STATISTICAL MODELS INFORMED BY DXA IMAGES SLIGHTLY OUTPERFORM T-SCORE IN THE PREDICTION OF HIP FRACTURE

Alessandra Aldieri (1), Margaret Paggiosi (2), Pinaki Bhattacharya (3), Richard Eastell (2), Alberto L. Audenino (1), Mara Terzini (1)

1. PolitoBIOMed Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Torino, Italy. 2. Department of Oncology and Metabolism University of Sheffield, UK. 3. Department of Mechanical Engineering and INSIGNEO Institute for in silico Medicine, University of Sheffield, UK.

Introduction

Fragility fractures due to osteoporosis represent a growing socio-economic burden as they cause a significant decrease in the quality of life, increase morbidity and mortality. As people live longer, the incidence of such fractures is increasing [1], which makes the need for an accurate fracture risk assessment tool urgent, in order to implement adequate preventive actions. In the clinical practice, osteoporosis severity is assessed based on T-score, derived from the areal Bone Mineral Density (aBMD) average measurement on DXA images. However, in spite of its correlation with fracture incidence, aBMD in not an accurate predictor of the risk of fracture [1]. Statistical models of femur shape and intensity built from CT scans have proved promising in improving hip fracture prediction [2]. Nevertheless, CT is not routinely performed for osteoporosis diagnosis, presenting a barrier to the adoption of such a methodology in clinical practice. In this context, this study aimed to assess the stratification accuracy of analogous shape and intensity statistical models developed from DXA images of the same cohort, which, instead, would be readily available.

Methods

DXA images of the proximal femur from a retrospective British postmenopausal women cohort [2] were employed to build Partial Least Square (PLS)-based statistical models. The cohort comprised 49 fracture and 49 control subjects, pair-matched for age, body weight and height. First, a Statistical Shape Model (SSM) was built through *Deformetrica* [3,4], which allowed to identify the main anatomical features of the population. Subsequently, the pixel-by-pixel BMD map was employed in order to build a Statistical Intensity Model (SIM). Eventually, the SSM and the SIM were merged into a unique Statistical Shape-Intensity Model (SSIM), gathering shape and density features together. PLS allowed identification of the space (PLS modes) of maximal covariance between the patients' shape and density features and their known fracture status. By projecting the original features onto the PLS modes, the PLS components could be obtained for shape, intensity and the two taken together. Logistic regression models based on the statistical models built were then implemented taking the PLS components as predictors of the subject-specific hip fracture status. A 10-fold cross-validation procedure was adopted to train and test the stratification accuracy of the predictive models in comparison with T-score.

Results

All the shape modes taken together could not explain more than 24% of the variability in the fracture risk, while 7 intensity modes explained more than 90% of it. A total of 31 shape-intensity modes explained 90% of the shape-intensity variability, whereas 3 modes, shown in Fig.1, were sufficient to explain 90% of the variability in the fracture risk.

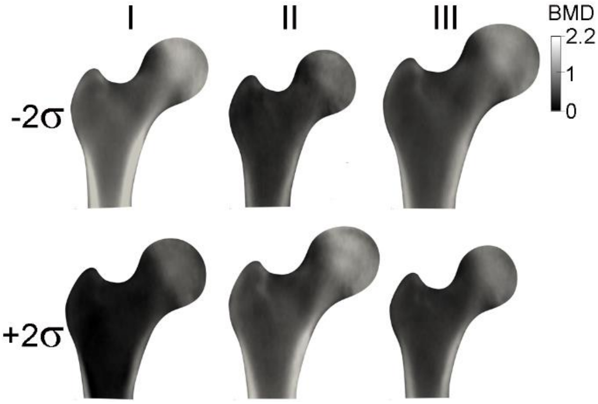

Figure 1: the first three SSIM modes shown as ±2σ deviations from the template. BMD is reported in g/cm^2.

The first three shape, intensity and combined components employed as predictors of the fracture risk in logistic regression models yielded Area Under Curve (AUC) values of 0.6, 0.77, 0.78 respectively. The shape-intensity model AUC was significantly higher ($p < 0.05$) than the T-score AUC, equal to 0.71.

Discussion

The aim of this study was to investigate whether hip fracture risk prediction could be improved taking advantage of all the information contained in a DXA scan, i.e., the proximal femur shape and the BMD pixel-by-pixel map. As expected, the stratification accuracy of the SIM was superior to that of the SSM. The inclusion of the shape within the SSIM did not improve significantly the SIM accuracy. As a whole, despite smaller than the increase obtained with CT-informed statistical models [3], a significant 7% increase in stratification accuracy was achieved with respect to T-score.

References

1. Carey et al., *Best Pract Res Clin Rheumatol*, 101775, 2022.
2. Aldieri et al., *ABME*, 50:211-221, 2022.
3. Durrleman et al., *Neuroimage*, 101: 35-49, 2014.
4. Aldieri et al., *Comput. Biol. Med*, 127: 104093, 2020.