Osteoporosis International (2025) 36:715–724
https://doi.org/10.1007/s00198-025-07421-4

ORIGINAL ARTICLE

Advancing trabecular bone score (TBS): clinical performance of TBS version 4.0 with direct correction for soft tissue thickness—the osteolaus study

Guillaume Gatineau[1,2] · Karen Hind[2,3,4] · Enisa Shevroja[1] · Elena Gonzalez-Rodriguez[1] · Olivier Lamy[1] · Didier Hans[1,2]

Received: 11 August 2024 / Accepted: 3 February 2025 / Published online: 4 March 2025

Abstract

Summary This study compared TBS v4.0, which uses DXA-derived tissue thickness corrections, with TBS v3, which adjusts using BMI. TBS v4.0 improved soft tissue adjustments and maintained fracture risk prediction equivalence with TBS v3, enhancing applicability across diverse body compositions/phenotypes. Direct tissue thickness adjustment increases TBS's utility in osteoporosis assessment and management.

Purpose This study aimed to compare trabecular bone score (TBS) version 4.0, which uses direct tissue thickness correction via DXA measurements, with TBS version 3, which adjusts for soft tissues using body mass index (BMI). The objective was to assess the performance of TBS v4.0 compared to v3, for bone health evaluation and fracture risk assessment across diverse body compositions.

Methods Data from the OsteoLaus cohort were analyzed. Associations between TBS, BMI, DXA-measured tissue thickness, visceral fat (VFAT), and android fat were examined using regression and correlation analyses. Machine learning, including Random Forest (RF) and SHapley Additive exPlanations (SHAP), explored TBS changes between versions. Five-year fracture risk was assessed using FRAX adjustment, and logistic regression.

Results TBS v3 correlated with BMI ($r=0.110$, $p<0$ .001), VFAT mass ($r=-0.162$, $p<0$ .001), and soft tissue thickness ($r=-0.165$, $p<0.001$). TBS v4.0 demonstrated weaker correlations with BMI ($r=-0.057$, $p>0.999$), VFAT Mass ($r=-0.067$, $p>0.779$), and soft tissue thickness ($r=-0.114$, $p=0.019$).

Differences between TBS versions were investigated with SHapley Additive exPlanations (SHAP) and explained by BMI, tissue thickness, VFAT, and gynoid fat. Logistic regression and Delong's test revealed no significant differences in *vertebral* fracture prediction between the two TBS versions ($p=0.564$). FRAX adjustments were highly consistent between versions ($r=0.994$, $p<0.001$), with no evidence of calibration bias ($p=0.241$).

Conclusion TBS v4.0 enhances the adjustment for regional soft tissue effects and results suggest comparable vertebral fracture risk prediction to TBS v3. Explainable AI provided insights into the contributions of BMI, tissue thickness, visceral fat, and gynoid fat to the observed changes between TBS versions. Incorporating direct tissue thickness adjustment improves TBS applicability across diverse body sizes and compositions.

Keywords Bone fragility · Bone microarchitecture · Fracture risk · Trabecular bone score · Osteoporosis

Introduction

The densitometric evaluation of bone health and fracture risk is a crucial component in the clinical assessment and management of osteoporosis. Traditionally, this evaluation has relied on the measurement of bone mineral density (BMD) using dual-energy X-ray absorptiometry (DXA). While BMD is clinically useful, it has limitations, particularly in accounting for the diverse factors influencing fracture risk. To address these limitations, trabecular bone score (TBS) was introduced over one decade ago, to provide a more comprehensive densitometric assessment of bone health. TBS (medimaps group SA, Plan-les-Ouates, Switzerland) assesses bone texture by evaluating gray scale variation in lumbar spine DXA images, serving as a surrogate measure

Extended author information available on the last page of the article

of bone microarchitecture that enhances fracture prediction independent of BMD, correlating closely with bone microarchitectural and mechanical properties [1–4]. TBS classifies the bone quality as normal, partially degraded, or degraded, alongside BMD which classifies bone quantity classification as normal, osteopenia, and osteoporosis. The combination of TBS and BMD is shown to enhance fracture risk assessment, treatment decision-making, and monitoring [5].

TBS reliably predicts fragility fracture independent of BMD and clinical risk factors (CRFs) and is an optional adjustment to the Fracture Risk Assessment Tool (FRAX) [6]. The 2023 update to the International Society for Clinical Densitometry (ISCD) position statement highlights the role of TBS in clinical practice, emphasizing its utility in assessing osteoporosis and fracture risk [7]. Furthermore, the integration of TBS into clinical workflows is supported by the recent expert position paper from the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the International Osteoporosis Foundation (IOF), under the auspices of the World Health Organization Collaborating Centre for Epidemiology of Musculoskeletal Health and Aging [5]. This paper provides comprehensive guiding statements on the use of TBS for fracture risk assessment in primary and secondary osteoporosis, treatment initiation, and monitoring, based on a series of systematic reviews.

Technically, deriving TBS from DXA images necessitates adjustment for soft tissue thickness anterior and posterior to the lumbar spine. These tissues possess distinct X-ray absorption properties, which can cause an underestimation of TBS values due to X-ray attenuation and associated noise. To some extent, this is also true for BMD measurements, as soft tissue interference can influence the accuracy of areal BMD assessments by affecting the attenuation of the X-ray beams used in dual-energy imaging. This interference may lead to variations in BMD values, particularly in individuals with higher levels of adipose tissue, as the X-ray path length and absorption properties are not solely determined by the bone itself but also by the overlying soft tissues. While the impact on BMD is typically less pronounced than on TBS, accounting for these factors is essential for improving the precision of both metrics in clinical assessments. Previous versions of the TBS algorithm have utilized body mass index (BMI) as a surrogate for regional soft tissue thickness [8]. While effective for most, this approach has some limitations for individuals whose tissue thickness is not accurately represented by BMI. Recent technological advancements in TBS measurements allow for direct adjustment of soft tissue thickness derived from the DXA. As such, the TBS software version 4.0 [currently TBS v4.0 (TBS Core Module v19.4.0)] accommodates variability in soft tissue thickness among individuals, which may not be completely captured by BMI. As such, TBS v4.0 is not restricted by the BMI range used for TBS v3 but instead, includes a soft tissue thickness restriction of 7–30 cm. This update aims to optimize the performance of TBS across a wider range of body sizes, improving the tool`s utility and inclusivity in osteoporosis assessment and management.

The objectives of this study were to evaluate the performance of TBS version 4.0, which incorporates direct tissue thickness corrections using DXA-derived measurements, compared to TBS version 3, which uses BMI as a surrogate for soft tissue thickness adjustment. Specifically, we investigated associations between DXA-measured tissue thickness, BMI, and body composition parameters (e.g., visceral fat and android fat) to understand factors contributing to differences between TBS versions. We also aimed to identify individuals whose TBS values exhibited significant changes (exceeding the least significant change, LSC) between versions due to these adjustments. Furthermore, the study compared the ability of TBS v4.0 and TBS v3 to predict fragility fractures in a population-based cohort of postmenopausal women, with an emphasis on evaluating clinical performance equivalence of TBS v4.0 across diverse body composition phenotypes.

Materials and methods

Study cohort

The OsteoLaus Study, a subset of the larger CoLaus|PsyColaus study, is a prospective population-based study of post-menopausal European women from Lausanne, Switzerland [9]. Details of the cohort, including inclusion and exclusion criteria can be found elsewhere (8). Data was collected longitudinally every 2.5 years over one decade, beginning in September 2010, with follow-ups in 2012, 2015, 2018, and 2020. This study utilized data from the visit 3, 5 years after the baseline visit. A BMI range restriction of 15–37 kg/m^2 as well as a soft tissue thickness range restriction of 7–30 cm were applied to ensure clinical validity. To ensure comparability between TBS versions, individuals with either out-of-BMI or out-of-soft tissue thickness ranges were further removed from the analysis. All assessments were performed at the Interdisciplinary Center of Bone Diseases, Lausanne University Hospital, Switzerland. The Ethics Committee for Human Research of Canton Vaud reviewed and approved the OsteoLaus Study. Participants provided written informed consent after receiving a detailed explanation of the study's objectives and funding.

Clinical assessments

At each visit, participants received a DXA scan of the lumbar spine (LS), hip, and total body (Lunar iDXA™, GE

Healthcare, Madison, WI, USA). TBS analysis at the lumbar spine (L1 to L4) was computed using both the TBS v3 and the pre-FDA TBS v4.0 (TBS Core Module v19.4.0) algorithms. The Core Module 19.4.0 algorithm represents the final pre-FDA version of TBS v4.0, incorporating a standardization parameter to ensure clinical usability and maintain comparable thresholds with earlier TBS v3 versions. Vertebral-level fracture exclusions for TBS calculations and lumbar spine BMD were based on artifact identification and in accordance with ISCD official positions (https://iscd.org/official-positions-2023/). Body composition variables were derived from the total body DXA scan, and included visceral adipose tissue mass (VFAT) mass (g), android, and gynoid fat percentage. At each visit, a clinical questionnaire recorded the participant medical history, fracture, and FRAX (Swiss) 10-year probability of fracture. Major osteoporotic fractures (MOF) were diagnosed, as non-traumatic fractures of the radius, humerus, hip, or vertebra, determined from questionnaire records and VFA. Prevalent fracture was defined as a fragility fracture that occurred between the baseline and visit 3 (year 5). Incident fracture was defined as a new fragility fracture occurring over the 5-year period after the visit 3 (year 5).

Fracture risk assessment

Incident fractures occurring within 2.5 years and 5 years from the visit 3 were included in the analysis. The FRAX MOF-adjusted by TBS was computed with both TBS versions at the 5-year visit. Pearson's product-moment correlation coefficient was used to assess the correlation strength between software versions and FRAX MOF. Deming Regression, Bland–Altman, and Student *t*-test analyses were conducted to assess the calibration of both FRAX adjustments and any differences.

The performance of TBS for future fracture prediction using v4.0 and v3 was evaluated using binary logistic regression. Covariates were age, BMI, and lumbar spine BMD T-score. A fivefold cross-validation technique was employed to train and evaluate both logistic regression models. The predictive performance of both software versions was assessed with odds ratios per standard deviation (SD) decrease in TBS. The overall performance of the models was assessed using AUROC and compared with Delong's test to ascertain any statistical difference between the AUROCs.

Statistical analysis

Statistical analyses were performed using R version 4.3.2 and Python version 3.9.1. TBS measurements correcting for BMI were obtained from the TBS iNsight v3.1.2 computation module, while TBS measurements corrected for abdominal soft tissue thickness were derived from ModuleCalcul v19.4.0, the final pre-FDA algorithm, implemented in TBS iNsight v4.0.

Differences in descriptive outcomes between women with and without fractures were assessed using Wilcoxon rank sum tests. Additionally, differences in incident fractures between those with and without prevalent fractures, defined as fractures that occurred between baseline and third (5-year) visit, were evaluated using Pearson's chi-squared test. Pearson's correlation coefficient was used to explore relationships between DXA-derived soft tissue thickness and body composition parameters, specifically visceral fat mass and android region fat percentage. The hypothesis was that directly measured soft tissue thickness would be a better estimator of body phenotype than BMI, correlating more strongly with android soft tissues. Polynomial regression, coefficient of determination, and correlation analyses were conducted to quantify and compare the associations between VFAT, android region fat, BMI, and soft tissue thickness.

Individual-level analysis was performed to identify participants whose TBS v4 values were discordant relative to TBS v3 (+/−5.3% LSC) [10]. The 5.3% LSC threshold was originally introduced in the 2005 ISCD Official Position paper as the minimum acceptable precision for an individual technologist when assessing lumbar spine BMD. Since the ISCD has not yet established a specific LSC threshold for TBS, we applied this widely recognized BMD LSC as a reference. This approach can be further supported by independent studies, such as the Manitoba cohort, which reported a comparable LSC for TBS (5.5%) [11]. While further validation is needed to define a TBS-specific LSC, this threshold provides a standardized approach to evaluate discordance between TBS v3 and v4 in clinical practice. The Random Forest (RF) machine learning model was used to identify explainable attributes of discordant cases, as a binary classification task, including lumbar spine tissue thickness, BMI, trunk fat mass, VFAT, trunk lean mass, android fat mass, android region fat percentage, gynoid fat mass, gynoid region fat percentage, android/gynoid fat percentage ratio, femoral neck BMD T-Score, and lumbar spine BMD T-Score. Model training and optimization involved fivefold cross-validation and grid-search for hyperparameter tuning. This enabled the computation of mean and 95% confidence intervals (CI) for performance metrics, including AUROC, precision, sensitivity, and specificity. Shapley Additive Explanations (SHAP) analysis was conducted with the trained RF models to explore feature importance and contributions for discordant cases.

Results

The study cohort from the third OsteoLaus visit included 1237 participants. From the 1237 participants, 1205 (97.4%) were within both the TBS v3 BMI validity range [15 kg m^{-2},

37 kg m^{-2}] and TBS v4 soft tissue thickness validity range [7 cm, 30 cm]. A total of 28 participants who were out of the BMI range still fell within the soft tissue thickness range, demonstrating that TBS v4 accommodates individuals previously excluded under BMI-based criteria. Four participants (0.3%) were outside both the BMI and soft tissue thickness valid ranges. From the 1205 participants with valid BMI and soft tissue thickness values, 14 were further excluded because of missing lumbar spine DXA data, which led to a study sample size of 1191.

From the 1191 women included in the analysis, 121 (10.2%) sustained a MOF over the 5-year period (Table 1). Among the 121 individuals who sustained a MOF, 20 were low trauma forearm fractures, 3 low trauma hip fractures, 10 low trauma humerus fractures, and 88 vertebral fractures. Vertebral fractures were defined as having at least two grade 1, one grade 2, or one grade 3 vertebra(e), following the Genant semiquantitative classification [12].

TBS v3 Age, weight, tissue thickness, body composition variables, TBS, and BMD differed between the fractured and non-fractured groups (Table 1). The proportion of women sustaining incident fractures in 5 years from the third visit was greater in the prevalent fracture group (32%) than non-prevalent fracture group (13%) ($p < 0.001$).

Table 1 OsteoLaus Visit 3 data characteristics on GE Lunar iDXA, grouped by 5-year prevalence of major osteoporotic fracture

Variable	Prevalent MOF (5 years)		
	No fracture $N = 1070$[1]	Fracture $N = 121$[1]	p-value[2]
Age (years)	68.7 (7.3)	72.8 (6.8)	< 0.001
Height (cm)	160.6 (6.5)	160.2 (7.2)	0.8
Weight (kg)	65.9 (11.1)	69.1 (12.6)	0.013
BMI (kg m^{-2})	25.6 (4.3)	26.9 (4.3)	0.001
LS thickness (cm)	20.7 (2.8)	21.8 (2.8)	< 0.001
Visceral fat mass (g)	826.0 (576.7)	985.0 (667.4)	0.013
Android fat (%)	0.42 (0.1)	0.44 (0.1)	0.027
Gynoid fat (%)	0.43 (0.1)	0.45 (0.06)	0.005
LS BMD T-score	− 0.1 (1.7)	− 1.4 (1.7)	0.003
FN BMD T-score	− 1.2 (0.8)	− 1.5 (0.9)	0.002
TH BMD T-score	− 0.8 (1.0)	− 1.2 (1.0)	< 0.001
TBS v3	1.326 (0.086)	1.282 (0.080)	< 0.001
TBS v4	1.301 (0.071)	1.262 (0.070)	< 0.001
5y incident MOF	108/814 (13%)	25/77 (32%)	< 0.001

[1]Mean (SD); *n*/*N* (%)

[2]Wilcoxon rank sum test; Pearson's chi-squared test

The grouping variable, Prevalent MOF (5 years), groups participants that sustained or not have major osteoporotic fractures between the baseline and the third visit

The 5y Incident MOF variable represents the proportion of individuals that sustained major osteoporotic fractures between the third visit and the last visit (10-year follow-up)

TBS v3 BMI range: 15 to 37 kg m.$^{-2}$

MOF major osteoporotic fracture, *BMI* body mass index (kg m^{-2}), *LS* lumbar spine, *FN* femoral neck, *TH* total hip, *TBS* trabecular bone score

Associations between DXA-derived soft tissue thickness and body composition variables

DXA-derived soft tissue thickness and BMI correlated positively with VFAT and android fat percentage, with stronger correlations for soft tissue thickness (Table 2). Second-order polynomial regression plots and coefficient of determination (Fig. 1A and B) demonstrated superior goodness of fit for tissue thickness compared to BMI for the prediction of VFAT (R^2/R^2 adjusted = 0.747/0.746 versus 0.575/0.575). Similarly, tissue thickness exhibited a superior goodness of fit (Fig. 1C and D), compared to BMI for prediction of android fat percentage (R^2/R^2 adjusted = 0.650/0.650 versus 0.620/0.619).

Associations between TBS, tissue thickness, and body composition variables

The present study data involved the GE Lunar iDXA scanner at visit 3. TBS v3 correlated with BMI ($r = 0.110$, $p < 0$ 0.001), VFAT mass ($r = -0.162$, $p < 0$ 0.001), and soft tissue thickness ($r = -0.165$, $p < 0.001$). TBS v4.0 demonstrated weaker correlations with BMI ($r = -0.057$, $p > 0.999$), VFAT mass ($r = -0.067$, $p > 0.779$), and soft tissue thickness ($r = -0.114$, p = 0.019). There were no associations with android fat percentage (Table 2). Linear regression models exhibited reduced associations between TBS v4.0 and soft tissues compared to TBS v3, with lower coefficients of determination for BMI (R^2/R^2 adjusted = 0.003/0.003 versus 0.0012/0.0011), LS soft tissue thickness (R^2/R^2 adjusted = 0.013/0.012 versus 0.0027/0.027), and visceral fat mass (R^2/R^2 adjusted = 0.004/0.004 versus 0.026/0.025) (Fig. 2).

Individual-level analysis

A total of 16.6% of participants had discordant TBS values between v3 and v4, with 13.6% showing lower TBS v4 values and 3.0% showing higher TBS v4 values compared to v3. For predicting an increase in TBS v4.0, where TBS v4.0 exceeds TBS v3.0 by > 5.3%, the Random Forest model achieved a mean Area Under the Curve (AUC), precision, sensitivity and specificity of 0.96 (95% CI, 0.94 to 0.97), 0.4 (95% CI, 0.00 to 1.00), 0.10 (95% CI, 0.00 to 0.14), and 1.00 (95% CI, 1.00 to 1.00), respectively. For predicting a decrease in TBS v4.0, where TBS v3.0 exceeds TBS v4.0 by > 5.3%, the model achieved a mean

Table 2 OsteoLaus Visit 3 correlations between TBS v3, TBS v4.0, and soft tissue measurements using the GE Lunar iDXA

	TBS v3	TBS v4	BMI	Android fat %	VFAT mass	LS tissue thickness
TBS v3		0.837 (*<0.001*)	0.110 (*0.027*)	−0.023 (*>0.999*)	−0.162 (*<0.001*)	−0.165 (*<0.001*)
TBS v4	0.837 (*<0.001*)		−0.057 (*>0.999*)	−0.029 (*>0.999*)	−0.067 (*>0.779*)	−0.114 (*0.019*)
BMI	0.110 (*0.027*)	−0.057 (*>0.999*)		0.787 (*<0.001*)	0.758 (*<0.001*)	0.908 (*<0.001*)
Android fat %	−0.023 (*>0.999*)	−0.029 (*>0.999*)	0.787 (*<0.001*)		0.796 (*<0.001*)	0.806 (*<0.001*)
VFAT mass	−0.162 (*<0.001*)	−0.067 (*0.779*)	0.758 (*<0.001*)	0.796 (*<0.001*)		0.864 (*<0.001*)
LS tissue thickness	−0.165 (*<0.001*)	−0.114 (*0.019*)	0.908 (*<0.001*)	0.806 (*<0.001*)	0.864 (*<0.001*)	

Computed correlation used Pearson method with pairwise-deletion

p-values were adjusted for multiple testing using the Bonferroni correction method

BMI body mass index, *VFAT* visceral fat, *LS* lumbar spine, *TBS* trabecular bone score

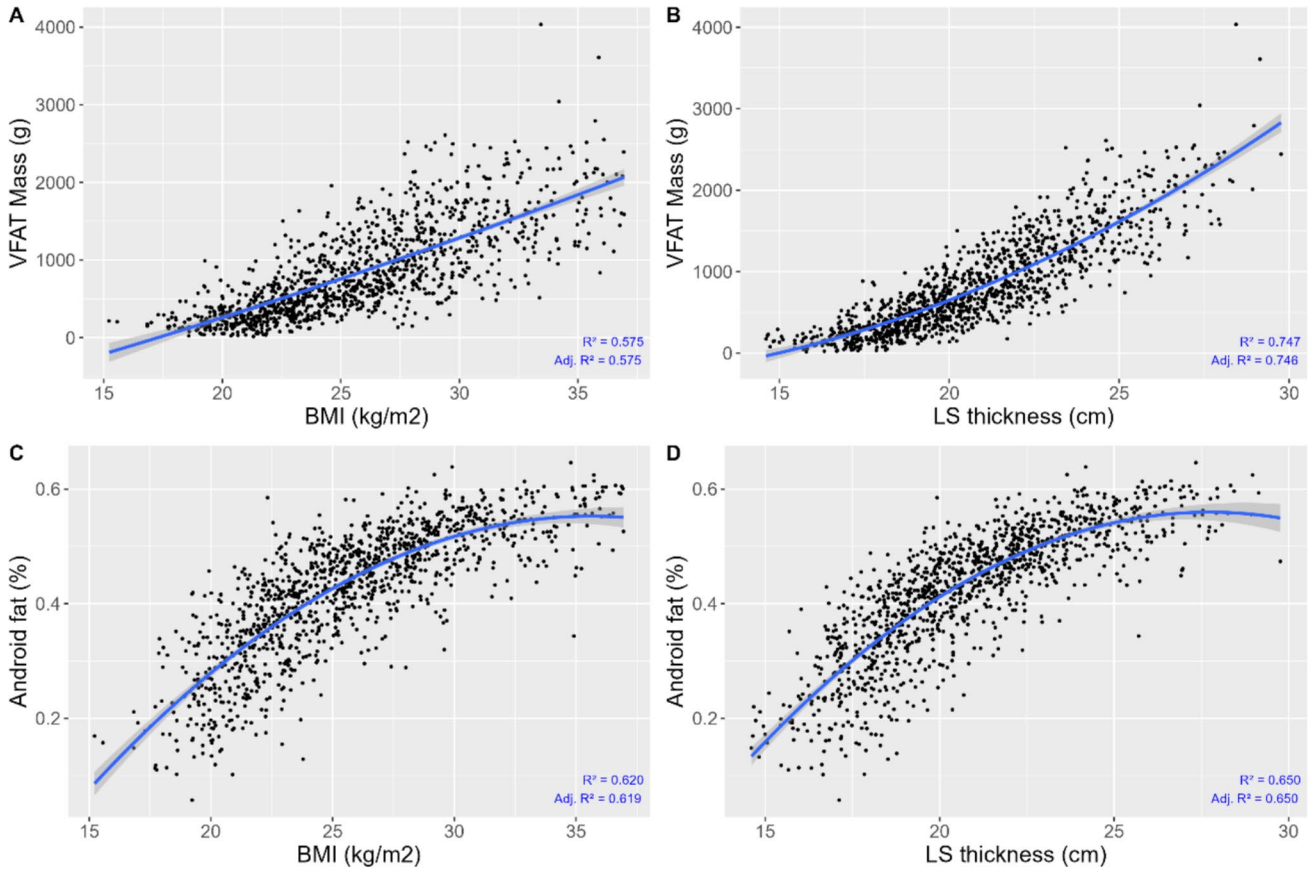

Fig. 1 OsteoLaus Visit 3 associations between body mass index, lumbar spine tissue thickness, visceral fat mass, and android fat percentage on GE lunar iDXA. BMI body mass index, VFAT visceral fat, LS lumbar spine, R^2 coefficient of determination, $Adj.R^2$ adjusted coefficient of determination. **A** Order 2 polynomial regression between VFAT Mass and BMI. **B** Order 2 polynomial regression between VFAT Mass and LS tissue thickness. **C** Order 2 polynomial regression between android fat percentage and BMI. **D** Order 2 polynomial regression between android fat percentage and LS tissue thickness

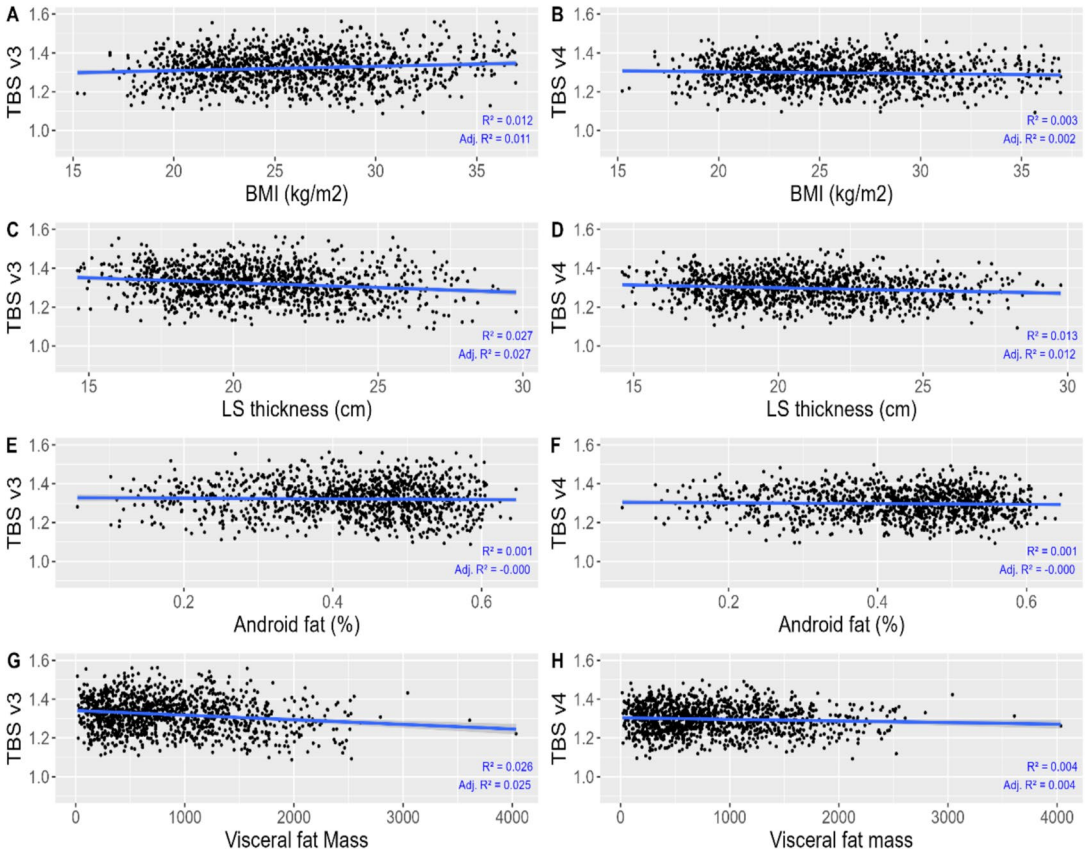

Fig. 2 OsteoLaus Visit 3 associations between TBS v3 and TBS v4.0 with soft tissues on GE Lunar iDXA. TBS trabecular bone score, BMI body mass index, VFAT visceral fat, LS lumbar spine. R^2 coefficient of determination, $Adj.R^2$ adjusted coefficient of determination. **A** Linear regression between TBS v3 and BMI. **B** Linear regression between TBS v4.0 and BMI. **C** Linear regression between TBS v3 and LS tissue thickness. **D** Linear regression between TBS v4 and LS tissue thickness. **E** Linear regression between TBS v3 and android fat percentage. **F** Linear regression between TBS v4 and android fat percentage. **G** Linear regression between TBS v3 and visceral fat mass. **H** Linear regression between TBS v4 and visceral fat mass

AUC, precision, sensitivity, and specificity of 0.90 (95% CI, 0.86 to 0.93), 0.90 (95% CI, 0.76 to 1.00), 0.26 (95% CI, 0.17 to 0.42), and 1.00 (95% CI, 0.99 to 1.00), respectively. The AUC values reported represent the performance of a classification model designed to predict whether TBS v4.0 values increased or decreased relative to TBS v3.0. These results highlight the model's ability to capture the contribution of key explanatory variables such as BMI, tissue thickness, VFAT, and gynoid fat. The low sensitivity can primarily be attributed to the small proportion of data representing positive instances.

The SHAP analysis proved more useful and identified explanatory variables for individuals with TBS discordance (Fig. 3). These individuals were more likely to have a BMI which was not reflective of their central tissue thickness and fat distribution. For example, cases with lower TBS v4.0 were explained by higher BMI with lower tissue thickness and central fat distribution (Fig. 3A). Conversely, cases with higher TBS v4.0 were associated with lower BMI and higher tissue thickness (Fig. 3B).

Fracture prediction

There were significant correlations between FRAX-adjusted by TBS using v3 and v4.0 ($r = 0.994$, $p < 0.001$). Deming Regression demonstrated great fit against identity, and the Bland–Altman plot paired with Student's *t*-test revealed good calibration with no significant difference between TBS v3 and TBS v4.0 ($p = 0.241$) (Fig. 4).

TBS v4.0 independently predicted fracture at 5 years after adjusting for age, BMI, and LS BMD T-score. For each SD decline in TBS v3, there was a significant 48% (OR 1.48; 95% CI, 1.17 to 1.89) increase in the odds of MOF. For each SD decline in TBS v4.0, there was a significant 57% (OR 1.57; 95% CI, 1.22 to 2.05) increase in the odds of MOF. The AUCs of the models accounting for TBS v3 and TBS v4.0 were 0.665 (95% CI; 0.51 to 0.83) and 0.671 (95% CI; 0.52 to 0.83). Delong's test revealed no significant difference ($p = 0.564$) between the logistic regression predictive models using TBS v3 and TBS v4.

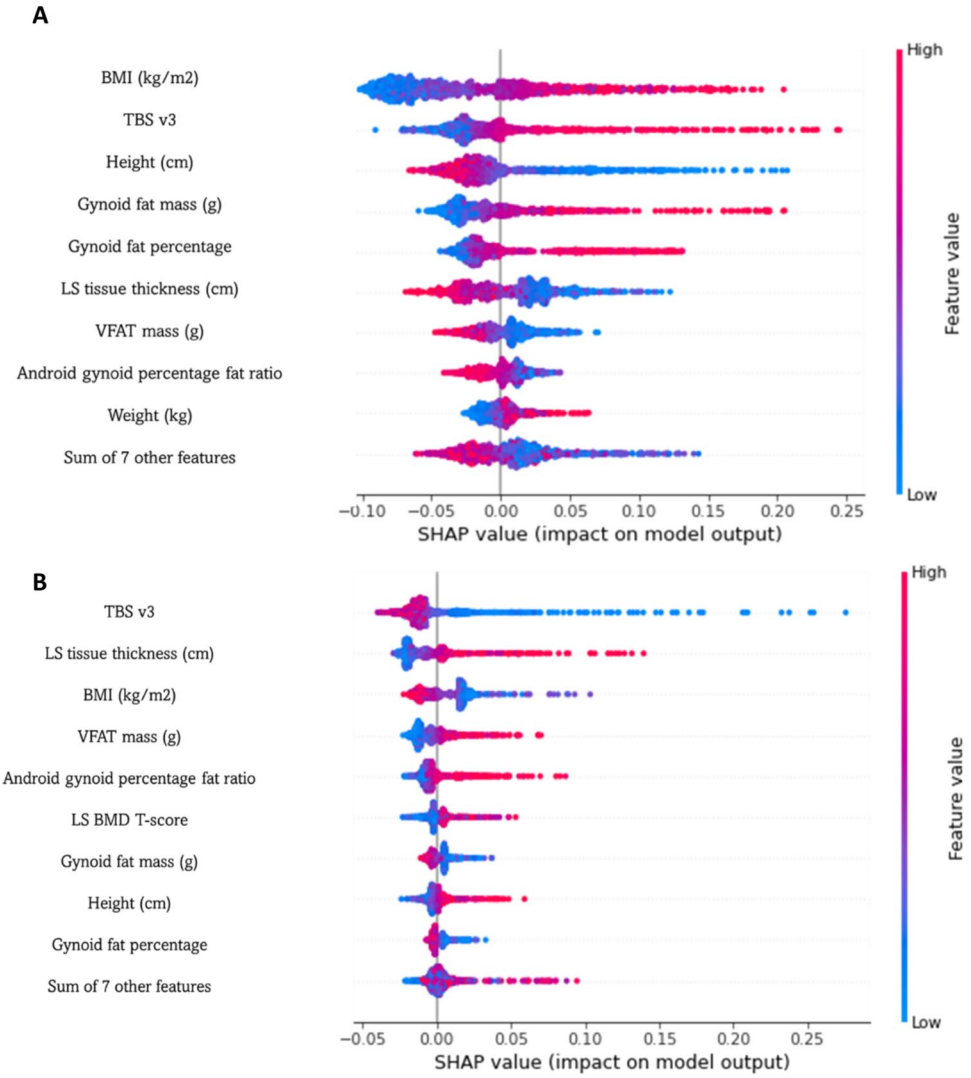

Fig. 3 OsteoLaus Visit 3 SHAP values for the iDXA TBS v4.0 change prediction. **A** SHAP values for the iDXA TBS v4.0 low change prediction. **B** SHAP values for the iDXA TBS v4.0 high change prediction. Each dot represents an instance in the dataset. The features are ordered from top to bottom by their importance. Positive SHAP values (towards the right) indicate that the feature contributes to a positive predicted value (TBS v4.0 change). Conversely, negative SHAP values (towards the left) indicate that the feature contributes to a decrease in the predicted value (no change in TBS v4.0). The color of each dot represents the feature value, with blue indicating low feature values and red indicating high feature values. TBS trabecular bone score, BMI body mass index, VFAT visceral fat, LS lumbar spine

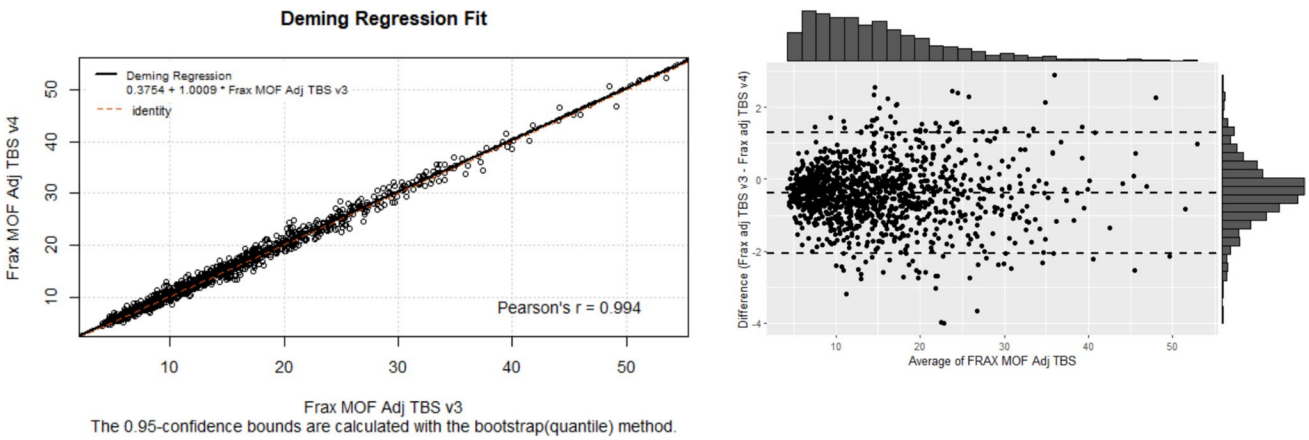

Fig. 4 OsteoLaus Visit 3 Deming regression and Bland–Altman plots of the FRAX MOF adjusted by TBS v4.0 and FRAX MOF adjusted by TBS v3

For FRAX adjusted by TBS (with ≥ 20% probability of MOF), TBS v4.0 demonstrated a slight decrease in accuracy over 2.5 years but a slight improvement over 5 years. The observed improvement with the BMD T-score adjusted by TBS reinforces the clinical benefits of TBS v4.0 when clinical risk factors are not included in fracture risk evaluations.

Comparing TBS and soft tissue interactions across osteolaus study visits

To further investigate the residual associations of soft tissues with TBS, additional analyses were conducted using data from the baseline ($n = 1475$) and Visit 2 ($n = 1349$), which involved a Hologic Discovery A DXA system, as well as Visits 4 ($n = 1104$) and 5 ($n = 944$), which involved a GE Lunar iDXA system).

For data collected using the Hologic DXA system, TBS v3 exhibited significant negative correlations with lumbar spine soft tissue thickness at baseline ($r = -0.294$, $p < 0.001$) and Visit 2 ($r = -0.332$, $p < 0.001$). In contrast, TBS v4 showed positive correlations with soft tissue thickness at both time points ($r = 0.201$ and $r = 0.185$, respectively; $p < 0.001$). Similar trends were observed for BMI, with TBS v3 showing negative correlations ($r = -0.181$ at baseline and $r = -0.221$ at Visit 2; $p < 0.001$) and TBS v4 showing positive correlations ($r = 0.203$ and $r = 0.182$, respectively; $p < 0.001$).

For Visits 4 and 5, data collected using the GE Lunar iDXA system showed further reductions in correlations between soft tissue thickness and TBS. At Visit 4, correlations were weak and not statistically significant for both algorithms (TBS v3, $r = -0.093$, $p = 0.072$; TBS v4, $r = -0.072$, $p = 0.621$). At Visit 5, correlations were negligible and non-significant (TBS v3, $r = -0.089$, $p = 0.235$; TBS v4, $r = -0.053$, $p = 1$). BMI correlations also showed neutral or weak positive trends across visits 4 and 5, with negligible associations for TBS v4.

Discussion

This study assessed the performance of TBS v4.0, which corrects for regional tissue thickness, in comparison to TBS v3, which corrects for BMI. The direct measurement of tissue thickness, as utilized by TBS v4.0, is particularly crucial given the known limitations of BMI. BMI, being an indirect measure, can often misrepresent actual soft tissue thickness, especially in individuals with diverse body composition phenotypes. By directly using DXA-measured tissue thickness over the region of interest, TBS v4.0 offers a more accurate evaluation, essential for clinical assessments and better patient outcomes. Our findings demonstrate that TBS v4.0 showed improved consistency by reducing the influence of regional fat tissues, evidenced by lower residual correlations with soft tissue measures compared to TBS v3. This reduction in variability leads to a more reliable assessment of bone quality, which is crucial for accurately predicting fracture risk.

Importantly, TBS v4.0 demonstrated fracture prediction performance comparable to TBS v3. The AUCs for models incorporating TBS v3 and TBS v4.0 were 0.665 and 0.671, respectively, indicating no significant difference in predictive performance. For each SD decline, TBS v3 was associated with a 48% increase in the odds of MOF (OR 1.48; 95% CI, 1.17 to 1.89), while TBS v4.0 was associated with a 57% increase (OR 1.57; 95% CI, 1.22 to 2.05). Delong's test showed insufficient evidence to reject that there is no difference in the AUC of the predictive models ($p = 0.564$).

These findings suggest that TBS v4.0 preserves the predictive ability of its predecessor for fracture risk assessment.

In a related study, Shevroja et al. evaluated the predictive performance of TBS adjusted for soft tissue thickness (beta version) for future MOF using the baseline visit of the OsteoLaus cohort, where DXA measurements were acquired on a Hologic Discovery A system. The study reported that for each SD decline in TBS v3.03, there was a 43% (OR 1.43; 95% CI, 1.12 to 1.83) increase in the odds of having MOF; whereas for each SD decline in TBSv4.0, there was a 54% (OR 1.54; 95% CI, 1.18 to 2.00) increase in the odds of MOF. These results align with the present findings obtained from a GE iDXA system, confirming the maintained fracture predictive abilities of TBS across different DXA devices.

The individual analysis of cases with discordant TBS results between software versions identified explanatory variables related to variations in body size phenotype, particularly when BMI does not accurately reflect central soft tissue thickness and fat distribution. TBS v3 utilized BMI as a surrogate for soft tissue thickness, converting it to a thickness-equivalent for adjustment. However, the BMI-to-tissue-thickness relationship implemented in TBS v3 may not accurately reflect the characteristics of the current cohort, introducing a systematic bias. This calibration bias likely resulted in systematically lower TBS v4 values in this cohort, as TBS v4 directly incorporates lumbar spine tissue thickness measurements from DXA, which are more closely associated with visceral and android fat distribution. In addition to the SHAP method, Random Forests thresholds from the trained model revealed that individuals with a BMI below 25.3 kg/m^2 and lumbar spine tissue thickness above 23.7 cm were more likely to experience a higher TBS v4.0. Conversely, those with a BMI above 27.2 kg/m^2 and lumbar spine tissue thickness below 20.5 cm may have a lower TBS v4.0 compared to TBS v3.

These findings underscore the importance of soft tissue thickness adjustments in providing a more nuanced understanding of bone health by accounting for regional tissue

distribution and its effects on absorptiometry. Clinically, the ability of TBS v4.0 to correct TBS measurements using direct tissue thickness from DXA, as opposed to BMI correction in TBS v3, suggests that this new method could lead to more accurate and individualized bone health evaluations. For DXA centers upgrading to TBS v4.0, patients with higher soft tissue thickness, visceral fat mass, and android fat mass may show increased TBS, while those with higher BMI, shorter stature, and increased gynoid fat may show decreased TBS.

The performance of TBS v4.0 for fracture prediction was consistent across different clinical scenarios, indicating that it preserves the predictive ability of its predecessor. Notably, FRAX adjusted by TBS v3 and TBS v4.0 were well calibrated and correlated ($r = 0.994$, $p < 0.001$), which supports clinical continuity when transitioning between software versions. This consistency ensures that TBS v4.0 can be integrated into existing clinical workflows without compromising fracture risk assessment.

The strengths of this study include the use of high-quality and well-defined population-based data, comprising both densitometric and body composition DXA scans, and fracture outcomes from the OsteoLaus cohort. Advanced machine learning techniques, including Random Forests and SHAP analysis, provided valuable insights into individual-level changes in TBS values. Additionally, extensive follow-up data supported fracture prediction analysis, making the cohort representative of the female post-menopausal population, a group disproportionately affected by osteoporosis and where TBS has substantial clinical value. However, this study has several limitations. It was assumed that the coefficients used in the TBS-adjusted FRAX calculation would be identical for TBS v4, which requires further confirmation despite the strong correlation of 0.994 between the two models. Additionally, the specific profile of the cohort, consisting of post-menopausal Swiss women aged 55 years and above, may limit the broader generalizability of the findings. The higher proportion of incident vertebral fractures (72.7%) compared to other major osteoporotic fracture (MOF) sites (27.3%) may limit the assessment of TBS's predictive performance. Further validation is needed to determine whether TBS v4 offers comparable MOF prediction to TBS v3, particularly in larger cohorts, men, individuals with type 2 diabetes, and hip fractures. Future studies should also assess its responsiveness to body weight changes and generalize these findings across diverse populations.

In conclusion, TBS 4.0 performed well across various analyses assessing its accuracy and comparative performance to TBS v3. The direct adjustment for abdominal soft tissue thickness represents an improvement over prior software versions, without adversely affecting its ability to classify and longitudinally predict major osteoporotic fractures. This update enhances the accuracy of TBS in evaluating bone quality and predicting fracture risk across diverse body composition phenotypes, thereby supporting more individualized and precise osteoporosis management.

Acknowledgements Special thanks to the radiology technicians of the CiMO / CHUV for data acquisition and to the study nurse for their assistance. Additionally, we extend our gratitude to Michele DeGruttola and Madeleine Davies, scientists at Medimaps Group, for their valuable comments and technical support. This work was supported by the Swiss National Science Foundation (SNSF 32473B_156978 and 320030_188886) and the Foundation of the Orthopedic Hospital of the Vaudois University Hospital, Lausanne, Switzerland.

Funding Open access funding provided by University of Lausanne.

Data Availability Some or all datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest Enisa Shevroja, Elena Gonzalez-Rodriguez, and Olivier Lamy declare that they have no conflicts of interest. Guillaume Gatineau and Karen Hind are employees of Medimaps Group SA, developers of TBS iNsight™ software. Didier Hans is co-owner of the TBS patent, has corresponding shares, and is part-time CEO at Medimaps Group.

References

1. Pothuaud L, Carceller P, Hans D (2008) Correlations between grey-level variations in 2D projection images (TBS) and 3D microarchitecture: applications in the study of human trabecular bone microarchitecture. Bone 42(4):775–787. https://doi.org/10.1016/j.bone.2007.11.018
2. Hans D, Barthe N, Boutroy S, Pothuaud L, Winzenrieth R, Krieg MA (2011) Correlations between trabecular bone score, measured using anteroposterior dual-energy X-ray absorptiometry acquisition, and 3-dimensional parameters of bone microarchitecture: an experimental study on human cadaver vertebrae. J Clin Densitom 14(3):302–312. https://doi.org/10.1016/j.jocd.2011.05.005
3. Winzenrieth R, Michelet F, Hans D (2013) Three-Dimensional (3D) Microarchitecture correlations with 2D projection image gray-level variations assessed by trabecular bone score using high-resolution computed tomographic acquisitions: effects of resolution and noise. J Clin Densitom 16(3):287–296. https://doi.org/10.1016/j.jocd.2012.05.001
4. Muschitz C, Kocijan R, Haschka J et al (2015) TBS reflects trabecular microarchitecture in premenopausal women and men

with idiopathic osteoporosis and low-traumatic fractures. Bone 79:259–266. https://doi.org/10.1016/j.bone.2015.06.007
5. Shevroja E, Reginster JY, Lamy O et al (2023) Update on the clinical use of trabecular bone score (TBS) in the management of osteoporosis: results of an expert group meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO), and the International Osteoporosis Foundation (IOF) under the auspices of WHO Collaborating Center for Epidemiology of Musculoskeletal Health and Aging. Osteoporos Int 34(9):1501–1529. https://doi.org/10.1007/s00198-023-06817-4
6. McCloskey EV, Odén A, Harvey NC et al (2016) A meta-analysis of trabecular bone score in fracture risk prediction and its relationship to FRAX. J Bone Miner Res 31(5):940–948. https://doi.org/10.1002/jbmr.2734
7. Goel H, Binkley N, Boggild M et al (2024) Clinical use of trabecular bone score: the 2023 ISCD official positions. J Clin Densitom 27(1):101452. https://doi.org/10.1016/j.jocd.2023.101452
8. Schacter GI, Leslie WD, Majumdar SR, Morin SN, Lix LM, Hans D (2017) Clinical performance of an updated trabecular bone score (TBS) algorithm in men and women: the Manitoba BMD cohort. Osteoporos Int 28(11):3199–3203. https://doi.org/10.1007/s00198-017-4166-1
9. Shevroja E, Marques-Vidal P, Aubry-Rozier B et al (2019) Cohort profile: the OsteoLaus study. Int J Epidemiol 48(4):1046–1047g. https://doi.org/10.1093/ije/dyy276
10. Shepherd JA, Lu Y, Wilson K et al (2006) Cross-calibration and minimum precision standards for dual-energy X-ray absorptiometry: the 2005 ISCD official positions. J Clin Densitom 9(1):31–36. https://doi.org/10.1016/j.jocd.2006.05.005
11. Leslie WD, Majumdar SR, Morin SN, Hans D, Lix LM (2017) Change in trabecular bone score (TBS) with antiresorptive therapy does not predict fracture in women: the manitoba BMD cohort. J Bone Miner Res 32(3):618–623. https://doi.org/10.1002/jbmr.3054
12. Genant HK, Wu CY, Van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res 8(9):1137–1148. https://doi.org/10.1002/jbmr.5650080915

Authors and Affiliations

Guillaume Gatineau[1,2] · Karen Hind[2,3,4] · Enisa Shevroja[1] · Elena Gonzalez-Rodriguez[1] · Olivier Lamy[1] · Didier Hans[1,2]

✉ Didier Hans
didier.hans@chuv.ch; didier.hans@ascendys.ch

Guillaume Gatineau
Guillaume.Gatineau@unil.ch

Karen Hind
khind@medimapsgroup.com

Enisa Shevroja
Enisa.Shevroja@chuv.ch

Elena Gonzalez-Rodriguez
Elena.Gonzalez-Rodriguez@chuv.ch

Olivier Lamy
Olivier.Lamy@chuv.ch

[1] Interdisciplinary Center of Bone Diseases, Bone and Joint Department, Rheumatology Unit, Lausanne University Hospital and University of Lausanne, Avenue Pierre Decker, 1011 Lausanne, Switzerland

[2] Medimaps Group SA, Plan-Les-Ouates, Geneva, Switzerland

[3] Wolfson Research Institute of Health and Wellbeing, Durham University, Durham, UK

[4] Faculty of Health and Medicine, Lancaster University, Lancaster, UK