Evaluation of the capability of the simulated dual energy X-ray absorptiometry-based two-dimensional finite element models for predicting vertebral failure loads

Yongtao Lu[1,2,*,#], Yifan Zhu[3,#], Matthias Krause[4], Gerd Huber[5], Junyan Li[6]

[1]Department of Engineering Mechanics, Dalian University of Technology, No. 2 Linggong Road, 116024 Dalian, China

[2]State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, No. 2 Linggong Road, 116024 Dalian, China

[3]Department of Engineering Mechanics, Shanghai Jiaotong University, No. 800 Dongchuan Road, 20024 Shanghai, China

[4]Department of Osteology and Biomechanics, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20251 Hamburg, Germany

[5]Institute of Biomechanics, TUHH Hamburg University of Technology, Denickestrasse 15, 21073 Hamburg, Germany

[6]Department of Design Engineering and Mathematics, School of Science and Technology, Middlesex University, The Burroughs, Hendon, NW4 4BT, London,UK

[#] The first two authors share the first authorship

Corresponding author:
Yongtao Lu, Ph.D
Department of Engineering Mechanics
Dalian University of Technology
No.2 Linggong Road
Dalian, 116024, China
Email: Yongtaolu@dlut.edu.cn

Number of words (Introduction to Discussion): 4101 (5000)

Number of tables: 1

Number of Figures: 4

Abstract

Prediction of the vertebral failure load is of great importance for the prevention and early treatment of bone fracture. However, an efficient and effective method for accurately predicting the failure load of vertebral bones is still lacking. The aim of the present study was to evaluate the capability of the simulated dual energy X-ray absorptiometry (DXA)-based finite element (FE) model for predicting vertebral failure loads.

Thirteen dissected spinal segments (T11/T12/L1) were scanned using a HR-pQCT scanner and then were mechanically tested until failure. The subject-specific three-dimensional (3D) and two-dimensional (2D) FE models of T12 were generated from the HR-pQCT scanner and the simulated DXA images, respectively. Additionally, the areal bone mineral density (aBMD) and areal bone mineral content (aBMC) of T12 were calculated. The failure loads predicted by the simulated DXA-based 2D FE models were more moderately correlated with the experimental failure loads ($R^2 = 0.66$) than the aBMC ($R^2 = 0.61$) and aBMD ($R^2 = 0.56$). The 2D FE models were slightly outperformed by the HR-pQCT-based 3D FE models ($R^2 = 0.71$). The present study demonstrated that the simulated DXA-based 2D FE model has better capability for predicting the vertebral failure loads than the densitometric measurements but is outperformed by the 3D FE model. The 2D FE model is more suitable for clinical use due to the low radiation dose and low cost, but it remains to be validated by further *in vitro* and *in vivo* studies.

Keywords: Vertebral failure; finite element analysis; DXA; prediction capability; BMD

1. Introduction

Vertebral fracture is a major clinical problem associated with low back pain and impaired quality of life **[1]**. Assessing the failure loads of vertebral bones is of great importance for the prevention and early treatment of bone fracture. Vertebral fractures in elderly people are strongly related to osteoporosis, which leads to the loss of bone mass and the deterioration of bone microarchitecture **[2]**. Currently, monitoring of the changes in the bone densitometric parameters such as bone mineral density (BMD) is the most important clinical approach for assessing the risk of bone fracture. The commonly used BMD measurements include the areal bone mineral density (aBMD) measured by dual energy X-ray absorptiometry (DXA) and the volumetric bone mineral density (vBMD) measured by quantitative computed tomography (QCT). However, QCT cannot be performed routinely due to its high radiation dose **[3]**. In addition, only approximately 50% of the variability in the vertebral failure load can be predicted by these BMD measurements, which cannot provide information about bone microarchitecture and BMD distribution **[4 - 6]**. By contrast, DXA can be used routinely and frequently because of its low radiation dose and low cost **[7]**. However, the aBMD obtained from DXA does not contain information about the material microarchitecture or any mechanical properties of the bone tissues. Therefore, it is necessary to develop advanced DXA-based techniques for the accurate prediction of bone failure loads that can be easily transferred into routine clinical use **[3, 8]**.

In recent years, the use of subject-specific finite element (FE) models to predict vertebral failure loads has attracted increasing attention, because the FE models account for the vertebral geometry, the BMD distribution and the mechanical properties of bone tissues **[9 - 12]**. Three-dimensional (3D) FE models have been demonstrated to be more reliable for predicting vertebral failure loads than aBMD **[13]** and vBMD **[14]**. However, it is very challenging to apply the subject-specific 3D FE models in clinical use due to the invasive QCT imaging and the complexity of 3D image segmentation that are required to construct the 3D FE models, and the high cost of performing the 3D FE simulations. Because of the low radiation dose and low cost associated with DXA scans and the high

efficiency of the construction of 2D FE models, DXA-based two-dimensional (2D) FE models have the potential for application in clinical use as an efficient tool to predict vertebral failure loads. However, no previous studies have evaluated the capability of the DXA-based 2D FE models for predicting the vertebral fracture risk.

The aim of the present study was to assess the capability of the simulated DXA-based 2D FE model for predicting vertebral failure loads by comparing its predictions with experimentally measured failure loads and by comparing its predictive power with those of the methods based on bone densitometric measurements and the quantitative computed tomography (QCT)-based 3D FE model.

2. Materials and methods

2.1. Specimen preparation, HR-pQCT imaging and mechanical testing

To validate the predictions of the simulated DXA-based 2D FE models, spinal segments were harvested, dissected, imaged and mechanically tested until failure. The detailed procedures of the dissection, HR-pQCT imaging and mechanical testing of the vertebral specimens are described in previous studies **[6, 9]**. Briefly, thirteen T11/T12/L1 spinal segments, which did not have any fracture or osteophytes, were harvested from postmenopausal female donors (mean age of 79.9 ± 7.9 years). The segments were scanned while frozen using a HR-pQCT scanner (XtremeCT, Scanco Medical AG, Bruettisellen, Switzerland) with an isotropic voxel size of 82.0 × 82.0 × 82.0 μm^3. The spinal facet joints were removed to allow for the loading transferred only through the vertebral bodies and failures of T11 and L1 were avoided by replacing all of the cancellous bones in T11 and L1 with polymethylmethalcrylate (PMMA) (see **Fig. 1** in **[6]**). The specimens were embedded in the metal cups with the application of a fixation frame to ensure that the mid-transverse planes of T12 were horizontal and in the neutral posture (no bending) [**9, 15**]. Then, the embedded specimens were mounted on the material testing machine (**Fig. 1e**). Failure loads of the T12 bodies were obtained using the loading scenario of a quasi-static compression via the intervertebral discs (IVD). The experimentally measured failure loads of T12 were used as the reference for validating

the predictions from the simulated DXA-based 2D FE models.

2.2. Finite element analysis and calculation of bone failure load

The 2D FE models, including the T12 vertebra and the adjacent IVDs, were created by converting each pixel in the simulated DXA images into a 2D 4-node plane stress element (PLANE182). The following steps were used to obtain the simulated DXA images. First, the HR-pQCT image data of each T12 vertebral body were rotated to align the spinal cranio-caudal and anterior-posterior axes along the Z- and Y-axes, respectively. The image voxel size was then coarsened to 1.002×1.002×1.002 mm^3 in order to match the resolution of a clinical lumbar DXA scan. Simulated DXA images were then obtained by projecting the 3D images onto the frontal plane of T12 (i.e., along the spinal anterior-posterior direction) (**Fig. 1a**). All of these image processing steps were performed using Amira (v5.4.3, FEI Visualization Sciences Group, France).

In the simulated DXA-based 2D FE models, heterogeneous material properties were defined for T12 using the following two steps. First, the grayscale image datasets were smoothed using a Gaussian filter (sigma = 1.2, support = 2.0) to reduce the influence of image noise. Second, the image grayscale values were converted into vBMD values based on the linear calibration equation provided by the HR-pQCT scanner. The vBMD values were further converted into bone ash density according to the relationship reported in the literature **[16]**. After matching the phantom type and anatomic site, the relationship of $\rho_{ash} = 0.877 \times \rho_{HA} + 0.079$ (ρ_{HA} is the HA-equivalent vBMD) was chosen. It should be noted that if clinical DXA images and aBMD values were available, the vBMD values could be obtained by dividing the aBMD by the subject-specific constant thickness [**3**].

Young's modulus of each bone element was calculated from the bone ash density based on the exponential density-modulus relationship reported in the literature **[16]**. Considering that some image pixels may have artificially high grayscale values that could lead to unrealistically high bone densities, an upper threshold value of 1200.00 mg/cm^3, which is the maximum bone ash density value **[16]**, was defined in the density-modulus relationship **[17]**. On the other hand, a lower threshold value of the bone ash density of 400.00 mg/cm^3 was adopted in the density-modulus relationship to avoid the

unrealistically low moduli in the FE models. Young's moduli for the elements with the bone ash density lower than 400.00 mg/cm^3 were set to 0.0104 MPa **[17]**. In summary, after matching the anatomic site (i.e., vertebra), the following exponential density-modulus relationship was used in the present study **[18]**:

$$E = \begin{cases} 0.0104 & \rho_{\text{ash}} < 400 \\ \text{a} \times \rho_{\text{ash}}{}^{\text{b}} & 400 \leq \rho_{\text{ash}} \leq 1200 \\ \text{a} \times 1200^{\text{b}} & \rho_{\text{ash}} > 1200 \end{cases} \quad (1)$$

where a and b are constants (a = 0.1127, b = 1.746 in the present study), E is Young's modulus (MPa) and ρ_{ash} is the bone ash density (mg/cm^3).

Poisson's ratio for the bone elements was set to 0.30. The material with the bone ash density lower than 400 mg/cm^3 was regarded as bone marrow, and the corresponding Poisson's ratio was set to 0.49 **[14]**. The heterogeneous FE models were generated by mapping the elastic modulus calculated at each image pixel onto the FE mesh using an in-house developed MATLAB (R2017a, MathWorks, Natick, Massachusetts, U.S.A.) code **[19]**.

The intervertebral discs were added into the 2D FE T12 models in order to enable the definition of consistent loading condition in the models (**Fig. 1b**). The IVDs in the 2D FE models were simplified as one material and no differentiation of the nucleus pulposus and annulus fibers was made in the 2D IVD models. An incompressible isotropic Mooney-Rivlin material model was used to describe the mechanical behavior of the 2D IVDs, with C_{10}, C_{01} and D were set to 0.10 MPa, 2.50 MPa and 0.30 MPa^{-1}, respectively **[20]**. The thickness of the IVDs was based on the average thickness of human IVDs, i.e., it was approximately 8.00 mm. The FE meshes of IVDs were created by converting each image pixel into PLANE182, and thus the IVDs were fully bounded with T12 at the interface. A mesh convergence study was performed by refining the PLANE182 elements until the predictions (failure loads) were not affected by the mesh size, resulting in approximately 5,128 elements per 2D FE spinal model. In the 2D FE models, a uniform displacement of 2.00 mm was applied on the topmost layer of the IVD, while all degrees of freedom were fixed for the nodes in the bottom layer. This boundary condition was defined because it can be easily applied and transferred into clinical use.

The failure load of T12 vertebra predicted from the simulated DXA-based 2D FE models was defined as the load under which at least 5% of the bone elements in the 2D model experience stress/strain that exceeds the failure threshold **[21]**. Because there is currently still no consensus on which failure criterion should be used for bone tissues, and to investigate the influence of the failure criterion on the 2D FE predictions, four different failure criteria were considered in the present study including the principal stress, the principal strain, the von Mises stress and the von Mises strain. The yield stresses in each bone element were related to Young's modulus using the empirical linear equations **[22]**:

$$S_t = 0.0039 \times E + 0.33 \quad (2)$$

$$S_c = 0.0062 \times E - 0.41 \quad (3)$$

where S_t is the tensile yield stress (MPa), S_c is the compressive yield stress (MPa) and E is Young's modulus (MPa).

The von Mises yield stress for bone tissues was defined as the average value of the tensile and compressive yield stresses. The tensile and compressive yield strains for bone tissues were set to 7300.00 με and 10400.00 με, respectively **[22]**. The von Mises yield strain was set to the average value of the tensile and compressive yield strains.

To investigate the influence of the failure criterion on the fracture initiation, the failure ratios in the 2D FE models were calculated using different failure criteria. The failure ratio using the failure criteria of principal stress (or strain) was defined as the larger value of the ratio of tensile stress (or strain) to tensile yield stress (or strain) and the ratio of compressive stress (or strain) to compressive yield stress (or strain), while the failure ratio using the failure criterion of von Mises stress (or strain) was defined as the ratio of the von Mises stress (or strain) to the von Mises yield stress (or strain). The region in the 2D model where the highest failure ratio occurred was considered the fracture initiation region. All of the DXA-based 2D linear FE models were solved using Ansys (Release 15.0, ANSYS, Inc., Canonsburg, PA, U.S.A).

The capability of the simulated DXA-based 2D FE models for predicting vertebral failure loads was assessed by comparing their prediction with those of the corresponding 3D FE models (**Fig. 1c**). The calculation of the failure loads of T12 from the 3D FE models

was performed as described in a previous study **[9]**. Briefly, the 3D FE models, including the T12 vertebral body and two adjacent IVDs, were generated from the HR-pQCT images. Quadratic wedge (C3D15) elements were defined for the cortex, and quadratic tetrahedral elements (C3D10) were defined for the trabecular bone and the IVD. A mesh convergence study was performed to ensure that the predicted failure loads were not affected by the mesh size, resulting approximately 35,874 elements per 3D FE spinal model. The anisotropic elastic-plastic-damage model **[23]** was used to simulate the mechanical behavior of bone elements until failure. The Mooney-Rivlin model was defined for the nucleus pulposus, and the fiber-reinforced hyperelastic model was chosen for the annulus fibrosus. The *in vitro* loading scenario was simulated, i.e., the bottom nodes from the inferior IVD were fully constrained, and the loading condition of a 4º forward bending followed by an axial displacement of 4.0 mm was applied on the cranial nodes of the superior IVD. The failure loads of T12 were computed from the 3D FE models as the maximal force obtained from the nonlinear FE analyses.

2.3. Measurements of bone densitometric parameters

The predictive power of the simulated DXA-based T12 FE model was compared to that of the aBMD and areal bone mineral content (aBMC) of T12. The aBMD and aBMC of T12 were calculated from the simulated DXA images (i.e., the projected images from the HR-pQCT) (**Fig. 1d**). To calculate the aBMD and aBMC of T12, the simulated DXA images were first smoothed using a Gaussian filter (convolution kernel = [3 3 3], standard deviation = 0.65) to reduce the influence of image noise. Then, the grayscale images were binarized using a threshold that was equal to 25.5% of the maximal grayscale value **[24]**, and bone masks (regions occupied by bone voxels) were defined in the binary images. The image threshold values applied were equivalent to an average BMD of 433.00 ± 14.00 mg HA/cm^3 (range from 401.00 mg HA/cm^3 to 447.00 mg HA/cm^3) and corresponded to the valley region between the two peaks in the BMD histograms. All of the segmentations were visually evaluated to ensure the proper application of the threshold values selected. Then, the HA-equivalent volumetric BMD (vBMD) values in the bone voxels (bone mask regions) were calculated from the CT grayscale values using the calibration law provided

by the manufacturer of the HR-pQCT scanner. The HR-pQCT scanner was calibrated weekly using the phantom provided by the manufacturer. The bone minerals in each bone pixel were calculated from the corresponding vBMD by multiplying the vBMD by the volume of the image voxel, i.e., 1.002×1.002×1.002 mm^3. Then, the aBMC of T12 was calculated as the total bone minerals over the masked bone regions, and the aBMD of T12 was obtained by dividing the aBMC of T12 by the total area of T12.

2.4. Statistical analysis

The normal distribution of the parameters was evaluated by the Shapiro-Wilk test and by visually inspecting the normal probability plots. If a normal distribution was fulfilled, the Pearson's correlation coefficients (r) were calculated to quantify the correlations among the failure loads predicted by the DXA-based 2D FE models using different failure criteria. Regression equations, coefficients of determination (R^2) and root mean squared errors (RMS) were computed to determine the linear correlations between the experimentally measured vertebral failure loads and the prediction from the simulated DXA-based 2D FE models, and between the 2D and 3D FE models. Statistical analyses were performed using MATLAB. The probability of type I error was set as alpha = 0.05, i.e., $p < 0.05$ was considered to be statistically significant.

3. Results

The mean ± standard deviation (SD) values of the failure loads of T12 predicted by the simulated DXA-based 2D FE models using the failure criteria of principal stress, von Mises stress, principal strain and von Mises strain were 540.00 ± 144.00 N, 460.00 ± 120.00 N, 952.00 ± 249.00 N and 792.00 ± 201.00 N, respectively. The vertebral failure loads predicted by the simulated DXA-based 2D FE models using different failure criteria were highly correlated with each other, and the Pearson's correlation coefficients (r) were all significant (all $r > 0.99$, $p < 0.001$) (**Table 1**). The distributions of the failure ratios and the fracture initiation regions calculated using different failure criteria were similar (**Fig. 2**). Therefore, in the following analysis, only the results from the failure criteria of the

principal strain are reported.

Linear correlations of the experimentally measured failure loads of T12 (F_{Exp}) with the aBMD, the aBMC and the failure loads predicted by the simulated DXA-based 2D FE models (DXA_F_{FE}) were all significant ($p < 0.005$). The failure loads predicted by the DXA-based 2D FE models (DXA-F_{FE}) were more moderately correlated with the experimental failure loads ($R^2 = 0.66$) than the aBMD ($R^2 = 0.56$) and the aBMC ($R^2 = 0.61$) (**Fig. 3 and Fig. 4a**). The DXA-based 2D FE models were slightly outperformed by the HR-pQCT-based 3D FE models ($R^2 = 0.71$ for the correlation with the experimental data). Moderate correlations were found between the failure loads predicted by the DXA-based 2D FE models and the HR-pQCT-based 3D FE models (HR-pQCT-F_{FE}) ($R^2 = 0.70$, $p < 0.001$) (**Fig. 4b**). Compared to the experimentally measured failure loads (2.09 ± 0.48 kN), the failure loads of T12 predicted by the FE models were 74% lower in the DXA-based 2D FE models (0.54 ± 0.14 kN, $p < 0.001$) and 12% lower in HR-pQCT-based 3D FE models (1.84 ± 0.47 kN, $p < 0.001$).

Using a computer with an i7 processor and 8G RAM, it typically took less than 15 minutes to perform the DXA-based 2D FE simulation, while the segmentation and simulation of the HR-pQCT-based 3D FE model required approximately 420 minutes (each calculation took approximately 190 minutes). The number of degrees of freedom was approximately 10,848 for the DXA-based 2D FE models and approximately 194,467 for the HR-pQCT-based 3D FE models.

4. Discussion

The goal of the present study was to assess the capability of a simulated DXA-based 2D FE model for predicting the vertebral failure loads by comparing its predictions with the experimentally measured vertebral failure loads and by comparing its predictive power with the predictive powers of the vertebral densitometric measurements and of the HR-pQCT-based 3D FE model. It was demonstrated that the simulated DXA-based 2D FE models are more reliable for predicting the failure loads of T12 ($R^2 = 0.66$) than the densitometric measurements including the aBMD ($R^2 = 0.56$) and the aBMC ($R^2 = 0.61$)

that are currently used in clinical practice. Although the 2D FE models are outperformed by the HR-pQCT-based 3D FE models (R^2 = 0.71) in predicting the failure loads of T12 **[9]**, the 3D approach requires the use of a high radiation dose and the construction of the 3D FE models has a high computational cost. By contrast, the DXA-based 2D FE modeling approach is highly efficient (requiring only a few minutes to run the simulation), requires the use of only a low radiation dose and has a low cost, making it more suitable for clinical use.

The present study is an extension of our previous study **[9]**, in which it was found that up to 71% of the variability in the vertebral failure loads can be predicted using the HR-pQCT-based 3D FE models including the T12 vertebral body and the adjacent intervertebral discs **[9]**. However, the main issue associated with the 3D models is the need for a high radiation dose and the long time that is normally needed to create and solve the 3D FE models, which pose considerable challenges for making the 3D modeling approach readily available in clinical use. Therefore, an efficient 2D FE modeling approach based on the simulated DXA images was developed in the present study. It should be noted that only a moderate correlation (R^2 = 0.70) was found between the simulated DXA-based 2D models and the HR-pQCT-based 3D FE models, implying that the 3D FE models contain some additional information that contributes to the 5% increase (**Fig. 4a**) in the prediction accuracy of vertebral failure loads.

It was demonstrated that the failure loads predicted by the simulated DXA-based 2D FE models are more moderately correlated with the experimentally measured failure loads than the densitometric measurements (aBMD and aBMC). This finding may be because the biomechanical features of T12 (including the heterogeneous mechanical properties, geometry and boundary conditions, etc.) that are important for the prediction of failure loads can be reflected in the 2D FE models to some extent **[25]**. By contrast, the densitometric measurements only contain the information regarding the average bone mineral density and bone mass and are not directly related to the mechanical behavior of the bones. Therefore, densitometric measurements have limited capability for predicting bone failure loads. The fact that the failure load is more moderately correlated with the

aBMC than with the aBMD may be because the bone failure load is a non-normalized parameter and can be influenced by the bone dimension. It should be noted that in addition to the 2D FE model, the trabecular bone score (TBS) can also be derived from the DXA images. TBS is a texture index and can provide information that is complementary to the information provided by BMD, motivating many investigations of its predictive capability in the recent years [**26, 27**]. Indeed, numerous studies have shown that lower TBS values are associated with increased risk for major osteoporotic fracture [**26**]. However, our previous study showed that the TBS is a poor surrogate for vertebral strength [**27**], suggesting that further research on the relationship between TBS and vertebral strength is necessary.

It should be noted that although several bone material models have been developed previously, there is still no consensus regarding which model can best describe the mechanical behavior of bone. Zysset et al. has developed a complex anisotropic elastic-plastic-damage model **[23]** to simulate the mechanical behavior of human vertebrae **[28, 29]**. On the other hand, Viceconti et al. has predicted the ultimate loads of the bone based on a linear elastic material model **[3, 30, 31]**. In the 2D FE models developed in the present study, Viceconti's approach is adopted. However, it is unclear which failure criterion should be used to accurately predict the bone failure loads. In previous studies **[3, 31, 32]**, the failure criteria of the principal stress, principal strain, von Mises stress and von Mises strain were all widely used. Therefore, these four failure criteria were assessed in the 2D FE models developed in the present study. It was found that the failure loads predicted by the DXA-based 2D FE models using different failure criteria are strongly correlated with each other, demonstrating that adoption of different failure criteria has a minimal influence on the results of the 2D FE models.

Several limitations of the present study need to be noted. First, the DXA-based 2D FE models are generated from the simulated DXA images, i.e., the 2D coarsened projections of the HR-pQCT images. The reasons for using the simulated DXA images are that the image datasets from our previous studies are used making it possible to validate the model and make comparisons with 3D models using these data. It should be

noted that in the present study, the vertebral posterior elements and surrounding tissues (ribs, etc.) were removed when projecting the HR-pQCT images, and consequently, the simulated DXA represents the best case condition for DXA imaging, which is expected to have lower quality in the clinical practice. Although the quality of the simulated DXA images was compared with the quality of the clinical images and it was found that aBMD can be simulated from HR-pQCT images of the distal radius [**33**], the comparison using the spinal segment has not been performed and furthermore, the results of the FE models obtained from the simulated and clinical spinal DXA images have not been evaluated. Therefore, in the future, the methodology developed in the present study should be validated directly using the clinical DXA scans. Second, the nonlinear behavior of bone prior to failure is not considered in the DXA-based 2D FE models. However, experimental data showed that bone is a brittle material **[34]** and plastic behavior has a minimal effect on the calculation of bone failure loads. Third, the sample size used in the present study is small (N = 13), and the bone samples are obtained only from old female donors (mean age of 79.9 ± 7.9 years), which may hinder the application of these findings to a wider range of vertebral bones in different conditions, in particular to younger individuals with higher BMD values. However, it is very challenging to harvest a sufficient number of vertebral specimens from young donors.

The present study is the first to assess the capability of simulated DXA-based 2D FE models for predicting the compressive failure loads of vertebral bodies. In conclusion, the present study showed that the simulated DXA-based 2D FE model is a better predictor than the densitometric measurements for predicting the compressive failure loads of vertebral bodies in elderly women with osteoporosis. Although the 2D FE model is not as capable as the 3D FE model for predicting the vertebral failure loads, the construction of the 2D model requires a markedly shorter period, less expertise and a much shorter computational time. Additionally, the DXA scan requires the use of a low radiation dose and incurs a low cost. However, only simulated DXA images were used in the present study, and this approach remains to be further validated for clinical applications by evaluating its performance *in vitro* and *in vivo* directly using clinical DXA images.

Conflict of interest

The authors declare there is no conflict of interest

Funding

This work was funded by the National Natural Science Foundation of China (11702057), the Chinese Fundamental Research Funds for the Central Universities (DUT18LK19) and the Open Fund from the State Key Laboratory of Structural Analysis for Industrial Equipment (GZ18104), Dalian University of Technology.

References

[1] Lips P, Cooper C, Agnusdei D, Caulin F, Egger P, Johnell O, et al. Quality of life in patients with vertebral fractures: Validation of the quality of life questionnaire of the European Foundation for Osteoporosis (QUALEFFO). Osteoporos Int 1999;10: 150-60.

[2] Krause M, Breer S, Mohrmann B, Vettorazzi E, Marshall RP, Amling M, et al. Influence of non-traumatic thoracic and lumbar vertebral fractures on sagittal spine alignment assessed by radiation-free spinometry. Osteoporos Int 2013;24:1859-68.

[3] Dall'Ara E, Eastell R, Viceconti M, Pahr D, Yang L. Experimental validation of DXA-based finite element models for prediction of femoral strength. J Mech Behav Biomed Mater 2016;63:17-25.

[4] Lochmüller EM, Bürklein D, Kuhn V, Glaser C, Muller R, Glüer CC, et al. Mechanical strength of the thoracolumbar spine in the elderly: Prediction from in situ dual-energy X-ray absorptiometry, quantitative computed tomography (QCT), upper and lower limb peripheral QCT, and quantitative ultrasound. Bone 2002;31: 77-84.

[5] Ebbesen EN. Thomsen JS, Beck-Nielsen H, Nepper-Rasmussen HJ, Mosekilde L. Lumbar vertebral body compressive strength evaluated by dual-energy X-ray absorptiometry, quantitative computed tomography, and ashing. Bone 1999;25: 713-24.

[6] Lu YT, Krause M, Bishop N, Sellenschloh K, Glüer C, Püschel K, et al. The role of patient-mode high-resolution peripheral quantitative computed tomography indices in the prediction of failure strength of the elderly women's thoracic vertebral body. Osteoporos Int 2014;26:237-44.

[7] Griffith JF, Genant HK. Bone mass and architecture determination: state of the art. Best Pract Res Clin Endocrinol Metab 2008;22:737-64.

[8] Maquer G, Lu YT, Dall'Ara E, Chevalier Y, Krause M, Yang L, et al. The Initial Slope of the Variogram, Foundation of the Trabecular Bone Score, Is Not or Is Poorly Associated With Vertebral Strength. J Bone Miner Res 2016;31:341-6.

[9] Lu YT, Maquer G, Museyko O, Pueschel K, Engelke K, Zysset P, et al. Finite element analyses of human vertebral bodies embedded in polymethylmethalcrylate or loaded via the hyperelastic intervertebral disc models provide equivalent predictions of experimental strength. J Biomech 2014;47:2512-6.

[10] Chevalier Y, Quek E, Borah B, Gross G, Stewart J, Lang T, et al., Biomechanical effects of teriparatide in women with osteoporosis treated previously with alendronate and risedronate: Results from quantitative computed tomography-based finite element analysis of the vertebral body. Bone 2010;46:41-8.

[11] Dall'Ara E, Pahr D, Varga P, Kainberger F, Zysset P. QCT-based finite element models predict human vertebral strength in vitro significantly better than simulated DEXA. Osteoporos Int 2012;23:563-72.

[12] Imai K, Ohnishi I, Yamamoto S, Nakamura K. In vivo assessment of lumbar vertebral strength in elderly women using computed tomography-based nonlinear finite element model. Spine 2008;33:27-32.

[13] Wang X, Sanyal A, Cawthon PM, Palermo L, Jekir M, Christensen J, et al. Prediction of New Clinical Vertebral Fractures in Elderly Men Using Finite Element Analysis of CT Scans. J Bone Miner Res 2012;27:808-16.

[14] Crawford RP, Cann CE, Keaveny TM. Finite element models predict in vitro vertebral body compressive strength better than quantitative computed tomography. Bone 2003;33:744-50.

[15] Lu YT, Rosenau E, Paetzold H, Klein A, Pueschel K, Morlock MM, Huber G. Strain changes on the cortical shell of vertebral bodies due to spine ageing: A parametric study using a finite element model evaluated by strain measurements. Proc Inst Mech Eng H 2013;227:1265 – 74.

[16] Knowles NK, Reeves JM, Ferreira LM. Quantitative Computed Tomography (QCT) derived Bone Mineral Density (BMD) in finite element studies: a review of the literature. J Exp Orthop 2016;3:36.

[17] Helgason B, Perilli E, Schileo E, Taddei F, Brynjólfsson S, Viceconti M. Mathematical relationships between bone density and mechanical properties: A literature review. Clin Biomech 2008;23:135-46.

[18] Easley SK, Jekir MG, Burghardt AJ, Li M, Keaveny TM. Contribution of the intra-specimen variations in tissue mineralization to PTH- and raloxifene-induced changes in stiffness of rat vertebrae. Bone 2010;46:1162-9.

[19] Chen Y, Dall'Ara E, Sales E, Manda K, Wallace R, Pankaj P, et al., Micro-CT based finite element models of cancellous bone predict accurately displacement once the boundary condition is well replicated: A validation study. J Mech Behav Biomed Mater 2017;65:644-51.

[20] Moramarco V, Perez Del Palomar A, Pappalettere C, Doblare M. An accurate validation of a computational model of a human lumbosacral segment. J Biomech 2010;43:334-42.

[21] Pistoia W, Van Rietbergen B, Lochmuller EM, Lill CA, Eckstein F, Ruegsegger P. Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images. Bone 2002;30:842-8.

[22] Morgan EF, Keaveny TM. Dependence of yield strain of human trabecular bone on anatomic site. J Biomech 2001;34:569-77.
[23] Schwiedrzik JJ, Zysset PK. An anisotropic elastic-viscoplastic damage model for bone tissue. Biomech Model Mechanobiol 2013;12:201-13.
[24] Klinck RJ, Campbell GM, Boyd SK. Radiation effects on bone architecture in mice and rats resulting from in vivo micro-computed tomography scanning. Med Eng Phys 2008;30:888-95.
[25] Wegrzyn J, Roux J, Arlot ME, Boutroy S, Vilayphiou N, Guyen O, et al., Role of Trabecular Microarchitecture and Its Heterogeneity Parameters in the Mechanical Behavior of Ex Vivo Human L-3 Vertebrae. J Bone Miner Res 2010;25:2324-31.
[26] Martineau P, Leslie WD. Trabecular bone score (TBS): Methods and applications. Bone 2017;104:66 -72.
[27] Maquer G, Lu YT, Dall'Ara E, Chevalier Y, Krause M, Yang L, Eastell R, Lippuner K, Zysset PK. The initial slope of the variogram, foundation of the trabecular bone score, is not or is poorly associated with vertebral strength. J Bone Miner Res 2016;31:341-6.
[28] Maquer G, Dall'Ara E, Zysset PK. Removal of the cortical endplates has little effect on ultimate load and damage distribution in QCT-based voxel models of human lumbar vertebrae under axial compression. J Biomech 2012;45:1733-38.
[29] Maquer G, Schwiedrzik J, Zysset PK. Embedding of human vertebral bodies leads to higher ultimate load and altered damage localisation under axial compression. Comput Methods Biomech Biomed Eng 2014;17:1311-22.
[30] Qasim M, Farinella G, Zhang J, Li X, Yang L, Eastell R, 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 2016; 27:2815-22.
[31] Li X, Viceconti M, Cohen MC, Reilly GC, Carré MJ, Offiah AC. Developing CT based computational models of pediatric femurs. J Biomech 2015;48:2034-40.
[32] Naylor KE, McCloskey EV, Eastell R, Yang L. Use of DXA-based finite element analysis of the proximal femur in a longitudinal study of hip fracture. J Bone Miner Res 2013;28:1014-21.
[33] Burghardt AJ, Kazakia GJ, Link TM, Majumdar S. Automated simulation of areal bone mineral density assessment in the distal radius from high-resolution peripheral quantitative computed tomography. Osteoporos Int 2009;20:2017-24.
[34] Keaveny TM, Guo XE, Wachtel EF, McMahon TA, Hayes WC. Trabecular bone exhibits fully linear elastic behavior and yields at low strains. J Biomech 1994; 27: 1127-36.

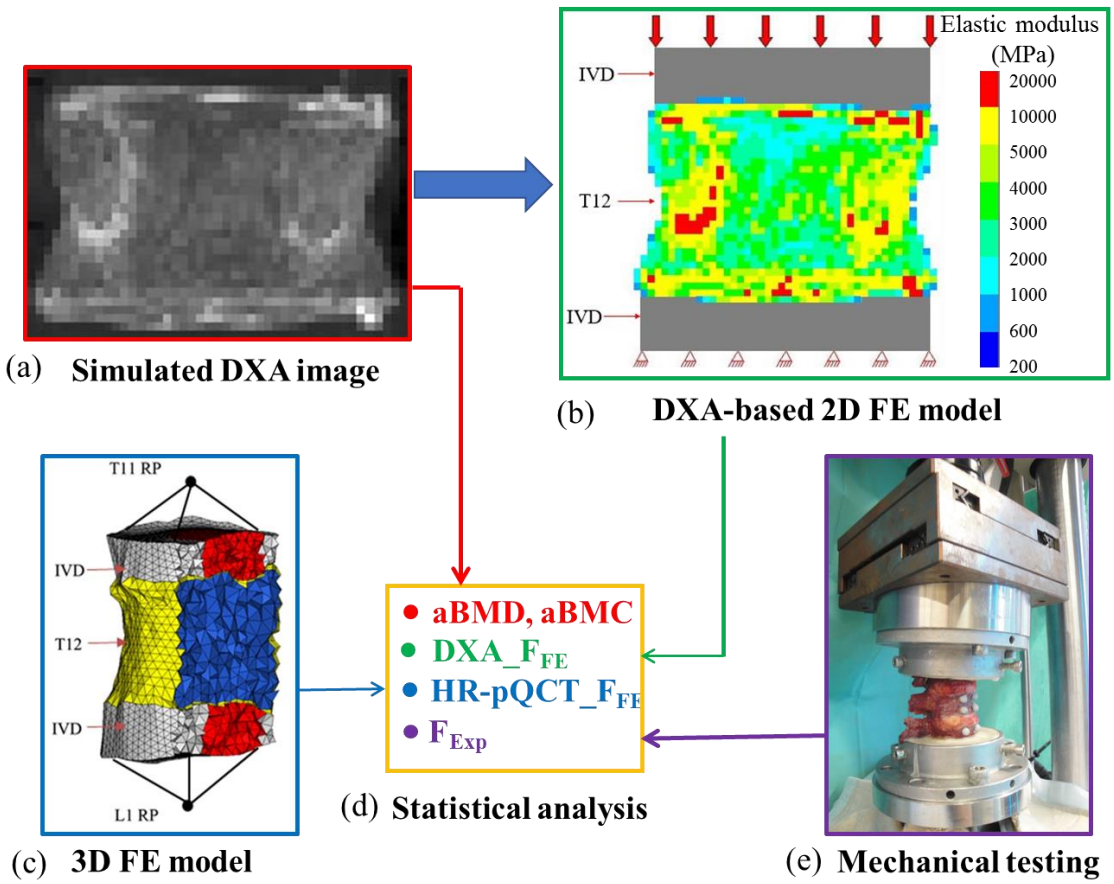

Fig. 1. Overview of the methods used in the present study: (a) aBMD and aBMC were calculated from the simulated DXA images of T12; (b) and (c) the simulated DXA-based 2D and the HR-pQCT-based 3D FE models were generated; (d) and (e) thirteen spinal segments (T11/T12/L1) were mechanically tested until failure (F_{Exp}) and statistical analysis was performed on these parameters.

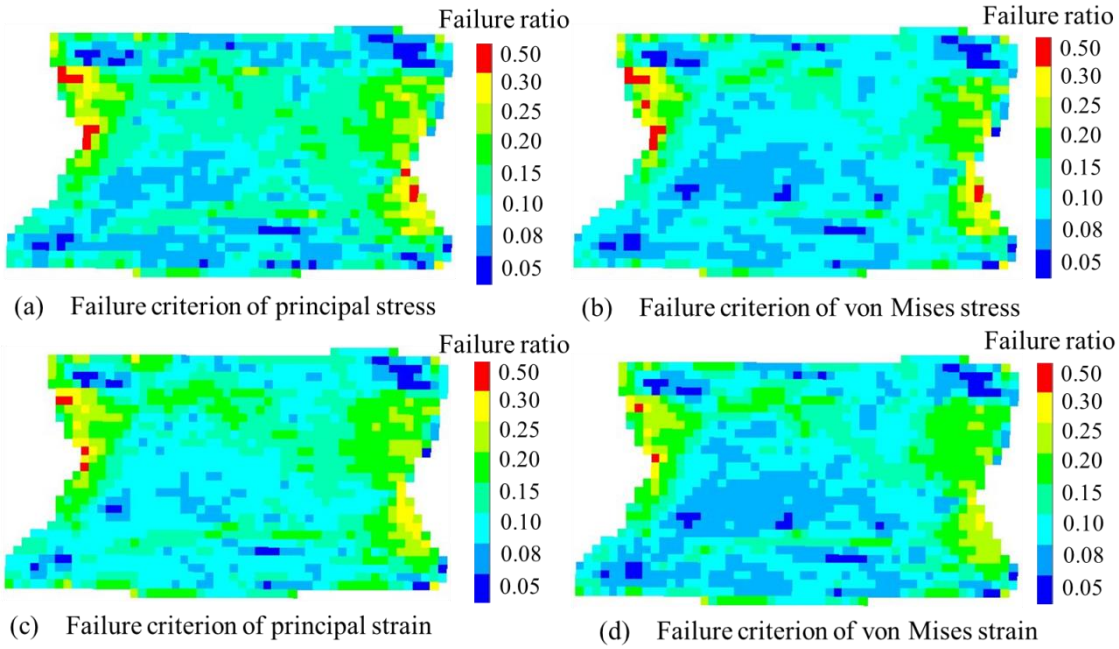

Fig. 2. Distribution of the failure ratios in the simulated DXA-based 2D FE models using different failure criteria: (a) principal stress, (b) von Mises stress, (c) principal strain and (d) von Mises strain.

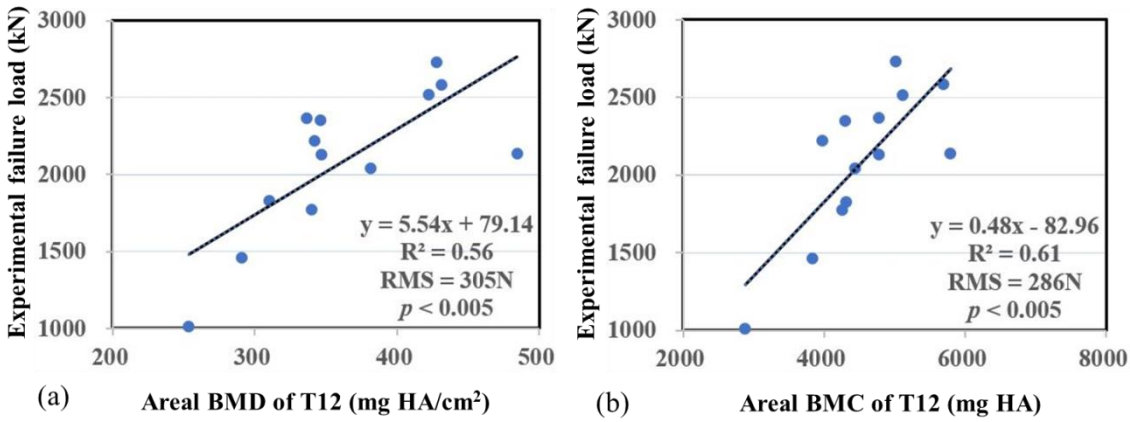

Fig. 3. Linear regressions of the experimentally measured failure loads of T12 as a function of (a) the aBMD of T12 and (b) the aBMC of T12.

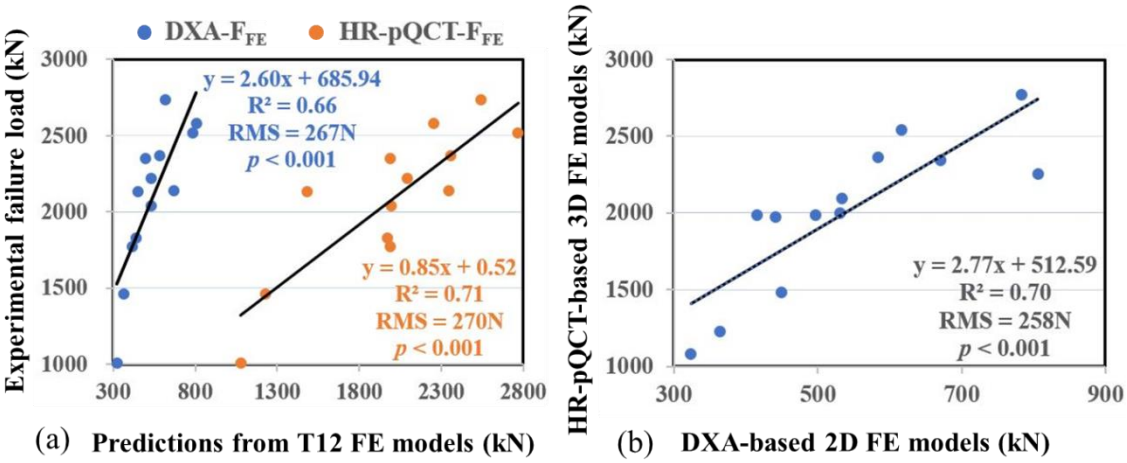

Fig. 4. (a) Linear regression of the experimentally measured failure loads of T12 as a function of the failure loads predicted by the FE models and (b) linear correlation between the failure loads of T12 predicted by the HR-pQCT-based 3D FE models (HR-pQCT-F_{FE}) and the simulated DXA-based 2D FE models (DXA-F_{FE}).

Table 1. Pearson's correlation coefficients (r) among the failure loads of T12 predicted by the simulated DXA-based 2D FE models using different failure criteria ($p < 0.001$)

	Principal stress	von Mises stress	Principal strain	von Mises strain
Principal stress	-	-		
von Mises stress	0.999	-		
Principal strain	0.999	0.999	-	
von Mises strain	0.997	0.995	0.995	-