



**Politecnico
di Torino**

Politecnico di Torino

Corso di Laurea Magistrale in Ingegneria Biomedica

A. A. 2021/2022

**DXA-based statistical shape-intensity
models for hip fracture prediction in
post-menopausal women**

Supervisori:

Prof. Cristina Bignardi

Dr. Alessandra Aldieri

Candidato:

Alessandro Celauro, 289337

Sessione di Laurea di Dicembre 2022

Abstract

Osteoporosis is a bone disease caused by an imbalance between bone deposition and resorption which affects bone microstructure and leads to a decreased bone mass. According to recent studies, over 200 million people worldwide have osteoporosis. There are approximately 9 million fractures worldwide per year due to osteoporosis, among which 1.6 million are hip fractures (75% of which affect women). Moreover, it is estimated that 1 in 3 females and 1 in 5 males over the age of 50 will have an osteoporotic fracture. Osteoporosis is particularly frequent in post-menopausal women for hormonal reasons. The current gold standard to diagnose osteoporosis is measuring the areal Bone Mineral Density (aBMD) through dual-energy X-ray absorptiometry (DXA); aBMD is then used to calculate the so-called T-score, an indicator based on the comparison between the screened subject's BMD and the BMD of young females aged in the range of 20-29 years. The World Health Organization (WHO) has chosen T-score as the ultimate parameter to discriminate osteoporotic subjects from non-osteoporotic ones; however, its performances in the fracture prediction field are limited: it is estimated that about one half of the subjects not classified as osteoporotic by the T-score did in fact fracture. Several alternatives to T-score have been investigated, like epidemiological models, structural parameters and FE models, all with their own limitations. Thus, more attention has been given to the so-called statistical models. This work was based on the development of the afore-mentioned statistical models starting from DXA images, which would be already clinically available. The aim of the work was to assess if they could support an enhanced identification of the risk of hip fracture with respect to the gold standard T-score. A retrospective cohort of 97 British post-menopausal women was herein considered. For these subjects, both CT and DXA images were available. Out of the 97 patients, 49 subjects had experienced a femoral fracture, while 48 had not. For each subject, the DXA image was segmented to extract the proximal femur's shapes. Hence, Statistical Shape Models (SSMs) based on Principal Component Analysis (PCA) and Partial Least Square (PLS) could be built. PCA aims to find new variables that maximize the feature's variance, while PLS maximizes the covariance with an external variable (which, in this work, is the fracture status). Furthermore, also PCA- and PLS-based Statistical Intensity Models (SIMs) were built, taking advantage of the available local aBMD maps. Then, the outcomes of the two kinds of model were combined and further processed through PCA and PLS to create two Statistical Shape-Intensity Models (SSIMs). The calculated PLS modes highlight that neck length, size and inclination are the shape features associated with fracture, while, as for intensity, a generally decreased density throughout the femur and a decreased cortical thickness are the main features associated with fracture. The logistic regression models built using the whole cohort as training set show that, in terms of Receiver Operating Characteristic (ROC) curves and Area Under Curve (AUC), the PLS-based SSIM show classification performances ($AUC \approx 0.85$) improved with respect to aBMD ($AUC \approx 0.74$). Eventually, the 2D-based models were compared to the 3D-based ones.

Contents

1. Introduction.....	1
1.1 Macrostructure and microstructure of human bone tissue	1
1.2 Osteoporosis: Pathophysiology, etiology and epidemiology	2
1.3 Osteoporosis: Diagnosis and available treatments	2
1.4 Osteoporosis: Fractures and fracture prediction	4
2. Materials and methods	7
2.1 Study population	7
2.2 Shape extraction	7
2.3 Deformetrica	9
2.4 Density extraction.....	11
2.5 PCA and PLS.....	12
2.5.1 Principal Component Analysis (PCA)	12
2.5.2 Partial Least Square (PLS)	13
2.5.3 Using PCA and PLS to build Statistical Models.....	13
2.6 SSMs, SIMs, SSIMs.....	14
2.6.1 SSMs.....	14
2.6.2 SIMs	14
2.6.3 SSIMs.....	14
2.7 Stratification accuracy assessment	15
3. Results	15
3.1 Statistical Shape Models	15
3.1.1 PCA	16
3.1.2 PLS.....	18
3.2 Statistical Intensity Models	20
3.2.1 PCA	20
3.2.2 PLS.....	21
3.3 Statistical Shape-Intensity Models.....	23
3.3.1 PCA	23
3.3.2 PLS.....	25
3.4 Accuracy assessment.....	27
4. Discussion	29
5. Conclusions.....	30
Bibliography	32

List of Figures

Figure 1: DXA raw image (left) and BMD map (right) of the same subject. Notice how part of the femur, present in the raw image, is cut out from the BMD map, hence the need to cut the segmentations too 8

Figure 2: overlapping of the mid-point of the femur profiles distal edges (left); overlapping of the shaft axes (right) 9

Figure 3: plot of the template (black) and a patient-specific contour (blue) reconstructed through the control points (red) and the patient-specific moment vectors (violet) 11

Figure 4: meshed template (left) and a zoomed-in part of said mesh (right) 12

Figure 5: Schematic representation of the process: shape (red) and intensity (blue) information is used to create statistical models through PCA and PLS and the corresponding components (t) are calculated; then, components are combined to build two more models based on shape and intensity together (purple); in the end, from the six models' components, six predictive models (PM) are built 15

Figure 6: Cumulative percentage of variance explained by the PCA shape modes 16

Figure 7: first 5 PCA shape modes shown as deformations of the template along each mode ... 17

Figure 8: Scatter plots of the first two (left) and three (right) PCA shape components for all the subjects. Fractured patients are shown in red, non-fractured patients in green 17

Figure 9: Cumulative percentage of (both shape's and fracture status') variance explained by the PLS modes 18

Figure 10: first 6 PLS shape modes shown as deformations of the template along each mode .. 19

Figure 11: Scatter plots of the first two (left) and three (right) PLS shape components for all the subjects. Fractured patients are shown in red, non-fractured patients in green 19

Figure 12: Cumulative percentage of variance explained by the PCA intensity modes 20

Figure 13: first 4 PCA intensity modes shown as variations of the template's Bone Mineral Density (BMD) along each mode 21

Figure 14: Scatter plots of the first two (left) and three (right) PCA intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green 21

Figure 15: Cumulative percentage of (both intensity's and fracture status') variance explained by the PLS intensity modes 22

Figure 16: first 6 PLS intensity modes shown as variations of the template's Bone Mineral Density (BMD) along each mode 22

Figure 17: Scatter plots of the first two (left) and three (right) PLS intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green 23

Figure 18: cumulative percentage of variance explained by the PCA modes 24

Figure 19: first 6 PCA shape-intensity modes shown as both deformations of the template and variations of its Bone Mineral Density (BMD) along each mode 24

Figure 20: Scatter plots of the first two (left) and three (right) PCA shape-intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green 25

Figure 21: Cumulative percentage of (both shape-intensity's and fracture status') variance explained by the PCA shape-intensity modes 26

Figure 22: first 6 PLS shape-intensity modes shown as both deformations of the template and variations of its Bone Mineral Density (BMD) along each mode 26

Figure 23: Scatter plots of the first two (left) and three (right) PLS shape-intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green 27

Figure 24: ROC curves for PCA models and Femoral Neck BMD (left) and PLS models and Femoral Neck BMD (right). Models built starting from 2D DXA images 27

Figure 25: ROC curves for PCA models (left) and PLS models (right). Models built starting from 3D CT scans 28

List of Tables

Table 1: Advantages of central DXA. Adapted from [10]..... 4

Table 2: AUC values for the logistic regression models built with the 2D and 3D results 29

1. Introduction

1.1 Macrostructure and microstructure of human bone tissue

The skeletal system provides support and protection to the human body and its soft tissues and, together with the muscular system, allows its movements. Bones have different shapes, dimensions and functions but they are all made up of two kinds of bone tissue:

- Cortical bone: or compact bone, it is the densely textured tissue that provides strength and rigidity to the bone and gives it its typical aspect, usually being the outer shell that encloses trabecular bone. Externally covered by a collagenous layer called periosteum.
- Trabecular bone: also known as cancellous or spongy bone, it is a network of bars and plates (known as “trabeculae”) that supports the outer shell of cortical bone. With its porous structure, trabecular bone also reduces the overall weight of the skeleton. Covered by a cellular layer called endosteum.

Proportions and disposition of these two structures of the bone tissue change between different kinds of bones (long, short, flat), so that each bone has the mechanical properties required for its functions [1].

From a microscopic point of view, the bone tissue is a mineralized collagenous extracellular matrix and the specialized cells it surrounds. These cells are:

- Osteoblasts: protein-secreting cells responsible for the synthesis, deposition and mineralization of the bone matrix.
- Osteocytes: osteoblasts encased in the same matrix they produced cannot divide or secrete new matrix and are called osteocytes. These abundant and long-lived cells form a network through which they act as mechanosensors, thus guiding the bone remodeling process [2].
- Bone lining cells: quiescent osteoblasts that cover the bone surfaces are called bone lining cells; these cells can reacquire their secretory activity if the bone physiological status requires it. They also limit the activity of osteoclasts [2].
- Osteoclasts: multinucleated cells located in resorption bays, osteoclasts remove bone (both the mineral part and the organic matrix) during growth or remodeling (bone resorption) [2].

These cells all work in synergy to build and maintain the bones: despite the mineral components, the bone tissue is alive, vascularized and innervated and capable of self-repairment and self-renewal; these processes are collected under the definition of “bone remodeling”, that is the process through which osteoclasts and osteoblasts respectively demolish and rebuild the bone in order to repair microdamages (typical of stiff materials) and adapt its mass and architecture to the loads it is subjected to (both in terms of directions and magnitude). Through this process, approximately 10% of bone is replaced each year in adults. Though adult bone shape and mass are partially determined by genetics, bone remodeling is guided mainly by mechanical loads on the bone, but also age and hormones play a key role in regulating the process [1].

1.2 Osteoporosis: Pathophysiology, etiology and epidemiology

When there is balance between bone resorption and bone deposition, bone remodeling maintains an efficient and mechanically optimized skeletal structure; but when this balance is compromised, several disfunctions and bone diseases can occur. One such disease is osteoporosis: defined as low bone mineral density caused by an altered bone microstructure, osteoporosis generates from an increase in bone resorption that leads to a decreased skeletal mass [3].

Osteoporosis can be primary or secondary:

- primary osteoporosis is related to aging and decreased sex hormones; bones' microarchitecture deteriorates, leading to loss of bone mineral density and increased risk of a fracture
- secondary osteoporosis is caused by other diseases or by some medications like glucocorticoids and anti-epileptics and is more frequent in men than in women [3]

Increasing age, low bodyweight, smoking, family history of osteoporosis, white or Asian race, early menopause, low levels of physical activity, and a personal history of a fracture are risk factors for osteoporosis, while osteoporosis itself increases the risk of a fracture. Black males and females have less osteoporosis than their white counterparts, but those diagnosed with osteoporosis have similar fracture risks [3].

According to recent studies [3], [4], over 200 million people worldwide have osteoporosis, and it is more common in females than in males: in the developed world, 2% to 8% of males and 9% to 38% of females are affected.

1.3 Osteoporosis: Diagnosis and available treatments

A physical exam rarely reveals any changes until osteoporosis is quite advanced and fractures occur [3]. That is why the World Health Organization (WHO) has analyzed the available techniques for measuring bone mass or bone loss in order to find the most suitable one for osteoporosis screening and assessment of fracture risk [5].

As of today, osteoporosis is diagnosed radiographically by measuring the bone areal mineral density (aBMD), in terms of grams of mineral per square centimeter (g/cm^2) [6], through dual-energy X-ray absorptiometry (DXA) [7]: DXA instruments generate X-rays at two different energies and calculate the bone mineral content in the scanned region (mainly spine, hip or forearm) through the differential attenuation of the X-ray beam at these two energies, all with a small effective dose [8].

The measured BMD at the femoral neck and lumbar spine is then used to calculate the so-called T-score and Z-score:

- T-score: defined as the difference between the patient’s BMD and mean BMD of young females aged in the range of 20-29 years (divided by the standard deviation of the reference population)

$$T\text{-score} = \frac{\text{Measured BMD} - \text{Young adult mean BMD}}{\text{Young adult population SD}}$$

According to the WHO, when the patient’s BMD is 2.5 standard deviations (SD) or more below the average value for young healthy women (which means T-score < -2.5), the patient is diagnosed with osteoporosis, while a T-score between -1 and -2.5 means osteopenia (lower BMD than normal, but not osteoporotic yet); finally, the patient is considered “normal” with a T-score higher than -1. T-score is used for postmenopausal women and men aged 50 years and older [6] [9].

- Z-score: similar to the T-score, it is obtained by comparing the patient’s BMD to a particular age, sex and ethnicity-matched adult reference population.

$$Z\text{-score} = \frac{\text{Measured BMD} - \text{Age matched mean BMD}}{\text{Age matched population SD}}$$

Patients with a Z-score < -2.0 are considered as “low bone mineral density for chronological age” or “below the expected range for age”, while those with Z-score > -2.0 are defined as “within the expected range for age”. Z-score is recommended by the International Society for Clinical Densitometry (ISCD) for premenopausal women, men less than 50 years of age and children [6] [9].

In healthy individuals without risk factors, experts recommend starting to screen women at the age of 65 years of age and men at the age of 70, but patients with risk factors or a high score on an osteoporosis risk assessment test should be screened sooner [3].

Pharmacological treatments of osteoporosis aim to reduce the risk of fracture by rebalancing bone deposition and bone resorption. According to [10], the available drugs can be classified in three groups based on their mechanisms of action:

- Anti-resorptive agents: these drugs reduce the activity of osteoclasts through different means in order to tip the scales in favor of bone deposition.
- Bone-formation drugs: PTH and PTHrP activate a signaling pathway that increases bone formation by promoting the differentiation and growth of osteoblasts (anabolic effect).
- Dual-action drugs: anti-sclerostin antibody removes sclerostin, an inhibitor of a signaling pathway, promoting bone formation and inhibiting bone resorption at the same time.

These drugs can be administered simultaneously (combination therapies) or sequentially (sequential therapies) [10].

1.4 Osteoporosis: Fractures and fracture prediction

The choice of central DXA (that is DXA scans of the central skeleton to measure BMD of the lumbar spine and hip [9]) among the many available techniques is based not only on its many advantages (see Table 1), but also (and foremost) on its good performances in fracture prediction: epidemiological studies [11] have demonstrated DXA's predictive capabilities.

Table 1: Advantages of central DXA. Adapted from [10]

Consensus that bone mineral density results can be interpreted using WHO T-scores
Proven ability to predict fracture risk
Basis of new WHO algorithm for predicting fracture risk
Proven for effective targeting of antifracture treatments
Good precision
Effective at monitoring response to treatment
Acceptable accuracy
Stable calibration
Effective instrument quality control procedures
Short scan times
Rapid patient set up
Low radiation dose
Availability of reliable reference ranges

However, being the best among the currently available techniques may not be enough: the study reported in [12] shows that T-score alone is not powerful enough to predict accurately whether a woman will or will not have a hip fracture, with more than half of the fractured patients (54%) having a T-score higher than -2.5 (which, according to WHO, means these patients were not osteoporotic when the fracture occurred). Refining the fracture prediction method is in fact a priority: worldwide, there are approximately 9 million fractures per year as a result of osteoporosis, among which 1.6 million are hip fractures (75% of which affect women) and it is estimated that 1 in 3 females and 1 in 5 males over the age of 50 will have an osteoporotic fracture. Women's, and more specifically postmenopausal women's, situation is more critical, since the reduced ovarian function that follows menopause and the consequent reduction of the estrogen levels lead to osteoporosis [1]. These fractures may have severe consequences: fractures of the spine or hip often lead to functional decline, disability (and thus reduced autonomy), chronic pain, reduced quality of life and increased morbidity and mortality; furthermore, individuals who have sustained a

fracture have about double the risk of a subsequent fracture as compared to the fracture-free population [4].

Overcoming T-score's limitations is the reason behind the investigation and, in some cases, adoption of several different methods and tools for fracture prediction.

The fracture risk assessment tool called FRAX and the Garvan fracture risk calculator are just two examples of epidemiological models in the field of fracture prediction:

- FRAX: developed by the University of Sheffield in collaboration with the World Health Organization, it estimates the individualized 10-year probability of hip and major osteoporotic fracture (hip, clinical spine, distal forearm or proximal humerus) by integrating eight clinical risk factors (prior fragility fracture, parental hip fracture, smoking, systemic glucocorticoid use, excess alcohol intake, body mass index, rheumatoid arthritis and other causes of secondary osteoporosis), age and sex. It is worth noticing that the BMD at the femoral neck is an optional input variable (which generally improves the performances of the tool) [13]. But even FRAX has its limitations: being an epidemiological model, it is highly population-dependent (both for calibration and validation), and it is considered a too-simple model which does not consider factors like the dose of exposure to many drugs, the lumbar spine BMD and the material or structural properties of bone [13]. Despite its limits, FRAX has been actually adopted in clinic prediction and it is constantly updated in terms of algorithms and additional risk factors [14].
- Garvan fracture risk calculator: similar to FRAX, it includes sex, age, prior fracture since age 50 (excluding major trauma), falls in last 12 months and weight (kg) to provide the predicted absolute risk of hip fracture and any osteoporotic fracture over 5- and 10-year time horizons. Although it incorporates fewer risk factors than the FRAX, it includes falls and counts the number of previous fractures [15]. Developed in Australia, it is still an epidemiological model and, as such, highly population dependent. However, several studies have contributed to assess its usefulness in other countries and its relative performance among other fracture predictors [15] [16].

To eliminate the high population dependency of these models, specific properties of the bone, like microstructure and geometry, have been investigated. Some of the chosen parameters are:

- Trabecular Bone Score (TBS): it is a bone texture index derived from lumbar spine DXA images [17] and carries information about the quality of trabecular bone microarchitecture [18]; multiple studies have demonstrated an association between reduced TBS and increased fracture risk [17] and the use of this tool has been already authorized in Europe, USA and Brazil, and yet not all the clinics that work with BMD have adopted it [18].
- Bone Material Strength index (BMSi): through a device called OsteoProbe, the Bone Material Strength index, in terms of resistance of cortical bone to microindentation, can be assessed. According to some studies, the cortical component seems to be an important determinant of bone strength, but the actual value of the BMSi as a fracture risk predictor is still debated and investigated upon [19].

- Hip Structural Analysis (HSA): it consists of a set of parameters which describe the proximal femur geometry and, like the TBS, are extracted from DXA images. However, HSA is not currently implemented because of its limitations: HSA-derived parameters represent discrete measures unable to completely describe femur shape, are highly cross-correlated and correlated with the BMD and T-score [20].
- Finite Element Method (FEM): an increasing computing power means more and more accurate FEM models are possible; by using CT scans, the proximal femur's 3D shape can be reconstructed, meshed, given accurate material properties and used to simulate different scenarios, from physiological loading conditions to traumas and falls. In [20], CT-derived 3D FEM models of proximal femurs are used to calculate patient-specific Femoral Strength (FS) values as the estimated failure loads in a sideways fall condition. However, CT is not routinely performed to assess osteoporosis because of its higher radiation dose compared to DXA, so the application of 3D FEM models and simulations in clinics is highly improbable.

In the last decades, attention in the fracture prediction field has shifted toward statistical models: by including information about the shape and/or the pointwise bone density, these models can use both geometry and material properties of the bone for prediction; statistical models basically consist of the average shape and/or intensity of the population and of the modes of variation of shape and/or intensity from those average values, so that each element of the training set can be described as the contribution of each mode (or at least the most relevant ones) to its deviation from the average [21].

These models can be classified as:

- Statistical Shape Models (SSM): they use the information about the shape of the bone to build a prediction model.
- Statistical Intensity Models (SIM): they use the information about the texture of the bone image (in terms of grayscale, BMD or other) to build the model.
- Statistical Shape-Intensity Models (SSIM): these models combine the information about shape and intensity in a single model in order to improve its performances. Studies building this kind of models [22], [23], [24] generally show that SSIMs are better fracture predictors than their SSMs and SIMs counterparts.

Several studies have already shown the high potential of statistical models: in [25] the proximal femur shape is obtained from 2D radiographic images and used to build a predictive model, then the results are compared to FRAX and BMD, with some promising outcomes; other works [26], [23], [24], [27] built SSMs and SIMs using 3D tomographic images and finite elements analysis [23] or 2D DXA images [26], [24], [27].

In the end, the aim is always the same: to create a model capable of predicting the fracture of subjects which were not part of the training set; in order to do so, the models must be fitted to new, unseen sets of images [21].

In study [20] the DXA images (2D) of 28 post-menopausal women are used to build SSMs, SIMs and SSIMs and their CT scans (3D) to build FE models to calculate their femoral strength as a comparison, since the actual fracture status of the subjects was unknown. Study [23] has a much larger cohort and fracture status of its subjects is known: three models (SSM, SIM, SSIM) are built and validated starting from a training set based on 3D CT images of 100 post-menopausal women and compared to aBMD (the actual gold standard); results show that the SIM and SSIM outperform aBMD and that SSIM has a higher predictive power than both SIM and SSM. Alas, as already mentioned, CT scans are not part of the clinical practice in osteoporosis diagnosis, hence the resistance in adopting this kind of models.

Statistical models based on 2D DXA scans could be a valid alternative that, by using the same data from the gold standard exam for osteoporosis diagnosis and requiring a lower computational power, would have a wider and easier adoption by clinics. The aim of this thesis work is to assess whether these 2D-based statistical models have lower or comparable performances to the 3D-based ones and thus if they can be a promising alternative to aBMD and T-score in the field of fracture prediction: three kinds of model (SSM, SIM, SSIM) are built from the 2D DXA scans of the same cohort of study [23] and the results, in terms of fracture prediction, are compared to those of the CT-based models of that same study.

2. Materials and methods

2.1 Study population

The study population chosen for this work was originally put together for another study [28] and then used also for studies [23] and [27]: 101 white women with and age ≤ 90 years and at least 5 years post-menopause were recruited; among these subjects, 49 had sustained a hip fracture following a low-energy trauma (the majority of fractures in post-menopausal women are low-trauma fractures and fall into the osteoporotic category [4]) less than 90 days prior to the recruitment, while 48 were chosen among the non-fractured population as control, matching the fractured ones in terms of body mass index (BMI), age, height and weight. Both QCT (Quantitative Computed Tomography) and DXA scans were performed on the subjects (the contralateral hip was scanned for the fractured ones). 4 subjects were excluded due to a lack of information about the fracture status, leaving a total of 97 subjects available for further analysis. Further information about this cohort is available in [28].

2.2 Shape extraction

Both DXA raw images and BMD maps of the proximal femur were available for all the subjects. In [27], the proximal femur's shape information is extracted from the DXA raw images by means of a semi-automatic interactive segmentation through the software Matlab; these segmentations

were made available for the current work but, since in this work the intensity (here meant in terms of BMD) is considered as well and the BMD maps are cut differently from the raw images (see Figure 1), some adjustments were made. In particular, most of the segmentations were just cut and rebuilt to fit the BMD maps, while a few of them were remade completely following the same procedure as in [27] (but working directly on the BMD maps, not on the raw images).



Figure 1: DXA raw image (left) and BMD map (right) of the same subject. Notice how part of the femur, present in the raw image, is cut out from the BMD map, hence the need to cut the segmentations too

Then, as required by the next step, the obtained contours were aligned; two possible alignments were explored:

- The first alignment was a simple overlapping of the mid-point of the femur profiles distal edges of all the contours through a rigid translation (see Figure 2)
- For the second alignment, an approximation of the shaft axis was found for every contour, which were then translated and rotated so that all the axis coincided (see Figure 2)

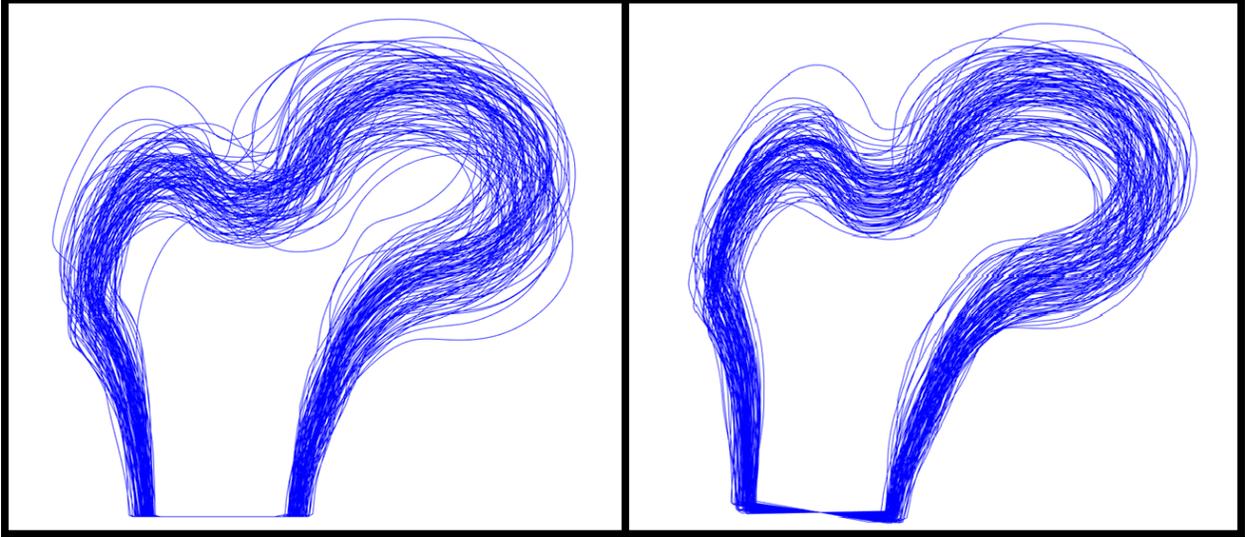


Figure 2: overlapping of the mid-point of the femur profiles distal edges (left); overlapping of the shaft axes (right)

In the end, the second alignment was chosen because it is based on an anatomical feature.

The overlapped shapes represent one of the inputs given to the software Deformetrica.

2.3 Deformetrica

Before statistical shape analysis can be carried out, shapes need to be represented in the same, coherent way. One such way is landmarking, which means locating a finite number of points (called landmarks) on the outline; landmarks should match between and within all the shapes. Nevertheless, the choice of landmarks positioning can be tricky: the manual choice of landmarks becomes impossible with relatively big datasets, while automatic positioning is not always possible. [29]

However, landmarking can be avoided: in this work, the mathematical framework called Deformetrica and developed by Durrleman et al. [30] [31], which uses currents (and not points) to describe shapes, was adopted. The current of a generic curve is the flux of a test vector field across that curve; this vector field ω varies in a vector kernel Hilbert space W generated by a Gaussian Kernel K_W with width λ_W , so that

$$K_W(x, y) = e^{-\frac{|x-y|^2}{\lambda_W^2}}$$

As ω varies, the current changes as well and these changes uniquely define the shape of the curve. This current can be seen as an infinite set of Dirac delta currents (one for each point of the curve) tangent to the curve outline; however, since the curve can also be described through a finite set of points, not all the infinite currents must be considered, and the overall flux can be approximated by a finite subset of currents:

$$S^i = \sum_k \delta_{x_k^i}^{a_k^i}$$

where $\delta_{x_k^i}^{a_k^i}$ is a Dirac delta current concentrated at the center of the curve segment x_k , while a_k^i is the tangent to the line.

Now that all the shapes are described in terms of currents, the forward modelling approach [32] proceeds with an alternate two-step minimization strategy [30] for the calculation of a mean shape, called template \bar{T} , and the functions Φ_i which deform the template into each i^{th} shape T^i of the initial set [30] according to:

$$T^i = \Phi^i \cdot \bar{T} + \varepsilon^i$$

where ε_i are some residuals accounting for features not captured by the deformation.

The function Φ_i is parametrized by a time-varying velocity field uniquely characterized by an initial vector speed v_0^i :

$$v_0^i(x) = \sum_k K_V(x_k, x) \beta_{x_k}^i$$

which, like ω , also belongs to a reproducible kernel Hilbert space V with kernel

$$K_V(x_k, x) = e^{-\frac{|x_k - x|^2}{\lambda_V^2}}$$

for any pair (x_k, x) where λ_V is the kernel width, the x are the nodes of the polygonal line and x_k the point position of the template Delta currents (also called control points). $\beta_{x_k}^i$ are the moment vectors centered at the position x_k of the template delta currents and $\Phi^i = f(\beta^i)$, so it is the moment vectors that drive the transformations of the template towards each shape and thus contain the specific shape information.

The open source code of Deformetrica, available at <http://www.deformetrica.org/>, takes the kernel widths λ_W and λ_V and the overlapped shapes as inputs and gives back the template, the control points, the moment vectors and the reconstructed shapes (as deformations of the template) as outputs. The input parameters control the resolution of the shape representation (λ_W) and the rigidity of the template deformation (λ_V) and can be chosen through a sensitivity analysis which, for this cohort, had already been performed by [27].

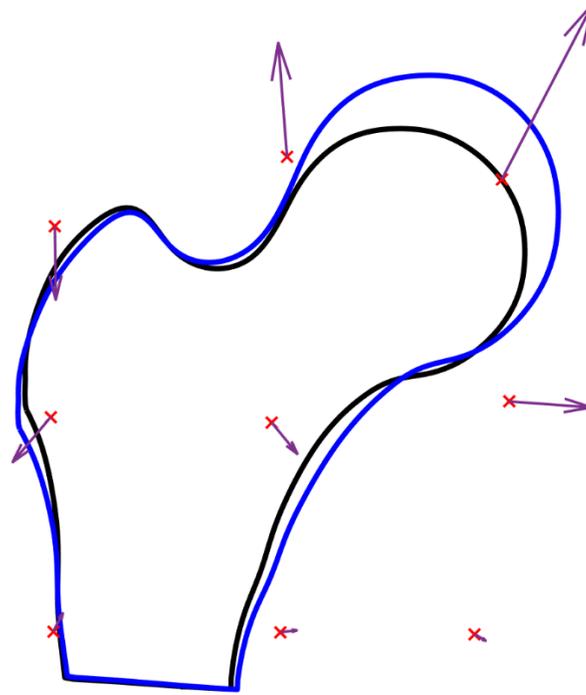


Figure 3: plot of the template (black) and a patient-specific contour (blue) reconstructed through the control points (red) and the patient-specific moment vectors (violet)

The moment vectors, which carry the shape information, are used to build the SSMs and the SSIMs.

2.4 Density extraction

Deformetrica's output is also useful for the extraction of BMD from the BMD maps: just like for the shape, density must be represented in the same way for all the subjects and simply taking the values of the pixels of the BMD maps is not ideal, since each femur has different shape and dimension and, thus, a different number of pixels; to overcome the issue, the template calculated by Deformetrica is given a triangular mesh (side $d = 0.5$ mm, Figure 4) through the software Hypermesh and then the same moment vectors given by Deformetrica are used to deform the template's mesh so that it fits every individual subject. In this way, every subject's mesh has the exact same number of nodes (and thus elements) of all the others, since they are just the same nodes but translated according to the moment vectors that hold the shape information of the subject.

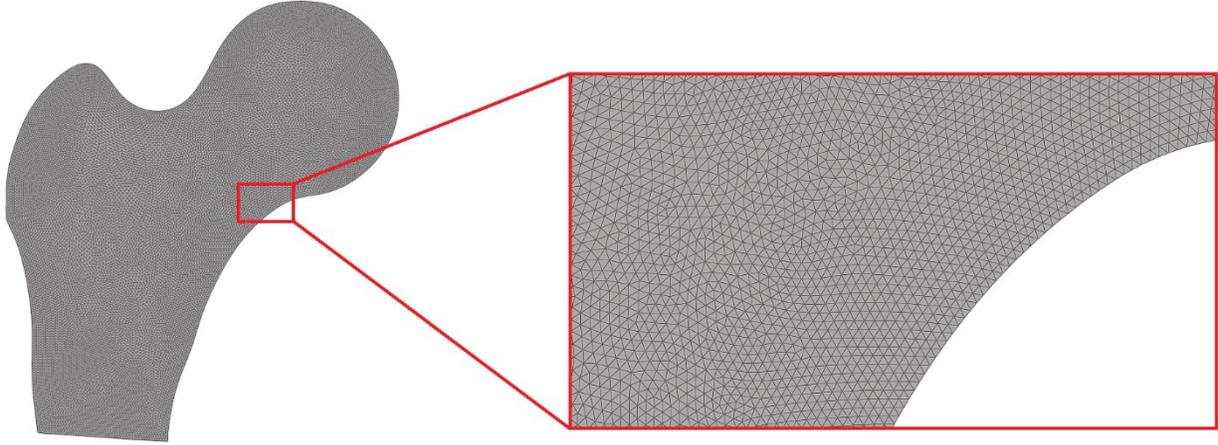


Figure 4: meshed template (left) and a zoomed-in part of said mesh (right)

The subject-specific mesh is then overlapped to the corresponding BMD map and each node is assigned with the BMD value of its closest pixel of the map. In order to avoid nodes with a too-low BMD value (artifacts, background pixels), a threshold is established and the nodes whose BMD is lower than the threshold are assigned the average BMD of the neighboring over-threshold nodes; this process is repeated as many times as it takes to eliminate all the under-threshold nodes.

Finally, each element of the mesh is assigned the average BMD of its three nodes.

The elements' BMD values are used to build the SIMs and the SSIMs.

2.5 PCA and PLS

As previously stated, statistical models consist of the average shape (SSM), intensity (SIM) or both (SSIM) of the dataset and of the main modes of variation of shape and/or intensity from those average values, so that each element of the training set can be described as the contribution of each mode (or at least the most relevant ones) to its deviation from the average [21]. In this work, the main modes of variation are calculated through two statistical techniques.

2.5.1 Principal Component Analysis (PCA)

This method aims to find new variables that are linear functions of those in the original dataset, that maximize variance and that are uncorrelated with each other.

In this work, the dataset of subjects and their moment vectors/elements' BMD values can be arranged in a $n \times p$ matrix X , where n is the number of observations (subjects) and p is the number of variables (moment vectors components/elements' BMD values), so that each column of X is a vector x_j of observations of the j^{th} variable. Xa is a linear combination of X 's columns, where a is a vector of constants, and maximizing its variance means finding the vector a that maximizes $var(Xa) = a'Sa$, where $S = \frac{X'^*X^*}{n-1}$ is the sample covariance matrix of X^* and X^* is the

matrix of the centered variables $\mathbf{x}_j^* = \mathbf{x}_{ij} - \text{mean}(\mathbf{x}_{ij})$. By imposing \mathbf{a} to be a unit-norm vector (so that $\mathbf{a}'\mathbf{a} = 1$), the problem becomes equivalent to maximizing $\mathbf{a}'\mathbf{S}\mathbf{a} - \lambda(\mathbf{a}'\mathbf{a} - 1)$, where λ is a Lagrange multiplier. Differentiating with respect to the vector \mathbf{a} , and equating to the null vector, produces the equation $\mathbf{S}\mathbf{a} - \lambda\mathbf{a}$, which is equivalent to $\mathbf{S}\mathbf{a} = \lambda\mathbf{a}$. This is now an eigenvector/eigenvalue problem, where \mathbf{a} is an eigenvector and λ the corresponding eigenvalue of the covariance matrix \mathbf{S} [33]. The matrix \mathbf{X} is thus decomposed as

$$\mathbf{X} = \mathbf{B}\mathbf{A}' + \mathbf{R}$$

where \mathbf{B} is the PCA scores (or components) matrix, \mathbf{A} is the PCA loadings (the modes) matrix and \mathbf{R} is the matrix of residuals.

Since the eigenvalues are the variances of the linear combinations defined by the corresponding eigenvector \mathbf{a} ($\text{var}(\mathbf{X}\mathbf{a}) = \mathbf{a}'\mathbf{S}\mathbf{a} = \lambda\mathbf{a}'\mathbf{a} = \lambda$), the eigenvectors associated to the higher eigenvalues explain most of the variance and a subset (and this is where the dimensionality reduction happens) of such eigenvectors can be chosen as a new orthonormal set of vectors to represent the data with. [33]

2.5.2 Partial Least Square (PLS)

The ultimate goal of this work is to use such subset of modes to predict the fracture of new subjects, but PCA does not consider the fracture status of the dataset's subjects and only calculates how much of the variability of the dataset's feature (shape and/or intensity) is explained by each mode. That is why Partial Least Square (PLS) could be the statistical technique that fits better the aim of this work: the modes are calculated considering also an external response variable (which, in this work, is always the fracture status). $\mathbf{X}^* = \mathbf{X}_C$ is still the matrix of centered variables (called predictors), while a second matrix \mathbf{Y}_C (called response) represents the external response variable; the space of maximal covariance between them is computed, which means that the weights vectors \mathbf{r} and \mathbf{s} are calculated so that $\max_{|\mathbf{r}'|, |\mathbf{s}'|=1} \text{cov}(\mathbf{X}_C\mathbf{r}, \mathbf{Y}_C\mathbf{s}) = \max_{|\mathbf{r}'|, |\mathbf{s}'|=1} \text{var}(\mathbf{X}_C\mathbf{r})\text{corr}(\mathbf{X}_C\mathbf{r}, \mathbf{Y}_C\mathbf{s})^2\text{var}(\mathbf{Y}_C\mathbf{s})$. The PLS modes \mathbf{t} and \mathbf{u} are then found by weighting \mathbf{X}_C and \mathbf{Y}_C by \mathbf{r} and \mathbf{s} within an iterative process which loops until all the PLS modes are extracted and where, at each step, the variance explained by \mathbf{r} and \mathbf{s} is removed from the original data (deflation). In the end, the two starting matrices are decomposed as

$$\begin{aligned}\mathbf{X}_C &= \mathbf{T}\mathbf{P}' + \mathbf{E} \\ \mathbf{Y}_C &= \mathbf{U}\mathbf{Q}' + \mathbf{F}\end{aligned}$$

where \mathbf{T} and \mathbf{U} are the PLS scores (or components) matrices, \mathbf{P} and \mathbf{Q} are the PLS loadings (the modes) matrices and \mathbf{E} and \mathbf{F} are the matrices of residuals. PLS loadings are not necessarily orthogonal. [34]

2.5.3 Using PCA and PLS to build Statistical Models

Building a statistical model means being able to represent the i^{th} subject's characteristic (shape, intensity or a combination of both), x^i , as

$$x^i = \bar{x} + \sum_{j=1}^m t_j^i p_j$$

where \bar{x} is the population's average characteristic, p_j are the m modes (PCA or PLS) which define the new space of representation of the original variables x where variance is maximized, and t_j^i are the components (PCA or PLS), which measure the relevance of each j^{th} mode in the representation of the i^{th} patient and are obtained by projecting the original variables onto the modes. [23]

2.6 SSMs, SIMs, SSIMs

By applying PCA and PLS to different matrices, six statistical models (three with PCA and three with PLS) are built.

2.6.1 SSMs

Statistical Shape Models are the result of the application of PCA or PLS to the moment vector matrix \mathbf{X}_β , where the i^{th} couple of columns is the two components of the moment vector β^i for all the 97 subjects of the dataset. PLS also requires an external response variable, which in this work is, as previously stated, the fracture status (here meant as a vector of zeros and ones). [20]

2.6.2 SIMs

The procedure to build the Statistical Intensity Models is basically the same but, instead of the moment vectors matrix, the input is the matrix of elemental BMD \mathbf{G} , where the i^{th} row contains the BMD values of all the elements of the mesh for the i^{th} subject of the dataset. As for PLS, the external response variable is still the fracture status. [20]

2.6.3 SSIMs

PCA and PLS scores (or components) contain the shape/intensity information for each subject of the dataset. This information can be combined and further processed with PCA and PLS to create two Statistical Shape-Intensity Models, which considers both shape and intensity at the same time. Calling \mathbf{B}^S and \mathbf{B}^G the PCA shape and intensity components matrices respectively, the new PCA input matrix \mathbf{B} can be assembled as

$$\mathbf{B} = \begin{pmatrix} \mathbf{W}_{PCA} \mathbf{B}^S \\ \mathbf{B}^G \end{pmatrix} = \begin{pmatrix} \mathbf{W}_{PCA} \mathbf{A}^{S'} \mathbf{X}'_\beta \\ \mathbf{A}^{G'} \mathbf{G}' \end{pmatrix}$$

where $\mathbf{W}_{PCA} = r\mathbf{I}$ is the weighting factor matrix and r is the ratio between the total variances in only intensity and only shape PCA modes.

Similarly, if \mathbf{T}^S and \mathbf{T}^G are the PLS shape and intensity components matrices respectively, the new PLS input matrix \mathbf{T} is calculated as

$$\mathbf{T} = \begin{pmatrix} \mathbf{W}_{PLS} \mathbf{T}^S \\ \mathbf{T}^G \end{pmatrix} = \begin{pmatrix} \mathbf{W}_{PLS} \mathbf{P}^{S'} \mathbf{X}'_{\beta} \\ \mathbf{P}^{G'} \mathbf{G}' \end{pmatrix}$$

where $\mathbf{W}_{PLS} = s\mathbf{I}$ is the weighting factor matrix and s is the ratio between the total variances in only intensity and only shape PLS modes.

\mathbf{B} and \mathbf{T} are then used to identify combined shape-intensity PCA and PLS modes. [20] [35]

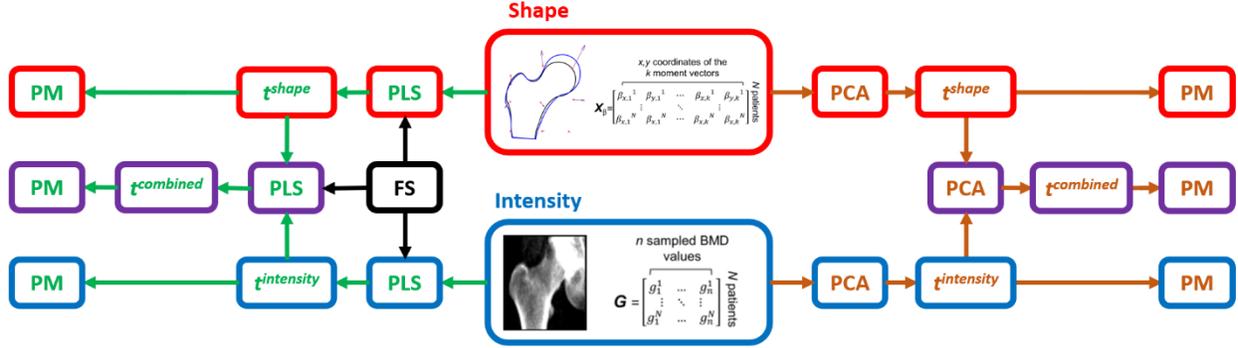


Figure 5: Schematic representation of the process: shape (red) and intensity (blue) information is used to create statistical models through PCA and PLS and the corresponding components (t) are calculated; then, components are combined to build two more models based on shape and intensity together (purple); in the end, from the six models' components, six predictive models (PM) are built

2.7 Stratification accuracy assessment

The final step of the whole process is using the models as predictors of the fracture status of the cohort and comparing their performances with those of the actual gold standard: one logistic regression model for each statistical model was built, together with a logistic regression model for the femoral neck BMD (here used as the gold standard). The logistic regression models were then compared by plotting the Receiver Operating Characteristic (ROC) curves and evaluating the Area Under Curve (AUC) for each curve [23]. Eventually, six other logistic regression models were built based on the statistical models obtained in [23] using the 3D CT scans of these same patients and the corresponding ROC curves and the AUC values were calculated.

3. Results

3.1 Statistical Shape Models

With 9 moment vectors for each of the 97 subjects and 2 components for each vector, a 97×18 matrix \mathbf{X}_{β} was assembled

3.1.1 PCA

A total of 17 shape modes and the corresponding eigenvalues were calculated through Principal Component Analysis, but 5 of them are already enough to explain more than 95% of the total variability. The cumulative explained variability can be seen in Figure 6.

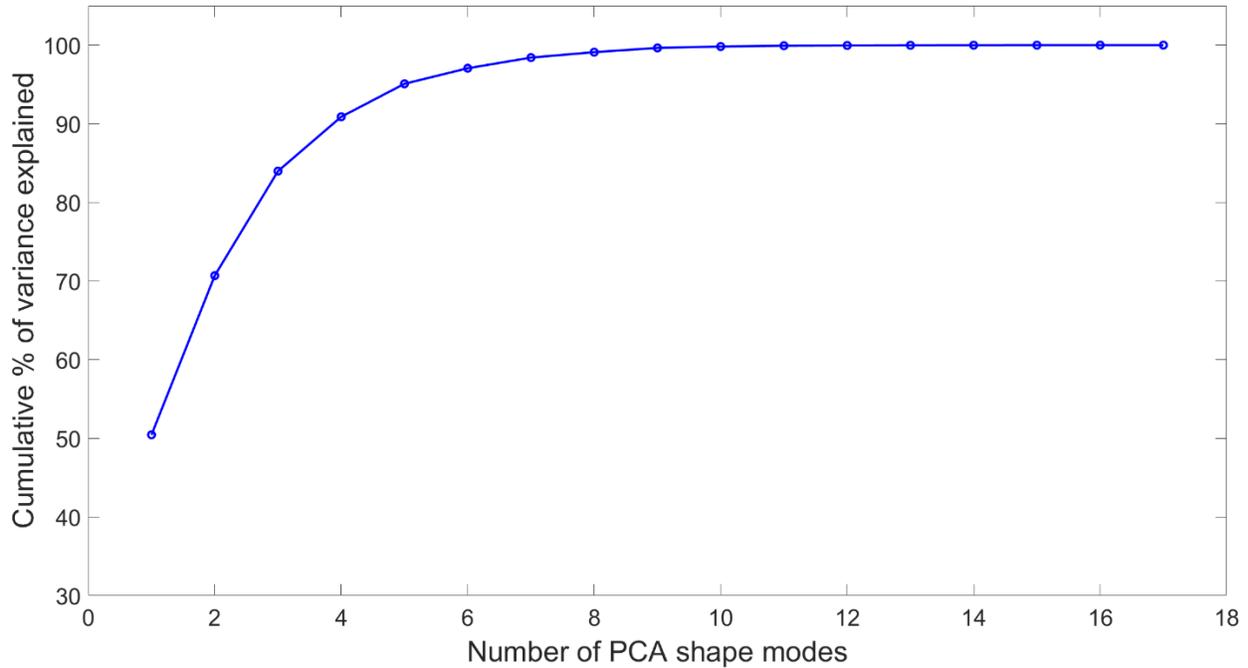


Figure 6: Cumulative percentage of variance explained by the PCA shape modes

In Figure 7, the first 5 modes are represented as deformations of $\pm 2\sqrt{\sigma^2}$ of the template, where σ^2 is the m^{th} PCA mode variance [20]: a simple visual analysis reveals that femur size (Mode 1), femoral neck length (Mode 2) and inclination (Mode 3) are the most prominent shape features.

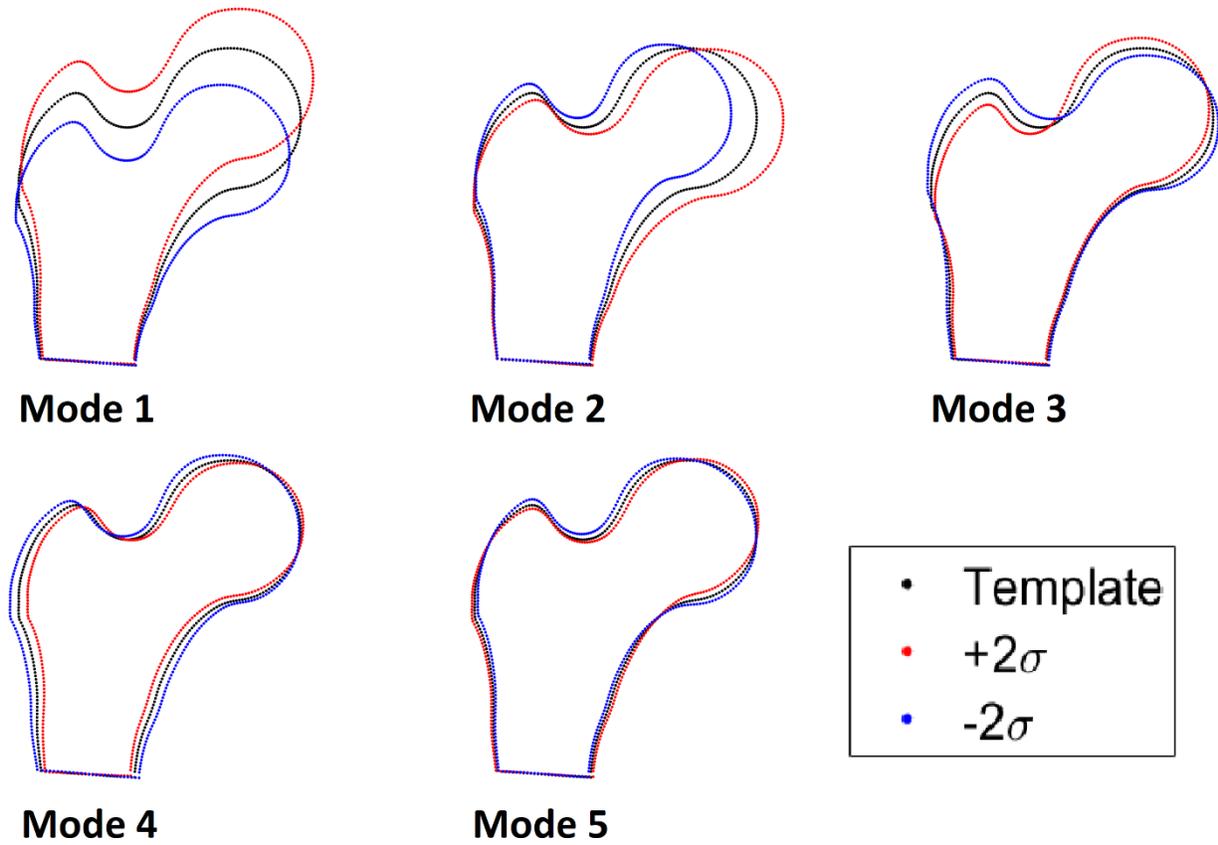


Figure 7: first 5 PCA shape modes shown as deformations of the template along each mode

In Figure 8, the first two or three PCA shape components of all the subjects are represented in two scatter plots, in which a color code allows to recognize fractured subjects (red) from non-fractured ones (green). Apparently, the PCA shape components are not able to separate and clusterize the two categories of subjects.

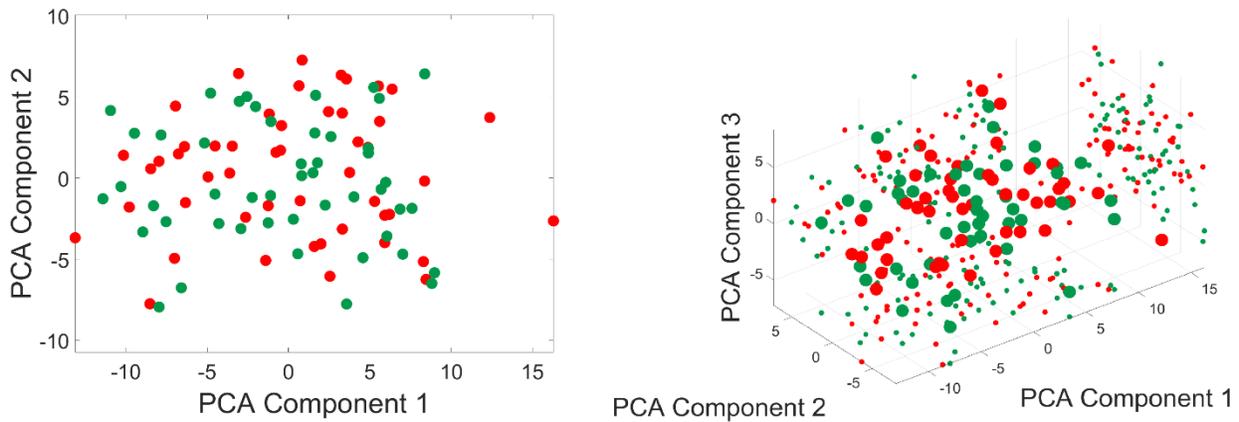


Figure 8: Scatter plots of the first two (left) and three (right) PCA shape components for all the subjects. Fractured patients are shown in red, non-fractured patients in green

3.1.2 PLS

Analogously, 17 shape modes were calculated through Partial Least Square, along the components and the percentage of explained variance of shape and fracture status (as seen in Figure 9). 6 modes explain 95.52% of the shape variance, but only 15.05% of the fracture status.

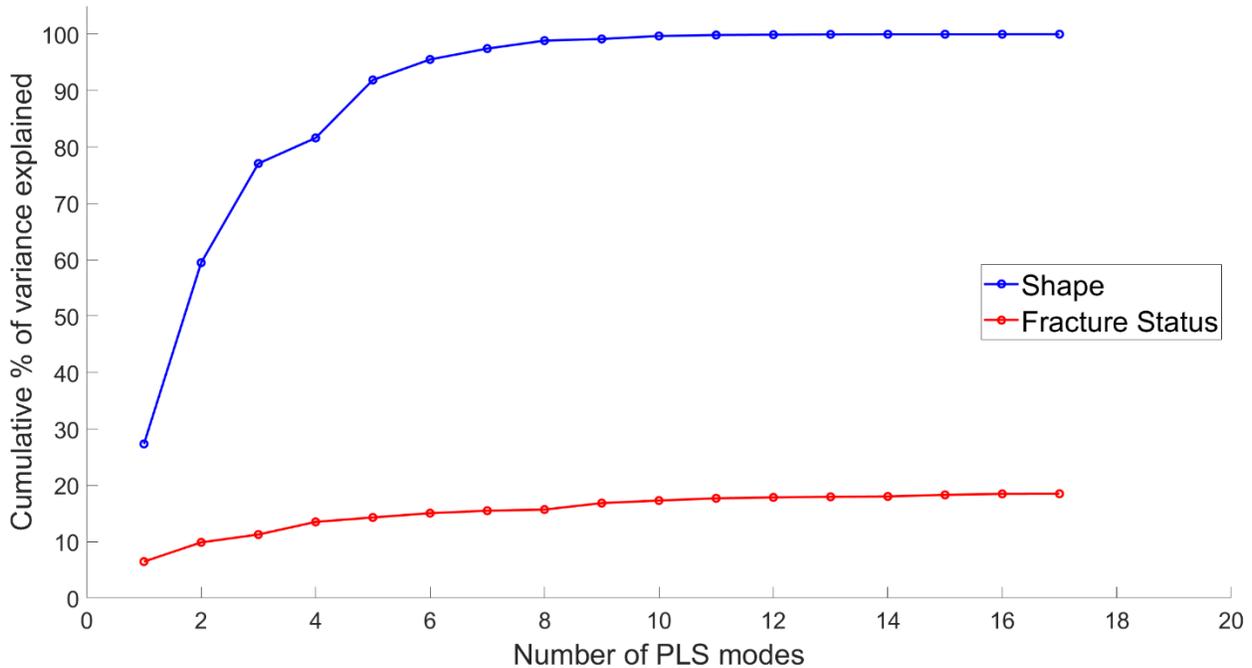


Figure 9: Cumulative percentage of (both shape's and fracture status') variance explained by the PLS modes

The first 6 modes are shown in Figure 10: PLS allows to find the shape features that are more correlated with fracture which, in this case, are femoral neck length (Mode 1), femur's size (Mode 2) and neck inclination (Modes 3 and 4).

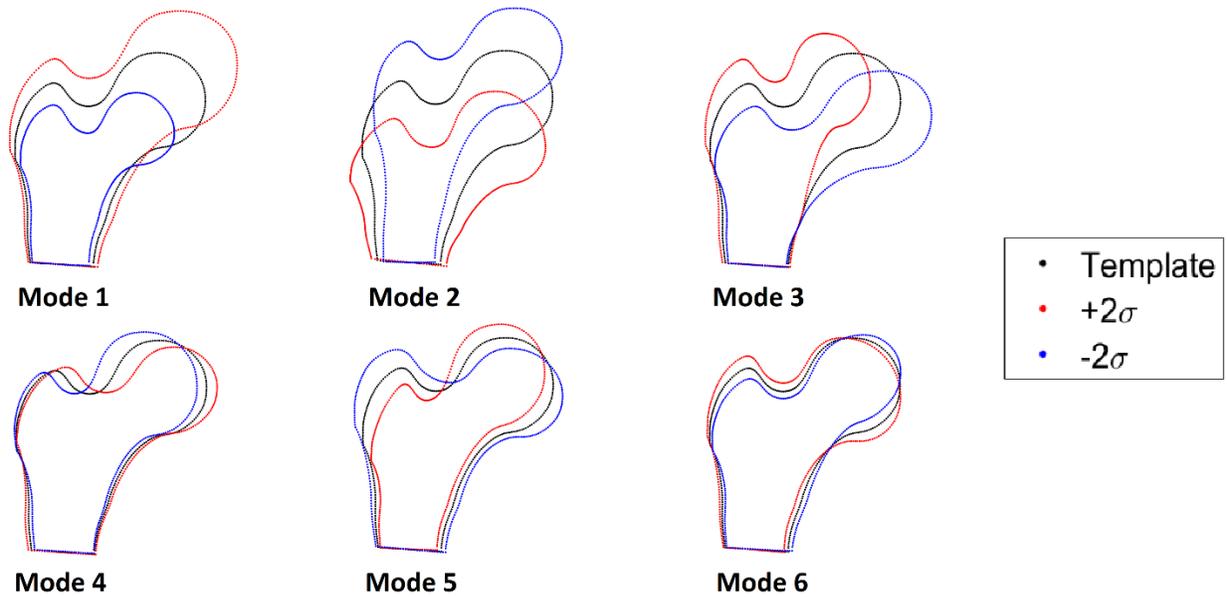


Figure 10: first 6 PLS shape modes shown as deformations of the template along each mode

In Figure 11, the first two or three PLS shape components of all the subjects are represented in two scatter plots, in which a color code allows to recognize fractured subjects (red) from non-fractured ones (green). Just as seen in Figure 8 for PCA, fractured subjects are not separated efficiently from the non-fractured ones, meaning that perhaps shape and fracture status are not correlated enough (as seen in Figure 9 from an explained variance perspective).

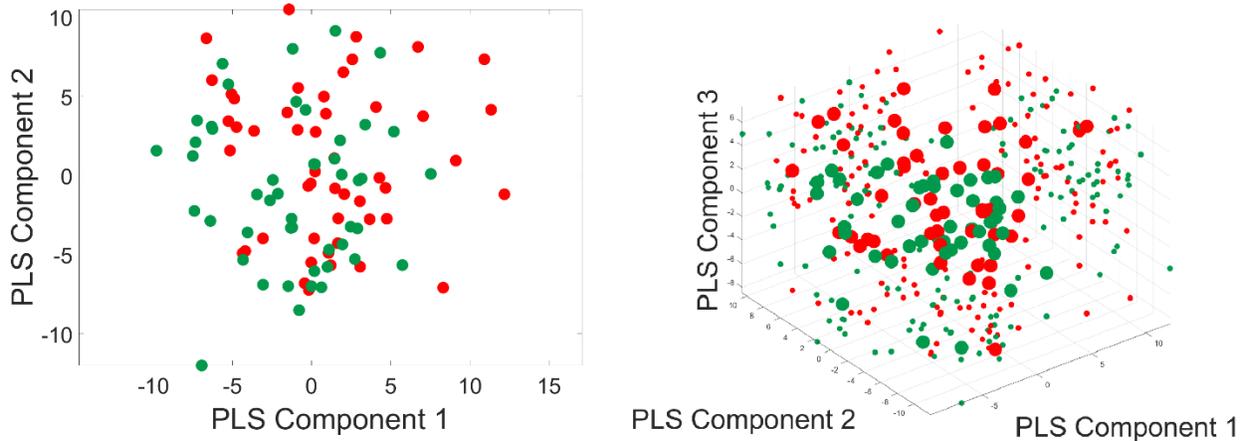


Figure 11: Scatter plots of the first two (left) and three (right) PLS shape components for all the subjects. Fractured patients are shown in red, non-fractured patients in green

3.2 Statistical Intensity Models

The template mesh (and thus all the subject specific meshes) used in this study counts 16072 nodes and 31564 triangular elements. As exposed in 2.4, a BMD value was assigned to each element of the mesh for all the subject specific meshes, hence the SIMs input matrix \mathbf{G} was a 97×31564 matrix where the i^{th} row contains the BMD values of all the elements of the mesh for the i^{th} subject of the dataset.

3.2.1 PCA

A total of 96 intensity modes and the corresponding eigenvalues were calculated through Principal Component Analysis, but 48 of them are already enough to explain more than 95% of the total variability. The cumulative explained variability can be seen in Figure 12.

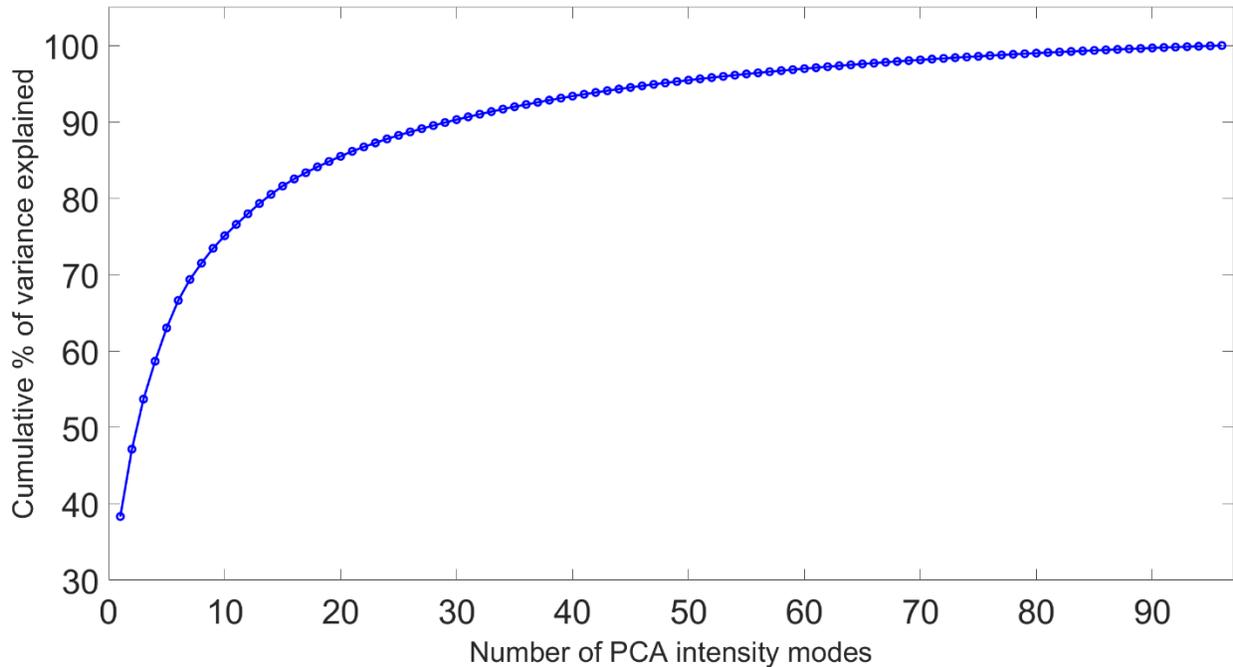


Figure 12: Cumulative percentage of variance explained by the PCA intensity modes

Figure 13 shows the first 4 PCA intensity modes as variations of the template's BMD: modes 1 and 2 suggest that the BMD of the entire proximal femur (mode 1) and especially of the femoral head (mode 2) are the most prominent intensity features.

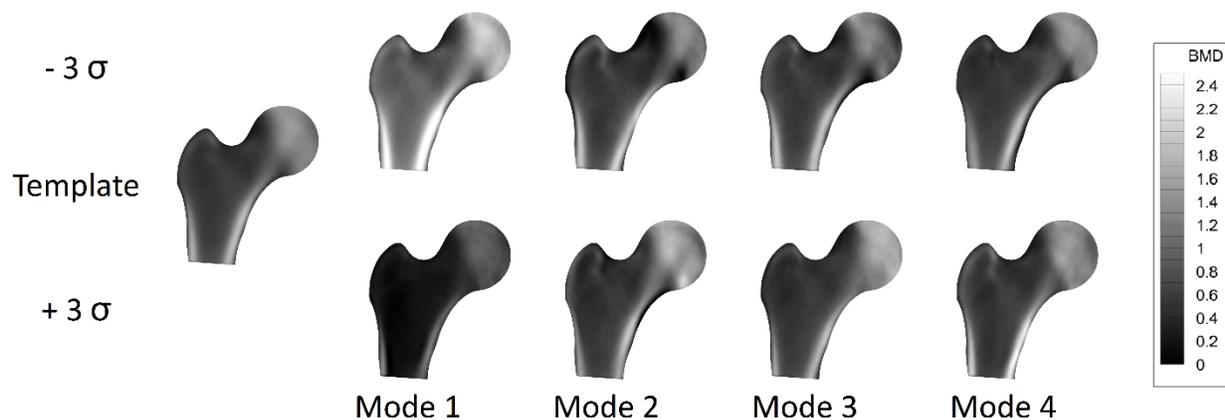


Figure 13: first 4 PCA intensity modes shown as variations of the template's Bone Mineral Density (BMD) along each mode

The scatter plots of the first two and three intensity components show that PCA and intensity are slightly better than shape at separating fractured subjects from the non-fractured ones, with the former having generally higher values of the first component than the latter (see Figure 14).

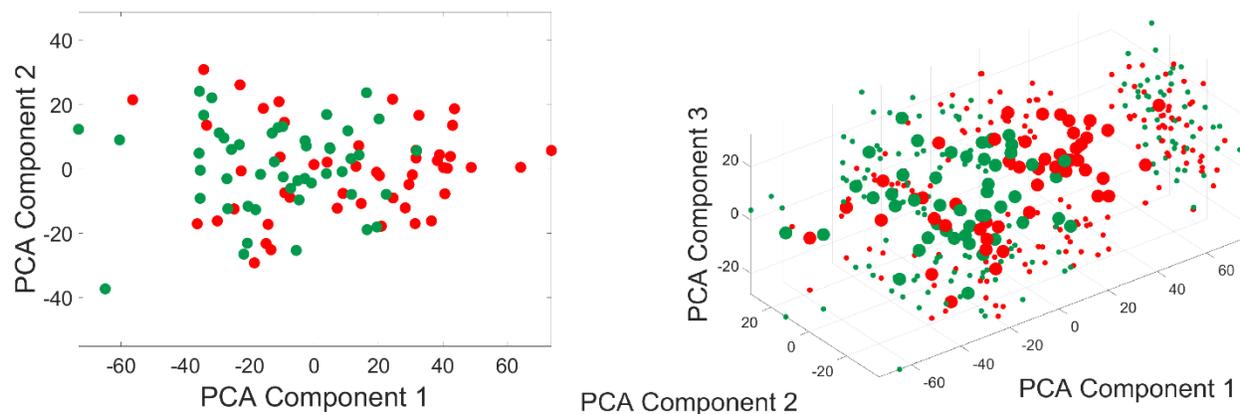


Figure 14: Scatter plots of the first two (left) and three (right) PCA intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green

3.2.2 PLS

Analogously, 96 intensity modes were calculated through Partial Least Square: 58 modes are needed to explain more than the 95% of intensity variance, while far less are necessary for the fracture status and 58 modes already explain 100% of its variance, as shown in Figure 15.

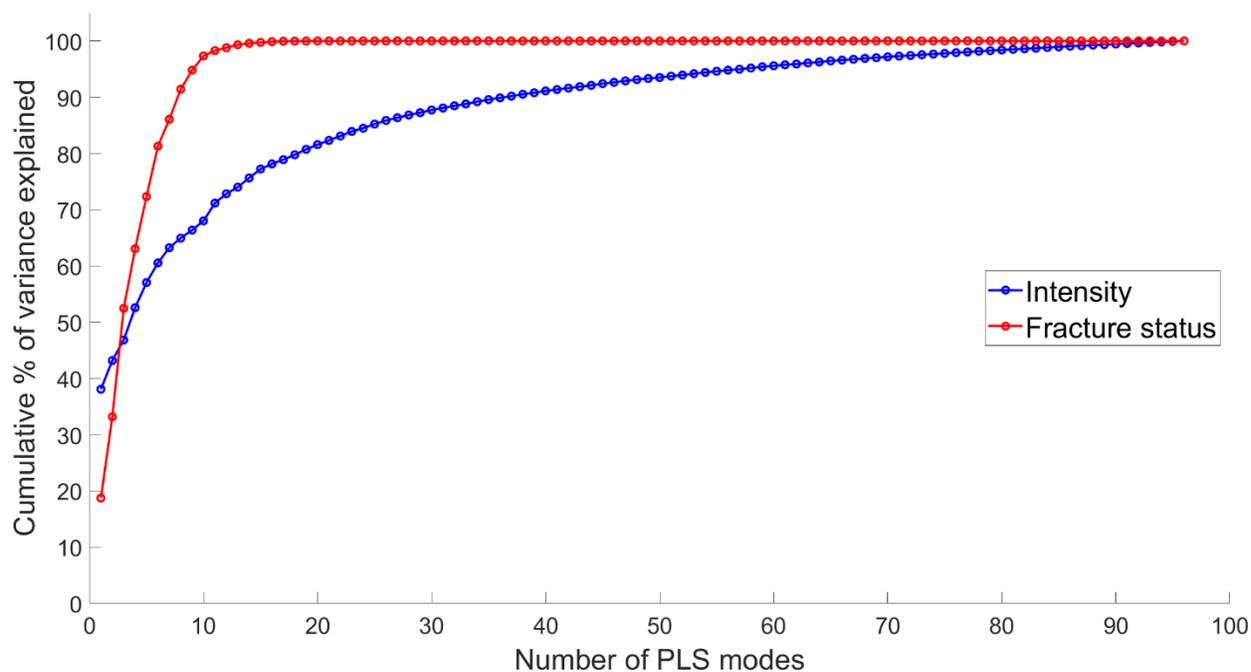


Figure 15: Cumulative percentage of (both intensity's and fracture status') variance explained by the PLS intensity modes

The first 6 intensity modes, as seen in Figure 16, show that the overall femoral BMD (modes 1 and 2) and a reduced cortical thickness (modes 4 and 5) are the intensity features that are correlated the most with fracture.

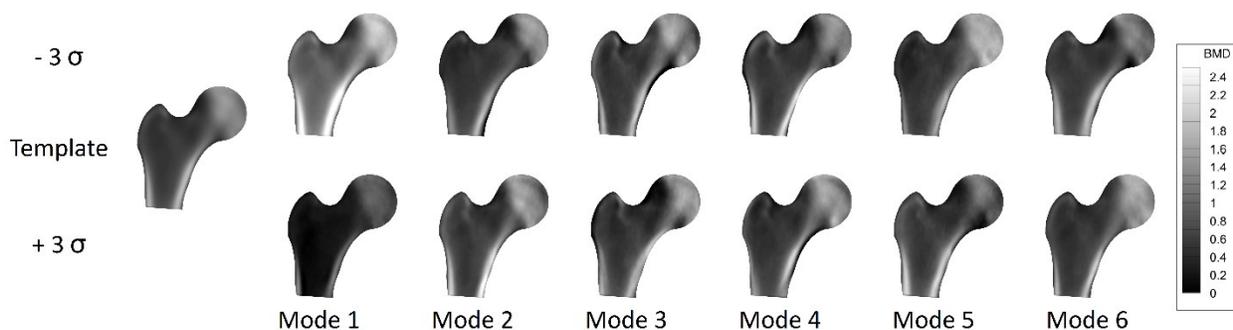


Figure 16: first 6 PLS intensity modes shown as variations of the template's Bone Mineral Density (BMD) along each mode

Finally, the scatter plots of the first two or three PLS intensity components show that the PLS-based SIM has slightly better performances than the PCA-based in distinguishing fractured and

non-fractured cases, with non-fractured subjects having generally lower first and second component than the fractured ones (see Figure 17).

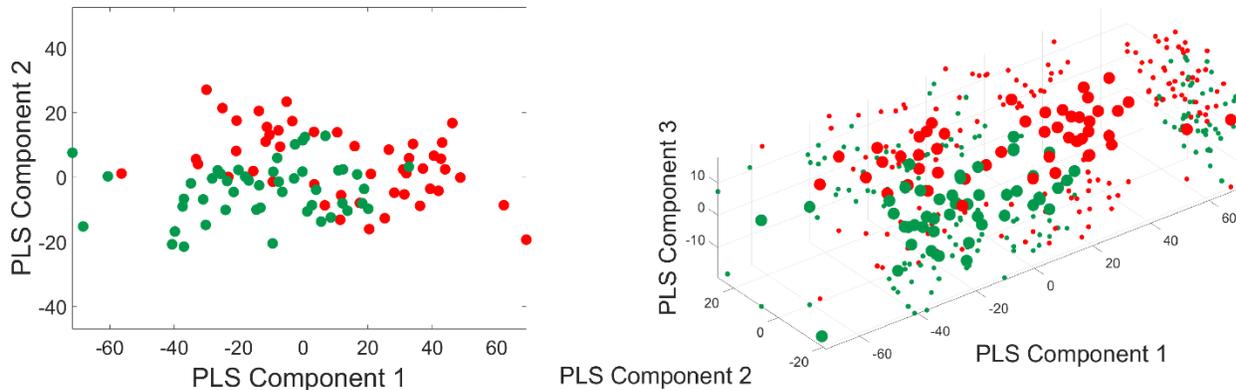


Figure 17: Scatter plots of the first two (left) and three (right) PLS intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green

3.3 Statistical Shape-Intensity Models

3.3.1 PCA

The 97×17 PCA shape components matrix \mathbf{B}^S and the 97×96 PCA intensity components matrix \mathbf{B}^G were combined, as shown in 2.6.3, into a single 97×113 matrix \mathbf{B} , which is then furtherly processed through PCA: 96 shape-intensity PCA modes were calculated, of which 34 explain more than 95% of the combined shape-intensity variance (see Figure 18).

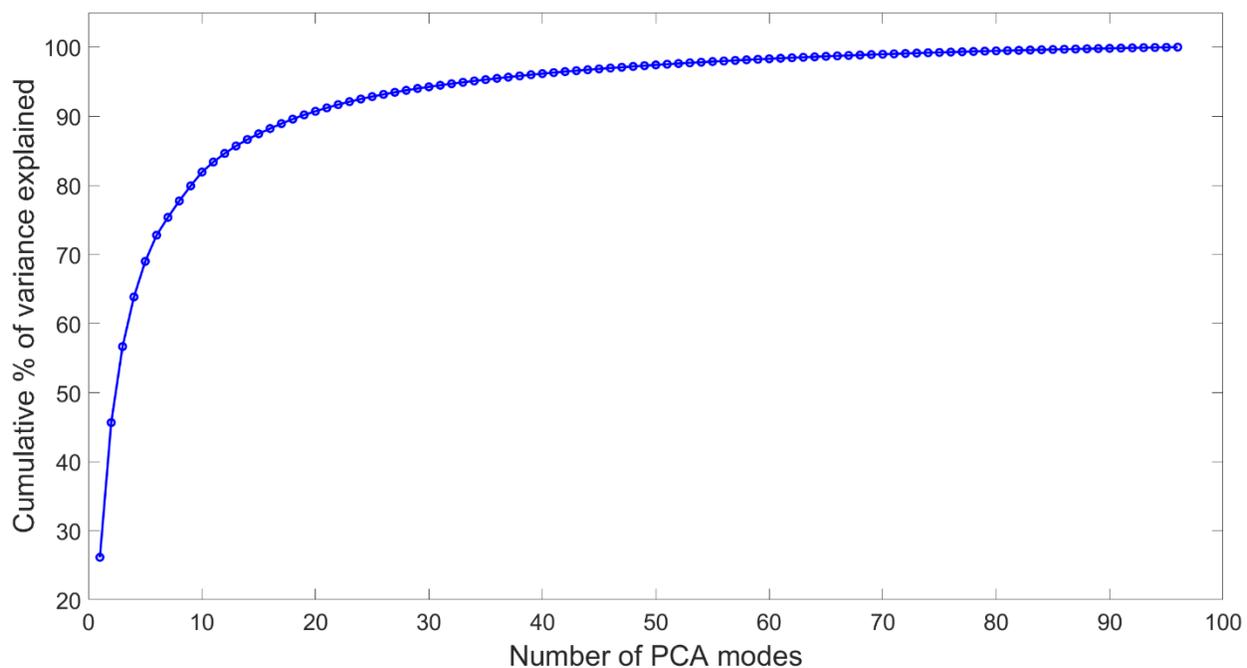


Figure 18: cumulative percentage of variance explained by the PCA modes

In Figure 19, the first 6 PCA shape-intensity modes are shown as both deformations of the template and variations of its BMD: as already seen from the PCA-based shape and intensity models, proximal femur’s size (mode 1), femoral neck’s size (mode 3) and inclination (mode 4), general femoral BMD (mode 2) and cortical thickness (modes 2 and 3) are the most prominent shape-intensity features.

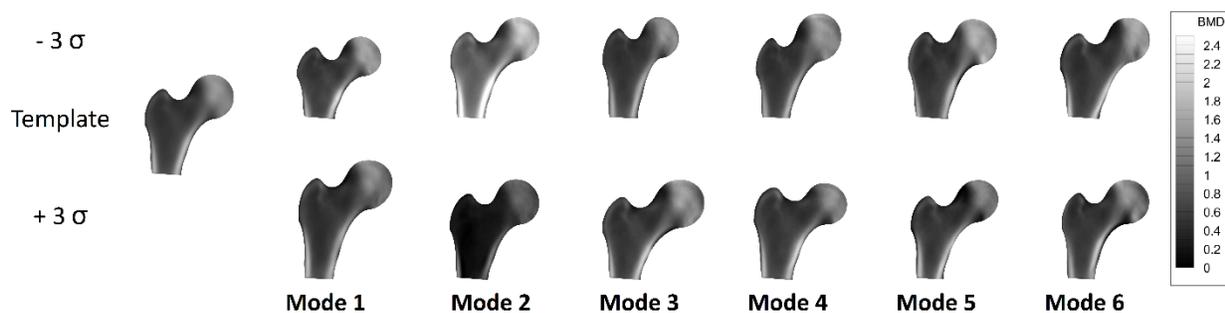


Figure 19: first 6 PCA shape-intensity modes shown as both deformations of the template and variations of its Bone Mineral Density (BMD) along each mode

The scatter plots of the first two and three PCA shape-intensity components (Figure 20) show that fractured patients have generally higher values of the second component, while non-fractured ones have generally lower values of both the first and the second component.

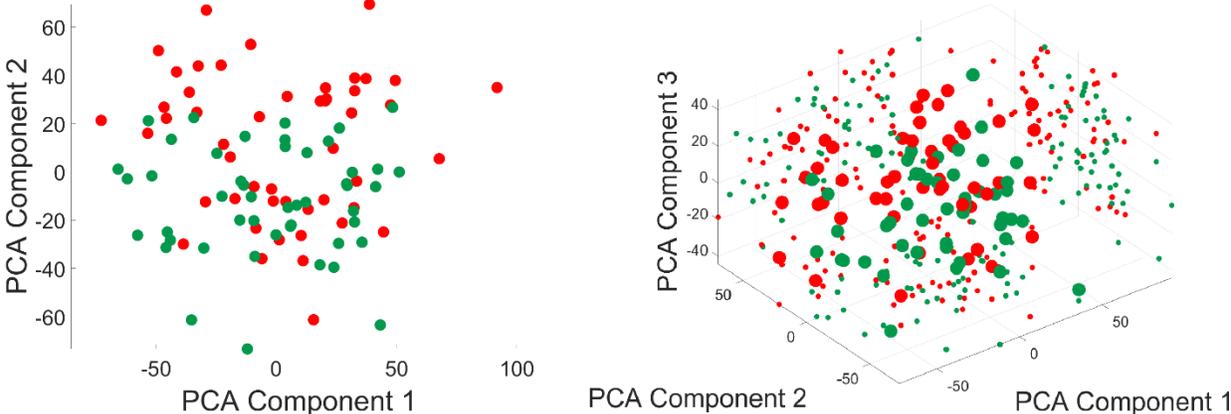


Figure 20: Scatter plots of the first two (left) and three (right) PCA shape-intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green

3.3.2 PLS

The 113×97 PLS shape-intensity input matrix \mathbf{T} was built by combining the 17×97 PLS shape components matrix \mathbf{T}^S and the 96×97 PLS intensity components matrix \mathbf{T}^G , then processed through PLS: of the 96 calculated modes, 9 already explain more than 95% of the shape-intensity variance and almost 93% of the fracture status' variance (Figure 21).

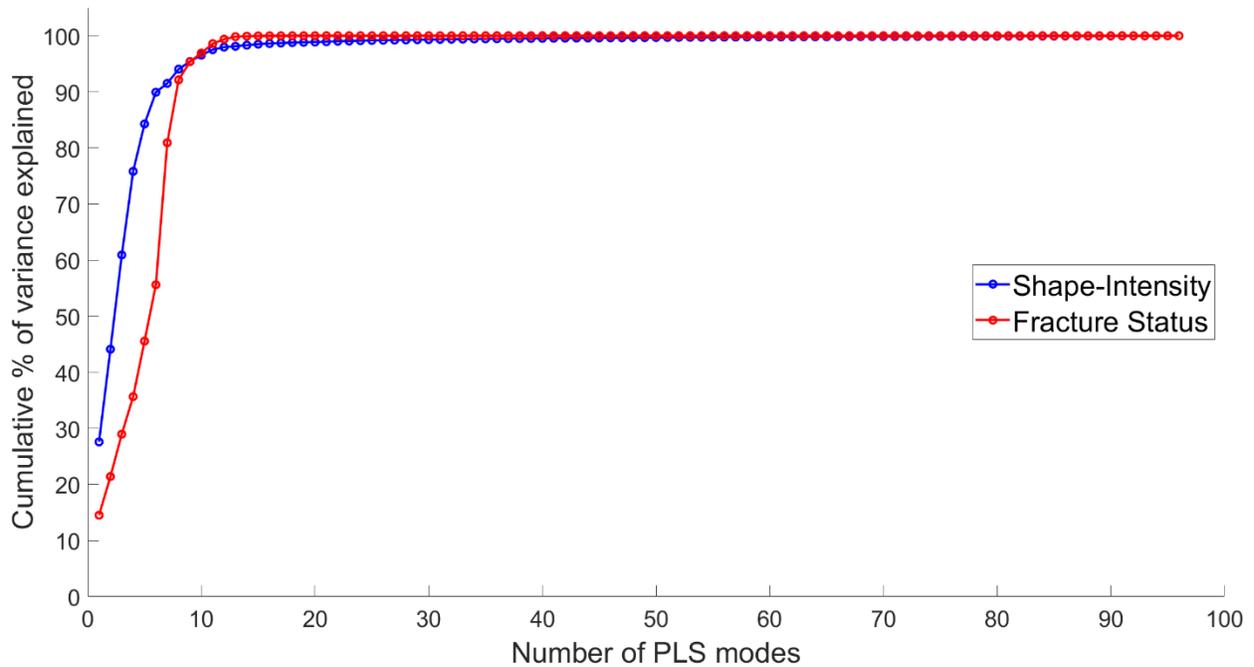


Figure 21: Cumulative percentage of (both shape-intensity's and fracture status') variance explained by the PCA shape-intensity modes

Figure 22 shows the first 6 PLS shape-intensity modes: once again, proximal femur's size (modes 1 and 2), femoral neck's length and inclination (modes 4 and 5), cortical thickness and the general femoral BMD are the shape-intensity features correlated to fracture.

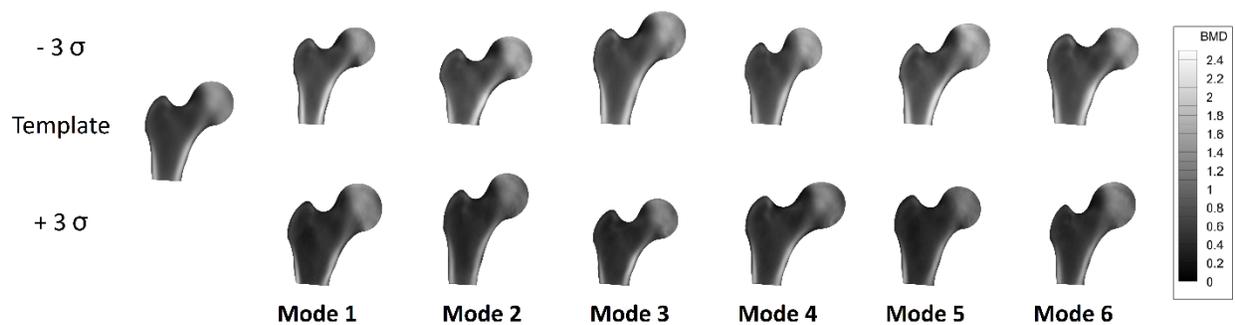


Figure 22: first 6 PLS shape-intensity modes shown as both deformations of the template and variations of its Bone Mineral Density (BMD) along each mode

The scatter plots in Figure 23 highlight that fractured patients have generally lower PLS shape-intensity components, while the fractured ones have higher components.

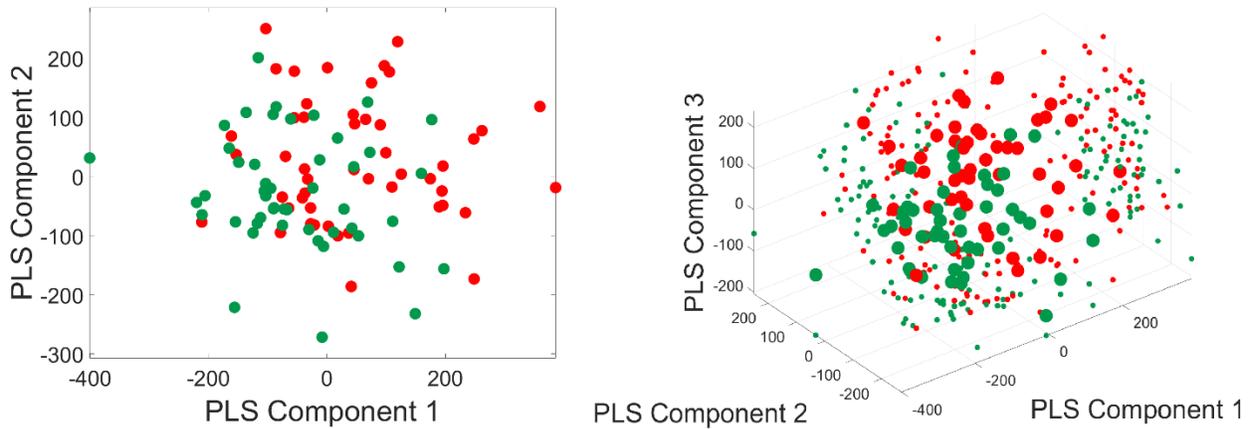


Figure 23: Scatter plots of the first two (left) and three (right) PLS shape-intensity components for all the subjects. Fractured patients are shown in red, non-fractured patients in green

3.4 Accuracy assessment

The regression models (one for each combination of feature and statistical technique) were built using the first 4 components of PCA or PLS and all the 97 subjects as training set. Although no subjects were available to form a test set (new or part of the original cohort), the predictive performance of the models could still be evaluated by using the same training set as test set and comparing the results with the performance of the femoral neck BMD. The Receiver Operating Characteristic (ROC) curves for PCA models and PLS models are shown in Figure 24.

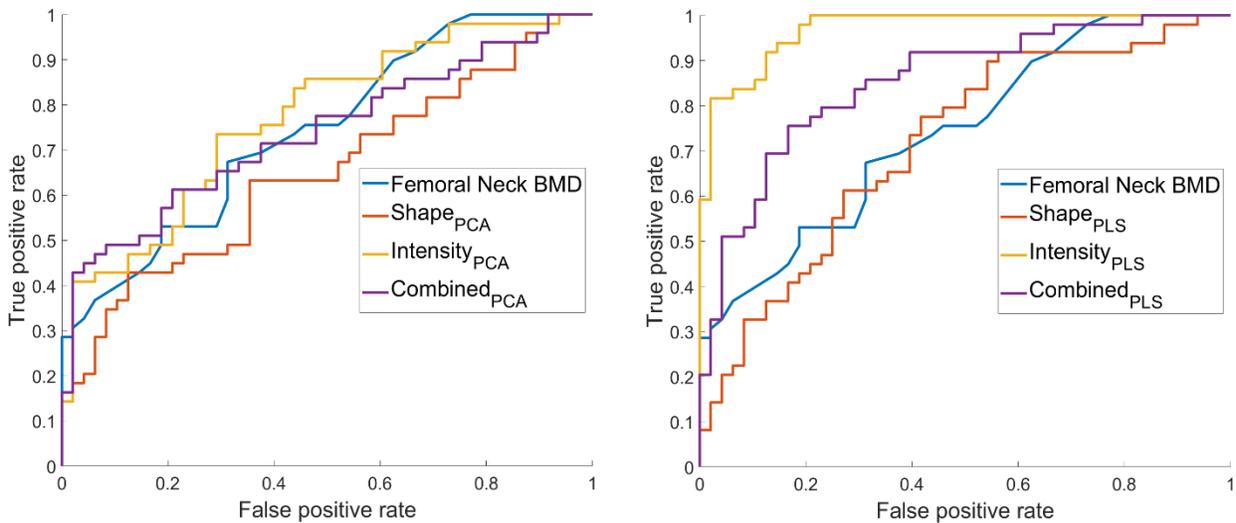


Figure 24: ROC curves for PCA models and Femoral Neck BMD (left) and PLS models and Femoral Neck BMD (right). Models built starting from 2D DXA images

A mere visual analysis of the curves is enough to see that PLS performs better than PCA for each kind of model and that SIMs are the best ones in the respective category (PCA or PLS), followed by SSIMs and SSMs.

A more accurate evaluation is possible by looking at the Area Under Curve (AUC) for each model: as shown in Table 2, both statistical shape models, with AUC equal to 0.65 and 0.72 for PCA and PLS respectively, have lower performances than the femoral neck BMD (AUC = 0.74), while all the other models are equal or better than the gold standard. It is worth mentioning though that these results could be influenced by the fact that the test set is the same set used to train the models, hence the performances of the statistical models could be higher than they would be if the test set was made up of subjects not belonging to the training set.

In Figure 25, the ROC curves of the models built using the 3D CT scans of the same cohort used in this study are shown, while the corresponding AUC values are listed in Table 2. These models too are built using the whole cohort as training set and thus the performances of some of them are unrealistically high (especially the SIM and SSIM built with PLS); anyway, these models are still enough to draw some conclusions: intensity-based models still perform better than the shape-based ones, while the PCA-based SSIM outperforms both the SSM and the SIM built with the same technique (a consideration of this kind cannot be made for the PLS-based ones), which is a different result than the 2D case. As for the comparison between the 2D-based models and the corresponding 3D-based ones, the latter generally outperforms the former in terms of AUC.

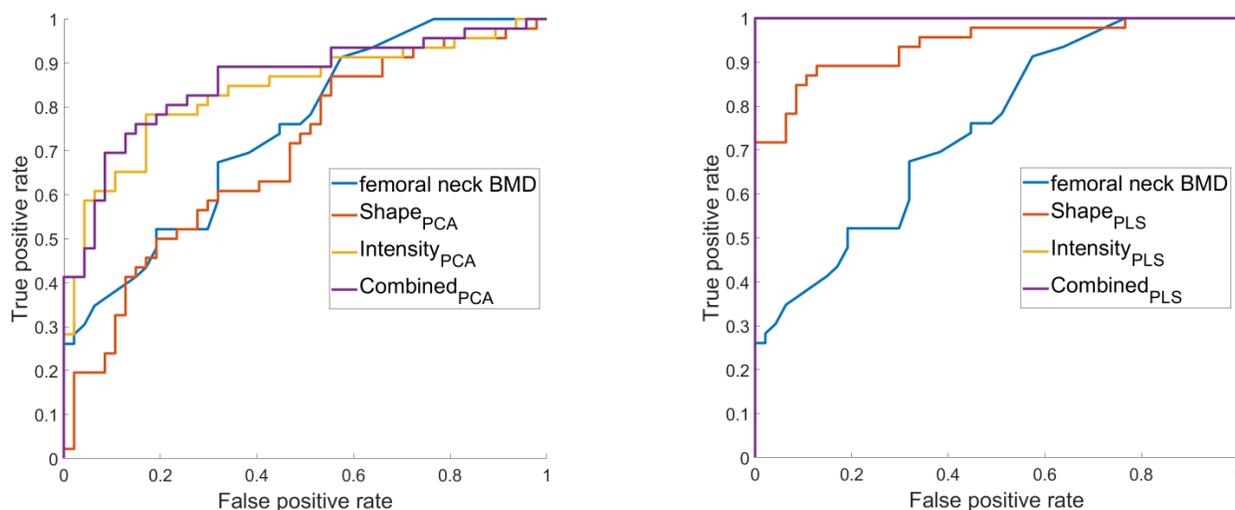


Figure 25: ROC curves for PCA models (left) and PLS models (right). Models built starting from 3D CT scans

Table 2: AUC values for the logistic regression models built with the 2D and 3D results

Model	AUC (2D)	AUC (3D)
Femoral neck BMD	0.74	0.75
Statistical Shape (PCA)	0.65	0.69
Statistical Shape (PLS)	0.72	0.94
Statistical Intensity (PCA)	0.77	0.84
Statistical Intensity (PLS)	0.97	1
Statistical Shape-Intensity (PCA)	0.74	0.86
Statistical Shape-Intensity (PLS)	0.85	1

4. Discussion

The results reported in the previous section allow to envision a promising future for DXA-based statistical models: while SSMs, although capable of finding the principal shape features, show generally low performances in terms of explaining and predicting fracture (which is coherent with the findings of [27]), SIMs demonstrate that intensity, in terms of bone mineral density (BMD) of the proximal femur, and fracture are highly correlated (Figure 15) and that, in terms of fracture prediction, they outperform the actual gold standard when built with either PCA or PLS (Figure 24). As for the SSIMs, combining shape and intensity information into a single model seems to reduce the predictive performances if compared to intensity models; this unexpected outcome could be due to the fact that, in this work, there is no test set and performances are evaluated on the same dataset the models are trained with, so the predictions of these models could be better (or, in the case of intensity, much better) than expected (same thing happens with the models built with 3D CT scans in [23], as shown in Figure 25); another reason could be the inability of 2D projections (like the BMD maps or the DXA raw images) to preserve the shape information, which could lead to both the poor results of shape-based models and the lower performances of SSIMs when compared to SIMs. Further investigation is required to ultimately assess whether including shape and thus building a SSIM is useful to improve fracture prediction or not, although it has already been demonstrated that CT-based SSIMs perform better than SSMs and SIMs [23], as partially shown in Figure 25 and Table 2. The comparison with the 3D-based models also reveals that these models generally perform slightly better than the 2D-based ones, but they still have the huge limit of using data from an imaging technique not adopted in clinics for osteoporosis diagnosis. Automatization of the whole process is another aspect that should be considered in future works: the building of the models and the prediction process are already completely automated, but the initial inputs still need human intervention to be prepared; as already stated in

2.2, the shape extraction through DXA images segmentation is still a semi-automatic process, mostly because the femoral head results partially overlapped with the pelvis and thus its boundaries are hardly recognizable, hence the need to manually rebuild it during segmentation; this arbitrary reconstruction of the femoral head could also be the cause of the lower performance of shape-based models. Moreover, this overlapping also causes some nuisances from an intensity perspective, since it leads to higher femoral head's BMD values. For these reasons, a possible future development could be eliminating the femoral head from the equation, while still preserving the femoral neck, and verifying if predictive performance does not change substantially. Lastly, finding a test set formed by patients not used during the models' training process would allow a real validation of the predictive models built in this study.

5. Conclusions

Osteoporosis is a bone disease which leads to a decreased bone resistance, and thus to a higher risk of fractures, and affects both men and women, although differently, typically later in life [1] [3]. With the currently aging population, osteoporosis is becoming a more and more relevant issue both in terms of quality of life of the affected people and economic burden on healthcare [3] [4]. That is why diagnosing and treating it as soon and accurately as possible is becoming fundamental. The current gold standard for osteoporosis diagnosis, as established by the World Health Organization, is measuring the areal bone mineral density (aBMD) of specific skeletal sections (spine, hip) by means of dual-energy X-ray absorptiometry (DXA) and comparing these BMD values to those of the average healthy population by calculating the so-called T-score [5] [6] [7] [8]. However, T-score has showed some weaknesses when used to predict fracture risk, with more than half of the fractured patients being classified as non-osteoporotic [12]. That is why many different alternatives to T-score have been investigated to enhance or replace it and improve fracture prediction: epidemiological models (FRAX [13] [14], Garvan [15] [16]), micro- (TBS [17] [18]) and macrostructural (BMSi [19], HSA [20]) bone features, Finite Element Method (FEM [20]) are just some examples, but they all have their own limitations and some of them are not even currently adopted in clinics. In the last decades, higher data availability and computational power have shifted researcher's attention toward statistical models: these models can use information about shape and/or bone density for fracture prediction [21] with promising outcomes, as proved by many studies [22] - [27]. In [23], statistical models were built from 3D CT scans of 97 post-menopausal women whose fracture status was known but, although the results were encouraging, these models had the great limit of using data from an imaging technique that is not currently adopted for osteoporosis diagnosis and fracture prediction; in this work, those same kinds of models were created, but using the DXA 2D images of those same 97 women as the starting point. A total of six statistical models were built (two SSMS, two SIMs, two SSIMs) using Principal Component Analysis (PCA), which highlights the most prominent features of shape and/or intensity, and Partial Least Square (PLS), which finds the main features of shape and/or intensity correlated with fracture. Results show that the main shape and intensity features can be isolated and used to predict fracture with moderate success, since both intensity and shape-intensity models

outperform femoral neck BMD, the current gold standard. In future works, a higher level of automatization of the process and a refinement of the data extraction from the medical images could lead to superior models that, by using the same images already used in clinics for osteoporosis diagnosis, could easily and rapidly enter the clinical routine to improve fracture prediction, thus easing the burden on healthcare due to osteoporotic fractures and enhancing the elderly people's quality of life.

Bibliography

- [1] H. Gray, *Gray's Anatomy*, London, England: Arcturus Publishing Ltd., 2015.
- [2] R. Florencio-Silva, G. Rodrigues da Silva Sasso, E. Sasso-Cerri, M. J. Simoes and P. S. Cerri, "Biology of Bone Tissue: Structure, Function, and Factors That Influence Bone Cells," *Biomed Res Int.*, 2015.
- [3] J. L. Porter and M. Varacallo, "Osteoporosis," *StatPearls. Treasure Island (FL): StatPearls Publishing*, 2022.
- [4] M. Lorentzon, H. Johansson, E. Liu, L. Vandenput, J. A. Kanis, N. C. Harvey and E. V. McCloskey, "Osteoporosis and fractures in women: the burden of disease," *Climacteric*, vol. 25, no. 1, pp. 4-10, 2022.
- [5] World Health Organization, "Prevention and management of osteoporosis : report of a WHO scientific group.," Geneva, Switzerland, 2000.
- [6] T. Sözen, L. Özişik and N. Ç. Başaran, "An overview and management of osteoporosis," *Eur J Rheumatol*, vol. 4, no. 1, pp. 46-56, 2017.
- [7] World Health Organization, "WHO scientific group on the assessment of osteoporosis at the primary health care level: summary meeting report," Brussels, Belgium, 2004.
- [8] M. A. Laskey, "Dual-energy X-ray absorptiometry and body composition," *Nutrition*, vol. 12, no. 1, pp. 45-51, 1996.
- [9] G. M. Blake and I. Fogelman, "The role of DXA bone density scans in the diagnosis and treatment of osteoporosis," *Postgrad Med J.*, vol. 83, no. 982, pp. 509-517, 2007.
- [10] B. Kim, Y. J. Cho and W. Lim, "Osteoporosis therapies and their mechanisms of action (Review)," *Exp Ther Med.*, vol. 22, no. 6, 2021.
- [11] D. Marshall, O. Johnell and H. Wedel, "Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures," *BMJ*, vol. 312, no. 7041, pp. 1254-1259, 1996.
- [12] S. A. Wainwright, L. M. Marshall, K. E. Ensrud, J. A. Cauley, D. M. Black, T. A. Hillier, M. C. Hochberg, M. T. Vogt, E. S. Orwoll and Study of Osteoporotic Fractures Research Group, "Hip fracture in women without osteoporosis," *J Clin Endocrinol Metab.*, vol. 90, no. 5, pp. 2787-2793, 2005.
- [13] E. V. McCloskey, N. C. Harvey, H. Johansson, M. Lorentzon, E. Liu, L. Vandenput, W. D. Leslie and J. A. Kanis, "Fracture risk assessment by the FRAX model," *Climacteric*, vol. 25, no. 1, pp. 22-28, 2022.

- [14] L. Vandenput, H. Johansson, E. V. McCloskey, E. Liu, K. E. Akesson and F. A. Anderson, "Update of the fracture risk prediction tool FRAX: a systematic review of potential cohorts and analysis plan," *Osteoporosis International*, vol. 33, pp. 2103-2136, 2022.
- [15] C. J. Crandall, J. Larson, A. LaCroix, J. A. Cauley, M. S. LeBoff and W. Li, "Predicting Fracture Risk in Younger Postmenopausal Women: Comparison of the Garvan and FRAX Risk Calculators in the Women's Health Initiative Study," *J Gen Intern Med.*, vol. 34, no. 2, pp. 235-242, 2019.
- [16] G. Todorov, S. Brook, N. Quah Qin Xian, S. Von Widekind, B. Freudenthal and A. N. Comminos, "Comparison of fracture risk calculators in elderly fallers: a hospital-based cross-sectional study," *BMJ Open*, vol. 12, 2022.
- [17] P. Martineau, W. D. Leslie, H. Johansson, N. C. Harvey, E. V. McCloskey, D. Hans and J. A. Kanis, "In which patients does lumbar spine trabecular bone score (TBS) have the largest effect?," *Bone.*, vol. 113, pp. 161-168, 2018.
- [18] A. C. Borba, N. L. Terra and S. V. C. Areosa, "Relationship between TBS, Bone Density and fractures in older men: integrative review," *Acta Ortopédica Brasileira*, vol. 30, no. 4, 2022.
- [19] P. Rufus-Membere, K. A. Anderson, K. L. Holloway-Kew, M. A. Kotowicz, A. Diez-Perez and J. A. Pasco, "Associations between bone material strength index and FRAX scores," *Research Square*, 2022.
- [20] A. Aldieri, M. Terzini, A. L. Audenino, C. Bignardi and U. Morbiducci, "Combining shape and intensity DXA-based statistical approaches for osteoporotic HIP fracture risk assessment," *Comput Biol Med.*, 2020.
- [21] N. Sarkalkan, H. Weinans and A. A. Zapdoor, "Statistical shape and appearance models of bones," *Bone*, vol. 60, pp. 129-140, 2014.
- [22] T. L. Bredbenner, R. L. Mason, L. M. Havill, E. S. Orwoll, D. P. Nicolella and Osteoporotic Fractures in Men (MrOS) Study, "Fracture Risk Predictions Based on Statistical Shape and Density Modeling of the Proximal Femur," *J Bone Miner Res*, vol. 29, no. 9, pp. 2090-2100, 2014.
- [23] A. Aldieri, P. Bhattacharya, M. Paggiosi, R. Eastell, A. L. Audenino, C. Bignardi, U. Morbiducci and M. Terzini, "Improving the Hip Fracture Risk Prediction with a Statistical Shape-and-Intensity Model of the Proximal Femur," *Ann Biomed Eng.*, vol. 50, no. 2, pp. 211-221, 2022.
- [24] N. Sarkalkan, J. H. Waarsing, P. K. Bos, H. Weinans and A. A. Zapdoor, "Statistical shape and appearance models for fast and automated estimation of proximal femur fracture load using 2D finite element models," *J Biomech.*, vol. 47, no. 12, pp. 3107-3114, 2014.
- [25] J. C. Baker-LePain, K. R. Luker, J. A. Lynch, N. Parimi, M. C. Nevitt and N. E. Lane, "Active Shape Modeling of the Hip in the Prediction of Incident Hip Fracture," *J Bone Miner Res.*, vol. 26, no. 3, pp. 468-474, 2011.

- [26] S. R. Goodyear, R. J. Barr, E. McCloskey, S. Alesci, R. M. Aspden, D. M. Reid and J. S. Gregory, "Can we improve the prediction of hip fracture by assessing bone structure using shape and appearance modelling?," *Bone*, vol. 53, no. 1, pp. 188-193, 2013.
- [27] F. Pagotto, "Use of DXA-based statistical shape models of the femur for hip fracture risk prediction," *Tesi di laurea magistrale, Politecnico di Torino*, 2022.
- [28] L. Yang, W. J. Udall, E. V. McCloskey and R. Eastell, "Distribution of bone density and cortical thickness in the proximal femur and their association with hip fracture in postmenopausal women: a quantitative computed tomography study," *Osteoporos Int.*, vol. 25, no. 1, pp. 251-263, 2014.
- [29] M. B. Stegmann, *Active Appearance Models: Theory, Extensions and Cases*, II ed., 2000.
- [30] S. Durrleman, X. Pennec, A. Trouvé and N. Ayache, "Statistical models of sets of curves and surfaces based on currents," *Med Image Anal.*, vol. 13, no. 5, pp. 793-808, 2009.
- [31] S. Durrleman, M. Prastawa, N. Charon, J. R. Korenberg, S. Joshi, G. Gerig and A. Trouvé, "Morphometry of anatomical shape complexes with dense deformations and sparse parameters," *Neuroimage*, vol. 101, pp. 35-49, 2014.
- [32] S. Durrleman, X. Pennec, A. Trouvé, P. Thompson and N. Ayache, "Inferring brain variability from diffeomorphic deformations of currents: An integrative approach," *Med Image Anal.*, vol. 12, no. 5, pp. 626-637, 2008.
- [33] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Phil. Trans. R. Soc. A.*, 2016.
- [34] R. Rosipal and N. Krämer, "Overview and Recent Advances in Partial Least Squares," *Lecture Notes in Computer Science*, 2006.
- [35] T. F. Cootes, G. J. Edwards and C. J. Taylor, "Active appearance models," *Lect. Notes Comput. Sci.*, pp. 484-498, 1998.

Ringraziamenti

Secondo un vecchio detto, “chi ben comincia, è a metà dell’opera”. Giustissimo, certo, ma nessuno parla mai dell’altra metà di quest’opera o del lavoro necessario per portarla a termine. Ebbene, questo elaborato è l’atto finale della mia “opera”, ovvero del mio percorso di studi quinquennale al Politecnico di Torino, e quest’ultima sezione è dedicata a chi l’ha resa possibile, in un modo o nell’altro, concedendomi di arrivare fin qui:

grazie alla Dr. Alessandra Aldieri, co-relatrice di questa tesi, che è senza ombra di dubbio la persona più paziente e disponibile che ho avuto il piacere di incontrare in questi anni di studio;

grazie alla Professoressa Cristina Bignardi, relatrice di questa tesi, che è riuscita, proponendomi questo lavoro, a venire incontro sia alle mie inclinazioni, sia alle mie necessità;

grazie ai colleghi del PoliTo e del PoliMi con cui ho condiviso il percorso dell’Alta Scuola Politecnica e con cui ho imparato cosa significa lavorare insieme per trasformare un’idea in realtà;

grazie ai miei nonni, a quelli che ci sono e che hanno sempre una parola dolce per me, e a quelli che non ci sono più e che spero siano fieri di me, ovunque siano;

grazie alla mia sorellina, che illumina sempre le mie giornate, anche le più buie;

e, ultimi ma decisamente non per importanza, grazie ai miei più grandi fan, ai miei più saggi consiglieri nonché i miei critici più severi, ovvero mamma e papà; senza il loro irremovibile supporto, il loro incommensurabile amore ma soprattutto il loro impeccabile esempio non solo non avrei mai raggiunto questo traguardo, ma sarei anche una persona diversa, e sicuramente non per il meglio.