# Quantitating Age-Related BMD Textural Variation from DXA Region-Free-Analysis: A Study of Hip Fracture Prediction in Three Cohorts

Mohsen Farzi, PhD [1,2,3]; Jose M. Pozo, PhD [3]; Eugene McCloskey, PhD [1,2]; Richard Eastell, PhD [1,2]; Nicholas C. Harvey, PhD [4,5]; Alejandro F. Frangi, PhD [3]; Jeremy Mark Wilkinson, PhD [1,2]

[1] Department of Oncology and Metabolism, The University of Sheffield, Sheffield S10 2RX, U.K.

[2] The MRC-Arthritis Research UK Centre for Integrated research into Musculoskeletal Ageing (CIMA), The University of Sheffield, Sheffield S10 2RX, U.K.

[3] Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), The University of Leeds, Leeds LS2 9JT, U.K.

[4] The MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton SO16 6YD, U.K.

[5] NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital, Southampton, UK

J. Mark Wilkinson and Alejandro F. Frangi contributed equally to this work. Corresponding author: Prof J Mark Wilkinson, Department of Oncology and Metabolism, University of Sheffield, The Medical School, Beech Hill Road, Sheffield, S10 2RX, United Kingdom

Email: [*j.m.wilkinson@sheffield.ac.uk*](mailto:j.m.wilkinson@sheffield.ac.uk)

Telephone +44 114 271 4705

Fax +44 114 2618775

# Abstract

The risk of osteoporotic fracture is inversely related to bone mineral density (BMD), but how spatial BMD pattern influences fracture risk remains incompletely understood. This study used a pixel-level spatio-temporal atlas of proximal femoral BMD in 13,338 white European women (age 20 to 97 years) to quantitate age-related texture variation in BMD maps and generate a ‘reference’ map of bone ageing. We introduce a new index, called *Densitometric Bone Age* (DBA), as the age at which an individual site-specific BMD map (the proximal femur is studied here) best matches the median ageing trajectory at that site in terms of the root mean squared error (RMSE). The ability of DBA to predict incident hip fracture and hip fracture pattern over five years following baseline BMD was compared against conventional region-based BMD analysis in a subset of 11,899 women (age 45 to 97 years), for which follow-up fracture records exist. There were 208 subsequent incident hip fractures in the study populations (138 femoral necks, 52 trochanteric, 18 sites unspecified). DBA had modestly better performance compared to the conventional femoral neck (FN), trochanteric (TR), and total hip (TOT) BMD in identifying hip fractures measured as the Area Under the Curve (AUC) using Receiver Operating Characteristics (ROC) analysis by 2% (95%CI -0.5% to 3.5%), 3% (95%CI 1.0% to 4.0%), and 1% (95%CI 0.4% to 1.6%), respectively. Compared to FN-BMD T-score, DBA improved the ROC-AUC for predicting trochanteric fractures by ~5% (95%CI 1.1% to 9.8%) with similar performance in identifying neck fractures. Compared to TR-BMD T-score, DBA improved the ROC-AUC for the prediction of neck fractures by ~3% (95%CI 1.1% to 4.9%) with similar performance in identifying trochanteric fractures. Our findings suggest that DBA may provide a spatially sensitive measure of proximal femoral fragility that is not captured by neck BMD or trochanteric BMD alone.

**Key Words:** Ageing, DXA, BMD, Hip Fractures, Osteoporosis

# Introduction

The assessment of bone quality by measuring bone mineral density (BMD) using dual-energy x-ray absorptiometry (DXA) is a cornerstone of osteoporosis management.(1,2) The inverse relationship between areal BMD (aBMD) and incident fracture,(3) combined with clinical risk factors,(4) is used to guide clinical management.(5–8) Data from several studies have shown that site-specific measurement of BMD provides the best prediction of fracture risk at that site.(9,10) At the hip, a large meta-analysis of several prospective studies showed a relative risk for hip fracture of 2.6 (95%CI 2.0 to 3.5) per standard deviation of decrease in femoral neck (FN) BMD.(3) However, data from the Study of Osteoporotic Fractures (SOF) also shows that almost half of all fragility hip fractures occur in individuals with a femoral neck T-score of >-1.5.(10,11) Further, FN-BMD does not capture all determinants of bone strength in the proximal femur. For example, trochanteric (TR) BMD is associated with intertrochanteric femoral fractures independent of FN-BMD,(12) consistent with the observation that pixel-summation and quantitation of BMD within one region of interest (ROI) give rise to better fracture prediction within that ROI versus BMD at another site,(13,14) even if the sites are contiguous. A further limitation of conventional BMD assessment is that the output metrics (T-score and Z-score) are not intuitive for patients, making the explanation of the results difficult for them to interpret and contextualise.

Alternative analytical approaches have been explored to address the low sensitivity of DXA in population-attributable fracture risk prediction. At the lumbar spine, trabecular bone score (TBS)(15,16) is an analytic method that measures the rate of local variations in grey-level from the 2-dimensional lumbar spine DXA image to provide an indirect index of 3-dimensional trabecular microarchitecture. Several studies, reviewed in a European consensus report by Harvey *et al*.,(17) have demonstrated that TBS is a predictor of vertebral fracture independent of aBMD and provides complementary information on vertebral bone quality in diseases associated with fragility fractures. However, at the femoral neck, the site of greatest disease burden in osteoporosis, studies have not yet identified an analytic approach that extracts greater clinically-useful information from conventional DXA images than aBMD. Hip structural analysis (HSA) uses the distribution of mineral mass in a line of pixels across the bone axis to measure geometric properties of cross-sections of bone (compiled to ~5mm thickness) at that region.(18) This approach provides estimates of bending strength, compressive strength and buckling strength at discrete regions of interest in the plane of the DXA image. Lower bone strength measured with HSA was significantly linked with a higher tendency to fracture but its clinical utility in adequately predicting femoral neck fractures beyond routine aBMD requires further study.(14,19–21) Yang *et al*.(22) used DXA-based structural engineering models to calculate lateral stress upon the hip with a sideways fall that provided better femoral neck fracture prediction than aBMD but did not better predict trochanteric fractures. Several machine-learning-based methods have also been proposed to quantitate texture features from pixel-level DXA scans to improve hip fracture prediction.(23–25) Despite their potential merits, the small number of fracture cases (n≤50) is a significant limitation in these studies.(23–25) Given the relatively low rate of hip fractures in the population (~2 to 5%), employing a discriminative learning approach is prone to overfitting errors.

We have previously reported a technique termed DXA Region-Free Analysis (DXA-RFA) to extract pixel-level BMD from DXA datasets that describes bone loss occurring around hip joint prostheses.(26–29) We have recently used this method to develop a calibrated and validated spatio-temporal ageing atlas of the native proximal femur.(30) In the present study, we aimed to 1) introduce a more intuitive index, coined “Densitometric Bone Age” (DBA), derived from quantification of textural BMD variation using the developed ageing atlas for the proximal femur; 2) determine whether DBA can better predict incident hip fractures versus conventional FN-BMD, TR-BMD, and total hip (TOT) BMD T-score, and 3) determine whether DBA can predict the anatomic pattern of incident hip fracture (femoral neck versus trochanteric).

# Patients and Methods

## Study Populations

The pseudo anonymised patient demographic, incident hip fracture and imaging data described in this study were accessed from UK Biobank (approval 17881; July 09, 2018), the OPUS study (31) and the MRC-Hip study (32) (Table 1). Ethics approval for these cohorts was obtained under institutional and national requirements, and all subjects provided written informed consent before participation. The participants comprised white women of European descent (MRC-Hip study n=5,018, aged 75 to 97 years, mean (SD)=80 (3.9) years; OPUS n=213, aged 20 to 39 years, mean (SD) = 32 (5.3) years and n=1,189, aged 55 to 79 years, mean (SD)=67 (7.1) years; UK Biobank n=6,918, age 45 to 80 years, mean (SD)=62 (7.3) years). All scans (n=13,338) were used to develop the spatio-temporal BMD ageing atlas, as detailed in our previous work (30) and to assess the relationship between DBA and FN-BMD T-score and Z-score. The precision of DBA as a quantitative tool was estimated by analysing 25 pairs of DXA scans in a subset of the OPUS cohort. The scan pairs were collected on the same day with patient repositioning between acquisitions.

The cohort used to evaluate the ability of DBA to predict incident hip fractures comprised (n=5,018) participants in the MRC-Hip study followed up for 5 years after baseline DXA measurement and (n=6,881) participants in the UK Biobank study followed up for a mean of 4.4 years. One hundred eighty-one participants in the MRC-Hip study and 27 in the UK Biobank study suffered an incident hip fracture during the follow-up period. In the MRC-Hip study, hip fracture types were also assessed by plain radiography, with 123 cases reported at the neck, 40 patients at the trochanteric region, and 18 instances without defined fracture pattern. In the UK Biobank (UKBB) study, hip fracture types were assessed using a combination of ICD-10 (S72.0, S72.1) hip fracture and OPCS4 procedure codes. Fifteen fractures were classified as in the neck, and 12 were trochanteric.

In this study, scans were collected either on a Hologic QDR 4500A in the MRC-Hip and OPUS studies or an iDXA Lunar GE scanner in the UKBB study. To amalgamate data from two manufacturers, linear calibration parameters were estimated using the quantile matching regression technique, as previously validated.(30) In brief, n = 406 white British women matched for age and Body Mass Index (BMI) were selected for each scanner. Next, at each pixel coordinate, the linear calibration parameters, i.e. the slope and the intercept, were estimated such that the BMD distribution in each group were matched between scanners.

# Statistical methods

## Densitometric Bone Age Estimation

The spatio-temporal atlas of BMD in the proximal femur was developed as described previously.(30) In brief, BMD maps were obtained for each DXA scan with an isotropic spatial resolution of 0.5×0.5 mm2 using either Hologic Apex v3.2 (Hologic Inc, Waltham, MA) or Lunar enCORE v16 (GE Healthcare, Madison, WI) proprietary software, respectively. Automatic segmentation of the proximal femur was performed by selecting 65 landmark points around the bone contour using the “Bone-Finder v.1.2.0” software developed by Lindner et al.(33) A standard template comprised of ~16,000 pixels was generated by averaging over all segmented femurs. To remove the morphological variation between scans, all DXA scans were then warped into the template using a thin-plate spline (TPS) registration method using in-house Matlab software v9.7.0.1190202 R2019b (Mathworks, Cambridge, MA, USA).(34) Age-specific BMD distribution at each individual pixel was estimated smoothly using the R-package VGAM.(35) The resulting atlas allows probabilistic estimation of age-specific pixel-level BMD at any given anatomic site within the proximal femur across the studied age range. Here, we define the “normal bone ageing” trajectory by estimating the median BMD map in the population (Figure 1). The underlying assumption of the bone ageing trajectory is that all subjects follow a consistent path across the chronological ageing spectrum but at a different speed due to relatively accelerated/decelerated rates of bone loss during ageing. With this definition, DBA is the age at which the root mean squared error (RMSE) between the median BMD map and the individual BMD map is smallest (or minimum) (Figure 2). Note that DBA depends only on the spatial texture of BMD maps rather than the chronological age.

## Precision Analysis

DBA is a quantitative measurement technique. A subset of the OPUS cohort (n=25) was scanned twice on the same day, with patient repositioning between scans to assess its precision. For each scan, DBA was computed independently. The coefficient of variation (CV) was then calculated as the root mean square standard deviation divided by the mean of paired measurements.(36) To visualise the agreement between measurements, Bland-Altman plots were employed.

## DBA and its Relationship with the FN-BMD T-score and Z-score

DBA is similar in principle to the conventional T-score and Z-score. An individual’s “score” is defined by comparison against a set of reference values from the population. Figure 3 demonstrates this analogy for aBMD at the femoral neck. Note that similar to FN-BMD T-score and Z-score, DBA is also a linear function of FN-BMD. Figure 3 shows the bone ageing distribution and trajectory for FN-BMD, with each point representing an individual subject from the population. Similarly, for n=16,000 pixels representing the whole proximal femur, the bone ageing trajectory would be a smooth non-linear curve in this high-dimensional space. When DBA is computed over n~16,000 pixels on the template, the spatial BMD texture also contributes to calculating bone age. The Pearson coefficient correlation *r* is reported to assess the relationship between DBA and FN-BMD T-score and Z-score. All statistical tests were performed in MATLAB v9.7.0.1190202 (R2019b) (Mathworks Inc, Cambridge, MA). A p-value of below 0.05 was considered to be significant.



## Fracture Prediction

The ability of baseline DBA to predict incident hip fractures was compared versus conventional FN-BMD), TR-BMD and TOT-BMD to determine whether the use of the full spatial resolution of DXA can help to capture bone strength beyond traditional region-based BMD values. We compared DBA versus FN-BMD, TR-BMD and TOT-BMD by classifying subjects into two groups, i.e. fractured versus fracture-free controls. To determine the sensitivity for discrimination of fracture types (femoral neck versus trochanteric), we repeated the experiments for each fracture type separately.

Evaluating classification performance is challenging due to the low proportion of individuals suffering an incident fracture during the follow-up period, termed *class imbalance*. To address this issue, the Precision-Recall Characteristics (PRC) plot(37) was used besides the Receiver Operating Characteristics (ROC) analysis. The Area Under the Curve (AUC) was reported for the ROC and the PRC plots. To determine the 95% confidence interval (CI) and statistical significance of any difference between the fracture versus fracture-free curve profiles, bootstrapping with n=1000 repetitions was employed.

## Fracture Patterns

ROC analysis was employed to classify fractured cases into neck versus trochanteric fractures to determine whether DBA can differentiate between fracture patterns. To further visualise spatially complex fracture-specific patterns, the component of the BMD texture pattern that is attributable to ageing alone, shown in Figure 1, must be removed. To cancel the ageing effect, BMD maps were normalised in respect of their DBA as follows: for each individual BMD map and at each pixel coordinate, the probability of observing a BMD value lower than the given pixel BMD among the population with a similar DBA is reported as a number between 0 and 1. Here, we refer to these normalised BMD maps as quantile maps. Next, the pixel-level changes in quantile maps were tested using a Mann Whitney U test between the fracture-free control group and the fractured cases. To account for the multiple testing issue, computed p-values were reported as q-values defined as the minimum false discovery rate (FDR) level for which a pixel is selected as significant, as described previously.(27) Regions with a q-value <0.05 were considered significant. All statistical tests were performed in MATLAB v9.7.0.1190202 (R2019b) (Mathworks Inc, Cambridge, MA).

# Results

## DBA Precision

Figure 4 shows the Bland-Altman plot for estimated DBA for each scan pair (*n*=25). The coefficient of variation was 2.3%. The mean difference in DBA between the first and second scan measurements was 0.5 years (95% CI -0.4 to 1.2 years).

## DBA relation with the Neck BMD T-score and Z-score

Figure 5 shows the estimated DBA versus the chronological age across the study cohorts (n=13,338). Each subject is represented with a single dot colour-coded by measured FN-BMD T-score category: osteoporotic (red; T-score), osteopenic (yellow; T-score), or normal (green; T-score). DBA was linearly correlated with both FN-BMD T-score (R2=-0.82; p-value<0.001) and Z-score (R2=0.78; p-value<0.001). T-score was inversely proportional to DBA (as DBA increases along the -axis, the T-score decreases, also demonstrated by the vertical colour variation from green to red among the scattered dots). The Z-score was directly proportionate to the chronological age minus DBA for each subject. Note that at Z-score=0, the dashed black line follows the solid blue identity line. As points deviate from the identity line, the Z-score increases for points below the identity line and decreases for points above the identity line.

Given the linear correlation between DBA and FN-BMD T-score, the corresponding cut-off DBA thresholds for T-scores -2.5 and -1 were 83.5 and 54.4 years, respectively. Figure 6 shows the confusion matrix for classifying subjects into osteoporotic, osteopenic, and normal using DBA versus the FN-BMD T-score. Few osteoporotic cases were misclassified as normal. This is consistent with observing a clear clustering demarcation line between normal (green dots) and osteoporotic (red dots) subjects in Figure 5. Most misclassifications were attributed to the osteopenic cases defined by either DBA or FN-BMD T-score (Figure 6). This is observed as yellow dots overlaid on the green and red dots in Figure 5. These observations suggest that spatial BMD patterns may provide discriminatory information in cases with intermediate FN-BMD.

## Fracture Prediction

Figure 2 provides a visual example of BMD maps for subjects who sustained a follow-up incident trochanteric fracture (a) and a control subject with similar neck BMD and age but remained fracture-free (b). The fractured subject had widespread bone loss in the trochanteric region. This texture variation is reflected in the 18-year difference in DBA between the fracture versus control subject, despite the same chronological age and FN-BMD. Tables 2 and 3 show the AUC for the corresponding ROC and PRC plots for quantitative analysis, respectively. In the ROC analysis, AUC=1 for an ideal classifier and AUC=0.5 for a random classifier (Figure 7). In the PRC analysis, AUC=1 for a perfect classifier and AUC=*p* for a random classifier, where *p* is the proportion of fractured cases among the cohort (Figure 8).

The ROC analysis for the prediction of both trochanteric and neck fractures suggests a slight increase of approximately 1% (95% CI 0.4% to 1.6%), 2% (95% CI -0.5% to 3.5%), and 3% (95% CI 1.0% to 4.0%) in AUC for DBA in comparison to TOT-BMD, FN-BMD and TR-BMD, respectively. Compared to FN-BMD, DBA improved the ROC-AUC for predicting trochanteric fractures by ~5% (95% CI 1.1% to 9.8%) with similar performance for detecting neck fractures. Compared to TR-BMD, DBA improved the ROC-AUC for predicting neck fractures by ~3% (95% CI 1.1% to 4.9%) with similar performance to detect trochanteric fractures (Table 2).

The PRC analysis suggested that DBA was more precise than FN-BMD, TR-BMD, and TOT-BMD in the prediction of incident fractures, but it was not statistically significant (Table 3).

## Fracture Patterns

Bone age, unlike FN-BMD T-score, was not dependent on the potential site of fracture (Tables 1 and 2). Figure 9 shows heat maps for the pixel-by-pixel difference between controls and fractured groups. The observed fracture-specific pixel BMD patterns were spatially complex. For neck fractures, the q-map shows a local pattern of bone deficiency that was most apparent in the same orientation as the principal tensile trabeculae first characterised in plain radiograph imaging by Singh et al.(38) For trochanteric fractures, widespread bone loss in the trochanteric region was observed.

# Discussion

We have examined DBA as a potential new marker of bone quality by quantitating age-related spatial texture variation in BMD maps using a recently developed spatio-temporal atlas of BMD in the proximal femur(30). We compared the ability of DBA to conventional region-based BMD measurements including FN-BMD, TR-BMD and TOT-BMD to predict hip fractures in a large cohort of n=11,899 white women from the MRC-Hip(32) and the UK Biobank studies(39). DBA showed the highest ROC-AUC, modestly improving the overall performance in comparison to TOT-BMD by 1% (95% CI 0.4% to 1.6%), TR-BMD by 3% (95% CI 1.0% to 4.0%), and FN-BMD by 2% (95% CI -0.5% to 3.5%).

Our findings confirm that region-specific BMD measurements are most sensitive to fractures occurring within the same anatomical site with relatively worse performance at other locations. FN-BMD was better than TR-BMD in identifying neck fractures whereas TR-BMD performed better for identifying trochanteric fractures, as measured as ROC-AUC (Table 2). DBA, however, was sensitive to both fracture types capturing texture patterns in the neck and trochanteric regions simultaneously. Compared to FN-BMD, DBA improved the ROC-AUC for predicting trochanteric fractures by ~5%, but with a similar performance for predicting neck fractures. Compared to TR-BMD, DBA improved the ROC-AUC for predicting neck fractures by ~3% with similar performance for the prediction of trochanteric fractures. Compared to TOT-BMD, DBA improved the ROC-AUC for the prediction of neck and trochanteric fractures by ~0.6% and ~1.4%, respectively.

Our results demonstrated that DBA is a precise quantitative tool, based, as tested by independent analysis of repeat scan acquisitions collected on the same day with patient repositioning between scans. Given its intuitive definition, DBA concept may facilitate patient communication and engagement in clinical practice. This concept is analogous to vascular age (40) or brain age (41) that have been proposed to express the risk of cardiovascular diseases or dementia, respectively. Note that the terminology bone age is not new; it is used by paediatricians to quantitate skeletal maturity in a child and is based on a comparison of a wrist radiograph with atlas patterns to assess the closure of the growth plates (42,43). To avoid confusion, the proposed concept is called densitometric bone age.

The overall prediction performance of a tool in detecting an event is a function of the prevalence of such events in the population. The population-attributable risk (PAR) for an incident hip fracture in this study was 52% (versus 28% in the SOF study(10)) for cohorts with a cut-off T-score of -2.5 and 85% (versus 51% in the SOF study(10)) for a more conservative cut-off point of -1.5. Unlike the SOF study (10) our findings suggest that a high proportion of proximal femoral fragility fractures (85%) may be attributed to low FN-BMD T-score. However, note that PAR depends not only on the excess risk imposed by low FN-BMD but also on the proportion of exposed subjects with a T-score below the cut-off threshold. In our cohort study, the prevalence of osteoporosis was 23.6% (versus 17.7% in the SOF study(10)) and 61.8% (versus 48.1% in the SOF study(10)) using the cut-off T-score of -2.5 and -1.5, respectively. The elevated PAR in our study may be attributed to the higher prevalence of subjects with a low FN-BMD T-score in our cohorts.

Here, we developed a reference ageing atlas of BMD textural variation in the whole populationand interpreted osteoporosis as a natural process of senescence. DBA was proposed as a surrogate for bone quality in the proximal femur by mapping individual BMD scans to the median ageing trajectory. Note that an alternative discriminative learning approach could also be adopted by generating two different ageing trajectories for the fractured cases and the fracture-free controls. We did not adopt this discriminative approach as only a small proportion of the population (~2%) experienced incident hip fractures. Moreover, whether a control subject would remain fracture-free cannot be guaranteed.

DBA does not *per se* predict specific fracture-specific patterns. However, analysing normalised BMD maps by their corresponding DBA suggests the potential for further improving fracture prediction by incorporating the observed spatially complex fracture patterns. For example, in Figure 9 those individuals who went on to have an incident femoral neck fracture had a baseline DBA texture pattern of BMD deficiency that was greatest in the distribution of the principal tensile trabeculae whilst those sustaining an incident trochanteric fracture had a baseline DBA texture pattern of BMD deficiency that was evident more diffusely in the trochanteric region. Distinct variation between observed fracture patterns attributed to the neck versus trochanteric fractures thus support the idea that identifying relevant texture features from BMD maps might facilitate fracture risk assessment.

This study also has limitations. First, the follow-up period for the cohort from the UK Biobank studies was relatively short, at 3 to 6 years. (39) Since fracture information were extracted from Hospital Episodes Statistics (HES) data that was truncated by May 31 2020, and variable baseline scan time, a variable follow-up period was inevitable to include the maximum number of participants. Secondly, while the atlas development methodology is generic and can be readily applied to other ethnicities, our current findings are only applicable to white European women.

In conclusion, the results of this study suggest the potential for improving fracture prediction by analysing spatial BMD texture patterns. We have shown that the proposed bone age concept is consistent to current diagnostic guidelines but provides a more intuitive reflection than both T-score and Z-score. Besides potential clinical value in facilitating patient communication, we showed that DBA is more precise than FN-BMD in identifying trochanteric fractures and may facilitate early-stage fracture risk screening.

# Acknowledgement

The work of M. Farzi was supported by a PhD Fellowship from the Medical Research Council/Arthritis Research UK Centre for Integrated Research into Musculoskeletal Ageing under Grant MR/P020941/1. The work of AF Frangi is partially funded by the Royal Academy of Engineering (CiET1819\19). This work was undertaken using data from the MRC-Hip study (MRC/G9518113) and OPUS study (sponsored by Eli Lilly, Sanofi-Aventis, Procter & Gamble Pharmaceuticals, Hoffman-La Roche, Pfizer, and Novartis). The authors thank the OPUS Steering Committee for permission to use its DXA scans.

# References

1. Kanis JA, Melton III LJ, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. J. Bone Miner. Res. 1994;9(8):1137–41.

2. Kanis J a., McCloskey E V., Johansson H, Oden A, Melton LJ, Khaltaev N. A reference standard for the description of osteoporosis. Bone. 2008;42(3):467–75.

3. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ. 1996;312(7041):1254–9.

4. Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. Lancet. 2002;359(9321):1929–36.

5. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R. Clinician’s Guide to Prevention and Treatment of Osteoporosis. Osteoporos. Int. [Internet]. 2014;25(10):2359–81. Available from: https://doi.org/10.1007/s00198-014-2794-2

6. Hip fracture: management (NICE Guideline CG124). London: National Institute for Health and Care Excellence (UK); 2017.

7. Compston J, Cooper A, Cooper C, Gittoes N, Gregson C, Harvey N, Hope S, Kanis JA, McCloskey E V., Poole KES, Reid DM, Selby P, Thompson F, Thurston A, Vine N, The National Osteoporosis Guidline Group (NOGG). UK clinical guideline for the prevention and treatment of osteoporosis. Arch. Osteoporos. Springer; 2017;12(1):43.

8. 2019 surveillance of osteoporosis: assessing the risk of fragility fracture (NICE guideline CG146). London: National Institute for Health and Care Excellence (UK); 2019.

9. Melton III LJ, Atkinson EJ, O’Fallon WM, Wahner HW, Riggs BL. Long-term fracture prediction by bone mineral assessed at different skeletal sites. J. Bone Miner. Res. 1993;8(10):1227–33.

10. Stone KL, Seeley DG, Lui L-Y, Cauley JA, Ensrud K, Browner WS, Nevitt MC, Cummings SR. BMD at Multiple Sites and Risk of Fracture of Multiple Types: Long-Term Results From the Study of Osteoporotic Fractures. J Bone Min. Res. 2003;18(11):1947–54.

11. Mai HT, Tran TS, Ho-Le TP, Center JR, Eisman JA, Nguyen T V. Two-Thirds of All Fractures Are Not Attributable to Osteoporosis and Advancing Age: Implications for Fracture Prevention. J. Clin. Endocrinol. Metab. 2019;104(8):3514–20.

12. Greenspan SL, Myers ER, Maitland LA, Kido TH, Krasnow MB, Hayes WC. Trochanteric bone mineral density is associated with type of hip fracture in the elderly. J. Bone Miner. Res. 1994;9(12):1889–94.

13. Greenspan S, Maitland L. Femoral bone loss progresses with age: a longitudinal study in women over age 65. J. Bone … [Internet]. 1994 [cited 2015 Apr 19]; Available from: http://onlinelibrary.wiley.com/doi/10.1002/jbmr.5650091216/abstract

14. Kaptoge S, Beck TJ, Reeve J, Stone KL, Hillier TA, Cauley JA, Cummings SR. Prediction of Incident Hip Fracture Risk by Femur Geometry Variables Measured by Hip Structural Analysis in the Study of Osteoporotic Fractures. J. Bone Miner. Res. 2008;23(12):1892–904.

15. Silva BC, Leslie WD, Resch H, Lamy O, Lesnyak O, Binkley N, McCloskey E V., Kanis JA, Bilezikian JP. Trabecular bone score: A noninvasive analytical method based upon the DXA image. J. Bone Miner. Res. 2014;29(3):518–30.

16. Pothuaud L, Carceller P, Hans D. Correlations between grey-level variations in 2D projection images (TBS) and 3D microarchitecture: Applications in the study of human trabecular bone microarchitecture. Bone. 2008;42(4):775–87.

17. Harvey NC, Glüer CC, Binkley N, McCloskey EV, Brandi M-L, Cooper C, Kendler D, Lamy O, Laslop A, Camargos BM, Reginster J-Y, Rizzoli R, Kanis JA. Trabecular bone score (TBS) as a new complementary approach for osteoporosis evaluation in clinical practice. Bone. 2015;78:216–24.

18. Beck TJ, Ruff CB, Warden KE, Scott WW, Rao GU. Predicting femoral neck strength from bone mineral data: a structural approach. Invest. Radiol. 1990;25(1):6–18.

19. Rivadeneira F, Zillikens MC, De Laet CEDH, Hofman A, Uitterlinden AG, Beck TJ, Pols HAP. Femoral Neck BMD Is a Strong Predictor of Hip Fracture Susceptibility in Elderly Men and Women Because It Detects Cortical Bone Instability: The Rotterdam Study. J. Bone Miner. Res. 2007;22(11):1781–90.

20. Johnson JT, Cherian KE, Kapoor N, Jebasingh FK, Asha HS, Mathai T, Nithyananth M, Oommen AT, Daniel AJ, Thomas N, Paul TV. Does hip structural analysis confer additional benefit to routine BMD assessment in postmenopausal women with hip fracture? A study from a tertiary center in southern India. Arch. Osteoporos. 2022;17(1):1–7.

21. Khoo BCC, Lewis JR, Brown K, Prince RL. Evaluation of a simplified hip structure analysis method for the prediction of incident hip fracture events. Osteoporos. Int. 2016;27(1):241–8.

22. Yang L, Peel N, Clowes JA, McCloskey E V, Eastell R. Use of DXA-Based Structural Engineering Models of the Proximal Femur to Discriminate Hip Fracture. J. Bone Miner. Res. 2009;24(1):33–42.

23. Hong N, Park H, Kim CO, Kim HC, Choi J-Y, Kim H, Rhee Y. Bone Radiomics Score Derived From DXA Hip Images Enhances Hip Fracture Prediction in Older Women. J. Bone Miner. Res. 2021;0(0):1–9.

24. Jazinizadeh F, Adachi JD, Quenneville CE. Advanced 2D image processing technique to predict hip fracture risk in an older population based on single DXA scans. Osteoporos. Int. Osteoporosis International; 2020;31(10):1925–33.

25. Boehm HF, Vogel T, Panteleon A, Burklein D, Bitterling H, Reiser M. Differentiation between post-menopausal women with and without hip fractures: Enhanced evaluation of clinical DXA by topological analysis of the mineral distribution in the scan images. Osteoporos. Int. 2007;18(6):779–87.

26. Morris RM, Yang L, Martín-Fernández MA, Pozo JM, Frangi AF, Wilkinson JM, Tower P. High-Spatial-Resolution Bone Densitometry with Dual-Energy X-ray Absorptiometric Region-free Analysis. Radiology. Radiological Society of North America; 2015;274(2):532–9.

27. Farzi M, Morris RM, Penny J, Yang L, Pozo JM, Overgaard S, Frangi AF, Wilkinson JM. Quantitating the effect of prosthesis design on femoral remodeling using high-resolution region-free densitometric analysis (DXA-RFA). J. Orthop. Res. 2017;35(10):2203–10.

28. Parker AM, Yang L, Farzi M, Pozo JM, Frangi AF, Wilkinson JM. Quantifying Pelvic Periprosthetic Bone Remodeling Using Dual-Energy X-Ray Absorptiometry Region-Free Analysis. J. Clin. Densitom. 2017;20(4):480–5.

29. Slullitel PA, Mahatma MM, Farzi M, Grammatopoulos G, Wilkinson JM, Beaulé PE. Influence of Femoral Component Design on Proximal Femoral Bone Mass After Total Hip Replacement: A Randomized Controlled Trial. JBJS. 2021;103(1).

30. Farzi M, Pozo JM, McCloskey E, Easte R, Harvey N, Wilkinson JM, Frangi AF. A Spatio-Temporal Ageing Atlas of the Proximal Femur. IEEE Trans. Med. Imaging. 2020;39(5):1359–68.

31. Glüer CC, Eastell R, Reid DM, Felsenberg D, Roux C, Barkmann R, Timm W, Blenk T, Armbrecht G, Stewart A, Clowes J, Thomasius FE, Kolta S. Association of five quantitative ultrasound devices and bone densitometry with osteoporotic vertebral fractures in a population-based sample: the OPUS Study. J. Bone Miner. Res. 2004;19(5):782–93.

32. McCloskey E V, Beneton M, Charlesworth D, Kayan K, DeTakats D, Dey A, Orgee J, Ashford R, Forster M, Cliffe J, Kersh L, Brazier J, Nichol J, Aropuu S, Jalava T, Kanis J a. Clodronate reduces the incidence of fractures in community-dwelling elderly women unselected for osteoporosis: results of a double-blind, placebo-controlled randomized study. J. Bone Miner. Res. United States; 2007 Jan;22(1):135–41.

33. Lindner C, Thiagarajah S, Wilkinson JM, Wallis GA, Cootes TF. Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE Trans. Med. Imaging. United States; 2013 Aug;32(8):1462–72.

34. Bookstein FL. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. IEEE Trans. Pattern Anal. Mach. Intell. 1989;11(6).

35. Yee TW. Quantile regression via vector generalized additive models. Stat. Med. 2004;23(14):2295–315.

36. Lodder MC, Lems WF, Ader HJ, Marthinsen AE, Van Coeverden SCCM, Lips P, Netelenbos JC, Dijkmans BAC, Roos JC. Reproducibility of bone mineral density measurement in daily practice. Ann. Rheum. Dis. 2004;63(3):285–9.

37. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One. 2015;10(3):1–21.

38. Singh M, Nagrath A, Maini P. Changes in trabecular pattern of the upper end of the femur as an index of osteoporosis. J. Bone Jt. Surg. 1970;52(3):457–67.

39. Littlejohns TJ, Holliday J, Gibson LM, Garratt S, Oesingmann N, Alfaro-Almagro F, Bell JD, Boultwood C, Collins R, Conroy MC, Crabtree N, Doherty N, Frangi AF, Harvey NC, Leeson P, Miller KL, Neubauer S, Petersen SE, Sellors J, Sheard S, Smith SM, Sudlow CLM, Matthews PM, Allen NE. The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. Nat. Commun. [Internet]. Springer US; 2020;11(1):1–12. Available from: http://dx.doi.org/10.1038/s41467-020-15948-9

40. Groenewegen KA, Den Ruijter HM, Pasterkamp G, Polak JF, Bots ML, Peters SAE. Vascular age to determine cardiovascular disease risk: A systematic review of its concepts, definitions, and clinical applications. Eur. J. Prev. Cardiol. 2016;23(3):264–74.

41. Wang J, Knol MJ, Tiulpin A, Dubost F, De Bruijne M, Vernooij MW, Adams HHH, Ikram MA, Niessen WJ, Roshchupkin G V. Gray matter age prediction as a biomarker for risk of dementia. Proc. Natl. Acad. Sci. U. S. A. 2019;116(42):21213–8.

42. Cole a J, Webb L, Cole TJ. Bone age estimation: a comparison of methods. Br. J. Radiol. 1988;61(728):683–6.

43. Pietka E, Gertych A, Pospiech S, Cao F, Huang HK, Gilsanz V. Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. IEEE Trans. Med. Imaging. 2001;20(8):715–29.

# Figure Legends

**![A picture containing chart

Description automatically generated]()**

**Figure 1.** Median spatial BMD maps for a population of white European women (n=13,338).(30) Note that the bone ageing trajectory is a continuum, and the six bone maps shown here at equal intervals of 15 years are for visual purposes.

Chart, diagram

Description automatically generated

**Figure 2**. A visual example of spatial BMD maps to differentiate between fractured and control subjects with similar neck BMD values. Panel (A) shows the bone map for a woman aged 75.8 years with femoral neck BMD of 0.5860 g/cm2 who experienced a trochanteric fracture following the baseline measurement. Panel (C) shows the bone map for a non-fracture subject with similar age (75.9 years) and femoral neck BMD (0.5900 g/cm2). Despite similar age and femoral neck BMD, the widespread trochanteric bone loss, which is not captured by neck BMD, resulted in a trochanteric fracture for the first subject. Panels (B) and (D) show the Root Mean Squared Error (RMSE) between the BMD maps in panels (A) and (C) and the median BMD maps in Figure 1, respectively. Densitometric bone age (DBA) is the age at which the RMSE is minimum. The associated DBA was 80 and 62 years for the top and bottom subjects, respectively.



**Figure 3**. The analogy between the T-score, Z-score, and densitometric bone age (DBA). The solid black line shows the average neck BMD; the red dashed lines show one standard deviation (SD). The blue dot represents an individual aged 60 years with neck BMD=0.63 g/cm2. DBA is the age at which the measured BMD equals the average BMD, i.e. 86.5 years. The BMD distribution is normal for areal neck BMD because of the pixel averaging in the neck region, and the median and mean trajectories overlapped here. When DBA is computed based on pixel BMD values, the BMD distribution is no longer normal, and the median trajectory is used instead.



**Figure 4**. Bland-Altman plot comparing estimated densitometric bone age before () and after () patient repositioning. μ= mean bias, SD= standard deviation. Dashed lines represent the 95% confidence interval (mean±1.96SD). The solid black line represents the overall difference (mean bias, μ).

****

**Figure 5**. Densitometric bone age (DBA) versus chronological age. Panel (A) shows the scattered plot for all cohort subjects. Each green, yellow, or red dot represents one subject from the study cohort categorised by T-score as normal (green), osteopenic (yellow), or osteoporotic (red), respectively. The solid blue line shows the line of equality or identity, and the dashed black line indicates the median DBA as a function of age. The dashed black line almost perfectly follows the blue identity line, demonstrating that DBA equals chronological age on average. Deviation of DBA from the chronological age is proportionate to the Z-score. The red and green dashed lines show Z-scores and , respectively. Panel (B) shows the distribution of fractured cases in relation to DBA and chronological age. Fractured cases were on average 10 years older in terms of DBA compared to the chronological age. In 72% of fractured cases (150/208), DBA was higher than the chronological age.

Text, table

Description automatically generated

**Figure 6**. Confusion matrix to assess the consistency between the densitometric bone age (DBA) and the neck aBMD T-score. Using the cut-off thresholds of 83.5 and 54.4 years for DBA, 80.6% of subjects were categorised in the same group as identified by the neck aBMD T-score of -2.5 and -1, respectively.

**Chart, line chart

Description automatically generated**

**Figure 7.** The Receiver Operating Characteristics (ROC) plots for prediction of fragility fractures. The black dashed line shows the performance of a random classifier, with the solid lines representing densitometric bone age (black), femoral neck BMD (red), trochanteric BMD (blue), and total hip BMD (green). See Table 2 for the reported Area Under the Curve (AUC) values for each graph. A higher AUC indicates better performance.

Chart

Description automatically generated

**Figure 8.** The Precision-Recall-Characteristics (PRC) plots for prediction of fragility fractures. The horizontal dashed black line shows the performance for a random classifier where its height equals the proportion of fractured cases in the population, with the solid lines representing densitometric bone age (black), femoral neck BMD (red), trochanteric BMD (blue), and total hip BMD (green). See Table 3 for the reported Area Under the Curve (AUC) values for each graph. A higher AUC indicates better performance.

**Chart

Description automatically generated with medium confidence**

**Figure 9**.Localising fracture-specific patterns using bone-age normalised BMD maps. Panels (a) and (b) show the difference in mean quantile maps between the fracture-free control groups (n=11,691) and neck fractured cases (n=138) and trochanteric fractured cases (n=52), respectively. Panels (c) and (d) show the corresponding statistical significance map using a Mann Whitney U test followed by False Discovery Rate (FDR) analysis. In panel (c), a local pattern of BMD deficiency was observed in the same orientation as the principal tensile trabeculae described in plain radiographs of the hip.(38) In panel (d), widespread BMD deficiency was observed in the trochanteric region. Panels (e) and (f) show the PP plot for the FDR analysis. In case of no significant difference, the solid black curve should follow the identity dashed line.

# Tables

**Table 1.** Characteristics of the Patient Populations Participating in This Study

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | N | **Gender** | **Age Span**  **[years]** | **Age Mean (SD) [years]** | **Body Mass Index**  **[kg/m2]** | **#Hip Fractures** |
| **UK Biobank Study** | 6,918 | Female | 45-80 | 62 (7.3) | 25.7 (4.7) | 27 |
| **MRC-Hip Study** | 5,018 | Female | 75-97 | 80 (3.9) | 26.6 (5.2) | 181 |
| **OPUS study-group1** | 1,189 | Female | 55-79 | 67 (7.1) | 26.2 (5.2) | - |
| **OPUS study-group2** | 213 | Female | 20-39 | 32 (5.3) | 24.2 (5.0) | - |

**Table 2.** Area Under the Curve for Receiver-Operating-Characteristics Analysis for the Prediction of Incident Fractures

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **#Controls** | **#Fractures** | **Densitometric Bone Age** | **Femoral Neck BMD** | **Trochanteric BMD** | **Total Hip BMD** |
| **All cohort**  **All fractures** | 11,691 | 208 | **0.799**  CI=[0.768-0.824] | 0.784  CI=[0.751-0.815] | 0.774  CI=[0.742-0.802] | 0.789  CI=[0.756-0.816] |
| **All cohort**  **Trochanteric fractures** | 11,691 | 52 | **0.839**  CI=[0.795-0.880] | 0.789  CI=[0.722-0.844] | 0.823  CI=[0.771-0.869] | 0.826  CI=[0.775-0.872] |
| **All cohort**  **Neck fractures** | 11,691 | 138 | 0.775  CI=[0.739-0.811] | **0.786**  CI=[0.747-0.821] | 0.746  CI=[0.704-0.785] | 0.768  CI=[0.731-0.805] |

**Table 3.** Area Under the Curve for Precision-Recall-Characteristics Analysis for the Prediction of Incident Fractures

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **#Controls** | **#Fractures** | **Densitometric Bone Age** | **Femoral Neck BMD** | **Trochanteric BMD** | **Total Hip BMD** |
| **All cohort All fractures** | 11,691 | 208 | **0.072**  CI=[0.053-0.097] | 0.070  CI=[0.053-0.093] | 0.063  CI=[0.046-0.085] | 0.069  CI=[0.051-0.092] |
| **All cohort**  **Trochanteric fractures** | 11,691 | 52 | **0.021**  CI=[0.012-0.037] | 0.019  CI=[0.010-0.031] | 0.020  CI=[0.011-0.039] | 0.019  [0.011-0.033] |
| **All cohort**  **Neck fractures** | 11,691 | 138 | 0.046  CI=[0.031-0.066] | **0.049**  CI=[0.035-0.068] | 0.036  CI=[0.025-0.052] | 0.046  CI=[0.031-0.066] |