Accepted Manuscript

The initial slope of the variogram, foundation of the trabecular bone score, is not or poorly associated with vertebral strength

Ghislain Maquer[1], Yongtao Lu[2,3], Enrico Dall'Ara[4], Yan Chevalier[5], Matthias Krause[6], Lang Yang[4], Richard Eastell[4], Kurt Lippuner[7], Philippe K. Zysset[1]

1 Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland
2 Department of Mechanical Engineering and INSIGNEO Institute for in silico Medicine, University of Sheffield, Sheffield, UK
3 Institute of Biomechanics, TUHH Hamburg University of Technology, Hamburg, Germany
4 Department of Human Metabolism and INSIGNEO Institute for in silico Medicine, University of Sheffield, Sheffield, UK
5 Klinikum Großhadern, Orthopaedic Department, Laboratory for Biomechanics and Experimental Orthopaedics, Munich, Germany
6 Department of Osteology and Biomechanics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
7 Osteoporosis Clinic, Inselspital, University Hospital and University of Bern, Bern, Switzerland

Corresponding author: Philippe K. Zysset
tel.: +41 31 631 59 25
fax: +41 31 631 59 60
E-mail address: philippe.zysset@istb.unibe.ch

Abbreviated title: Vertebral strength predictions by textural analysis

Abstract word count: 267

Attachments: 3 figures and 1 table

source: https://doi.org/10.7892/boris.70967 | downloaded: 17.4.2024

Disclosure page

The authors have no conflict of interest to report.

Abstract

Trabecular bone score (TBS) rests on the textural analysis of DXA to reflect the decay in trabecular structure characterising osteoporosis. Yet, its discriminative power in fracture studies remains incomprehensible as prior biomechanical tests found no correlation with vertebral strength. To verify this result possibly due to an unrealistic set-up and to cover a wide range of loading scenarios, the data from three previous biomechanical studies using different experimental settings was used. They involved the compressive failure of 62 human lumbar vertebrae loaded 1) via intervertebral discs to mimic the *in vivo* situation ("full vertebra"), 2) via the classical endplate embedding ("vertebral body") or 3) via a ball joint to induce anterior wedge failure ("vertebral section"). HR-pQCT scans acquired prior testing were used to simulate anterior-posterior DXA from which areal bone mineral density (aBMD) and the initial slope of the variogram (ISV), the early definition of TBS, were evaluated. Finally, the relation of aBMD and ISV with failure load (F_{exp}) and apparent failure stress (σ_{exp}) was assessed and their relative contribution to a multi-linear model was quantified via ANOVA. We found that, unlike aBMD, ISV did not significantly correlate with F_{exp} and σ_{exp}, except for the "vertebral body" case (r^2=0.396, p=0.028). Aside from the "vertebra section" set-up where it explained only 6.4% of σ_{exp} (p=0.037), it brought no significant improvement to aBMD. These results indicate that ISV, a replica of TBS, is a poor surrogate for vertebral strength no matter the testing set-up, which supports the prior observations and raises *a fortiori* the question of the deterministic factors underlying the statistical relationship between TBS and vertebral fracture risk.

Keywords: Biomechanics, Fracture risk assessment, Osteoporosis, Trabecular bone score (TBS), Vertebral strength.

Introduction

Osteoporosis is a metabolic bone disease characterised by bone loss and impaired architecture[1]. Those alterations induce an increased likeliness of bone failure, even during daily activities[2]. Consequently, millions of fractures are attributable each year worldwide to osteoporosis[3]. Dual energy X-ray absorptiometry (DXA), the clinical standard to evaluate fracture risk, is based on the measure of the areal bone mineral density (aBMD), which explains bone strength only partially[4] and strongly overlaps between patients with and without fractures[5]. Indeed, the trabecular structure, known to affect bone fragility, is not accounted for by DXA measurements, but textural approaches such as trabecular bone score (TBS) have been proposed in order to overcome this limitation[6].

TBS evaluated on lumbar DXA improves fracture prediction[7-9] and the monitoring of glucocorticoid therapy[10] compared to aBMD alone. Its integration in FRAX, an epidemiological tool including population-based clinical risk factors (age, body mass index, prior fracture…), is now considered[11,12]. However, the reasons for such predictive ability are not clear as TBS exhibits contradictory relationships with trabecular micro-architectural parameters[13,14]. More intriguing is the absence of correlation between lumbar TBS and *in vitro* vertebral strength observed by Roux et al.[15].

This surprising result may be due to the chosen biomechanical testing set-up: the authors embedded the vertebral endplates in a stiff material before uniaxial compression. This configuration circumvents the impact of disc degeneration and forward flexion on vertebral strength, but prevents common osteoporotic fractures such as endplate and anterior wedge failures[16]. Our hypothesis is that a more physiological testing set-up could improve the relations between TBS and vertebral strength. Accordingly, the aim of this study was to

provide a more comprehensive understanding of the relationship between the initial slope of the variogram (ISV), the textural component of TBS, and vertebral strength. For this purpose, 62 human vertebrae were evaluated by ISV on simulated DXA projections and their strength measured from a set-up similar to Roux et al.'s[15] or along two more realistic loading scenarios mimicking, either an anterior wedge fracture or a loading via the intervertebral discs.

Materials and methods

Implementation of the initial slope of the variogram

Published details of the original calculation of TBS[6,13,17] were used to implement what is renamed here "initial slope of the variogram" (ISV) for copyright reasons. ISV is also based on an isotropic experimental variogram, extensively used in geostatistics to evaluate terrain roughness[18] or the heterogeneity of a landscape's vegetation[19] from an aerial picture. This variogram is built by computing the mean squared difference $\gamma(h)$ between the intensity of pixels located at an increasing lag distance (h). The larger γ, the more heterogeneity in the image. Parameters such as nugget (subpixel variance), range (lag distance for which the pixels are not anymore related) and sill (variance towards which the variogram tends) are image-specific and can be determined after fitting exponential or spherical models to the variogram data (kriging). ISV is the slope at the origin of this variogram on a log-log scale and characterises how quickly intensities vary between nearby pixels. In practice, ISV would be computed on the selected region of interest (ROI) of raw lumbar DXA scans[20], where the point resolution is different from the line spacing (respectively 0.901 mm and 1.008 mm for a QDR 4500A, Hologic, Bedford, MA). To deal with the anisotropic pixel size, ISV interpolates the intensity of the pixels. An isotropic variogram is then built by averaging the

intensity pairs according to their orientation. Finally, ISV is computed as the least square regression line of the first 5 points of the variogram[17]. ISV and TBS are based on the same textural principles, but may not be rigorously equivalent. Yet, the high similarities between the two measures are shown in the appendices.

Imaging and biomechanical testing of vertebral samples

In order to cover a wide range of loading scenarios, the data from three previous studies using different biomechanical settings was used[21,22,23]. However, their initial objectives, sample selection and inclusion criteria were different.

- The “full vertebra” specimens

In the first experiment (Fig. 1A) and after approval from the Ethics Committee of the Hamburg Chamber of Physicians, 13 T_{11}-T_{12}-L_1 spinal segments from female donors (80 ± 8 years) were scanned by a high-resolution peripheral quantitative computed tomography system with posterior elements (HR-pQCT, 82µm voxel size, XtremeCT, Scanco Medical AG, Zurich, Switzerland). The specimens were stripped from all soft tissues, but the intervertebral discs. The facet joints were also removed to focus the load on the vertebral body. T_{11} and L_1 were filled with polymethylmethacrylate (PMMA) and partially embedded in PMMA discs. The spinal segments were fixed to a servo-hydraulic testing machine (MTS Bionix 858.2, Eden Prairie, MN, US). After preconditioning, a quasi-static uniaxial compression (6 mm/min) was applied on each spinal segment with a 4° angle until anterior failure of the middle vertebral body (T_{12}). Details of the study can be found elsewhere[21].

- The “vertebral body” specimen

For the second configuration (Fig. 1B), 12 vertebral bodies (L_1 to L_5) from four male donors

(66 ± 15 years) were deprived from all soft tissues and posterior elements. Their endplates were embedded in PMMA discs and they were scanned by HR-pQCT (82µm voxel size, XtremeCT, Scanco Medical AG, Zurich, Switzerland). For the mechanical testing, the PMMA discs were rigidly coupled to a servo-hydraulic testing apparatus (5560 Series Table Model Systems, Instron, Norwood, MA, USA). After preconditioning, a quasi-static uniaxial compression (5 mm/min) was performed on each specimen until failure. This study received authorization from the Ethics Committee of the Medical University of Vienna. More details can be found in a previous publication[22].

- The "vertebral section" specimens

For the last experiment (Fig. 1C), 37 vertebral bodies (T_{12} to L_4) from male (7) and female (3) donors (64 ± 12 years) were isolated. Their endplates were trimmed in parallel surfaces perpendicular to the cranio-caudal axis (300 CP, Exakt Gmbh, Germany). The vertebral sections were scanned with HR-pQCT (82µm voxel size, XtremeCT, Scanco Medical AG, Zurich, Switzerland) for the densitometric and ISV analyses. After scanning each specimen was placed between the sand-blasted loading plates of a servo-hydraulic apparatus (Mini-Bionix, MTS system, U.S.A.). After preconditioning, a quasi-static compression (5 mm/min) was performed on each section until anterior wedge failure was induced thanks to the free rotation of the loading plate allowed by a ball joint and the shift of the load application 5-10% of the vertebral width anteriorly. These tests have been conducted after approval by the Ethics Committee of the Medical University of Vienna. Other details are also available elsewhere[23].

Mechanical, densitometric and textural variables

In biomechanics, strength is usually a structural parameter defined as the largest force a bone

sustains before failure (failure load F_{exp} [N]). In mechanics, strength is a material parameter corresponding to the apparent failure stress (σ_{exp} [MPa]) defined by a ratio of the failure load and mean cross-sectional area of the sample (CSA). For completeness, the predictions of both definitions of vertebral strength were evaluated for ISV, aBMD, but also for bone mineral content (BMC), total volumetric BMD including both cortical, and trabecular bone (vBMD). The HR-pQCT data of each specimen was rotated to match the coordinate systems of the image with the cranio-caudal and anterior-posterior axis respectively along Z and Y. This data was coarsened to match the resolution of a lumbar DXA scan (0.901*0.901*1.008 mm^3) and CSA (volume/height) as well as vBMD were evaluated. Anterior-posterior DXA scans were then simulated by projecting the BMD values along the frontal plane of each sample (Fig. 3). These simulated DXA, similar to those previously used to mimic clinical DXA[24,25] or to test the robustness of TBS[26], were used to measure aBMD, ISV and BMC. The image processing was performed with Medtool (www.dr-pahr.at).

Statistical analysis

A statistical analysis was conducted after logarithmic transformation (Python-Statsmodels). First, the relation of ISV, aBMD, BMC, vBMD with F_{exp} and σ_{exp} was assessed via root-mean-square-errors (rmse) and coefficients of determination (r^2). Then, aBMD and ISV were combined in multi-linear models after evaluation of their multi-collinearity via the variance inflation factor (VIF)[27]. VIF quantifies the inflation of the regression coefficient due to correlation between predictors and is used to avoid over-fitting of multi-linear models. Parameters were considered independent if VIF < 4[28]. An analysis of variance (ANOVA) was performed on the multi-linear models to evaluate the relative contribution of aBMD and ISV to the strength predictions. Their contribution was given as a percentage corresponding

to their relative sum of squares over the total sum of squares.

Results

Value of ISV as a surrogate for vertebral strength

The correlations between ISV and F_{exp}, respectively σ_{exp}, were non-significant, except in the "vertebral body" case (F_{exp}, r^2=0.396, p=0.028) for which the errors reached 24 to 50% of the actual failure load (rmse = 1.55 kN, $3.057 \leq F_{exp} \leq 6.289$ kN). In fact, ISV is the explanatory variable related to the largest errors in the strength predictions, around 30% more than aBMD. The aBMD performed reasonably well, even when the posterior elements were included in the simulated DXA ("full vertebra"). It correlated significantly against with F_{exp} ($0.587 < r^2 < 0.694$, $p < 0.05$) and σ_{exp} ($0.47 < r^2 < 0.570$, $p < 0.05$). In comparison, BMC provided significant, but lower correlations (F_{exp}: $0.533 < r^2 < 0.672$, $p < 0.05$; σ_{exp}: $0.228 < r^2 < 0.422$, $p < 0.05$) and the correlations observed with vBMD were generally higher than with the other predictors (Table 1).

Relative contribution of aBMD and ISV to the strength predictions

VIF evaluation confirmed that aBMD and ISV were independent enough to be included in multi-linear models (VIF < 1.55). Without surprise, aBMD contributed highly and significantly to the explained variance of the models explaining more than 60% and 50% of the variance in F_{exp} and σ_{exp}, for all testing set-ups. In most cases, the improvements due to ISV were not significant ($p > 0.05$), with the exception of the failure stress of the "full vertebra" ($p = 0.037$), where only 6.4% of the variance was explained by ISV (Fig. 3).

Discussion

Presented as an estimate of the trabecular structure, TBS should be related to vertebral strength. Interestingly, the only study on the matter shows otherwise[15]. To verify this unexpected result, the initial slope of the variogram (ISV), derived from early descriptions of TBS[6,13,17], was evaluated on simulated DXA for 62 human vertebrae compressed to failure via three distinct biomechanical tests. At best, ISV poorly predicts vertebral strength.

A major strength of this work is that it is not limited to the uniaxial compression of vertebral bodies via rigid endplate embedding. More physiological testing scenarios such as loading via a rotating plate, to induce anterior wedge[23], or via intervertebral discs, to allow endplate failure[29], have been investigated. All three configurations led to the same conclusion: ISV is poorly related to vertebral strength and induces large root-mean-square errors, even when a significant correlation is reached. Combining aBMD and ISV as a multi-linear model did not prove more successful. Independently of the loading condition, the contribution of ISV to the model predictions was poor or non-significant. Whether strength was considered a failure load or an apparent failure stress did not have much impact on those results either. Another strong point is that the geometry of the vertebra differs in the three experiments. Cortical bone induces high contrasts on DXA that influence ISV. Still, regardless of the presence of the posterior elements and/or endplates, ISV did not predict vertebral strength. Finally, the study benefits from a higher power than the previous work. Indeed, 62 human vertebral samples have been crushed in total and our largest cohort alone includes more than twice the number of specimens tested by Roux et al. in a prior study (only 16 samples).

Based on the same textural principles as ISV, TBS does not predict vertebral strength either[15]. Of course, the purpose of TBS is to assess fracture risk, not vertebral strength, but a link to the mechanical properties is clearly missing. The weak, but significant, correlation

between TBS and vertebral stiffness found by Roux et al.[15] was sometimes provided as a justification[11,30]. Yet, bone stiffness is the least accurate variable measured in a biomechanical test and is highly correlated with strength in well-controlled experiments[31]. It is unlikely for TBS to share a solid relationship with one and not with the other. In fact, this weak correlation with stiffness cannot be trusted: unlike strength that is solely due to bone failure, stiffness is a composite measure affected by each element of the structure being tested. In other words, if vertebral stiffness were to be calculated from Roux et al.'s experiments, the actual deformation of the embedding material on the endplates must be accounted for.

Furthermore, if TBS estimates trabecular morphology[32], which is sometimes not so clear[13,14], it should at least predict the mechanical properties of trabecular bone. This crucial point is yet to be demonstrated for TBS[33], but constitutes the cornerstone of other textural methods. The "sill variance" is also computed from the isotropic variogram and reflects the global heterogeneity of an image, when TBS focuses on the rate of variation between nearby pixels. The sill variance improves the fracture discrimination at the hip compared to aBMD alone[34], which is sound since it describes around 80% of the variance in elastic modulus and strength of trabecular samples[35].

To improve the fracture risk assessment, alternative exist. Several parameters excluded from FRAX, such as fall history, also contribute to fracture risk[36]. Then, fracture remains a mechanical event that occurs when the load applied on a bone is larger than its strength, for which finite element analyses (FEA) are the best surrogate[37]. Indeed, FEA generated from DXA images can predict hip fracture for diverse loading scenarios (fall, stance)[38,39], an essential aspect of fracture etiology[40] that is omitted by densitometric and textural measures. Finally, trabecular anisotropy is the second best determinant of trabecular stiffness after bone

volume fraction[33]. It can be retrieved *in vivo* from HR-pQCT scans at the radius or the tibia, and allows accurate non-invasive evaluation of its strength via FEA[41].

Two points deserve clarification. First, ISV is not exactly TBS, or rather the contrary, as TBS may have been modified since its original description from which ISV was defined[13]. Fortunately, the modifications must have been minimal since ISV, based on Pothuaud's early description of TBS[6,17], and the latest version of TBS at the time of the study (1.8.2) compare well on clinical DXA (Appendix 1). Just as TBS, ISV is independent from lumbar aBMD[42] (VIF << 4) and relatively insensitive to noise[26] (Appendix 2). Just as TBS[13,14], ISV seems to be more affected by regions of high contrast located near the contour of the vertebra rather than its trabecular core (Appendix 3). The second point concerns the use of simulated DXA, which is not strictly equivalent to the clinical modality. Still, clinical and simulated aBMD measures correlate highly (Appendix 4).

In conclusion, the initial slope of the variogram was not appreciably related to vertebral strength and brought no meaningful improvements to strength predictions provided by aBMD alone, independently of the loading set-up. Given the similitudes between ISV and TBS, this result concurs with prior observations showing the lack of relationship between TBS and vertebral strength[15]. Unlike other textural variables[34,35], there is also no evidence that TBS predicts the mechanics of trabecular bone either, peculiar for an index of trabecular structure. The rationale behind the statistical relationship between TBS and fracture risk remains *a fortiori* unknown.

Acknowledgments

The authors would like to acknowledge Claus-Christian Glüer (Christian-Albrechts Universität, Germany), Klaus Püschel (University Medical Center Hamburg-Eppendorf, Germany), Michael Morlock, Gerd Huber and Kay Sellenschloh (Hamburg University of Technology, Germany) for providing parts of the data used in this study. This work was partly supported by the German Federal Ministry of Education and Research (BMBF, BioAsset 01EC1005D).

Study design: GM, PZ; Study conduct: GM, PZ; Data collection: ED, YC, YL, MK, KL; Data analysis: GM, ED, YC, YL, LY, PZ; Data interpretation: GM, KL, RE, PZ; Drafting manuscript: GM; Revising manuscript: GM, ED, YC, YL, LY, MK, RE, KL, PZ; Approving manuscript: GM, ED, YC, YL, LY, MK, RE, KL, PZ; GM and PZ are responsible for the integrity of the data.

References

1 Consensus Development Conference. Osteoporosis prevention, diagnosis, and therapy. NIH Consens Statement. 2000;17:1-45.

2 Dempster DW. The contribution of trabecular architecture to cancellous bone quality. J. Bone Miner. Res. 2000;15:20-23.

3 Johnell O, Kanis J. Epidemiology of osteoporotic fractures. Osteoporosis Int. 2008;16 (Suppl. 2):S3-S7.

4 Lochmüller EM, Burklein D, Kuhn V, Glaser C, Muller R, Gluer CC, Eckstein F. Mechanical strength of the thoracolumbar spine in the elderly: prediction from in situ dual-energy X-ray absorptiometry, quantitative computed tomography (QCT), upper and lower limb peripheral QCT, and quantitative ultrasound. Bone. 2002;31(1):77-84.

5 Stone KL, Seeley DG, Lui LY, Cauley JA, Ensrud K, Browner WS, et al. BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. J. Bone Miner. Res. 2003;18(11):1947-1954.

6 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):775-787.

7 Krueger D, Fidler E, Libber J, Aubry-Rozier B, Hans D, Binkley N. Spine trabecular bone score subsequent to bone mineral density improves fracture discrimination in women. J. Clin. Densitom. 2014;17(1):60-65.

8 Silva BC, Leslie WD, Resch H, Lamy O, Lesnyak O, Binkley N, et al. Trabecular bone score: a noninvasive analytical method based upon the DXA image. J. Bone Miner. Res. 2014;29(3):518-530.

9 Touvier J, Winzenrieth R, Johansson H, Roux JP, Chaintreuil J, Toumi H, et al. Fracture Discrimination by Combined Bone Mineral Density (BMD) and Microarchitectural Texture Analysis. Calcif. Tissue Int. 2015:1-10.

10 Paggiosi MA, Peel NF, Eastell R. The impact of glucocorticoid therapy on trabecular bone score in older women. Osteoporos Int. 2015;26(6):1773-80.

11 Iki M, Fujita Y, Tamaki J, Kouda K, Yura A, Sato Y, et al. Trabecular bone score may improve FRAX® prediction accuracy for major osteoporotic fractures in elderly Japanese men: the Fujiwara-kyo Osteoporosis Risk in Men (FORMEN) Cohort Study. Osteoporosis Int. 2015:1-8.

12 McCloskey EV, Odén A, Harvey NC, Leslie WD, Hans D, Johansson H, Kanis JA. Adjusting Fracture Probability by Trabecular Bone Score. Calcif. Tissue Int. 2015:1-10.

13 Bousson V, Bergot C, Sutter B, Levitz B, Cortet B. Trabecular bone score (TBS): available knowledge, clinical relevance, and future prospect. Osteoporis Int. 2012;23(5):1489-1501.

14 Bousson V, Bergot C, Sutter B, Thomas T, Bendavid S, Benhamou C L, et al. Trabecular Bone Score: Where are we now? Joint Bone Spine. 2015. doi:10.1016/j.jbspin.2015.02.005

15 Roux JP, Wegrzyn J, Boutroy S, Bouxsein ML, Hans D, Chapurlat R. The predictive value of trabecular bone score (TBS) on whole lumbar vertebrae mechanics: an ex vivo study. Osteoporis Int. 2013;24(9):2455-2460.

16 Maquer G, Schwiedrzik J, Huber G, Morlock MM, Zysset PK. Compressive strength of elderly vertebrae is reduced by disc degeneration and additional flexion.

J. Mech. Behav. Biomed. Mater. 2015;42:54-66.
17 Pothuaud L. Method for determining a three-dimensional structure from a two-dimensional image, in particular a bone structure. Patent US 7609867 B2. Issued Oct 27 2009.
18 Smith MW. Roughness in the Earth Sciences. Earth-Sci. Rev. 2014;136:202-225.
19 Garrigues S, Allard D, Baret F, Weiss M. Quantifying spatial heterogeneity at the landscape scale using variogram models. Remote Sens. Environ. 2006;103(1):81-96.
20 Hans D, Barthe N, Boutroy S, Pothuaud L, Winzenrieth R, Krieg MA. 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. 2011;14(3):302-312.
21 Lu Y, Maquer G, Museyko O, Puschel K, Engelke K, Zysset PK, et al. Finite element analyses of human vertebral bodies embedded in polymethylmethalcrylate or loaded via the hyperelastic intervertebral disc models provide equivalent predictions of experimental strength. J. Biomech. 2014;47(10):2512-2516.
22 Chevalier Y, Charlebois M, Pahr D, Varga P, Heini P, Schneider E, Zysset PK. A patient-specific finite element methodology to predict damage accumulation in vertebral bodies under axial compression, sagittal flexion and combined loads. Comput. Methods Biomech. Biomed. Eng. 2008;11:477-487.
23 Dall'Ara E, Schmidt R, Pahr D, Varga P, Chevalier Y, Patsch J, et al. A nonlinear finite element model validation study based on a novel experimental technique for inducing anterior wedge-shape fractures in human vertebral bodies in vitro. J. Biomech. 2010;43(12):2374-2380.
24 Burghardt AJ, Kazakia GJ, Link TM, Majumdar S. Automated simulation of areal bone mineral density assessment in the distal radius from high-resolution peripheral quantitative computed tomography. Osteoporosis Int. 2009:20:2017-2024.
25 Dall'Ara E, Pahr D, Varga P, Kainberger F, Zysset PK. QCT-based finite element models predict human vertebral strength in vitro significantly better than simulated DEXA. Osteoporosis Int. 2012;23(2):563-572.
26 Winzenrieth R, Michelet F, Hans D. 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. 2013;16(3):287-296.
27 Lin D, Foster DP, Ungar LH. VIF regression: A fast regression algorithm for large data. J. Am. Stat. Assoc. 2011;106(493):232-247.
28 Peat J, Barton B. Continuous Data Analyses: Correlation and Regression, in Medical Statistics: A Guide to Data Analysis and Critical Appraisal, Blackwell Publishing Inc., Malden, Massachusetts, USA. 2005. doi: 10.1002/9780470755945.ch6
29 Maquer G, Schwiedrzik J, Zysset PK. Embedding of human vertebral bodies leads to higher ultimate load and altered damage localisation under axial compression. Comput. Methods Biomech. Biomed. Eng. 2014;17(12):1311-1322.
30 Iki M, Tamaki J, Sato Y, Winzenrieth R, Kagamimori S, Kagawa Y, Yoneshima H. Age-related normative values of trabecular bone score (TBS) for Japanese women: the Japanese Population-based Osteoporosis (JPOS) study. Osteoporosis Int. 2015;26(1), 245-252.

31 Fyhrie DP, Vashishth D. Bone stiffness predicts strength similarly for human vertebral cancellous bone in compression and for cortical bone in tension. Bone. 2000 Feb;26(2):169-73.
32 Muschitz C, Kocijan R, Haschka J, Pahr D, Kaider A, Pietschmann P, et al. TBS reflects trabecular microarchitecture in premenopausal women and men with idiopathic osteoporosis and low-traumatic fractures. Bone. 2015;79:259-266.
33 Maquer G, Musy SN, Wandel J, Gross T, Zysset PK. Bone volume fraction and fabric anisotropy are better determinants of trabecular bone stiffness than other morphological variables. J. Bone Miner. Res. 2015;30(6):1000-1008.
34 Dong XN, Pinninti R, Lowe T, Cussen P, Ballard JE, Di Paolo D, Shirvaikar M. Random field assessment of inhomogeneous bone mineral density from DXA scans can enhance the differentiation between postmenopausal women with and without hip fractures. J. Biomech. 2015;48(6):1043-1051.
35 Dong XN, Shirvaikar M, Wang X. Biomechanical properties and microarchitecture parameters of trabecular bone are correlated with stochastic measures of 2D projection images. Bone. 2013;56(2):327-336.
36 Lippuner K, Johansson H, Kanis JA, Rizzoli R. FRAX® assessment of osteoporotic fracture probability in Switzerland. Osteoporosis Int. 2010;21(3):381-389.
37 Zysset PK, Dall'Ara E, Varga P, Pahr D. Finite element analysis for prediction of bone strength. BoneKEy reports, 2013;2.
38 Naylor KE, McCloskey EV, Eastell R, Yang L. Use of DXA-based finite element analysis of the proximal femur in a longitudinal study of hip fracture. J. Bone Miner. Res., 2013;28(5):1014-1021.
39 Yang L, Palermo L, Black DM, Eastell R. Prediction of Incident Hip Fracture with the Estimated Femoral Strength by Finite Element Analysis of DXA Scans in the Study of Osteoporotic Fractures. J. Bone Miner. Res. 2014;29(12):2594-2600.
40 Dall'Ara E, Luisier B, Schmidt R, Pretterklieber M, Kainberger F, Zysset PK, Pahr DH. DXA predictions of human femoral mechanical properties depend on the load configuration. Med. Eng. Phys. 2013;35(11):1564-1572.
41 Varga P, Dall'Ara E, Pahr D, Pretterklieber M, Zysset PK. Validation of an HR-pQCT-based homogenized finite element approach using mechanical testing of ultra-distal radius sections. Biomech. Model. Mechanobiol. 2011;10(4):431-444.
42 Hans D, Goertzen AL, Krieg MA, Leslie WD. Bone microarchitecture assessed by TBS predicts osteoporotic fractures independent of bone density: the Manitoba study. J. Bone Miner. Res. 2011;26:2762-2769.
43 Khoo BCC, Brown K, Cann C, Zhu K, Henzell S, Low V, et al. Comparison of QCT-derived and DXA-derived areal bone mineral density and T-scores. Osteoporosis Int. 2009;20(9):1539-1545.
44 Lu Y, Krause M, Bishop N, Sellenschloh K, Glüer CC, Püschel K, et al. The role of patient-mode high-resolution peripheral quantitative computed tomography indices in the prediction of failure strength of the elderly women's thoracic vertebral body. Osteoporosis Int. 2015;26(1):237-244.

Appendices

1. ISV correlates highly with TBS on clinical data

TBS (TBSiNsight, v.1.8.2.0, MedImaps, Pessac, France) was analysed on lumbar DXA scans (QDR4500A, Hologic, Bedford, MA) of 16 female patients without deformities, calcifications or fractures (61 ± 15 yrs). Their body mass index (BMI: 23.4 ± 4.24 $kg.m^{-2}$) fitted the range suggested by MedImaps. ISV was evaluated on BMD maps generated from the raw data of the scans (R files) via a Matlab script developed at the University of Sheffield[38,39]. ISV correlated significantly with TBS ($r^2 = 0.745$, $p < 0.001$) despite being calculated on BMD maps. As the amount of abdominal tissues can affect TBS, its calculation is sometimes adjusted to BMI[8]. Indeed, combining the BMI of the patients to ISV further enhanced the correlation ($r^2 = 0.845$, $p < 0.001$).

2. Just as TBS, ISV is relatively insensitive to noise

Unlike their clinical analogues, simulated DXA are almost noiseless. A Gaussian white noise (SD = 0.9) was added to the simulated DXA of the full vertebrae to evaluate the effect of a poor signal-to-noise ratio on ISV. ISV values decreased due to noise (-7.34 ± 1.78 %), but correlated highly with ISV values from the clean images ($r^2 = 0.954$, $p < 0.001$).

3. The contour of the vertebra has more influence on ISV than the core

TBS and ISV reflect how quickly intensities vary from one pixel to another. Since the cortical regions feature the strongest gradients of intensity, they may dominate their calculation[13,14]. To probe in this direction, $ISV_{contour}$ was computed after erosion of the contour of each "full vertebra" by a single pixel. For ISV_{core}, the same amount of pixels, corresponding to 10.6 ± 0.9 % of the initial area, was removed from the trabecular core. ISV decreased less after deletion from the core (-26.50 ± 7.68 %) than after contour erosion (-34.28 ± 9.89 %) and correlated better with ISV_{core} ($r^2 = 0.779$, $p < 0.001$) than with $ISV_{contour}$ ($r^2 = 0.616$, $p < 0.001$). It appears that the contour of the vertebra has more influence on ISV than the trabecular core. Should TBS share the same sensitivity, inaccurate definition of the region of interest could affect its diagnosis.

4. aBMD values from simulated and clinical DXA correlate highly

Previous works established that DXA could be simulated from QCT images. Their degree of concordance with the clinical scans was evaluated based on the aBMD values. At the wrist[24] and hip[43], simulated and clinical aBMD measures were highly correlated ($r^2 > 0.82$). A similar procedure was used to verify our simulated images. For the "full vertebra" group, DXA scans were simulated for L_1 vertebrae from the HR-pQCT data of 11 $T_{11}/T_{12}/L_1$ spinal segments used in this study. Their aBMD was then compared to values obtained from lumbar DXA scans (Lunar Prodigy, Lunar Corporation, Madison, WI) performed on the intact cadavers before extraction of those segments[44]. For the "vertebral body" dataset, the 12 simulated images were directly compared to DXA scans (Hologic QDR 4500A Scanner, Hologic, Waltham, USA) performed on the freshly extracted lumbar spines[22]. In both cases, a strong and significant correlation was found between clinical and simulated aBMD measures ($r^2 = 0.711$, $p = 0.0011$ for the "full vertebra" group and $r^2 = 0.775$, $p < 0.0001$ for the "vertebral body" group). No clinical DXA had been performed for the "vertebral section" samples. Those correlations were lower than reported for other sites[24,43], which can be due to dissimilar DXA and HR-pQCT scanning positions in case of the "full vertebra" cohort and to the presence of posterior elements during the DXA scanning of the "vertebral body" specimens.

Figures

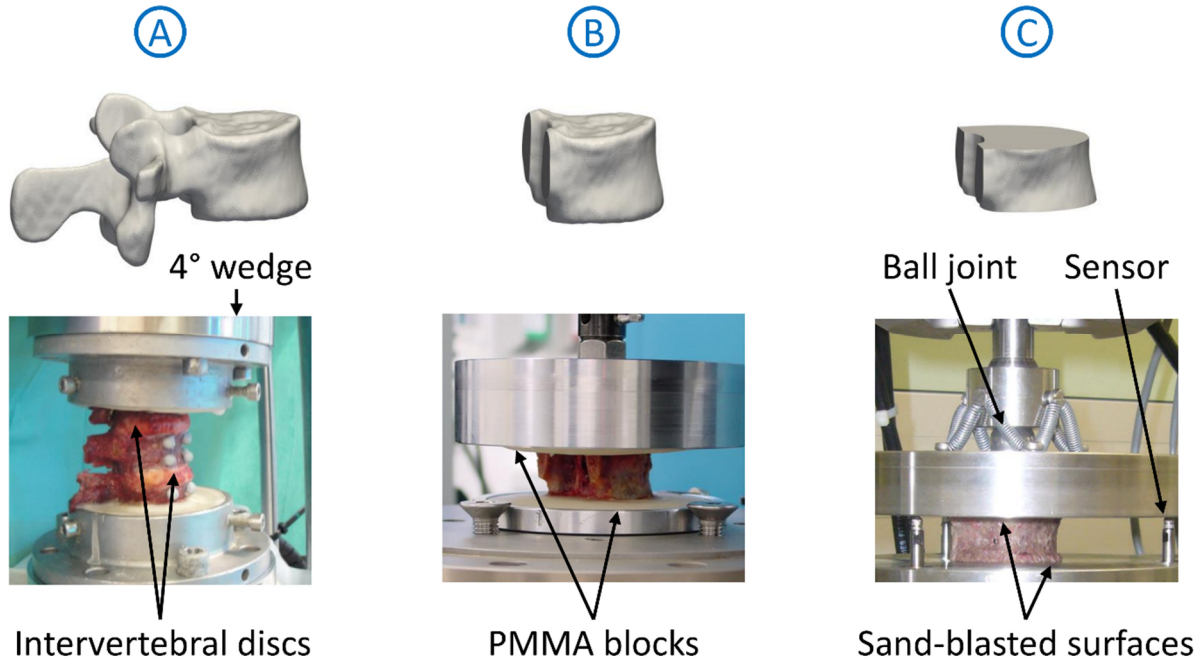

Fig.1. Three experimental set-ups for vertebra testing. In configuration A, the loads are transmitted to the vertebra via the intervertebral discs (“full vertebra”). In set-up B, the endplates of the vertebral body have been embedded in a stiff resin (“vertebral body”). Design C induces wedge failure in a vertebral section by allowing rotations of the loading plate (“vertebral section”).

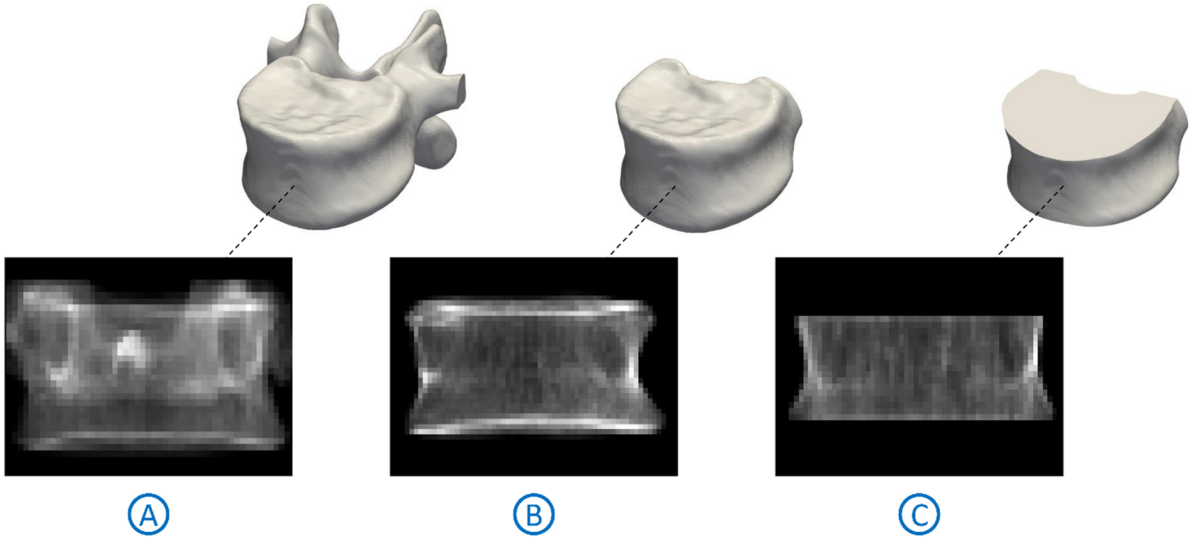

Fig.2. ISV was computed on DXA images simulated by projecting the HR-pQCT data coarsened to DXA resolution on the frontal plane of the “full vertebra” (A), “vertebral body” (B) and “vertebral section” (C).

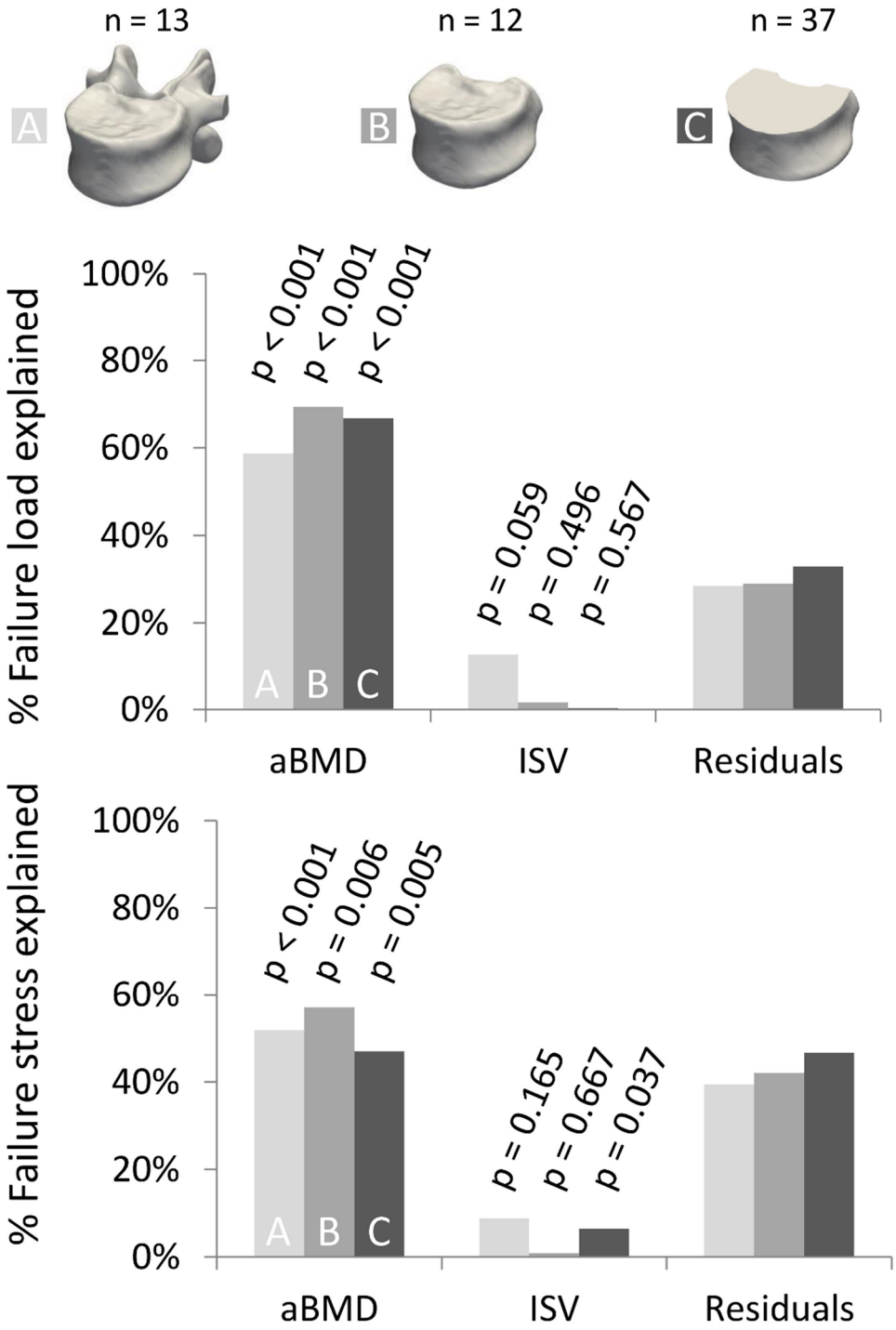

Fig.3. Relative contribution of aBMD and ISV to the variance explained by the multi-linear models for the “full vertebra” (A), “vertebral body” (B) and “vertebral section” (C). The contribution of ISV is minimal and mostly not significant.

Table

Table 1. Coefficients of determination (r^2) after log transformation and root-mean-square error (rmse) between explanatory variables and vertebral strength. Two significance levels * ($p < 0.05$) and ** ($p < 0.001$) are displayed. Rmse is provided in the unit of the dependent variable: kN for the failure load (F_{exp}), MPa for the apparent failure stress (σ_{exp}).

		A. Full vertebra		**B. Vertebral body**		**C. Vertebral section**	
		F_{exp}	σ_{exp}	F_{exp}	σ_{exp}	F_{exp}	σ_{exp}
ISV	r^2	0.186	0.134	0.396*	0.309	0.014	0.011
	rmse	0.407	0.373	1.550	1.318	1.638	1.227
aBMD	r^2	0.587*	0.518*	0.694**	0.570*	0.667**	0.470**
	rmse	0.323	0.308	0.937	0.968	0.979	0.888
BMC	r^2	0.533*	0.369*	0.672*	0.422*	0.643**	0.228*
	rmse	0.337	0.346	1.097	1.169	0.891	1.070
vBMD	r^2	0.714**	0.703**	0.707**	0.718**	0.487**	0.707**
	rmse	0.275	0.250	0.953	0.773	1.196	0.624

ISV: initial slope of the variogram, aBMD: areal bone mineral density, BMC: bone mineral content, vBMD: volume bone mineral density