visualize the L1 vertebra, then moving it down to L4, following the on-screen and audible instructions from the EchoStudio software (Echolight S.p.a., Lecce, Italy). Femoral neck scans were performed by placing the echographic transducer parallel to the head-neck axis of the femur to visualize the typical proximal femur profile, including the interfaces of femoral head, neck and trochanter. Once the target interfaces were identified, the operator held the image for 40 s according to the indications provided by the device.

All acquisitions were performed by operators who had received the specific training according to the guidelines of the device manufacturer and had at least 3 months of previous experience in REMS acquisitions.

Parameters

BMD (g/cm^2) and T-score values, the latter derived from the standard reference database for Caucasians, were included in the subsequent statistical analysis.

REMS software also produces a proprietary parameter, named the Fragility Score, used as an indicator of skeletal fragility. For the calculation of the FS, ultrasound radiofrequency signals reflected by the bone target are processed through spectral and statistical analyses [27, 28]. An algorithm embedded in EchoStudio automatically identifies the bone interface and the region of interest (ROI). In detail, the analysis of FS involved the following steps: (1) automatic segmentation of vertebral or femoral interfaces, (2) automatic selection of a ROI for each segmented interface, (3) classification of each radiofrequency segment as "frail" or "not frail", (4) for each considered ROI, calculation of the FS values, defined as the percentage of the analysed RF portions that were previously classified as "frail", (5) calculation of the FS value as average of all the ROI values. A graphical abstract of the FS is shown in Fig. 1.

Then, the spectral profile of the patient ROI is classified as frail or not-frail by comparing it with reference spectral models previously built from fractured and non-fractured patients [27, 28]. From this comparison, the system automatically provides a quantitative estimation through a dimensionless parameter ranging from 0 (maximum similarity with the healthy model) to 100 (maximum similarity

with the fractured model). Based on the manufacturer's instructions, the bone quality of a patient can be classified as normal when the FS value is included within the range indicated by the green code, as reduced when the FS value is included within the range indicated by the yellow code, and as low when the FS value is included within the range indicated by the red code.

By combining the information obtained from the FS and the T-score values, a 3 × 5 matrix was generated. To correctly account for "borderline" cases, a tolerance of ± 0.3 to the T-score thresholds was applied [10]. Similar risk ranges were subsequently merged, generating 7 final risk classes. When FS and T-score values were measured through lumbar REMS scanning, patients were stratified into 7 risk classes (from R1 to R7), with a related probability range of MOF. Analogously, when the REMS scan was performed at the femoral neck, the relationship between the corresponding parameters generated the 7 risk classes along with the fracture probability range for hip fractures.

The combination of the REMS T-score and FS parameters to generate the 7 risk classes is illustrated in Table 1.

Statistical Analysis

The analysis was conducted separately at lumbar spine and femoral neck. Parameters derived from lumbar spine scans were matched with data on MOF, while parameters derived from femoral neck acquisitions were matched exclusively with data on hip fractures.

Continuous variables, including patients' age, height, weight, body mass index (BMI), REMS T-scores and FS are presented as mean ± SD. All continuous variables were tested to assess the normality by Shapiro–Wilk test. A non-parametric Mann–Whitney test was performed to assess statistically significant differences between patient groups with or without incident fragility fractures. A DeLong test was performed to compare the areas under the receiver operating characteristic curves (AUCs) [29].Discrete variables (7 risk classes, number of patients enrolled, number of fractured and non-fractured patients) were reported as absolute frequencies. In order to assess the distribution of fractured patients across the 7 risk classes and to evaluate the presence of an ordered trend, a Chi-squared test followed by a Cochran-Armitage trend test were performed. Subsequently, to compare the likelihood of experiencing a fragility fracture for individuals classified in a specific risk class *vs*. those classified in all the remaining classes, class-specific unadjusted logistic regression models were used. For each risk class, a separate unadjusted logistic regression model was fitted by comparing patients assigned to that specific class with patients assigned to all the remaining classes combined. Therefore, class-specific odds ratios (ORs) should be interpreted as the relative odds of fracture associated with belonging to a given class compared with belonging to any other class. Finally, to evaluate the contribution of each risk class to the overall frequency distribution between fractured and non-fractured individuals, standardized residuals from the Chi-squared test were examined.

In all the statistical tests listed, a p-value < 0.05 was considered statistically significant.

All statistical analyses were conducted using MedCalc® (version 19, MedCalc Software, Mariakerke, Belgium), and MATLAB® (MathWorks, Natick, MA, USA) software.

Results

Patient Characteristics T-Score and Fragility Score in Fracture Occurrence Prediction

A total of 4443 female patients was recruited. During the follow-up period, 426 participants (9.6%) had incomplete follow-up data or were lost to follow-up, leaving 4017 women who completed the study. At total of 5458 scans was obtained from this cohort of patients and was used for the subsequent analysis. Among these, 2744 were scans of the lumbar spine, whereas 2714 were scans of the femoral neck. During the follow-up period, 418 scans performed at the lumbar spine (15.2%) and 275 scans performed at the femoral neck (10.1%) were associated to patients who experienced MOF or hip fragility fractures, respectively. Specifically, $n = 1441$ underwent REMS examinations of both anatomical sites, $n = 1303$ underwent REMS examination at lumbar spine only, and $n = 1273$ underwent REMS examination at femoral neck only.

Table 2 summarizes the general baseline characteristics of the participants and the bone-related statistics of examinations grouped by anatomical site of the scan (lumbar spine

Table 1 Combination of T-score (normal, osteopenic and osteoporotic) and fragility score (normal, reduced and decreased) classes and the generated specific risk classes (R1-R7)

		T-score classification						
		Normal		Osteopenic			Osteoporotic	
		≥ −0.7	]− 1.0;− 0.7[	]− 1.3; − 1.0[	[− 2.2; − 1.3[	[− 2.5; − 2.2[	[− 2.8; − 2.5[	< − 2.8
Fragility Score classification	Normal	R1	R2	R2	R3	R4	R4	R5
	Reduced	R2	R3	R3	R4	R5	R5	R6
	Decreased	R3	R4	R4	R5	R6	R6	R7

Table 2 General baseline characteristics of the participants and scan characteristics of the patients assessed based on anatomical location (lumbar spine or femoral neck). The reported values are expressed as mean ± standard deviation (SD). An independent non-parametric Mann–Whitney test was used to compare the groups of patients without and with fragility fractures occurred during the follow-up period. MOF: major osteoporotic fractures

		Parameter		
General population		Age (years)		65.5 ± 10.1
		Weight (kg)		63.7 ± 10.3
		Height (cm)		158.4 ± 6.9
		BMI (kg/m^2)		25.4 ± 4.0
Anatomical site	Parameter	Patients without fragility fracture	Patients with fragility fracture (MOF/hip)	*p* value
Lumbar spine	N° of scans	2326	418	
	REMS T-score	− 2.1 ± 1.1	− 2.5 ± 1.1	< 0.0001
	Fragility score	42.5 ± 16.9	54.0 ± 17.3	< 0.0001
Femoral neck	N° of scans	2439	275	
	REMS T-score	− 2.0 ± 1.0	− 2.4 ± 0.9	< 0.0001
	Fragility score	36.6 ± 16.9	49.2 ± 15.9	< 0.0001

or femoral neck). As expected, the mean REMS T-score values were significantly lower, and the mean FS measured were significantly higher in scans from patients who experienced a fragility fracture during the follow-up compared to those who did not, regardless the anatomical site of the scan.

ROC Curve Analysis

A De Long test was performed to compare the differences in the AUC between a model including T-score + Fragility Score and a model including T-score alone at both anatomical sites (lumbar spine and femoral neck). The results showed a statistically significant difference ($p<0.001$) indicating the Fragility Score provides incremental discriminative information beyond the REMS-derived T-score.

Fracture Risk Analysis

In Fig. 2, for women analysed at lumbar spine and femoral neck, the percentage of scans associated to patients who experienced an incident fracture during follow-up is highlighted as a proportion of the total number of scans in each risk class (R1-R7).

The Chi-squared test showed a statistically significant *p* value ($p<0.0001$) for both lumbar spine and femoral neck scans, indicating that the distribution of the fractured patients differed across the 7 risk classes. Likewise, the Cochran-Armitage trend test showed high values at lumbar spine ($Z = 47.83$; $p<0.0001$) and femoral neck ($Z = 75.41$; $p<0.0001$), demonstrating a strong positive monotonic trend in both anatomical sites. Furthermore, Class-specific binary logistic regression analyses confirmed this pattern, revealing progressively increasing odds ratios from R1 to R7 for both lumbar spine (Table 3) and femoral neck (Table 3). At lumbar spine, the analysis of class-specific OR revealed a clear and progressive gradient of fracture risk across the seven risk classes. The low-risk classes (R1–R2), in fact, showed a significantly reduced risk compared with the remaining population collocated in the other risk classes, with OR ranging from 0.35 to 0.58. Conversely, the high-risk classes (R5–R7) exhibited a progressively increased risk, with OR moving from 1.04 (R5) to 1.91 (R7). When examining the analysis at femoral neck, similar results were found: while low-risk classes (R1–R3) were associated to a significantly reduced risk (OR ranging from 0.20 to 0.57), the high-risk classes (R5–R7) were associated to a progressively increased odds of fracture, with OR ranging from 1.41 (R5) to 2.26 (R7).

Standardized residuals were used to identify which risk classes contributed most to the overall Chi-squared

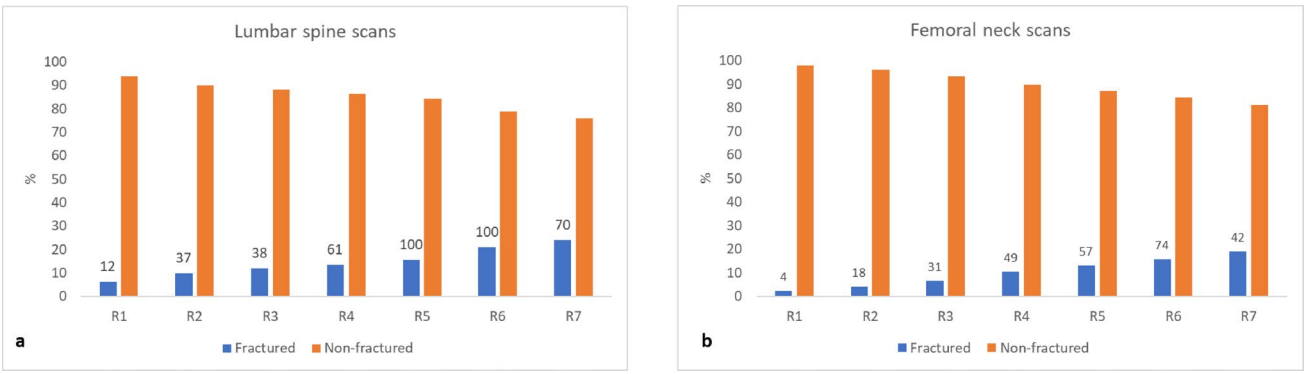

Fig. 2 Frequency charts depicting the percentage of patients without and with an incident fracture during follow-up (in orange and blue, respectively) for each risk class across the two patient cohorts are presented as follows: women analysed for lumbar spine (**a**), women analysed for femoral neck (**b**)

Table 3 For lumbar spine and femoral neck scans the odds ratio (OR) with the corresponding 95% confidence interval (CI) and *p* value were reported within each risk class (from R1 to R7)

Lumbar spine scans				Femoral neck scans		
Risk class	Odds ratio	95% CI	*p* value	Odds ratio	95% CI	*p* value
R1	0.35	0.19 to 0.63	< 0.001	0.20	0.074 to 0.545	< 0.01
R2	0.58	0.40 to 0.82	< 0.01	0.32	0.20 to 0.52	< 0.0001
R3	0.73	0.51 to 1.04	> 0.05	0.57	0.39 to 0.843	< 0.01
R4	0.85	0.63 to 1.14	> 0.05	1.01	0.73 to 1.4	> 0.05
R5	1.04	0.81 to 1.32	> 0.05	1.41	1.03 to 1.921	< 0.05
R6	1.64	1.27 to 2.10	< 0.0001	1.89	1.4 to 2.5	< 0.0001
R7	1.91	1.42 to 2.55	< 0.0001	2.26	1.6 to 3.2	< 0.0001

Table 4 Standardized residuals from the Chi-squared frequency distribution at lumbar spine and femoral neck. Residual values greater than + 2 indicate more fractured or non-fractured patients than expected (*i.e.,* under the assumption that no differences exist among the seven risk classes), whereas values lower than − 2 indicate fewer than expected (*i.e.,* under the assumption that no differences exist among the seven risk classes). Residuals between − 2 and + 2 fall within the range of expected frequencies (*i.e.,* under the assumption that no differences exist among the seven risk classes)

Lumbar spine			Femoral neck	
Risk class	Standardized residual (fractured patients)	Standardized residuals (non-fractured patients)	Standardized residual (fractured patients)	Standardized residuals (non-fractured patients)
R1	− 3.23	1.37	− 3.20	1.08
R2	− 2.63	1.11	− 4.14	1.39
R3	− 1.50	0.64	− 2.46	0.83
R4	− 0.93	0.39	0.05	− 0.02
R5	0.24	− 0.10	1.87	− 0.63
R6	3.25	− 1.38	3.78	− 1.27
R7	3.83	− 1.62	4.11	− 1.38

association. Negative residuals below −2 indicate fewer fractured patients than expected under the hypothesis of no association between risk class and fracture occurrence, whereas positive residuals above +2 indicate more fractured patients than expected. At the lumbar spine (Table 4), low-risk classes R1 and R2 showed less fractured patients than expected, as indicated by strongly negative residuals (−3.23 and − 2.63), whereas high-risk classes R6 and R7 showed an excess of fractures (with respect to the assumption that no differences exist among the seven risk classes). A similar pattern was observed at the femoral neck (Table 4), with low-risk classes (R1-R3) showing less fractured patients (residuals ranging from −3.20 to −2.46) than expected and high-risk classes (R6-R7) showing more fractured patients than expected. These patterns support the ability of the REMS-derived stratification to separate lower- and higher-risk groups.

Discussion

Currently, the most used clinical technologies are not able to provide a comprehensive diagnosis of osteoporosis that might predict a reliable occurrence of fragility fractures in a clinically sustainable framework [9, 30]. Bone mineral density, in fact, still represents the main driver of treatment decision, and it is generally used to identify people with osteoporosis or prone to be osteoporotic. Nevertheless, several studies have demonstrated that BMD alone does not reliably predict whether a patient will develop a fragility fracture [31]. As a direct consequence, a high percentage of patients remain underdiagnosed, receiving a treatment only when the clinical outcome is already compromised [32]. In order to encompass the multifactorial nature of osteoporosis, different predictive tools have been developed, including, but not limited to the FRAX questionnaire [33], Qfracture algorithm [34], Garvan Fracture-Risk calculator (Garvan) [35] and Derived Fracture Risk Assessment (DeFRA) [36]. This strategy has certainly improved the ability to diagnose osteoporosis and to identify individuals who are at higher risk of fragility fracture, but an evident gap still exists [30, 37]. A significative amount of fragility fractures, in fact, still occurs in individuals who are not classified at high-risk; furthermore, none of these tools include information about bone quality and skeletal fragility, which play a crucial role in fracture risk. For this, the interest for a new technology that can provide an accurate diagnosis without further complications has increasingly grown over the past few years. Therefore, the present study aimed to introduce a novel composite stratification risk analysis

comprising quantitative information from both bone mass (REMS T-score) and bone quality (REMS FS) to generate an integrated and reliable measure of fracture susceptibility. To date, this is the first study evaluating the REMS fracture risk stratification as a parameter for prediction of fragility fractures.

This prospective observational multicentre analysis included data from female patients who underwent REMS scans at lumbar spine and/or femoral neck and who were monitored over a 5-year period for the occurrence of fragility fractures. At the time of recruitment, patients were examined, the REMS parameter were annotated (T-score, FS and risk stratification derived by the combination of the two REMS parameters) and close monitoring (every 6 months in-person or by telephonic interview) for the occurrence of any fragility fracture was conducted. A recent study performed by Pisani et al. [16] demonstrated the higher performance of single parameters (REMS FS and T-score) in predicting fragility fractured compared to DXA T-score, which is currently recognized as a clinically available indicator to predict fracture risk. In addition, FS consistently demonstrated superior predictive capacity compared with both REMS and DXA T-score values, with an ability to identify above the cut-off a nine-fold and a six-fold higher risk to develop a MOF or hip fracture, respectively [17]. These findings further support the hypothesis that BMD alone has a limited fracture prediction ability, emphasizing the need for reliable parameters that estimate fracture risk independently from or in addition to BMD [38, 39].

The stronger association of FS with fracture outcomes should be interpreted as a confirmation that BMD and FS provide complementary information on different domains of skeletal fragility: the former reflects bone mass, whereas the latter reflects bone-quality-related features. The rationale of the present stratification approach is therefore to integrate both of them into a clinically interpretable classification of fracture risk.

Based on this rationale, the combination of REMS T-score and FS was used to define 7 risk classes (here named from R1 to R7), each associated to a progressively increasing probability of fracture. Therefore, based on this assumption, at the time of scan examination, patients were stratified twice across 7 different risk classes (once for lumbar spine and once for femoral neck examination) according to the measured values of REMS T-score and FS. Then, according to the information reported about MOF or hip fractures during the 5 years follow-up period, the percentages of fractured patients across each risk class were annotated. The statistical analysis confirmed the clinical relevance of the proposed risk classification at both anatomical sites. The Chi-squared test demonstrated that the distribution of fractured patients differed significantly across the seven risk classes, indicating a non-random association between class assignment and fracture occurrence. The Cochran–Armitage trend test further showed a highly significant and progressive increase in fracture prevalence from the lowest (R1) to the highest (R7) risk class for both the considered anatomical sites.

From a clinical perspective, the class-specific odds ratios clearly identified two opposite risk profiles: patients belonging to the low-risk classes (R1–R2/R3) exhibited a significantly reduced odds of fracture, whereas those in the high-risk classes (R6–R7) showed a markedly increased fracture risk. This finding supports the ability of the classification system to effectively stratify patients according to clinically meaningful levels of fracture susceptibility. For example, patients classified as R1 (a low-risk class) showed an OR = 0.35 at lumbar spine and an OR = 0.20 at femoral neck, which correspond to a 65% and 80% reduced odds of developing a fragility fracture compared with patients in all other risk classes. On the other hand, patients classified as R6 (a high-risk class) showed an OR = 1.64 at lumbar spine and OR = 1.89 at femoral neck, which correspond to a 64% and 89% increased odds of developing a fragility fracture compared with patients in all other risk classes. The analysis of standardized residuals provided additional insight into the contribution of each class to this overall pattern. Low-risk classes consistently showed fewer fractures than expected under the hypothesis of no association between risk class and fracture occurrence, whereas high-risk classes were characterized by an excess of fractured patients. Intermediate classes are expected to show less marked separation because fracture risk is biologically continuous. This is consistent with the absence of a clear-cut threshold between normal and pathological skeletal conditions [40, 41].

Taken together, these findings support REMS 5 years fracture risk classification. The strong agreement between global association tests, trend analysis, odds ratio estimates and residual diagnostics highlights the internal consistency of this stratification and supports its potential clinical relevance for identifying individuals at progressively increasing risk of fracture. Compared with questionnaire-based fracture risk calculators (e.g., FRAX, QFracture, Garvan, DeFRA, etc.), the REMS approach provides an automatic stratification based on skeletal parameters acquired during a single echographic examination. Importantly, it focuses on a shorter time horizon (5 years with respect to the 10 years of the current standards), which may be clinically relevant for near-term prediction of fracture risk and treatment prioritization.

The present study has some limitations. First, the cohort was composed of women referred for densitometric assessment in clinical settings, and therefore may not be representative of the general population. This may partly explain the

relatively high fracture rates observed during the follow-up. Second, only female patients were included in the analysis. Although women (especially post-menopausal women) represent the group at highest risk for osteoporotic fragility fractures, the incidence of fragility fractures in men is increasing as well, with mortality and disability rates even higher than in women [42], therefore deserving inclusion in studies dedicated to the evaluation of fracture risk prediction tools. Third, REMS examinations were not paired with contemporaneous DXA scans at the same anatomical sites, preventing a direct head-to-head comparison with DXA-based fracture risk assessment. Fourth, the absence of measurements at both sites in all participants prevents a complete within-patient comparison between lumbar and femoral risk classification. Fifth, detailed information on anti-osteoporotic treatment at baseline and during follow-up was not collected. Finally, the study did not include a formal comparative analysis against established fracture risk calculators such as FRAX, DeFRA, QFracture or Garvan.

Since osteoporosis is becoming prevalent also in middle-income countries, future studies will be specifically dedicated to populations of those areas [43].

Conclusions

Osteoporosis and related fragility fractures show a multifactorial nature that cannot be fully captured by current predictive tools. This 5 years prospective longitudinal multicentre study supports the clinical consistency of a REMS composite approach for fracture risk stratification based on the integration of REMS BMD T-score and Fragility Score. The seven predefined classes R1-R7 showed a progressive association with incident fractures, with lumbar spine-derived classes evaluated against major osteoporotic fractures and femoral neck-derived classes evaluated against hip fractures.

These findings support the role of REMS in 5 years estimation of fracture risk by combining information related to bone quantity and bone quality within a single echographic examination.

Author Contributions Conceptualization: CC, MLB, SG; Methodology: CC, MLB, SG; Formal analysis and investigation: CC, MLB, GG, AG, MM, SG; Writing—original draft and preparation: CC, MLB, SG; Writing—review and editing: CC, MLB, GG, AG, MM, SG; Supervision: CC, MLB, SG.

Funding No funding was received for conducting this study.

Data Availability All relevant data will be available upon reasonable request and under a dedicated agreement to the corresponding author.

Declarations

Conflict of interest Nothing to declare.

Ethical Approval The study protocol was approved by the Ethics Review Board of the participating institutions, and it was conducted in accordance with the ethical standards of the Declaration of Helsinki (1964).

Human and Animal Rights and Informed Consent No animals were used in the study. Informed consent for participation and publication has been obtained from all the participants included in the study.

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

References

1. NIH Consensus Development Panel on Osteoporosis Prevention D and T (2001) osteoporosis prevention, diagnosis, and therapy. JAMA J Am Med Assoc 285:785–795. https://doi.org/10.1001/jama.285.6.785
2. Kanis JA, McCloskey EV, Johansson H et al (2013) European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int 24:23–57. https://doi.org/10.1007/s00198-012-2074-y
3. Salari N, Ghasemi H, Mohammadi L et al (2021) The global prevalence of osteoporosis in the world: a comprehensive systematic review and meta-analysis. J Orthop Surg Res 16:609. https://doi.org/10.1186/s13018-021-02772-0
4. Compston JE, McClung MR, Leslie WD (2019) Osteoporosis. Lancet 393:364–376. https://doi.org/10.1016/S0140-6736(18)32112-3
5. Sornay-Rendu E, Duboeuf F, Chapurlat RD (2024) Postmenopausal women with normal BMD who have fractures have deteriorated bone microarchitecture: a prospective analysis from The OFELY study. Bone 182:117072. https://doi.org/10.1016/j.bone.2024.117072
6. Kanis JA (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. Osteoporos Int 4:368–381. https://doi.org/10.1007/BF01622200
7. Link TM (2012) Osteoporosis imaging: state of the art and advanced imaging. Radiology 263:3–17. https://doi.org/10.1148/radiol.12110462
8. Docaj A, Carriero A (2024) Bone health: quality versus quantity. J Pediatr Orthop Soc North Am 7:100054. https://doi.org/10.1016/j.jposna.2024.100054
9. Diez-Perez A, Brandi ML, Al-Daghri N et al (2019) Radiofrequency echographic multi-spectrometry for the in-vivo

assessment of bone strength: state of the art—outcomes of an expert consensus meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). Aging Clin Exp Res 31:1375–1389. https://doi.org/10.1007/s40520-019-01294-4
10. Di Paola M, Gatti D, Viapiana O et al (2019) Radiofrequency echographic multispectrometry compared with dual X-ray absorptiometry for osteoporosis diagnosis on lumbar spine and femoral neck. Osteoporos Int 30:391–402. https://doi.org/10.1007/s00198-018-4686-3
11. Cortet B, Dennison E, Diez-Perez A et al (2021) Radiofrequency Echographic Multi Spectrometry (REMS) for the diagnosis of osteoporosis in a European multicenter clinical context. Bone 143:115786. https://doi.org/10.1016/j.bone.2020.115786
12. Tomai Pitinca MD, Fortini P, Gonnelli S, Caffarelli C (2021) Could radiofrequency echographic multi-spectrometry (<scp>REMS</scp>) overcome the limitations of <scp>BMD</scp> by <scp>DXA</scp> related to artifacts? A series of 3 cases. J Ultrasound Med 40:2773–2777. https://doi.org/10.1002/jum.15665
13. Caffarelli C, Tomai Pitinca MD, Al Refaie A et al (2022) Could radiofrequency echographic multispectrometry (REMS) overcome the overestimation in BMD by dual-energy X-ray absorptiometry (DXA) at the lumbar spine? BMC Musculoskelet Disord 23:469. https://doi.org/10.1186/s12891-022-05430-6
14. Semeraro A, Chialà A, Carafa A et al (2026) Very short-term monitoring of Romosozumab longitudinal effects in a cohort of postmenopausal women by means of radiofrequency echographic multi-spectrometry (REMS) technology. Aging Clin Exp Res. https://doi.org/10.1007/s40520-026-03391-7
15. Caffarelli C, Forcignanò R, Muratore M, Gonnelli S (2026) Longitudinal monitoring of denosumab-associated bone changes by REMS in postmenopausal women with ER-positive breast cancer receiving aromatase inhibitors. Aging Clin Exp Res. https://doi.org/10.1007/s40520-026-03402-7
16. Pisani P, Conversano F, Muratore M et al (2023) Fragility score: a REMS-based indicator for the prediction of incident fragility fractures at 5 years. Aging Clin Exp Res. https://doi.org/10.1007/s40520-023-02358-2
17. Al Refaie A, Mondillo C, Cavati G et al (2025) Utility of REMS-derived fragility score and trabecular bone score in evaluating bone health in type 2 diabetes mellitus. Diagnostics 15:2877. https://doi.org/10.3390/diagnostics15222877
18. Bischoff E, Vladeva S, Kirilov N, Bischoff F (2026) Fragility score in radiographic axial spondyloarthritis assessed with radiofrequency echographic multi-spectrometry (REMS). Life 16:1121. https://doi.org/10.3390/life16071121
19. Fuggle NR, Reginster J-Y, Al-Daghri N et al (2024) Radiofrequency echographic multi spectrometry (REMS) in the diagnosis and management of osteoporosis: state of the art. Aging Clin Exp Res 36:135. https://doi.org/10.1007/s40520-024-02784-w
20. Ye C, Ebeling P, Kline G (2025) Osteoporosis. Lancet 406:2003–2016. https://doi.org/10.1016/S0140-6736(25)01385-6
21. Seeman E, Delmas PD (2006) Bone quality—the material and structural basis of bone strength and fragility. N Engl J Med 354:2250–2261. https://doi.org/10.1056/NEJMra053077
22. Rubin KH, Friis-Holmberg T, Hermann AP et al (2013) Risk assessment tools to identify women with increased risk of osteoporotic fracture: complexity or simplicity? A systematic review. J Bone Miner Res 28:1701–1717. https://doi.org/10.1002/jbmr.1956
23. Beaudoin C, Moore L, Gagné M et al (2019) Performance of predictive tools to identify individuals at risk of non-traumatic fracture: a systematic review, meta-analysis, and meta-regression. Osteoporos Int 30:721–740. https://doi.org/10.1007/s00198-019-04919-6
24. El-Hajj Fuleihan G, Chakhtoura M, Cauley JA, Chamoun N (2017) Worldwide fracture prediction. J Clin Densitom 20:397–424. https://doi.org/10.1016/j.jocd.2017.06.008
25. Conversano F, Franchini R, Greco A et al (2015) A novel ultrasound methodology for estimating spine mineral density. Ultrasound Med Biol 41:281–300. https://doi.org/10.1016/j.ultrasmedbio.2014.08.017
26. Casciaro S, Peccarisi M, Pisani P et al (2016) An advanced quantitative echosound methodology for femoral neck densitometry. Ultrasound Med Biol 42:1337–1356. https://doi.org/10.1016/j.ultrasmedbio.2016.01.024
27. Pisani P, Greco A, Conversano F et al (2017) A quantitative ultrasound approach to estimate bone fragility: a first comparison with dual X-ray absorptiometry. Measurement 101:243–249. https://doi.org/10.1016/j.measurement.2016.07.033
28. Greco A, Pisani P, Conversano F et al (2017) Ultrasound fragility score: an innovative approach for the assessment of bone fragility. Measurement 101:236–242. https://doi.org/10.1016/j.measurement.2016.01.033
29. De Long ER, DLDM, C-PD (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
30. de Oliveira MA, Moraes R, Castanha EB et al (2022) Osteoporosis screening: applied methods and technological trends. Med Eng Phys 108:103887. https://doi.org/10.1016/j.medengphy.2022.103887
31. Aspray TJ (2015) Fragility fracture: recent developments in risk assessment. Ther Adv Musculoskelet Dis 7:17–25. https://doi.org/10.1177/1759720X14564562
32. McClung MR (2005) The relationship between bone mineral density and fracture risk. Curr Osteoporos Rep 3:57–63. https://doi.org/10.1007/s11914-005-0005-y
33. Kanis JA, Johnell O, Oden A et al (2008) FRAX™ and the assessment of fracture probability in men and women from the UK. Osteoporos Int 19:385–397. https://doi.org/10.1007/s00198-007-0543-5
34. Hippisley-Cox J, Coupland C (2009) Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. BMJ 339:b4229. https://doi.org/10.1136/bmj.b4229
35. Nguyen ND, Frost SA, Center JR et al (2008) Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks. Osteoporos Int 19:1431–1444. https://doi.org/10.1007/s00198-008-0588-0
36. Adami G, Biffi A, Porcu G et al (2023) A systematic review on the performance of fracture risk assessment tools: FRAX, DeFRA, FRA-HS. J Endocrinol Invest 46:2287–2297. https://doi.org/10.1007/s40618-023-02082-8
37. Fuggle NR, Curtis EM, Ward KA et al (2019) Fracture prediction, imaging and screening in osteoporosis. Nat Rev Endocrinol 15:535–547. https://doi.org/10.1038/s41574-019-0220-8
38. Pisani P, Renna MD, Conversano F et al (2016) Major osteoporotic fragility fractures: risk factor updates and societal impact. World J Orthop 7:171. https://doi.org/10.5312/wjo.v7.i3.171
39. Sun X, Chen Y, Gao Y et al (2022) Prediction models for osteoporotic fractures risk: a systematic review and critical appraisal. Aging Dis 13:1215. https://doi.org/10.14336/AD.2021.1206
40. Baek Y-H, Cho SW, Jeong HE et al (2021) 10-Year fracture risk in postmenopausal women with osteopenia and osteoporosis in South Korea. Endocrinol Metab 36:1178–1188. https://doi.org/10.3803/EnM.2021.1215
41. Reid IR, McClung MR (2024) Osteopenia: a key target for fracture prevention. Lancet Diabetes Endocrinol 12:856–864. https://doi.org/10.1016/S2213-8587(24)00225-0
42. Haentjens P, Magaziner J, Colón-Emeric CS et al (2010) Meta-analysis: excess mortality after hip fracture among older women

and men. Ann Intern Med 152:380–390. https://doi.org/10.7326/0003-4819-152-6-201003160-00008
43. Maaloui G, Gannagè-Yared MH, Ezzedine J et al (2007) Middle East and North Africa consensus on osteoporosis. J Musculoskelet Neuronal Interact 2:131–143

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