Computers in Biology and Medicine 198 (2025) 111214

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Identification of statistical critical area to discriminate osteoporotic hip fractures in women

Nicole Morando [a,b], Carlos Ruiz Wills [a], Simone Tassani [a,*]

[a] BCN MedTech, Universitat Pompeu Fabra, Department of Engineering, C/ Roc Boronat, 138, 08018, Barcelona, Spain
[b] Politecnico di Torino, Torino, Italy

ARTICLE INFO

Keywords:
Hip fracture
Fracture discrimination
Random Field Theory (RFT)
Statistical Parametric Map (SPM)
3D finite element analysis

ABSTRACT

Osteoporotic hip fracture is a worldwide health problem, but its understanding is still out of reach. Finite Element models are often implemented to study the phenomenon, but the analysis of simulation's results is in discussion. The simple identification of maximum stress or strain might be misleading and only partially related to the development of the fracture.

The aim of the present study is to identify regions with statistically significant differences between fractured and control patients using a rigorous methodology based on Random Field Theory.

A cohort of 90 osteoporotic female subjects was used: 45 fractured and 45 controls. 3D FE models were built from Dual-energy X-ray Absorptiometry (DXA) acquisitions. The cohort included both neck and trochanteric fractures.

Areas with statistical differences were selected through Random Field Theory. The suitability of the selected elements for the discrimination of a fracture event was validated through the area under the curve (AUC) methods, and binary logistic regression with leave-one-out validation.

The FE models elements identified in such a way were below 7 % of the total elements. Major Principal Stress in the selected elements showed an AUC up to 0.95. Patients were classified with an accuracy of up to 84.2 %.

The methodology explored focused the analysis on specific points. This approach not only allowed for reaching a relevant classification power, but also suggested a specific bone remodeling process, including reduction of variability and interacting behavior between cortical and trabecular bone.

In conclusion, a novel approach to finite element model analysis is presented, showing good classification power and extraction of information about bone remodeling in osteoporotic subjects.

1. Introduction

Osteoporotic hip fracture represents a great social and economic burden in developed countries [1]. The fractures are often associated with physical disability, reduction of quality of life, and increased mortality. Due to changes in population demographics, the number of men and women with osteoporosis is rising. Especially postmenopausal women are at a significantly higher risk of developing osteoporosis due to hormonal changes that accelerate bone loss. In the EU the number passed from 27.5 million in 2010 to 32 million in 2019, and it is still increasing [2,3]. In addition, the incidence of osteoporotic hip fractures has an evident impact on health care costs for pre- and post-fracture treatments, resulting in many days spent in the hospital. In the same years, the economic burden of the incident and prior fragility fractures rose from 37 billion euros in 2010 [3] to 57 billion euros in 2019 [2].

Because the incidence of osteoporotic hip fracture is closely related to the growth of the elderly population, the identification of rigorous methodology for early diagnosis and identification of mechanisms of the pathology constitutes an important clinical aim to properly select the appropriate treatment for every patient. Therefore, this study aims to present a novel methodology for the identification of the statistically critical areas associated with fracture discrimination.

1.1. Diagnostic tools for assessment of osteoporosis and fracture risk

The onset of osteoporosis is closely related to bone density loss, causing the weakening of the bone, and making it more susceptible to fracture risk. Several techniques are commonly used by clinicians to

* Corresponding author.
E-mail address: simone.tassani@upf.edu (S. Tassani).

https://doi.org/10.1016/j.compbiomed.2025.111214
Received 26 November 2024; Received in revised form 11 October 2025; Accepted 13 October 2025
Available online 17 October 2025

measure the amount of mineral content (BMC) and density (BMD) in the bone [4].

Dual-energy X-ray absorptiometry (DXA) is the most widely used technique to detect osteoporosis because it is easy, quick, and minimally invasive. The amount of radiation dose used is extremely small, and therefore compatible with screening use and making DXA the gold standard for diagnosing osteoporosis. Nevertheless, DXA measures areal, not volumetric, BMD, introducing limitations in the analysis of the pathology.

By contrast, Quantitative Computed Tomography (QCT) provides a method to assess osteoporosis in relation to a direct volumetric BMD measurement [5]. The three-dimensional information obtainable from QCT allows for the development of patient-specific Finite Element Models (FEM). This combination of methodologies is widely employed in clinical studies to evaluate the influence of factors such as age and gender on the investigated groups [6] and to analyze the effect of drug therapy on the treatment of osteoporosis [7].

However, compared to DXA, the QCT technology uses a higher dose of radiation, is less available for clinicians, and more expensive; therefore, it cannot be used as a routine diagnostic tool.

To overcome this limitation, nowadays it is possible to use statistical shape and appearance models to obtain 3D shape reconstruction and BMD distribution starting from 2D DXA images [8]. This approach also allows the definition of patient-specific FEM for the prediction of osteoporotic fractures. DXA-based FEMs have not been evaluated in clinical studies as deeply as QCT models, but have the potential to enter the clinical routine settings [9].DXA-based FEM brings with it the advantages of a minimally invasive procedure to evaluate not only tissue condition but also biomechanical descriptors of bone fracture obtained from simulation [10].

1.2. Statistical identification of the critical region

Once the FEM is developed and the distribution of stresses is obtained, the issue of data analysis comes into place. A common selection of critical areas is based on the identification of maximum stresses and strains. However, these regions might be only partially related to the development of the fracture [11].

Moreover, by virtue of the local nature of the fracture [12] not all the anatomical volume (proximal femur) takes part in the fracture, and using a mean value would be inappropriate, although surrounding areas, far from the fracture line, might actually be determinant in the development of the fracture. For these reasons, the visual identification of regions of elements does not represent a reliable criterion to discriminate different critical regions between investigated osteoporotic groups.

On the other hand, statistical analysis based on the evaluation of FEM results in each element is not a valid option since the Bonferroni correction required in virtue of the multiple comparison problem comes at the cost of increasing the false negatives and reducing statistical power [13].

Random Field Theory (RFT) [13], along with its topological extension based on Statistical Parametric Mapping (SPM), offers a way to address the multiple comparison problem by accounting for the correlation between nearby elements in the finite element (FE) model, which are spatially related due to the physiological nature of the signal. In other words, the application of RFT takes the multiple comparison problem from discrete to continuous contexts. An example of the application of RFT is reported in the literature, especially concerning gait analysis [14]. The application of RFT can be divided into three different steps:

1. Estimation of smoothness with the effect of blurring data reducing the number of independent observations [13].
2. Use of the smoothness to calculate the Expected Euler characteristic E[EC] at different thresholds. The threshold corresponds to the probability of finding an above-threshold region in a smooth Gaussian field, approximately associated with the probability of family-wise error E[EC] $\approx P^{FWE}$ [13].
3. Use of the EC to calculate a proper threshold, which lets us reject the null hypothesis (equality of the averages between fracture and control) with an error α, specified by the user [13].

At the end of this procedure, a list of probability p-values, one for each threshold-surviving cluster, is obtained.

This work aims to present a novel and rigorous methodology. The main objective is to identify the statistically critical areas associated with fracture discrimination using RFT applied to DXA-based FEM. We applied this methodology over a reference dataset and presented the preliminary results.

2. Materials and method

2.1. Study cohort

The methodology was applied to 90 female subjects, who had taken part in a previous study [10]. All subjects were post-menopausal osteoporotic subjects: 45 had suffered from fracture located in the proximal femur (neck or trochanter) and 45 characterized by the lack of previous fracture were used as control. Therefore, the study is divided into two groups. One for the study of neck fracture (26 fractured and 26 control), and the other for the study of trochanteric fracture (19 fractured and 19 control). Data of fractured patients were obtained from DXA images before the fracture event.

The use of clinical data, including DXA 2D images for the reconstruction of 3D FEM using the software 3D-Shaper® (version 2.6, 3D Shaper Medical, Barcelona, Spain), was evaluated and approved by the ethics committee of the University Hospital Mutua de Terrassa [10].

Three-dimensional distribution of BMD was also obtained from 3D-Shaper® and implemented in the FEM. Both cortical and trabecular bone were considered as linear isotropic elastic materials. The Young modulus was obtained from the volumetric bone mineral density using an empirical relationship as implemented in Ref. [10]. The boundary conditions used for lateral fall simulations were described in the literature [10]. Briefly, the femur distal part was fully constrained, the trochanter was fixed in the direction of the force, and a patient-specific fall force was applied at the femoral head (Fig. 1 - a) and computed as in Equation (1) [15]:

$$F_{fall} = \sqrt{2*g*h_c*k_{tissue}*m} \quad (1)$$

In the formula g is the gravity acceleration field (9.81 m/s^2), h_c is the height of the patient to its center of gravity, m is the mass of the patient and k_{tissue} is the trochanter soft tissue stiffness. The soft tissue stiffness for women was considered 71 N/mm based on the study of Robinovitch et al. [16]. The values, per element, of Major Principal Stress (MPS) and Major Principal Strain (MPE) were used for statistical analysis [20]. The unit of measurement of MPS and MPE are $N/\ mm^2$ (MPa) and mm/mm, respectively.

For all subject-specific FE femur models, for both groups (fractured and controls), MPS and MPE were extracted by trabecular and cortical bone (Fig. 1 - b).

Each subject-specific model consists of the same number and type of elements (126,800 hexahedral elements) and the same number of nodes (132,120). This means that all models have one-to-one correspondence, i.e., element 1 of model 1 is built by the same nodes that element 1 of model 2. Such correspondence makes it possible to consider a unique structural geometry as a reference, where each element will correspond to the value of a specific variable (MPS, MPE) for each osteoporotic patient.

In addition, the elements on which the boundary conditions had been applied were excluded from the analysis (Fig. 1 - c). These modifications to the FE original geometry led to a reduction of the number of elements

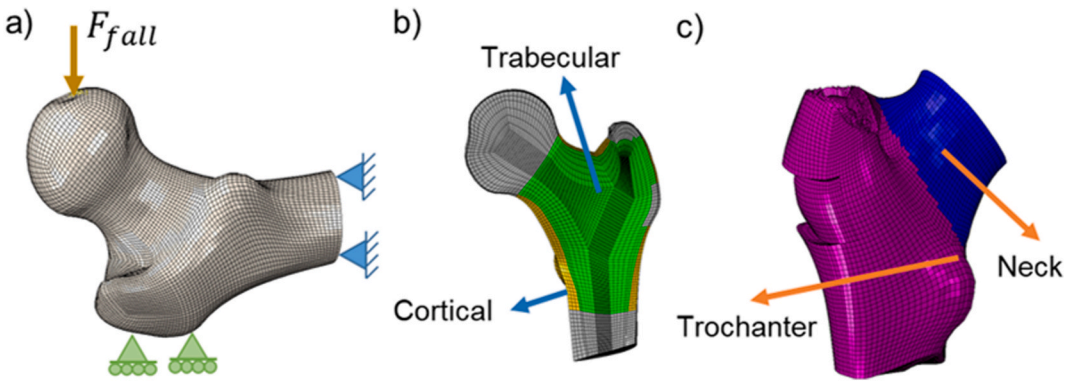

Fig. 1. (a) Boundary conditions applied to FE model. (b) Bone tissue: trabecular and cortical. (c) Zone of analysis. Areas in proximity to boundary conditions were removed to create the zone of analysis.

(trabecular tissue = 68327, cortical tissue = 16353, total = 84680).

2.2. Data organization

The results concerning BMD, MPS, and MPE extracted by simulations represent the input data provided to SPM to conduct statistical analysis between the two recruited groups (fractures and controls). The software Spm1d (version M.0.4.10) [17] was used to obtain one-dimensional SPM.

The software, open source, and implemented in MATLAB by Todd Pataky,[1] requires experimental data described as "1D continua" (1D trajectories) to make statistical inferences under RFT [18]. FEM output 3D data is not organized in a stack of bi-dimensional images, and this makes the application of RFT challenging. This problem can be overcome by the implementation of a 1D trajectory.

2.2.1. Creation of a one-dimensional trajectory

The implementation of a one-dimensional field is obtained by replacing the 3D FE geometry with a 1D ordered sequence of elements. Because each element belonging to the FE femur model consists of 8 nodes, only one node for each element has been chosen for the creation of the path. As a general criterion of choice, for each element, the junction node corresponds to the node in the first position. In the case that different elements have the same junction node, the developed algorithm (implemented in MATLAB) selects another reference node, among the potential 7 remaining, for the second repetition. And so on for any subsequent repetitions. At the end of this procedure, to each element in which the variable of interest (MPS, MPE, and BMD) is calculated, there is a corresponding node in the path. The list of nodes, node coordinates, elements, and variables of interest were extracted from FE simulations, previously implemented in Abaqus 2018 (Dessault Systèmes).

To reduce the risk of spatial discontinuity, the sequence of reference nodes in the path follows a logical order based on the minimum distance between them. The risk of discontinuity, due to different types of tissue, was indeed avoided, achieving two distinct trajectories: one for trabecular tissue and one for cortical tissue.

Starting from the spatial coordinate of each reference node, the algorithm proceeds in steps:

1. Arbitrary choice of the first node of the path. It represents the origin of the one-dimensional trajectory.
2. Calculation of Euclidean distance between the origin node of the path and all reference nodes. The node with the minimum distance from the origin will be the second node of the trajectory.
3. Calculation of Euclidean distance between the second node in the path and all remaining reference nodes. The node with the minimum distance from the second node will be the third node in the trajectory.
4. The algorithm is repeated until the path containing all reference nodes is created.
5. Adding up all the distances between a reference node and its previous node in the path, the total distance from the origin is obtained (Fig. 2).

At the end of the 5th step, each patient for each investigated group will be represented by a one-dimensional ordered sequence of elements, valid for all subjects, being each patient-specific FEM based on the same mesh. A continuum of values of maximum absolute stress and strain, associated with each element of the trajectory, will be created for each patient.

2.2.2. Smoothing

Before conducting tests, the 1D continuum, for each patient, is smoothed using the full width at half-maximum (FWHM). The spatial smoothness essentially increases the signal-to-noise ratio and allows using the RFT for thresholding.

Given the critical role of smoothing in the analysis, 6 different conditions were tested. Starting from the absence of smoothing (original data), the smoothing Gaussian kernel was increased with the step of 5 elements up to a maximum of 25 (about 2.5 cm).

2.3. Statistical analysis of 1D continua: Two-sample t-test

Two-sample (independent) t-tests, performed using "spm1d", were applied to examine the effects of MPS, MPE, and BMD in fractures and control groups, taking into consideration the region of fracture (neck or trochanter) and type of tissue (trabecular and cortical). For a 1D dataset, the goal is to quantify the probability that smooth, random 1D continua would produce a test statistic continuum whose maximum exceeds a particular test statistic value [19].

Test outputs are reported in an SPM. In particular, where SPM{t} is above the critical RFT threshold, one or more clusters of elements are recognized as significant, identifying the critical regions. Each cluster is defined by a specific p-value, equal to or less than alpha (specified by the user at 0.025 since for each kind of fracture, neck, and trochanter, we are testing two variables, MPS and MPE), for which the null hypothesis is rejected, and the cluster of elements will be considered as statistically significant. Otherwise, if the 1D test statistic field SPM{t} does not reach the critical threshold, no supra-threshold clusters will be generated, and the null hypothesis will not be rejected.

[1] https://spm1d.org/.

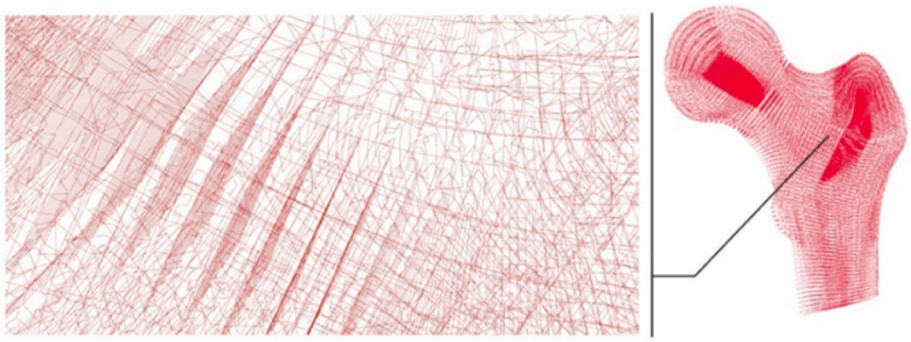

Fig. 2. 1D trajectory of the FE femur model.

2.4. Second-level analysis: Two-way ANOVA with repeated measures

Once the critical region is identified for each parameter and each type of fracture, averages and standard deviations of MPS, MPE, and BMD were computed over the selected FEM elements. A two-way ANOVA with repeated measures was performed to analyze possible statistically significant interactions between factors on the dependent variable. The two factors analyzed were "Group" (between-subjects factor – Control or Fractured) and "Type of tissue" (within-subjects factor – Cortical or Trabecular). Analyses were performed using SPSS (version 23.0; IBM Corp., Armonk, NY, United States).

2.5. Receiver Operating characteristic (ROC) curves and classification

The results of the second-level analysis were finally compared to the analysis performed over the whole neck and trochanteric regions. The capability of discriminating fractures from control subjects was tested by performing a ROC analysis and computing the area under the curve (AUC) for the parameters identified.

The input features were all the significant variables identified in the second-level analysis. Variables representing significant interactions were calculated by multiplying the values computed in the two levels of the within factor (cortical and trabecular tissue) and comparing the results between fracture and control subjects.

Finally, a binary logistic regression with leave-one-out validation was applied to perform patient discrimination. Stepwise feature selection was used to identify the variables producing the best classification in terms of accuracy. This also allowed for the identification of the most relevant variables for fracture identification.

The analysis was implemented in Matlab (R2024a) and repeated for each smoothing level, and the results were compared.

3. Results

3.1. Smoothing selection based on the classification results

The whole process of analysis described in the Material and Methods sessions 2.3 to 2.5 was repeated 6 times to test the effect of the smoothing level over the final classification result. Classification accuracy for neck and trochanter fractures is shown in Fig. 3. Different levels of smoothing varied the accuracy of the classifier from 71 % to 84 % for the classification of trochanter fractures, and from 65 % to 81 % for the classification of neck fractures.

For the sake of clarity, we will here present detailed results only for the smoothing that gives the best classification results for both trochanteric (smoothing 5) and neck (smoothing 20) fractures. Extended results are shown in the supplementary materials.

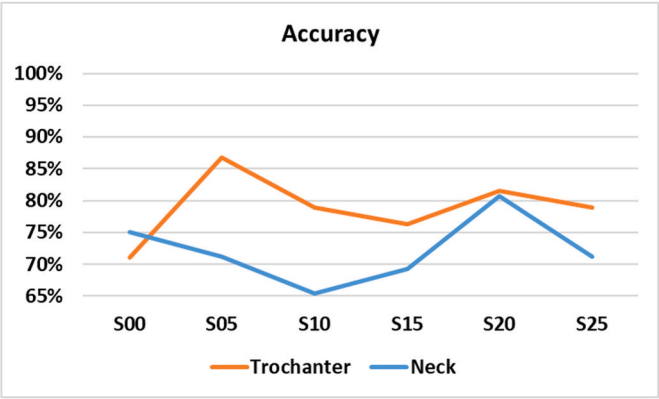

Fig. 3. Classification accuracy is shown for neck and trochanter fractures for the six different levels of smoothing.

3.2. Two sample t-test

For each comparison, different SPMs have been obtained about the different variables of interest (MPS/MPE/BMD), type of tissue (trabecular/cortical), and type of fracture (neck/trochanter). An example in Fig. 4. The graphs show the variation of the dependent variable along the order sequence of elements in the path. The thick black line depicts the test statistic continuum SPM{t}, while the red dashed line is the

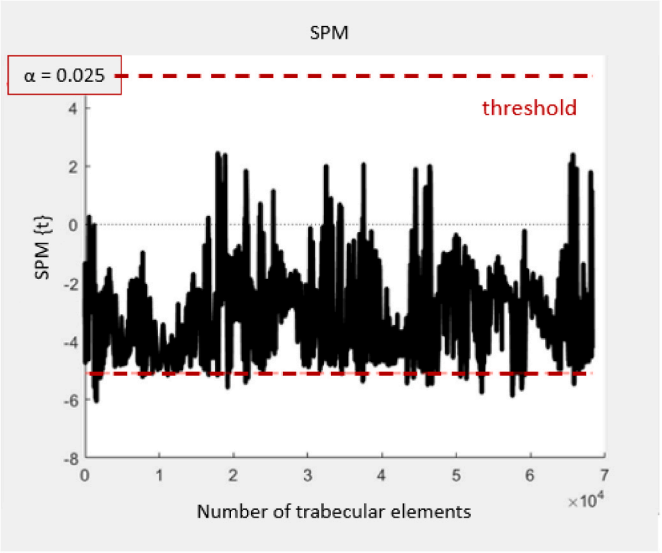

Fig. 4. SPM obtained by two-sample *t*-test between Neck fracture patients and controls concerning MPS variable in trabecular bone.

critical threshold t^* at α value.

Test outputs of each SPM, concerning test statistic computation and statistical inference, are reported in Table 1. The highlighted cells identify the tests in which the null hypothesis is rejected.

The tests, concerning the variable MPS, produced significant clusters in relation to cortical and trabecular tissue, both comparing fractured neck patients with controls and fractured trochanter patients with controls.

By contrast, the tests concerning the variable MPE, produce significant clusters in both tissues only comparing fractured trochanter patients with controls. These results show how both MPS and MPE in the critical zones of fractured subjects are lower compared to controls. The significant elements over which conducting the analysis are indeed a small percentage of the total elements of the model (Table 2). BMD distribution was also tested in both groups and tissues, but no statistically significant differences were observed.

3.3. 3D visualization

The 3D visualization, reported in Fig. 5, provides statistically significant elements, belonging to the trabecular, cortical, and both tissues, respectively. Only the 3D visualizations of the proximal femur that result from tests that have produced significant clusters are shown. Significant elements are marked in red.

3.4. Second-level analysis

Two-way ANOVA was run over the elements identified as significant by the SPM. Values of stress and strain were averaged in each cluster and were tested together with their standard deviation. The difference between groups was found to be significant by definition since the test was run over the elements identified as critical regions by the SPMs. Differences in stress and strain between trabecular and cortical bone are also straightforward, as the different tissues have distinct load capacities and distributions. Therefore, cortical bone always exhibits higher MPS and lower MPE. This analysis aimed to study the interaction between the two factors and the variability of the results in terms of standard deviation within the critical area.

Interactions between tissue and group were significant in all cases, and always suggested a reduced difference in MPS and MPE, and their standard deviation, in the fractured group compared to the control one (Fig. 6). In neck fractures, the SPM identified only elements with critical MPS. Differences in the standard deviation of MPS values were always significant in both neck and trochanter fractures, between groups, tissues, and interactions, with cortical bone always showing higher variability compared to the trabecular bone, and control subjects presenting higher variability than fractured ones. Differences in the standard deviation of strain values were significant in trochanter fractures, again with cortical bone, and in general control subjects, showing a higher standard deviation. The interaction was also significant for the averaged values of stress in both neck and trochanter fractures. In the latter, average values of strain were also significant. In all analyzed cases $p < 0.0005$.

Table 1
SPM statistical Inference.
SPM {t} statistical inference in relation to compared group (fractured/controls), zone of fracture (neck/trochanter), type of tissue (trabecular/cortical) on variable MPS and MPE. Significant differences ($\alpha<0.025$) between the fracture and control groups are in **bold**. Description of the variables: FWHM – The estimated full-width at half maximum of 1D Gaussian Kernel which produces the same smoothness of observed residuals in the case it is convolved with 1D Gaussian continua; Resels – Resolution elements that represent the total number of independent processes in the continuum; z* – Critical Random Field Theory statistical threshold; n – It defines how many supra-threshold clusters have been generated.

	SPM{t} inference	Neck		Trochanter	
		Trabecular	Cortical	Trabecular	Cortical
MPS	FWHM	51.305	60.5497	20.1654	28.8526
	Resels	1.33E+03	270.059	3.39E+03	685.5450
	z*	**5.094**	**4.6**	**5.6727**	**5.1216**
	n	49	9	288	27
MPE	FWHM	64.4779	67.7822	23.5079	28.1326
	Resels	1.06E+03	241.2433	2.91E+03	581.2476
	z*	5.024	4.565	**5.62**	**5.0643**
	n	0	0	2	15

Table 2
Number and percentage of statistically significant elements.

		MPS		MPE	
		Neck Fracture	Trochanter fracture	Neck Fracture	Trochanter fracture
		Control	Control	Control	Control
Significant elements (n)	Trabecular	502	4747	0	12
	Cortical	95	371	0	118
	total	**597**	**5118**	**0**	**130**
Significant elements (%)	Trabecular	0.73 %	6.95 %	0.00 %	0.01 %
	Cortical	0.58 %	2.27 %	0.00 %	0.72 %
	total	**0.71 %**	**6.04 %**	**0.00 %**	**0.15 %**

3.5. ROC curves and classification

ROC analysis was performed, and the AUC was computed for each significant variable. Interactions were computed as the product of the two tissues' MPS and MPE values. Results are summarized in Table 3.

Binary logistic regression was applied to perform subject discrimination. The trochanter fracture showed a better classification (Table 4). For the best smoothing coefficient, the classification of neck fractures identified only three variables, which were always the same with very similar coefficients in all the training sets: Std. Dev of trabecular MPS, Interaction between the MPSs, and Interaction between the MPS Std. Dev. Between 2 and 4 features, depending on the training set, were required to classify trochanteric fracture, but the three most used were Std. Dev of cortical MPS, trabecular MPS, and Std. Dev of trabecular MPS.

A two-dimensional representation of the classification procedure for both neck and trochanteric fractures is presented in Fig. 7.

4. Discussion

The research question investigated in this study concerns the application of a rigorous methodology to identify statistically different regions between controls and fracture patients. The methodology provides a quantitative and qualitative approach that allows the statistical analysis of distributions obtained by finite element models. Moreover, the proposed approach overcomes the limitation of the multiple comparison problem that would arise from testing all the elements of a FEM. Finally, visual identification of the specific regions involved in fracture discrimination is given. However, it is important to remember that the study was performed over a relatively small cohort of 90 subjects, which restrict the generalizability of the results. Future validation over a larger dataset is necessary.

4.1. Interpretation of statistically significant stress differences

The results obtained, resumed in the tables of test outputs (Table 1), suggest how the dependent variable MPS is dominant compared to the variable MPE, in agreement with Ruiz Wills el al. [10], but also presenting an effect of MPS over a larger anatomical area compared to MPE. It is also observed that both MPS and MPE always present reduced values in the significant regions of the fractured patients. This might seem a

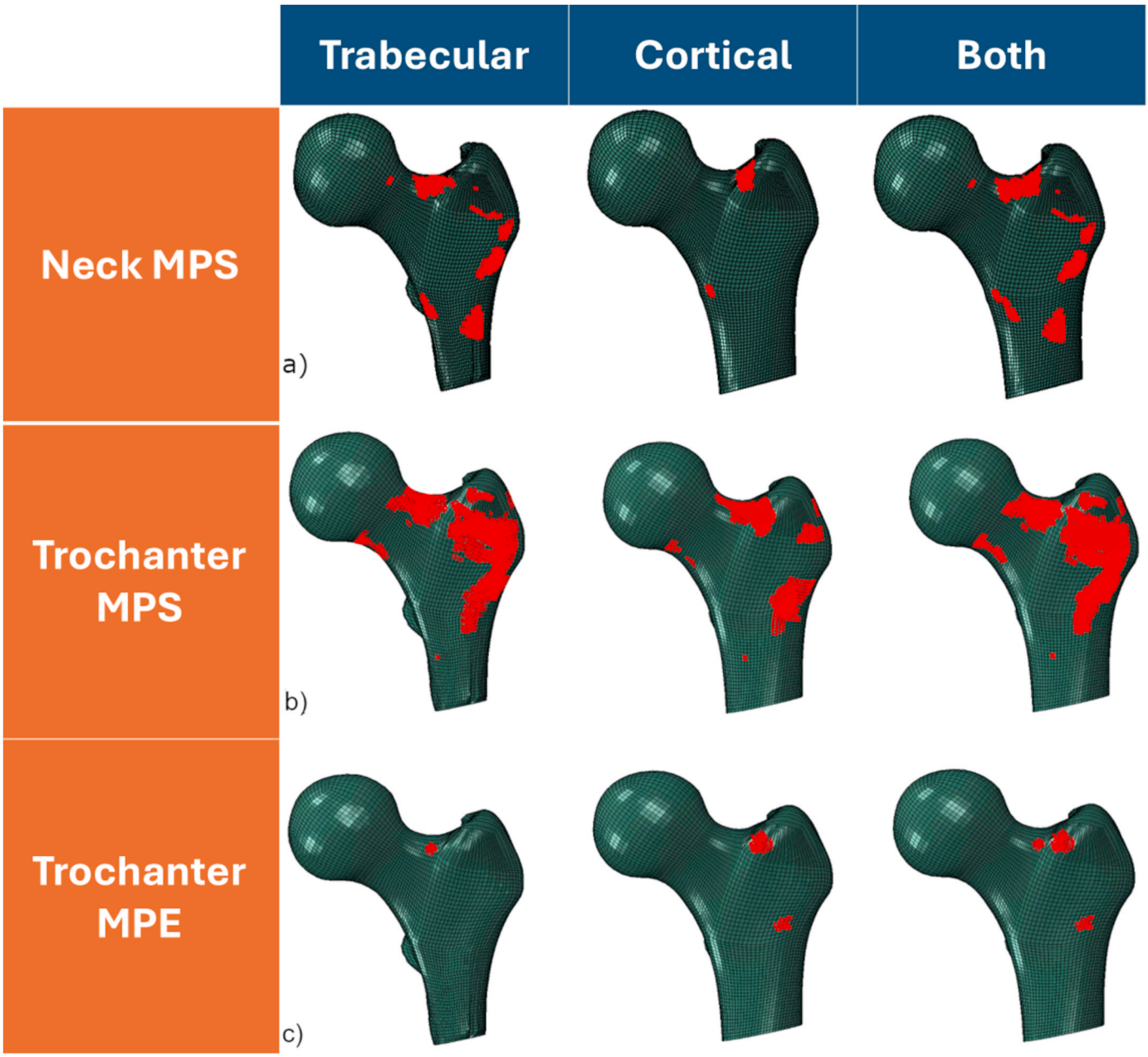

Fig. 5. 3D visualization of statistically significant elements. Critical regions are shown for a) neck fracture in relation to MPS variable; b) trochanter fracture in relation to MPS variable; c) trochanter fracture in relation to MPE variable, concerning trabecular tissue (left), cortical tissue (center), and both tissues (right) respectively.

counterintuitive result since higher MPS is usually related to a higher risk of fracture [20]. However, it might be highlighted that the identified critical clusters *are not necessarily related to the fracture region but are regions where the differences between the groups are statistically significant*. These regions appear to be small compared to the total modeled region, emphasizing the importance of analyzing the local behavior of the bone (Table 2). In the case of a neck fracture, fewer than 1 % of the elements are enough to correctly classify about 81 % of the subjects. It is confirmation that not the whole area examined can be considered significant as a critical region. Although this area includes regions usually subjected to the risk of fractures (neck and intertrochanter), other areas can play a key role. In the comparison between trochanter fracture patients and controls, MPS variable was identified as significant in about 6 % of the elements. Palanca et al. highlighted in their experimental study the importance of local bone distribution in fracture classification [21]. The results, therefore, demonstrate that it is possible and suggested to reduce the analyzed regions to a subset of elements that revealed a statistical difference between the investigated groups. This result is in line with the study of Ruiz et al. [22] where manually-selected reduced risk areas allowed for a better classification of the patients.

In addition, it is observed that the percentage of statistically significant elements belonging to trabecular tissue is higher than those belonging to cortical tissue in relation to MPS variable. However, the opposite situation is presented for MPE variable, i.e., the elements with statistical significance were higher for cortical than for trabecular tissue.

4.2. *Visual patterns and their anatomical relevance*

Three-dimensional visualizations of the critical regions are reported in Fig. 5. The image obtained by comparison between neck fracture

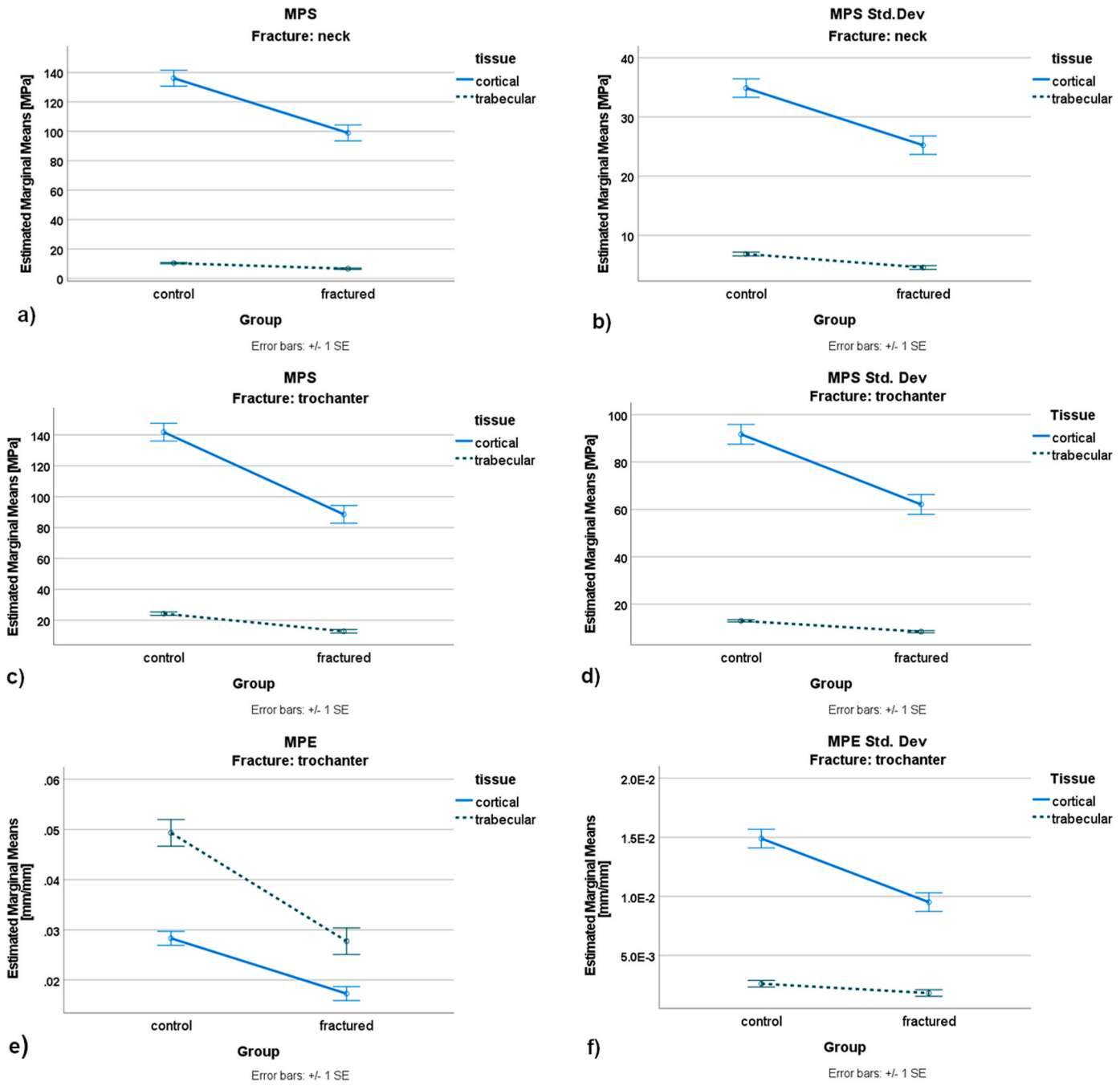

Fig. 6. Interactions between group and tissue are shown. MPS and its standard deviation (respectively left and right columns) are presented for the neck (a,b) and trochanter (c,d) fractures. Interactions of MPE and its standard deviation for trochanter fracture are also shown (e,f).

Table 3
ROC-AUC values by the zone of fracture and type of tissue.

	Neck		Trochanter	
	Trabecular	Cortical	Trabecular	Cortical
MPS	0.88	0.82	0.95	0.92
MPS Std.Dev	0.83	0.80	0.94	0.87
Interaction MPS	0.85		0.94	
Interaction MPS Std.Dev	0.83		0.91	
MPE	–	–	0.90	0.90
MPE Std.Dev	–	–	0.69	0.87
Interaction MPE	–		0.91	
Interaction MPE Std.Dev	–		0.78	

Table 4
Confusion matrix for the classification based on binary logistic regression. The average result of all the final results of the test-set are presented for Neck and Trochanter fractures.

	Observed		Predicted		
			Group		Percentage Correct
			control	fracture	
Neck	Group	control	21	5	80.8
		fracture	5	21	80.8
	Overall Percentage				**80.8**
Trochanter	Group	control	16	3	84.2
		fracture	2	17	89.5
	Overall Percentage				**86.8**

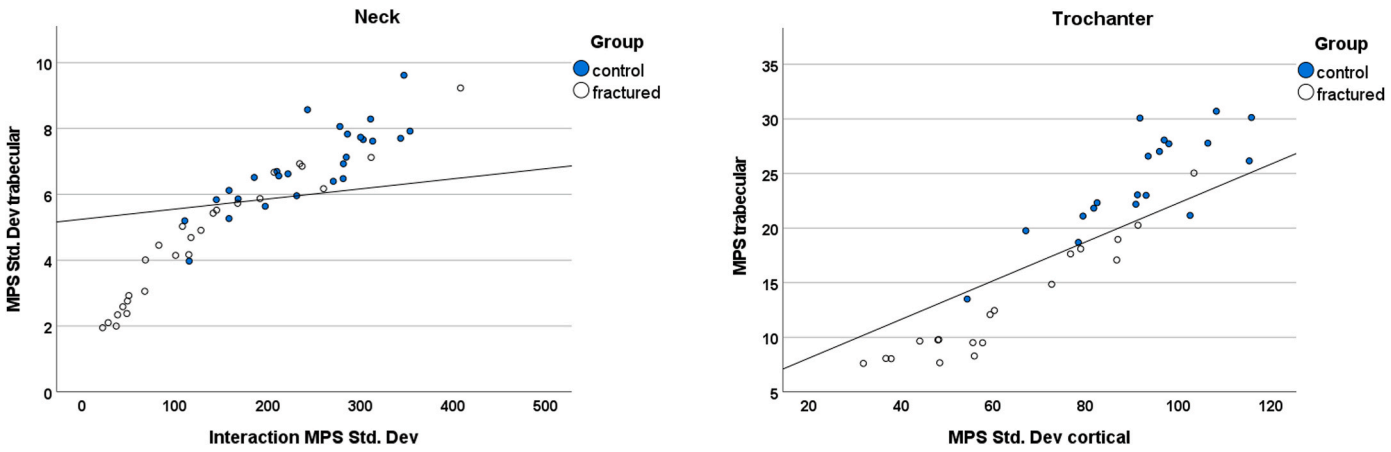

Fig. 7. Two-dimensional representation of the classification procedure. The plotted line is the linear law based on the parameters defined by the Stepwise procedure. Performances of the classification law in separating the fractured from control subjects are shown for both neck (on the left) and trochanter (right) fractures.

patients and controls on MPS shows a considerable number of significant elements identified in the intertrochanter area and also below the greater trochanteric area. It means that the future clinical analysis that must be performed to identify the risk of neck fracture must not necessarily be based on elements belonging only to the neck region. Statistically significant elements are evident in both trabecular and cortical tissue.

A similar situation was also reported for the risk of trochanteric fracture (Fig. 5 b). MPS distribution resulted significantly different, not only in the inter-trochanteric area. The upper part of the neck and the lower part of the proximal femur were involved. In particular, the significant regions of cortical tissue were identified mainly outside of the trochanteric region.

For MPE distribution, only two clusters of elements were identified as significant in the trabecular tissue (Table 1), and they were both in the neck region, although a trochanteric fracture was analyzed. In the cortical tissue, both the neck and trochanter regions showed critical zones, but the neck region was still the bigger one.

4.3. Insights on bone tissue variability and remodeling

The second level analysis allowed us to identify two effects that, to the knowledge of the authors, were never examined before: variability of the results within the tissues and interaction between tissues.

The clusters of elements identified as critical by the SPMs, showed different distributions of values. However, the standard deviation of those distributions was also significantly different (Fig. 6). Reduced variability of the mechanical response in the tissues is an interesting biological insight into the remodeling of bone perpetuated by osteoporosis. This result is presented together with the significant interactions identified in this study. Both average values and their variability decrease in the fracture subjects, but this decrease is not the same for trabecular and cortical bone. MPS values decrease more in cortical bone, both in average and standard deviation, therefore reducing the difference between the two tissues in the fractured subjects.

4.4. Classification implications from ROC and logistic regression

Sixteen out of eighteen of the variables computed in the critical region showed an AUC higher than 0.8, and in the case of the trochanter fracture, 4 variables showed an AUC higher than 0.9 (Table 3). The possibility to compute MPS and MPE, their variability and interaction, only in the critical region presents a high potential for the discrimination of patients at risk of fracture. In particular, the higher AUC values are related to MPS (0.95) and its standard deviation (0.94) in the trabecular bone and the interaction between the two tissues (interaction MPS 0.94, interaction MPS Std.Dev 0.91).

The features selected for the classification of the subjects can help us in understanding what generates the risk of fracture. In the case of neck fracture classification, the leave-one-out validation presents an accuracy of 80.8 % and identifies as best predictor the standard deviation of MPS in trabecular bone, the interaction of MPSs, and the interaction of their standard deviation. Decreased MPS standard deviation, combined with the decreased difference between the variability of MPS in cortical and trabecular bone, is associated with the fractured subjects. Even though the MPS of trabecular bone showed the highest AUC (0.88), the best combination selected by the step-wise feature selection does not include this parameter.

Higher classification accuracy was reached in the identification of subjects presenting trochanter fracture, where it was possible to correctly classify 86.8 % of the cases. Decreased MPS in the trabecular bone, combined with decreased MPS standard deviation in the cortical bone, is associated with fractured subjects.

It has been reported that FE modelling provides higher discriminative and predictive power than aBMD as a singular metric for hip fracture risk [23,24]. Grassi et al. [23], using standard AUROC analysis, the clinical benchmark for predictions, showed that FE-derived strength (AUROC up to 0.78, $p = 0.017$) is significantly better than aBMD (AUROC = 0.72), a critical finding validated even when FE models were derived from clinically available 2D DXA scans.

On the other side, the application of machine learning algorithms in the prediction of bone fractures is of great interest [25,26]. However, these techniques generally require high numerosity to present good results and avoid overfitting. Nishiyama et al. [24] used a Support Vector Machine (SVM) model with tenfold cross-validation to classify subjects with prior fractures. The combined model of volumetric BMD (vBMD) and FE estimates achieved a high 91.4 % classification accuracy (AUROC = 0.94), showcasing the strong potential of combining biomechanical simulation outputs with powerful machine learning algorithms. Nishiyama achieved very good results, but using QCT images as a basis for the development of finite element models, which limits the applicability of this approach in clinics.

The techniques presented in our study, such as Random Field Theory (RFT), have shifted the focus from a global hip strength score to identifying the statistically significant localized stress concentrations that initiate fracture, showing that accuracy gains are achieved by understanding where and how the bone is most vulnerable. Given the limited sample size, it was not the objective of the present study to introduce a in deep study about the use of machine learning. We introduced a step of classification using simple linear techniques to present a complete pipeline of analysis and allow an insight into the variables that can play the biggest role in fracture prediction.

4.5. Clinical applicability

Nowadays, clinical screening for osteoporosis is based on DXA images, where 2D mineral density distribution is observed in macro-areas, like the femoral neck and the great trochanter, to estimate the risk of fracture. The present study introduced a different clinical approach, focusing the analysis on reduced and statistically significant regions of interest and examining the interactive relationship between trabecular and cortical tissue. The approach has the potential to be applied in clinics since it is based on DXA images, a routine screening technique in most hospitals. It can be implemented during follow-up to assess the specific changes of density and stress distribution in the critical areas. Also, the effectiveness of drugs could be evaluated by measuring the changes in densities in specific areas, not generally in the neck and trochanteric regions. Our results are a step forward for the development of a precise and clinically actionable predictive tool derived from standard 2D DXA images.

5. Conclusions

In conclusion, this study introduces the use of RFT and SPM in the analysis of FEM for the study of osteoporotic fractures. RFT makes it possible to overcome the multiple comparison problem that occurs when a lot of dependent tests are performed, as in this instance, allowing the identification of regions where the difference between the groups is statistically significant. In the fractured group, those regions showed reduced MPS and MPE, reduced variability, and reduced difference in the behaviour of cortical and trabecular bone.

The dataset used in this study is the same as that presented in Ruiz Wills et al. [10,22], nonetheless, the innovative protocol used to statistically identify critical regions allowed for an increase in both the ROC-AUC (up to 0.95) and the classification accuracy (up to 86.8 %), which in the cited studies reached a maximum of 0.91and 79 % respectively. Most importantly, it allowed us to study the mechanical behavior of trabecular and cortical tissue and the anatomical identification of critical regions. The use of subject-specific FEM obtained from CT scans was widely presented in the literature, showing AUC values, generally higher than those obtained using only BMD, and ranging from 0.7 to 0.86 [27–29]. To the authors' knowledge, only one study presents an AUC of 0.95 analyzing FE-strength in multiple loading conditions [30]. The present study obtained results comparable to the best results obtained in the literature using models generated from DXA images, which is a screening imaging technique. We introduced a different concept of analysis that can be applied in the clinic on a large population during clinical routine follow-ups.

The results presented here are limited to the reduced sample size; however, they underline the importance of the right selection of the critical region and of the variability of the parameters within the selected region. It should be clear nowadays that focusing the analysis only on the regions of maximum stress or strain is not the best way to assess the problem of fracture.

These results, if confirmed in the following studies in larger cohorts, might allow us to study osteoporosis, focusing attention on the critical regions, and could give new insight into the progression of the pathology and remodeling of the tissue.

CRediT authorship contribution statement

Nicole Morando: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Carlos Ruiz Wills:** Writing – review & editing, Supervision, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Simone Tassani:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Ethics statement

The authors here declare that the work described has not been published previously except in the form of a preprint, an abstract, a published lecture, academic thesis or registered report. The article is not under consideration for publication elsewhere. The article's publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out. If accepted, the article will not be published elsewhere in the same form, in English or in any other language, including electronically, without the written consent of the copyright-holder.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Office and Grammarly proofing tools in order to check grammar and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Simone Tassani reports financial support was provided by Pompeu Fabra University Engineering School. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Funds from the Spanish Government (CEX2021-001195-M/ AEI /10.13039/501100011033 and DENSI3D-TSI-100101-2013-109) are acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2025.111214.

References

[1] L. Sànchez-Riera, N. Wilson, Fragility fractures & their impact on older people, Best Pract. Res. Clin. Rheumatol. 31 (2) (Apr. 2017) 169–191, https://doi.org/10.1016/J.BERH.2017.10.001.
[2] C. Willers, et al., Osteoporosis in Europe: a compendium of country-specific reports, Arch. Osteoporosis 17 (1) (2022), https://doi.org/10.1007/s11657-021-00969-8.
[3] E. Hernlund, et al., Osteoporosis in the European Union: medical management, epidemiology and economic burden: a report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA), Arch. Osteoporosis 8 (1–2) (Dec. 2013), https://doi.org/10.1007/s11657-013-0136-1.
[4] T.R. Nagy, A.-L. Clair, Precision and accuracy of dual-energy X-ray absorptiometry for determining in vivo body composition of mice, Obes. Res. 8 (5) (Aug. 2000) 392–398, https://doi.org/10.1038/oby.2000.47.
[5] H.K. Genant, J.E. Block, P. Steiger, C.-C. Glueer, R. Smith, Quantitative computed tomography in assessment of osteoporosis, Semin. Nucl. Med. 17 (4) (1987) 316–333, https://doi.org/10.1016/S0001-2998(87)80024-7.
[6] T.M. Keaveny, et al., Age-dependence of femoral strength in white women and men, J. Bone Miner. Res. 25 (5) (May 2010) 994–1001, https://doi.org/10.1359/jbmr.091033.
[7] T.M. Keaveny, et al., Femoral bone strength and its relation to cortical and trabecular changes after treatment with PTH, alendronate, and their combination as assessed by finite element analysis of quantitative CT scans, J. Bone Miner. Res. 23 (12) (Dec. 2008) 1974–1982, https://doi.org/10.1359/jbmr.080805.
[8] L. Grassi, S.P. Väänänen, M. Ristinmaa, J.S. Jurvelin, H. Isaksson, Prediction of femoral strength using 3D finite element models reconstructed from DXA images: validation against experiments, Biomech. Model. Mechanobiol. 16 (3) (2017) 989–1000, https://doi.org/10.1007/s10237-016-0866-2.

[9] K.E. Naylor, E. V McCloskey, R. Eastell, L. Yang, Use of DXA-based finite element analysis of the proximal femur in a longitudinal study of hip fracture, J. Bone Miner. Res. 28 (5) (May 2013) 1014–1021, https://doi.org/10.1002/jbmr.1856.
[10] C. Ruiz-Wills, et al., 3D patient-specific finite element models of the proximal femur based on DXA towards the classification of fracture and non-fracture cases, Bone 121 (December 2018) (2019) 89–99, https://doi.org/10.1016/j.bone.2019.01.001.
[11] S. Tassani, M. Pani, J. Noailly, M.A. Gonzalez Ballester, Trabecular fracture Zone might not be the higher strain region of the trabecular framework, Front Mater 5 (2018), https://doi.org/10.3389/fmats.2018.00006.
[12] S. Tassani, G. Matsopoulos, F. Baruffaldi, 3D identification of trabecular bone fracture zone using an automatic image registration scheme: a validation study, J. Biomech. 45 (11) (2012) 2035–2040 [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0021929012002886. (Accessed 16 May 2014).
[13] M. Brett, W. Penny, S. Kiebel, Introduction to Random Field Theory, 2003, https://doi.org/10.1016/B978-012264841-0/50046-9.
[14] E. De Pieri, D.E. Lunn, G.J. Chapman, K.P. Rasmussen, S.J. Ferguson, A. C. Redmond, Patient characteristics affect hip contact forces during gait, Osteoarthr. Cartil. 27 (6) (2019) 895–905, https://doi.org/10.1016/j.joca.2019.01.016.
[15] M.L. Bouxsein, P. Szulc, F. Munoz, E. Thrall, E. Sornay-Rendu, P.D. Delmas, Contribution of trochanteric soft tissues to fall force estimates, the factor of risk, and prediction of hip fracture risk, J. Bone Miner. Res. 22 (6) (2007), https://doi.org/10.1359/jbmr.070309.
[16] S.N. Robinovitch, T.A. Mcmahon, Wilson, C. Hayes, Force Attenuation in Trochanteric Soft Tissues During Impact from a Fall, Orthopaedic Research Society, 1995.
[17] T.C. Pataky, One-dimensional statistical parametric mapping in python, Comput. Methods Biomech. Biomed. Eng. 15 (3) (2012) 295–301, https://doi.org/10.1080/10255842.2010.527837.
[18] R.J. Adler, The Geometry of Random Fields, Society for Industrial and Applied Mathematics, 2010, https://doi.org/10.1137/1.9780898718980.
[19] T.C. Pataky, rft1d: smooth one-dimensional random field upcrossing probabilities in python, J. Stat. Software 71 (7) (2016) 1–22, https://doi.org/10.18637/jss.v071.i07.
[20] D. Endo, et al., Reduced cortical bone thickness increases stress and strain in the female femoral diaphysis analyzed by a CT-based finite element method: implications for the anatomical background of fatigue fracture of the femur, BoneKEy Rep. 13 (Dec. 2020), https://doi.org/10.1016/j.bonr.2020.100733.
[21] M. Palanca, E. Perilli, S. Martelli, Body anthropometry and bone strength conjointly determine the risk of hip fracture in a sideways fall, Ann. Biomed. Eng. 49 (5) (2021), https://doi.org/10.1007/s10439-020-02682-y.
[22] R.W. C, et al., Relative Fragility of Osteoporotic Femurs Assessed with DXA and Simulation of Finite Element Falls Guided by Emergency X-rays, 2020, pp. 1–8.
[23] L. Grassi, et al., 3D finite element models reconstructed from 2D Dual-Energy X-Ray Absorptiometry (DXA) images improve hip fracture prediction compared to areal BMD in osteoporotic fractures in men (MrOS) Sweden cohort, J. Bone Miner. Res. 38 (9) (2023), https://doi.org/10.1002/jbmr.4878.
[24] K.K. Nishiyama, M. Ito, A. Harada, S.K. Boyd, Classification of women with and without hip fracture based on quantitative computed tomography and finite element analysis, Osteoporos. Int. 25 (2) (2014), https://doi.org/10.1007/s00198-013-2459-6.
[25] D. Senanayake, et al., Classification of fracture risk in fallers using Dual-Energy X-Ray Absorptiometry (DXA) images and deep learning-based feature extraction, JBMR Plus 7 (12) (2023), https://doi.org/10.1002/jbm4.10828.
[26] S.H. Kong, et al., A novel fracture prediction model using machine learning in a community-based cohort, JBMR Plus 4 (3) (2020), https://doi.org/10.1002/jbm4.10337.
[27] A.L. Adams, et al., Osteoporosis and hip fracture risk from routine computed tomography scans: the fracture, osteoporosis, and CT utilization Study (FOCUS), J. Bone Miner. Res. 33 (7) (Jul. 2018) 1291–1301, https://doi.org/10.1002/jbmr.3423.
[28] I. Fleps, et al., Finite element derived femoral strength is a better predictor of hip fracture risk than aBMD in the AGES Reykjavik study cohort, Bone 154 (Jan) (2022), https://doi.org/10.1016/j.bone.2021.116219.
[29] M. Qasim, et al., Patient-specific finite element estimated femur strength as a predictor of the risk of hip fracture: the effect of methodological determinants, Osteoporos. Int. 27 (9) (2016) 2815–2822, https://doi.org/10.1007/s00198-016-3597-4.
[30] C. Falcinelli, et al., Multiple loading conditions analysis can improve the association between finite element bone strength estimates and proximal femur fractures: a preliminary study in elderly women, Bone 67 (2014) 71–80, https://doi.org/10.1016/j.bone.2014.06.038.