JBMR Plus, 2025, **9**, ziaf075
https://doi.org/10.1093/jbmrpl/ziaf075
Advance access publication: April 28, 2025
Review

Critical evaluation of 3D-DXA and 3D-Shaper: methodological limitations and their implications

Tristan Whitmarsh*

Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, United Kingdom

*Corresponding author: Tristan Whitmarsh, Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, United Kingdom (tw401@cam.ac.uk).

Abstract

3D-DXA, as implemented in the software tool 3D-Shaper, is a software method that generates a 3D reconstruction of the proximal femur from a single 2D DXA image by registering a statistical model. Implementations of 3D-DXA aim to provide estimates of trabecular, cortical, and structural parameters, similar to those derived from quantitative computed tomography (QCT). As the inventor and developer of the software methods upon which 3D-DXA is built, I have been observing its adoption and widespread use with increasing concern. This article provides a critical evaluation of the methodological limitations inherent to 3D-DXA and discusses their implications for research and patient care. The primary issue is that the limited visibility of the cortex in a DXA image prevents 3D-DXA from accurately deriving cortical parameters. Instead, the software relies on predictions based on overall BMD rather than direct cortical measurements. This may lead to results that do not reflect actual cortical measurements. Additional concerns include the population bias due to the statistical model being derived from a specific demographic, and limited reconstruction accuracy by using single-view DXA images. These limitations have likely resulted in incorrect measurements and research outcomes, which have largely gone unrecognized due to the use of inappropriate performance assessment metrics and the absence of multiple comparison corrections in studies involving 3D-DXA. Despite these limitations, 3D-DXA has received regulatory approval in various countries, potentially compromising the accuracy of clinical diagnoses and treatment decisions. By highlighting these issues, this article aims to inform clinicians, researchers, and regulatory bodies about the significant limitations of 3D-DXA. It underscores the urgent need for a reevaluation of its use in research and clinical settings to prevent misinterpretation of results and to ensure patient safety.

Keywords: 3D-DXA, 3D-Shaper, quantitative computed tomography, bone mineral density, dual-energy X-ray absorptiometry, cortical parameter mapping, hip structure analsyis

Lay summary

This review critically examines the 3D-DXA software, also known as 3D-Shaper, which aims to estimate trabecular, cortical, and structural parameters of the proximal femur from a single DXA image. 3D-DXA has several inherent limitations that may lead to inaccurate measurements, potentially affecting research outcomes and patient treatment decisions. Understanding these issues is crucial for researchers and clinicians in order to avoid misinterpretations that could impact our knowledge of bone physiology, drug treatment effects, and patient safety.

Introduction

3D-DXA is a software method that registers a 3D statistical deformable model onto a single 2D DXA image to generate a 3D model of the proximal femur.[1] The model is constructed from a set of QCT scans of a Spanish Caucasian population of 81 women and 30 men with a mean age of 56.2 ± 12.1 yr, ranging between 30 and 84 yr.[2] The parameters of the statistical model subsequently describe the main variation in shape and density distribution of this population. In an iterative process, the parameters of the model, as well as the location, orientation, and size, are searched in an attempt to have the projection of the model resemble the DXA image (Figure 1). From the resulting 3D model, cortical parameters are then measured across the bone surface, along with trabecular volumetric BMD (vBMD) values within. A mean absolute error of 0.33 mm for cortical thickness and 72 mg/cm^3 for cortical density was reported when comparing 3D-DXA with QCT, with correlation coefficients greater or equal to 0.86.[2]

Although other methods have been proposed for the 3D reconstruction of bone structures from DXA images using a statistical model,[4,5] it is the software method first published in 2010[6] that was later commercialized as 3D-DXA by Galgo Medical SL (Barcelona, Spain), a spin-off company of the Pompeu Fabra University (Barcelona, Spain). This software was later extended to measure the cortical thickness and cortical BMD from the volumetric reconstructions,[2] which is now commercialized as 3D-Shaper by 3D-Shaper Medical SL, a further spin-off company of Galgo Medical SL, providing both a service and a software. The software is licensed to DMS Imaging to be sold as 3D-DXA with their Stratos/Medix DXA devices. Fujifilm is selling the same DXA devices by DMS under the Fujifilm brand name FDX Visionary DXA. Also Imex Medical sells these devices as the Elipse series and Radiología SA sells them as Radioscore - DR. These all include the option of adding 3D-DXA.

Received: October 7, 2024. **Revised:** March 24, 2025. **Accepted:** April 14, 2025

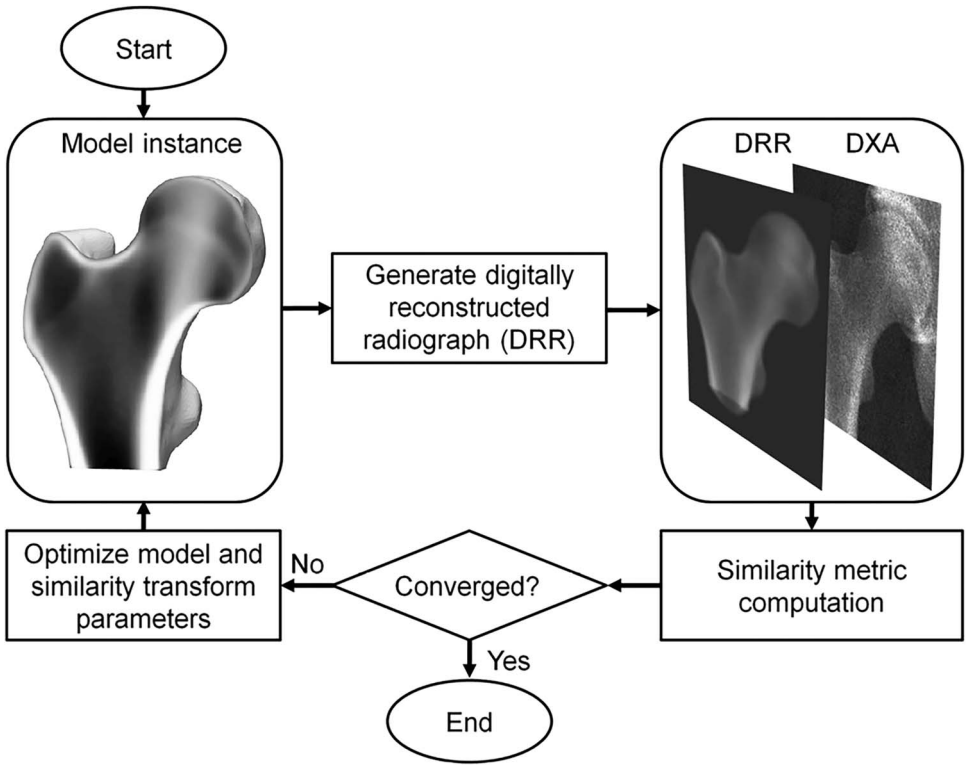

Figure 1. Flowchart illustrating the iterative process of 3D-DXA used to reconstruct the 3D proximal femur shape and density distribution from a single 2D DXA image. At each iteration, a projection is generated from the deformed density model instance, known as a digitally reconstructed radiograph (DRR). The DRR is then compared with the real DXA image by calculating a metric of similarity like the mean squared error. The model parameters and similarity transform (rotation, translation, and scale) are then adjusted in a way that attempts to increase the similarity in the next iteration. This process continues until the process has converged whereby the similarity between the DRR and the DXA image does not change beyond a predetermined threshold. Adapted from Whitmarsh et al.[3]

Having developed the original software method,[1] I possess a deep understanding of its inherent limitations that may not be apparent to other users or researchers. Although I have explained the main limitation with this software in a letter to the editor of Bone,[7] 3D-DXA continues to be sold and used throughout the world. 3D-Shaper has received approval from regulatory authorities in the EU, Japan, Thailand, Argentina, and a 510(k) clearance from the FDA, authorizing its use for diagnosing and making treatment decisions in patients. Given these endorsements, it is critical and urgent to address and clarify some of the misconceptions about this software, thereby enhancing understanding of its limitations.

Methodological limitations

In this section, I will explain some of the limitations of the methodology that 3D-DXA is based on, as well as specific implementation details that may contribute to further inaccuracies.

Cortical parameters are not measured

The main issue lies with a lack of visibility of the cortex in DXA images for 3D-DXA to derive its cortical parameters from. To illustrate this, I would like to revisit a figure from the letter to the editor of Bone (Figure 2). The DXA image clearly shows that at most locations on the bone surface, there is no discernible cortex to derive the cortical parameters from. This applies to the contour of the bone projection (A) and is especially true where 2 opposing cortices are perpendicular to the X-ray detector (B). This is evident in this image from the GE iDXA scanner with a pixel size of 0.3 × 0.25mm, but becomes even more pronounced in images using older GE Prodigy DXA devices, which produce considerably lower resolution images with a pixel size of 0.6 × 1.05mm, but are also supported by the 3D-Shaper software.

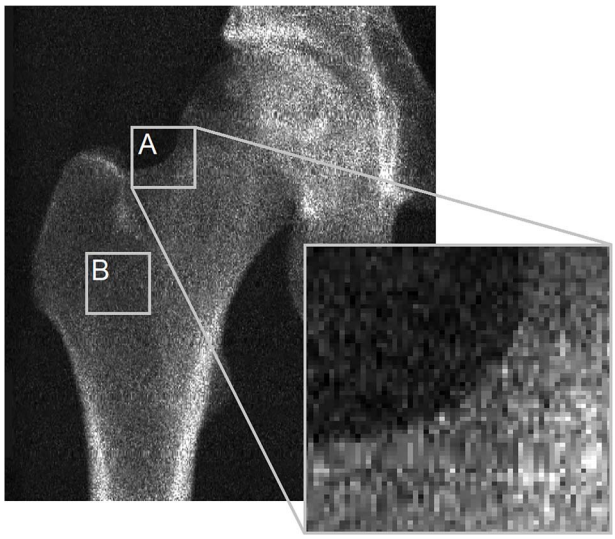

Figure 2. Example DXA hip scan of a young adult acquired using the GE iDXA scanner. This figure was previously published in my earlier work.[7]

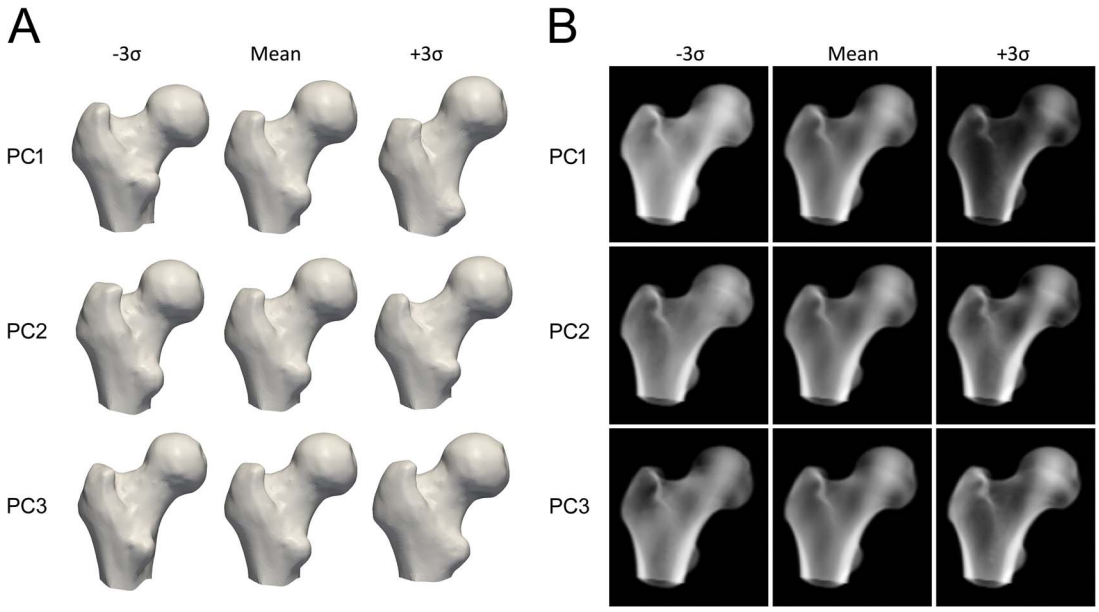

Figure 3. The mean and the first 3 principal components (PCs) of the shape model (A) and projections of the density model (B), varying between −3 and +3 standard deviations (σ). This figure was adapted from Whitmarsh et al.[3]

Instead of measuring the cortical parameters directly from the DXA image, 3D-DXA registers a statistical model of the shape and density distribution onto the DXA image,[3] and extracts the cortical parameters from this model.[2] This model is constructed from a set of calibrated QCT scans, in which the Hounsfield units of the voxels in a standard CT scan are converted to BMD values using a calibration phantom.

To construct the model, the proximal femur in each QCT scan is first segmented to generate a surface mesh. A mathematical technique, such as principal component analysis, is then applied to the aligned surface points, resulting in an average shape and a set of principal components (PCs) that describe the variations in shape, ordered by their importance.

To construct the statistical model of the density distribution, the QCT volumes are deformed to the mean shape, and the same statistical method is applied to the BMD values in the volumes, generating a mean volume and a set of PCs that describe the variation in BMD across the volume.

Each PC represents the way in which the bone's shape or density can vary (Figure 3). The model parameters are numerical factors that determine how much of each component is added to the mean model. Adjusting these parameters modifies the shape or density distribution, producing new model instances. The parameters are constrained to typically 2.5 to 3 standard deviations of its variation to make sure that the bone model always looks realistic according to the set of QCT scans it is built from. To create a new instance of the model, the density volume is deformed to match the new shape, using a Thin Plate Splines transformation calculated with a subset of the shape points.

3D-DXA subsequently generates a 3D reconstruction by iteratively changing the model parameters and generating corresponding model instances, as well as rotating and translating the model, until the projection of the model matches with the DXA image according to a predetermined similarity metric threshold (Figure 1).

The cortical thickness and cortical density are then measured from the instance of the model using a method equivalent to a deconvolution approach proposed for QCT.[8] Therefore, the cortical parameters are not measured from the DXA image, but from a parametric model that was registered onto this image.

Thus, when maps of the cortical parameters are presented by 3D-DXA, with values at locations where no cortex could possibly have been measured,[9–13] these maps were merely derived from the model instance. Although 3D-Shaper Medical clarifies in a response to the letter to the editor of Bone[14] that these parameters are estimated rather than measured, a more accurate description would be that 3D-DXA predicts these values using a complex and potentially error-prone statistical method. Given the limited visibility of the cortex in DXA images, the software primarily relies on overall bone density and is constrained by the statistical characteristics of the Spanish population on which the model is based. The broader implications of these limitations are explored further in the review.

Number of model parameters used

There is no single parameter that increases the cortical thickness. Instead, the representation of the cortex depends on a linear combination of the average 3D volume and a subset of the PCs of the density model. In Figure 4 we can see an example of how the mean and first 3 PCs can generate a new model instance. In this example, the model parameters, which represent the contributions of the PCs, have values of 2, 3, and −1.5, respectively. While PC1 increases or decreases the density everywhere, the other PCs change the density distribution in nondescriptive ways. This figure also illustrates how the cortical parameters are subsequently estimated from this new volume by fitting a smoothed stair-step model. Selective and independent adjustments to cortical thickness, cortical density, or trabecular density, if achievable at all, would require a complex combination of multiple PCs and consequently, many model parameters.

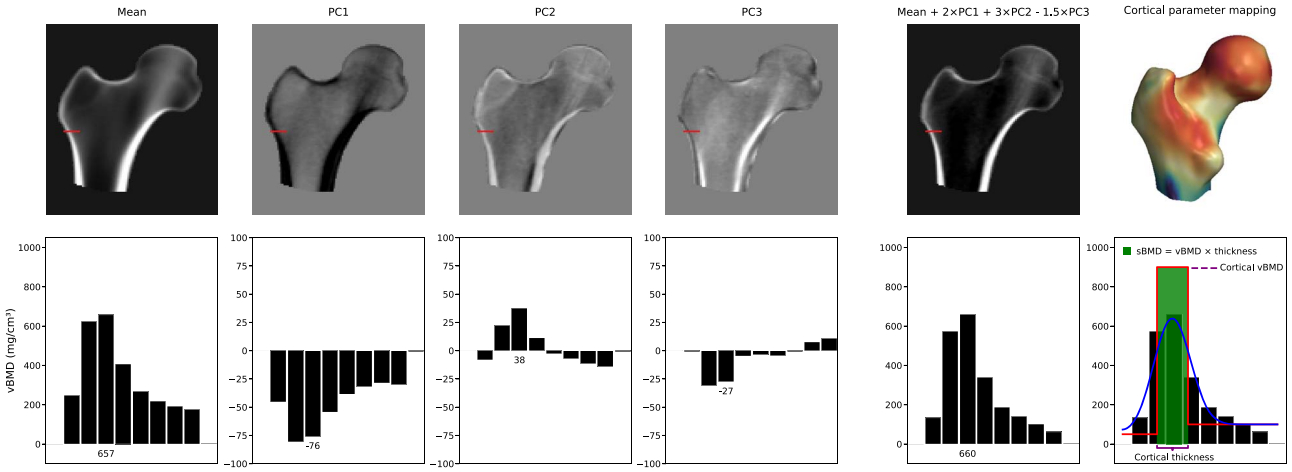

Figure 4. Example of generating an instance of a density model and measuring the cortical parameters from this volume. The top panels show cross-sections of the density model with the mean, PC1, PC2, PC3, and the new model instance generated by a linear combination of these volumes. Below them we see the values of the voxels taken from a line in the volume indicated by the red line in the cross-section. In the far left we can see how a stair step model (red) is fitted to the voxel values by smoothing the model (blue). This results in cortical thickness and cortical vBMD values that can be combined into a surface BMD (sBMD) value (green). When this is done at every point on the bone surface a color coded cortical surface map can be generated as shown above. Abbreviation: PC, principal component.

Typically, a subset of PCs is retained that describe the main modes of variation. This reduces the variation of the BMD distribution and surface points to a reduced set of parameters. How many model parameters are needed is typically determined by calculating the number of shape and density model parameters that describe 95% of the variations within its population, or by determining the "elbow" in the cumulative variance plot. However, more accurate is to use Horn's parallel analysis,[6] which mathematically assesses which PCs should be considered noise and can therefore be discarded.

Although the number of parameters that are used is a critical element in any statistical modeling approach, to my knowledge, this number, or the method to derive this number, has never been disclosed by 3D-Shaper Medical. If the number of parameters is not appropriately large, as determined by Horn's parallel analysis, it may not be able to accurately represent the full range of variation in the femur morphology of the input population. Furthermore, if this number of parameters is not the same in research and clinical applications, there will be a disconnect in the reported and clinical accuracy of 3D-DXA.

Potential population bias

Any 3D bone model generated by 3D-DXA is derived from statistics on the variations of the input population. In the case of 3D-Shaper, the statistical model is built from a Spanish population of 81 women and 30 men with a mean age of 56.2 ± 12.1 yr [30-84 yr][2] without treatments or diseases that impact bone metabolism.[15] It is clear that femur morphology differs between men and women, but demographics also play an important role. For example, the femur bones of Caucasian populations are significantly different from those of Asian populations, as documented in previous studies.[16]

A model built from a Caucasian population will not be able to generate a reconstruction that fits every subject in an Asian population. This is because the parameters of a statistical model are constrained to 3 standard deviation around the average femur model of a Spanish population. For instance, 1 study reports that White women have a femoral neck cortical thickness of 1.84±0.03 mm, while for Korean women it is 2.41±0.71 mm, as measured by QCT.[17] Consequently, a Korean woman with an average cortical thickness of 2.41 mm would fall outside the range of variation for a White population when constrained to 3 standard deviations (1.75-1.93 mm). This example illustrates the importance of having the model represent the target population.

Although the measures by 3D-DXA correlate with QCT in a Japanese population,[18] this is simply because all parameters correlate with areal BMD. It does not mean that patient-specific reconstructions were obtained. In particular, the estimated cortical parameters may deviate substantially from the true values, especially regarding their spatial distribution. Even if the predicted femoral shape is faulty, the density values, including cortical thickness and cortical BMD, will still mirror the aBMD, as long as the projections of the statistical model largely overlap with the bone in the DXA images. A person with low aBMD will have a low vBMD, a low cortical BMD, and thinner cortex compared with a patient with high aBMD in any population. Furthermore, Sone et al.[18] reported only correlations and there may be significant bias when applying a model developed from a Spanish population to a Japanese population, which could have clinical consequences. Despite this fundamental limitation, 3D-DXA is currently being marketed in Asia, including countries such as Japan, Thailand, and India.

A statistical model needs to be trained on the population to which it is applied. This could mean building separate models for each gender, ethnicity, and potentially treatment type, or by including all these subgroups in 1 model, provided that enough parameters are retained to capture the full range of variability. This principle is well established for machine learning and artificial intelligence tools,[19–22] but is even more critical for statistical modeling approaches, where models are actively constrained by the input population. Nonetheless, 3D-DXA has been used in studies with populations that obviously differ from the model population, including professional dancers,[9] football players and swimmers,[23] young women with obesity,[24] Black women,[25] patients with high bone mass,[11] with adult growth hormone deficiency[26] with psoriatic disease,[27] with type 2 diabetes,[28] with primary

hyperparathyroidism,[29] with acromegaly,[30] with down syndrome,[13] after sleeve gastrectomy,[31] men with spinal cord injury,[32] Australian middle-aged and older men with low bone mass,[33] and perhaps most concerning of all, pediatric cancer survivors.[34]

Limited reconstruction accuracy

3D-DXA is a highly complex method prone to errors at multiple stages. These include inaccuracies from QCT calibration, deformable registration for building statistical models, deformation of the density model to the shape instance, discrepancies in BMD and resolution between QCT and DXA, and simplification of the model projection by an isometric projection instead of the fan beam projection. Each of these factors can introduce biases in unpredictable ways, potentially skewing the results. In addition, specific implementation details, such as building the statistical model from a very small set of 111 adult Spanish subjects (which limits the variance in shape, density, and cortical thickness), further restrict the applicability of this software tool.

Furthermore, it is not clear whether 3D-Shaper can accurately read the proprietary data files from GE and Hologic devices, particularly with respect to correctly applying the calibration and correction factors for varying body compositions. These potential inaccuracies may influence 3D-Shaper results when applied to different populations or as population characteristics drift over time.

It stands to reason that, if the proximal femur shape and orientation is not perfectly matched with the DXA image, within a submillimeter accuracy, it should not be possible to extract the cortical thickness within a submillimeter accuracy, nor the cortical density. Unfortunately, it is not possible to generate a perfect reconstruction from just a single 2D DXA image as an early study with 3D-DXA already showed.[35] Here, it was shown that adding a second view reduces the shape error from 1.3 to 0.9 mm, and the BMD error from 4.4% to 3.2%, indicating a clear sub-optimal reconstruction from a single view. In contrast, computed tomography reconstructs a volume using hundreds of projections acquired from all angles around the femur. QCT, therefore, does allow for the independent measurement of cortical and trabecular parameters.

3D-DXA can produce femur models with a completely different morphology depending on where the model parameter search converged, in particular if the correct rotation was not recovered. This can vary greatly, as evidenced by a study using 3D-DXA involving same-day repeated DXA scans, which revealed differences in predicted strength of up to 62%.[36] This study also reported element-by-element BMD differences reaching 30% ± 50%, which likely also reflects a low repeatability accuracy for the cortical parameters, although these numbers were not provided. In a separate study, the correlation between 3D-Shaper and QCT density values was also reported to be low ($r^2 = 0.48$),[37] further indicating a limited subject-specific reconstruction accuracy.

In a study on measuring the structural parameters of the femur using 3D-DXA, a correlation coefficient of $r = 0.86$ was reported for the femoral neck axis length[38] when correlating the measurements derived from the 3D-DXA volume with the ground truth measurements from CT. This is worse than when measured directly in the DXA image ($r = 0.90$,[39]). Also, the femoral neck shaft angles produced by 3D-DXA are of limited accuracy ($r = 0.71$,[38]). An obvious reason for the lack of this accuracy is because the femoral head is excluded from the reconstruction process using a mask to prevent the overlap of the hemipelvis from interfering with the reconstruction, although this could be resolved by adding a second model of the hemipelvis.[5]

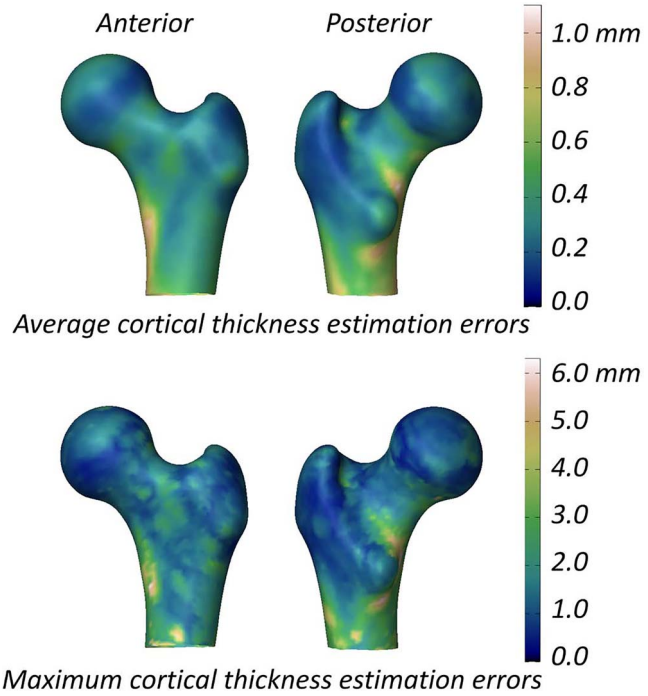

Figure 5. The color coded map of the mean (top) and maximum (bottom) cortical thickness estimation errors of 3D-DXA with respect to the same subject QCT scans. Adapted from Figure 8 in Humbert et al.[2]

The study evaluating the ability of 3D-DXA to measure the cortical parameters reports a mean (± standard deviation) cortical thickness difference between 3D-DXA and QCT of 0.04 ± 0.21 mm at the femoral neck and −0.07 ± 0.15 mm at the trochanter.[2] This study also presented a figure showing the mean and maximum absolute cortical thickness estimation errors across the femur model surface (Figure 5), with a mean error greater than 0.2 mm and maximum error greater than 1 mm across much of the femur surface. In comparison, a QCT study measuring cortical thickness changes following an 18-mo alendronate treatment reported a 1.4% increase, corresponding to an approximate 0.018 mm increase in cortical thickness.[40] Data compiled from 3 clinical trials on teriparatide (TPTD) indicated that cortical thickness increased by 0.035 mm[41] following an 18-mo treatment. Given that the inherent error in 3D-DXA exceeds the magnitude of typical treatment-induced changes, its measurements lack the sensitivity required to reliably distinguish genuine therapy-driven effects from methodological noise. In other words, if an individual's cortical bone appears to change, it remains unclear whether this arises from genuine physiological change, or from the inherent 3D-DXA measurement error.

Limited benefit over areal BMD

In the statistical density model that 3D-DXA uses, the first model parameter accounts for the vast majority of the variation, and increasing its value leads to an increase in both the overall density and all cortical parameters.[7] Given that there

is little information about the cortex in the DXA image, 3D-DXA primarily relies on the overall density and thus on the first model parameter. As a consequence, the cortical and trabecular parameters are inherently correlated. Although there may be some contribution from the inferior medial and lateral cortices of the shaft that are visible in the DXA image, the parameters returned by 3D-DXA predominantly reflect the total areal BMD (aBMD). This means that if aBMD increases, the 3D-DXA analysis will show simultaneous increases in trabecular BMD, cortical BMD, and cortical thickness. Indeed, a recent study presented at the 2024 European Calcified Tissue Society Congress[42] and the 2024 annual meeting of the ASBMR[43] confirms that 3D-Shaper parameters are highly correlated to aBMD and therefore do not provide additional fracture prediction information.

Furthermore, while the 3D-DXA reconstruction captures the overall femoral silhouette visible in the DXA image, its ability to recover the neck axis length and neck shaft angle is limited. As a consequence, the reconstructed shape mainly reflects the general bone size, which is already reflected by aBMD and does not add much to fracture strength prediction. Illustrative of this is the finding that strength predictions from 3D-DXA correlated with QCT ($r^2 = 0.88$), but were not statistically better than when predicted only by femoral neck aBMD from the DXA image ($r^2 = 0.87$).[37] A study by 3D-Shaper Medical shows a similar correlation ($r^2 = 0.86$) but a comparison with aBMD was excluded.[44] In a different study on a Japanese population, measures from 3D-Shaper were not statistically better at predicting hip fracture than total hip aBMD.[45]

There have been other articles published on a similar method of registering a 3D statistical model onto a 2D DXA image and predicting femoral strength,[46,47] with some showing an improvement over aBMD in fracture risk prediction.[48,49] Although it remains subject to many of the same methodological limitations, this method may offer some advantages over 3D-DXA, such as the use of an additional pelvic model to help recover the full proximal femur shape.

If 3D-DXA fails to outperform aBMD alone, it certainly cannot surpass a combination of aBMD and HSA parameters, which are directly measured from DXA images. This highlights the fundamental limitations of 3D-DXA and calls into question its justification for use in clinical or research settings. Nonetheless, 3D-Shaper Medical now offers a 3D-DXA based finite element analysis as a service.

Methods of validation

Despite the inherent limitations of 3D-DXA, numerous publications report positive findings on its accuracy. This discrepancy can be attributed to the use of inappropriate validation methods[21,22] that tend to obscure the true limitations of the software. In the following sections, I will discuss the issues with these validation methods and how they may contribute to a misleading perception of 3D-DXA's accuracy, reliability, and clinical utility.

Wrong performance assessment metrics

3D-Shaper Medical compared the surface mesh generated by 3D-DXA from a DXA image with the surface mesh manually generated from the same subjects' QCT scan, resulting in a mean point-to-surface distance of 0.93 mm.[2] They also compared the cortical parameters, finding a mean absolute error of 0.33 mm for cortical thickness and 72 mg/cm^3 for cortical density. However, there is no frame of reference to determine whether these reported errors are acceptable, which means these metrics do not necessarily validate the effectiveness of the software.

Another way in which 3D-DXA was evaluated was through the use of correlations. Correlation coefficients of 0.86, 0.93, 0.95, and 0.91 were reported for trabecular, cortical, integral vBMD, and cortical thickness, respectively. However, these strong correlations are primarily due to the fact that all cortical parameters in both 3D-DXA models and QCT scans correlate strongly with overall density. Patients with a higher than average aBMD in their DXA scans will, on average, also exhibit increased trabecular, cortical, and integral vBMD, along with a thicker cortex. These correlation results do not provide conclusive evidence of the accuracy of the software in generating patient-specific reconstructions and, in fact, may create a misleading impression of its performance.

Correlations are also used to evaluate population-based studies using 3D-DXA. Similar effects can be seen when a 3D-DXA analysis shows a significant correlation between its parameters and the factors tested or across different populations. This is again because these parameters are inherently correlated with overall BMD, and thus directly reflect changes or differences in aBMD. The findings may appear sensible and can closely mirror the true changes. For example, exercise increases both overall density and cortical bone mineral content,[50] leading to seemingly sensible results from a 3D-DXA analysis when aBMD increases due to exercise. However, these do not reflect the actual cortical parameters themselves, but merely the underlying changes or differences in the overall aBMD. Consequently, these studies can create a misleading impression of the effectiveness of 3D-DXA, giving a false sense of its ability to accurately measure cortical parameters.

A more appropriate evaluation would involve comparing the cortical parameters with a basic baseline model, such as a blind estimator. Here, a blind estimator refers to a simple, naive method of estimation that does not account for individual variability and instead applies an average value from the entire population to all subjects. For example, a blind estimator for cortical thickness would use the mean cortical thickness map from a group of individuals and apply it uniformly to every case, regardless of their specific characteristics. This approach serves as a baseline, or minimum standard, against which more advanced methods should be evaluated. Any sophisticated method should, at the very least, outperform this basic estimate. To my knowledge, such an evaluation has not yet been published.

Most publications on the 3D-DXA software showcase only successful reconstructions, which can create a false sense of confidence in the software's reliability. To quantitatively assess the reliability of 3D-DXA, an evaluation of the failure rate can be performed. In a previous study,[51] although not explicitly stated, only 80 out of 173 subjects were retained after excluding reconstructions identified as inaccurate through comparison of the model projections with the DXA images, yielding a failure rate exceeding 50%. To enable such an assessment, 3D-Shaper would need to provide a side-by-side display of the model projection and the DXA image. A more robust evaluation, however, would compare renderings and cross-sectional views of the resulting volumes with the ground-truth QCT scans and assess the failure rate in an independent evaluation.

In some instances, merely observing significant changes or differences is presented as evidence that 3D-DXA works while

disregarding whether the actual effects are correct. In response to my letter to the editor, significant differences between treatment groups were presented as evidence of 3D-DXA's efficacy. However, in this assessment, the authors provide 2 different results for cortical BMD changes after TPTD treatment, one showing a significant increase[52] and the other a nonsignificant decrease.[15] Neither of these reflects the actual change that one would expect in TPTD treatment, as I explain in Section 4.1.

Absence of multiple comparison correction

The concern in this section is not primarily related to the 3D-DXA software itself, but rather to the service provided by 3D-Shaper Medical and the subsequent publications. The critical issue lies in the lack of multiple comparison corrections when reporting changes or differences in the 3D-DXA parameters, particularly in the cortical parameter maps.

Applying multiple comparison corrections is crucial when analyzing data with multiple distinct measurements. For 3D-Shaper, the software generates 71 bone parameters, with each representing a separate statistical test. Without adjusting for these multiple comparisons, such as with a Bonferroni correction, the likelihood of obtaining significant results by chance increases, leading to false conclusions. Despite this, studies using 3D-Shaper often report significant changes or differences without applying any form of multiple comparison correction,[10,24,27,28] creating an unwarranted sense of confidence in the results. This can be misleading when certain parameters are measured but not reported, a form of p-hacking bias,[53] where significant results are highlighted while ignoring the increased risk of false positives due to the large number of tests conducted.

A similar issue occurs when studies using 3D-Shaper present color-coded maps showing changes or differences in cortical parameters on the bone surface or BMD values in the volume. In these cases, each vertex or voxel represents an individual statistical test, with significance often determined by a simple t-test at each point.[9–13,28,52,54,55] This approach highlights seemingly significant regions in the color-coded maps, creating the illusion of localized changes or differences where there are none. When presenting changes or differences in the volume through color-coded cross-sectional views, the statistical significance at the voxel level is generally not provided at all.[9,11,30,52,54–56] In the same way this suggests real changes when, in fact, they may simply be a result of random variation. Consequently, these color-coded maps can give an overly optimistic impression of the impact of treatments or interventions, falsely suggesting efficacy where there may be none.

The issue of false positives due to inadequate multiple comparison corrections has been a persistent challenge in the field of neuroimaging.[57] To address this, the neuroimaging community has developed solutions that can also be applied to studies examining cortical bone changes on the bone surface and density changes in the voxels of QCT scans. Poole et al.[58] describe how SurfStat (https://www.math.mcgill.ca/keith/surfstat/) can be used to test whether differences in cortical parameters at each point on the bone surface are statistically significant, with random field theory applied for multiple comparison correction. Similar packages exist for voxel-based analysis, such as the Statistical Parametric Mapping library (https://www.fil.ion.ucl.ac.uk/spm/) and the FMRIB Software Library (https://fsl.fmrib.ox.ac.uk). The use of these tools would improve the statistical rigor of studies using 3D-DXA, although the observed changes or differences would still predominantly reflect the correlation of all parameters with aBMD.

Implications

The use of 3D-DXA has wide-reaching implications in both research and clinical practice. The following sections will delve into these challenges.

Misleading results in drug efficacy studies

In most studies, the results of 3D-DXA align with what is expected in the cortex, as cortical parameters in healthy individuals are typically correlated with aBMD. This often leads to seemingly sensible results. However, this approach becomes flawed when the changes do not follow conventional patterns. The method is particularly problematic in drug trials where normal bone modeling and remodeling are altered, and the usual proportional increases or decreases in all parameters may not occur.

An illustrative example is a study on TPTD, where 3D-DXA indicated an increase in all cortical parameters, including a significant increase in cortical vBMD compared with placebo.[52] In contrast, using an equivalent cortical measurement technique applied to QCT, cortical BMD was shown to significantly decrease following the same 18-mo TPTD treatment, which was seen consistently in data from 3 different clinical trials.[41] This decrease is attributed to an increased rate of remodeling, which results in greater cortical porosity, as also observed using micro-CT.[59]

In an earlier study, 3D-Shaper Medical did report a decrease in cortical BMD in response to TPTD.[15] Although this was noted by the authors as a decrease, the change was not statistically significant. A more recent independent study also found no significant changes in cortical BMD in response to TPTD treatment using a 3D-DXA analysis.[60] It may be possible that there is some influence from the cortex of the shaft, which is partly visible in the DXA image. Due to the nature of 3D-DXA, an aBMD decrease in this region might have been projected across the entire femur model, including areas where no cortex is visible, since the model parameters influence the density distribution globally. However, this did not lead to the correct significant decrease in cortical BMD observed in QCT studies.

The study that reported an increase in all cortical parameters with TPTD also evaluated abaloparatide in parallel, showing an even greater increase in all parameters for the latter.[52] Follow-up studies using 3D-DXA analysis also reported significant increases in cortical BMD after abaloparatide treatment.[55,56] Given that abaloparatide shares a similar mechanism of action with TPTD, increasing the rate of bone remodeling, one would reasonably expect it to also lead to a decrease in cortical BMD. In fact, a study using QCT observed a decrease in cortical BMD following a similar period of abaloparatide treatment in 35 patients,[61] although the reduction was not statistically significant. This suggests that the 3D-DXA studies not only produced inaccurate changes but also likely resulted in an overly favorable assessment of abaloparatide's effects.

A study by Lewiecki et al.[54] further illustrates the limitations of 3D-DXA in measuring discordant cortical changes. Their 3D-DXA analysis reported increases in all cortical parameters in response to romosozumab. In contrast, a previous QCT analysis found no increase in cortical BMD

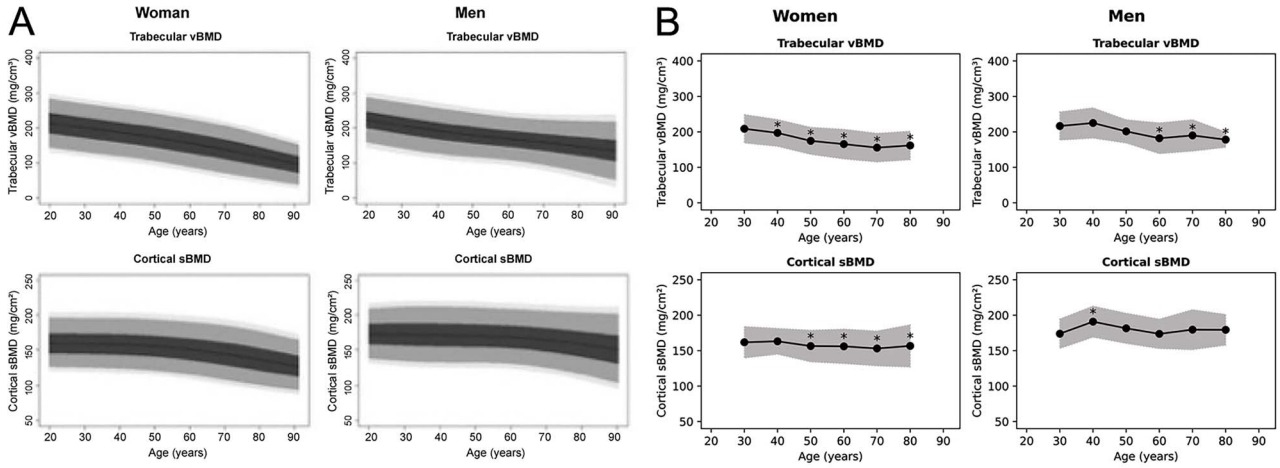

Figure 6. Reference data for a Spanish (A) and Argentinian (B) population, derived from[64] and[65] respectively. *indicates significant differences compared with decade 20–30.

following the same treatment.[62] The authors noted: "It is unclear why data obtained by QCT and DXA-based 3D-SHAPER for romosozumab treatment on hip integral and trabecular vBMD were similar across the studies but differed for cortical vBMD." Given that the letter to the editor of Bone seen by some of the same authors had already explained and predicted this discrepancy,[7] it appears that there persists a misunderstanding regarding the capabilities and limitations of 3D-DXA. As a result, this misunderstanding has likely again led to an overly optimistic assessment of the drug's efficacy.

These studies further disseminate potentially incorrect results by presenting color maps of cortical and trabecular bone changes. 3D-DXA cannot measure localized or focal changes, partly because each model parameter affects the density distribution globally, and also because these changes are not visible on a DXA image. The failure to apply multiple comparison corrections in these studies further undermines the results, as the seemingly significant regions in the maps are, in fact, much less significant, leading to an overly optimistic assessment of various drugs.

Patient management

While the 3D-Shaper software can produce 71 measures, the clinically approved version produces only the global trabecular vBMD and cortical surface BMD (sBMD) as well as an associated T-score and Z-score. It is not clear from what population the T- and Z-score are derived, although in a webinar by 3D-Shaper Medical Caucasian reference data are noted.[63] Reference plots have previously been produced for a Spanish[64] and Argentinian[65] population, which are both reproduced in Figure 6. These plots, however, appear distinct from each other and from the figures produced by the 3D-Shaper software. 3D-Shaper Medical, DMS Imaging, and their distributors do not disclose this population, or how these data have been derived. However, it is confirmed by Toyo Medic Co., Ltd., the distributor of 3D-Shaper in Japan, that the 3D-Shaper software approved for the Japanese market does not use reference data from a Japanese population. Considering the large difference in cortical and trabecular parameters between Caucasian and Asian populations,[17] applying these T- and Z-scores in Japan does not seem prudent without further validation. Accordingly, the validity of these reference values should be independently confirmed for each population in which the software is applied.

While the 3D-Shaper or commercial 3D-DXA software does not provide a treatment decision automatically, promotional material suggests the use of a table to aid the clinician in deciding on a treatment following a 3D-DXA analysis. This table, which was reproduced in Table 1, provides, for various treatment types, symbols (−, =, +, ++, +++) to represent how the treatment affects the trabecular vBMD or cortical sBMD. The symbols are based on threshold which, where available, are provided in Table 2. Although the threshold values are rather arbitrary or not specified, the table may be largely correct from knowledge on QCT studies. However, these data are said to have been derived from the 3D-DXA studies with references added to the various treatment types. Since 3D-DXA has particular issues when the cortical parameters do not change in line with the overall density, such as with a TPTD treatment, I would like to highlight the provided effects from this drug. I have provided the symbols for each of the references provided by 3D-Shaper Medical, as well as a different study that was not mentioned, in Table 3. They are quite varied and different from the symbols 3D-Shaper Medical provided in the table. It is therefore not clear how the symbols were derived, which puts into question their validity when used in conjunction with a 3D-DXA analysis.

In a webinar[66] 3D-Shaper Medical suggested a potential clinical use where 3D-Shaper resulted in a low cortical sBMD and very low trabecular vBMD. Since TPTD, according to the table, improves the trabecular bone more than the cortical bone, TPTD was said to be the best treatment. However, based on this table, a clinician would naturally chose romosozumab, denosumab, or abaloparatide over teribaratide if neglecting vertebral aBMD.

3D-DXA will likely only be used alongside regular DXA T-score and Z-score assessments. Thus, the danger of missing a high-risk individual will be minimum. However, due to the error associated with 3D-DXA, the cortical sBMD and trabecular vBMD will always be slightly higher or lower. In patients on the border of treatment, this may influence the clinician to give anti-osteoporosis drugs where this would not be indicated by aBMD alone.

A study on trabecular bone score (TBS) shows that this additional assessment significantly influences treatment

Table 1. The table provided by 3D-Shaper Medical, which is said to help prescribe a treatment best suited to a patients' needs. Only the references for teriparatide are provided here.

	2D DXA	3D-DXA (3D-Shaper)	
	(Total hip) aBMD	Trabecular vBMD	Cortical sBMD
Alendronate	+	+	+
Denosumab	++	++	++
Zoledronic acid	+	+	++
Teriparatide[15,52,55,56*]	+	++	=
Abaloparatide	++	++	++
Romosozumab	+++	+++	+++

*the addition of[55] is likely an error since this study does not include teriparatide. Abbreviations: aBMD, areal BMD; sBMD, surface BMD; vBMD, volumetric BMD.

Table 2. Reported effects of teriparatide treatment from a 3D-DXA analysis in studies referenced by 3D-Shaper Medical, as well as one additional study not included in their references.

	2D DXA	3D-DXA (3D-Shaper)	
	(Total hip) aBMD	Trabecular vBMD	Cortical sBMD
Teriparatide[52,56] **(18 mo)**	++ (3.3%)	++/+++ (9%)	+ (1.8%)
Teriparatide[15] **(24 mo)**	= ($p > .05$)	++/+++ (5.9%)	= ($p > .05$)
Teriparatide[60] **(24 mo)**	= ($p > .05$)	++/+++ (> 14%)	= ($p > .05$)

Abbreviations: aBMD, areal BMD; sBMD, surface BMD; vBMD, volumetric BMD.

Table 3. Threshold values used to determine the symbols indicating how a particular treatment affects aBMD, trabecular vBMD, and cortical sBMD, where available.

Symbol	Description from[15] and[63]		
	(Total hip) aBMD	Trabecular vBMD	Cortical sBMD
+++	$\Delta >$ n/a	$\Delta >$ n/a	$\Delta >$ n/a
++	$2\% < \Delta \leq$ n/a	$4\% < \Delta \leq$ n/a	$2\% < \Delta \leq$ n/a
+	$0.5\% < \Delta \leq 2\%$	$0.5\% < \Delta \leq 4\%$	$0.5\% < \Delta \leq 2\%$
=	$-0.5\% \leq \Delta \leq 0.5\%$	$-0.5\% \leq \Delta \leq 0.5\%$	$-0.5\% \leq \Delta \leq 0.5\%$
−	$-2\% \leq \Delta < -0.5\%$	$-4\% \leq \Delta < -0.5\%$	$-2\% \leq \Delta < -0.5\%$
−−	$\Delta < -2\%$	$\Delta < -4\%$	$\Delta < -2\%$

Abbreviations: aBMD, areal BMD; sBMD, surface BMD; vBMD, volumetric BMD.

decisions in secondary osteoporosis.[67] In that study, 21%-25.5% of patients with unremarkable BMD measurements had poor bone quality indicated by TBS, which changed the treatment decision. A similar effect could occur with 3D-DXA, potentially leading to unintended consequences, such as overprescription. If the table was used, this may also sway the clinician to chose one drug over the other.

It is also stated by 3D-Shaper Medical that patients can be monitored to determine whether the selected treatment does indeed have the expected effect on the cortex and trabecular compartment. However, considering that errors of 3D-DXA are greater than expected changes, and that repeatability is low,[36] it cannot be ascertained whether the changes seen in a follow-up 3D-DXA assessment are due to real effects or merely the inherent error of 3D-DXA. If a clinician trusts these results, they may be swayed to unnecessarily change treatment.

Although the clinical version of 3D-DXA does not provide information about the cortical sBMD in the various anatomical regions, it does display a color coded 3D model of the sBMD across the bone surface. 3D-Shaper Medical subsequently suggests this can be used to detect local fragility.[66] Unfortunately, 3D-DXA cannot reconstruct the patient-specific cortical map. I will refer again to Figure 2 for an intuitive explanation for this. Thus, if a 3D-DXA cortical map suggests local deficiencies, these may be statistically plausible outputs of the model, but do not correspond to the patient's actual anatomy. This again may lead to an unnecessary or inappropriate treatment decision.

The 3D-Shaper software has now also received FDA clearance through the 510(k) premarket notification process.[68] 3D-Shaper Medical obtained this using a correlation study comparing the cross-sectional area, cross-sectional moment of inertia, section modulus (Z), Buckling Ratio (BR), cortical sBMD, trabecular vBMD and integral vBMD to similar measurements from the Hip Structural Analysis (HSA) software by Hologic Inc., which had previously received FDA clearance.

Notably, the neck axis length and neck shaft angle were not included in this evaluation, and thus their use has not been cleared by the FDA. It should be noted that The International Society for Clinical Densitometry guidelines recommend that HSA parameters should not be used to assess the risk of hip fracture, with the exception of hip axis length.[69] Furthermore, this regulatory approval does not include a T-score or Z-score to compare the measurements with reference data. This raises concerns about the clinical applicability of 3D-Shaper in the United States.

Finally, 3D-Shaper parameters were considered to be substantially equivalent to the HSA option for the Hologic QDR X-ray Bone Densitometers. The FDA is not able to confirm

whether 3D-Shaper can be used with other DXA scanners, leaving the question of compatibility with other devices in the United States unresolved.

Discussion

In summary, 3D-DXA produces a variety of bone parameters that appear highly detailed, but primarily reflect overall aBMD rather than measuring distinct cortical or trabecular properties. While the software may yield results that correlate with true values in populations where cortical and trabecular changes are proportional to aBMD, it falls short in capturing patient-specific measurements or localized changes. This limitation is especially pronounced in scenarios such as anabolic therapy, where cortical and trabecular bone parameters may change inversely and 3D-DXA fails to reflect these true changes observed by QCT. This raises significant concerns about its use in routine clinical care or as a reliable research tool.

3D-DXA was initially developed to better diagnose osteoporosis by providing an estimate of the integral vBMD.[70,71] By not attempting to separately assess the cortical and trabecular compartments, the integral vBMD produced by 3D-DXA may still represent a valid estimate. The methodology was later extended to attempt to improve fracture risk estimations by analyzing the model parameters directly, since these parameters fully describe the morphology of the statistical model.[51,72–74] However, the key limitations of 3D-DXA, as detailed in this review, remain and undermine its accuracy and reliability.

Companies commercializing 3D-DXA, including 3D-Shaper Medical, DMS Imaging, and Fujifilm, did not provide answers to key technical and methodological questions, including the clinically important issue of which population is used to calculate the T-scores and Z-scores. Consequently, this study relies on publicly available information and my own experience in developing the methodology. I encourage readers to seek clarifications from these companies or their representatives regarding aspects of 3D-DXA and 3D-Shaper that have not been disclosed.

Evaluations from professional societies could provide valuable guidance by providing independent assessments of the validity of 3D-DXA and establishing official recommendations regarding its use. Although an update on the practice guideline for DXA by the International Working Group on DXA Best Practices states regarding 3D-DXA that "more evidence is needed to make recommendations for the application of these novel imaging techniques in clinical practice,"[75] I believe there is now sufficient evidence to base a recommendation on, and I hope this review will be helpful in that regard.

In conclusion, given the fundamental limitations of 3D-DXA, it is my professional assessment that the cortical and trabecular parameters provided by 3D-DXA should not be used for research purposes and are not suitable for clinical applications such as diagnosis, monitoring, or treatment decision support.

Author contributions

Tristan Whitmarsh (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Visualization, Writing—original draft, Writing—review & editing).

Funding

No funding was received for this study.

Conflicts of interest

The author is the co-inventor of a patent related to the methodology underlying the 3D-DXA software. The author has been involved in discussions regarding the intellectual property and origins of the commercial 3D-DXA software code, statistical models, and promotional material. The author is not affiliated with 3D-Shaper Medical or other companies involved in the commercial exploitation of 3D-DXA and has not received equity, royalties, or other financial compensation. The author has lectured for UCB in educational fora and received research grant support from Amgen Inc. and Lilly. The views expressed in this article are solely those of the author and are based on a critical evaluation of the methodology and available scientific evidence.

Data availability

All data presented in this article are available from the author upon request.

References

1. Whitmarsh T. *3D Reconstruction of the Proximal Femur and Lumbar Vertebrae from Dual-Energy X-Ray Absorptiometry for Osteoporotic Risk Assessment* PhD thesis. Universitat Pompeu Fabra; 2012.
2. Humbert L, Martelli Y, Fonolla R, et al. 3D-DXA: assessing the femoral shape, the trabecular macrostructure and the cortex in 3D from DXA images. *IEEE Trans Med Imaging*. 2017;36(1):27–39. https://doi.org/10.1109/TMI.2016.2593346
3. Whitmarsh T, Humbert L, De Craene M, Del Rio Barquero LM, Frangi AF. Reconstructing the 3D shape and bone mineral density distribution of the proximal femur from dual-energy X-ray absorptiometry. *IEEE Trans Med Imaging*. 2011;30(12):2101–2114. https://doi.org/10.1109/TMI.2011.2163074
4. Ahmad O, Ramamurthi K, Wilson KE, Engelke K, Prince RL, Taylor RH. Volumetric DXA (VXA): a new method to extract 3D information from multiple in vivo DXA images. *J Bone Miner Res*. 2010;25(12):2744–2751. https://doi.org/10.1002/jbmr.140
5. Väänänen SP, Grassi L, Flivik G, Jurvelin JS, Isaksson H. Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image. *Med Image Anal*. 2015;24(1):125–134. https://doi.org/10.1016/j.media.2015.06.001
6. Whitmarsh T, Humbert L, De Craene M, et al. 3D bone mineral density distribution and shape reconstruction of the proximal femur from a single simulated DXA image: an in vitro study. In: Dawant BM, Haynor DR, eds. *Medical Imaging 2010: Image Processing*. Vol. 7623. SPIE; 2010:76234U.
7. Whitmarsh T. Concerns regarding the use of 3D-DXA. *Bone*. 2021;149:115939. https://doi.org/10.1016/j.bone.2021.115939
8. Treece GM, Gee AH. Independent measurement of femoral cortical thickness and cortical bone density using clinical ct. *Med Image Anal*. 2015;20(1):249–264. https://doi.org/10.1016/j.media.2014.11.012
9. Freitas L, Amorim T, Humbert L, et al. Cortical and trabecular bone analysis of professional dancers using 3D-DXA: a case–control study. *J Sports Sci*. 2018;37(1):82–89. https://doi.org/10.1080/02640414.2018.1483178
10. Gracia-Marco L, García-Fontana B, Ubago-Guisado E, Vlachopoulos D, García-Martín A, Manuel M. Analysis of bone impairment by 3D-DXA hip measures in patients with primary hyperparathyroidism: a pilot study. *J Clin Endocrinol Metab*. 2019;105(1):175–184.
11. Orduna G, Humbert L, Fonolla R, et al. Cortical and trabecular bone analysis of patients with high bone mass from the

Barcelona osteoporosis cohort using 3-dimensional dual-energy X-ray absorptiometry: a case-control study. *J Clin Densitom*. 2018;21(4):480–484. https://doi.org/10.1016/j.jocd.2017.05.012
12. Gifre L, Humbert L, Muxi A, et al. Analysis of the evolution of cortical and trabecular bone compartments in the proximal femur after spinal cord injury by 3D-DXA. *Osteoporos Int*. 2017;29(1):201–209. https://doi.org/10.1007/s00198-017-4268-9
13. Hoyos MG, Humbert L, Salmón Z, Riancho JA, Valero C. Analysis of volumetric BMD in people with down syndrome using DXA-based 3D modeling. *Arch Osteoporos*. 2019;14(1):98. https://doi.org/10.1007/s11657-019-0645-7
14. Beck B, Harding A, Weeks B, Humbert L, Winzenrieth R. Response to "concerns regarding the use of 3D-DXA". *Bone*. 2021;149:115936. https://doi.org/10.1016/j.bone.2021.115936
15. Winzenrieth R, Humbert L, Di Gregorio S, Bonel E, García M, Del Rio L. Effects of osteoporosis drug treatments on cortical and trabecular bone in the femur using DXA-based 3D modeling. *Osteoporos Int*. 2018;29(10):2323–2333. https://doi.org/10.1007/s00198-018-4624-4
16. Cummings SR, Cauley JA, Palermo L, et al. Racial differences in hip axis lengths might explain racial differences in rates of hip fracture. *Osteoporos Int*. 1994;4(4):226–229. https://doi.org/10.1007/BF01623243
17. Kim KM, Brown JK, Kim KJ, et al. Differences in femoral neck geometry associated with age and ethnicity. *Osteoporos Int*. 2010;22(7):2165–2174. https://doi.org/10.1007/s00198-010-1459-z
18. Sone T, Humbert L, Lopez M, Winzenrieth R, Ohnaru K. Assessment of femoral shape, trabecular and cortical bone in Japanese subjects using DXA-based 3D modelling. *J Bone Miner Res*. 2022;37(S1):S1–S385.
19. Varoquaux G, Cheplygina V. Machine learning for medical imaging: methodological failures and recommendations for the future. *npj Digital Med*. 2022;5(1):48. https://doi.org/10.1038/s41746-022-00592-y
20. Hadjiiski L, Cha K, Chan H-P, et al. AAPM task group report 273: recommendations on best practices for AI and machine learning for computer-aided diagnosis in medical imaging. *Med Phys*. 2023;50(2):e1-e24. https://doi.org/10.1002/mp.16188
21. Huisman M. When AUC-ROC and accuracy are not accurate: what everyone needs to know about evaluating artificial intelligence in radiology. *Eur Radiol*. 2024;34(12):7892–7894. https://doi.org/10.1007/s00330-024-10859-5
22. Gallifant J, Bitterman DS, Celi LA, et al. Ethical debates amidst flawed healthcare artificial intelligence metrics. *npj Digital Med*. 2024;7(1):243. https://doi.org/10.1038/s41746-024-01242-1
23. Amani A, Bellver M, del Rio L, et al. Femur 3D-DXA assessment in female football players, swimmers, and sedentary controls. *Int J Sports Med*. 2022;44(06):420–426. https://doi.org/10.1055/a-1928-9824
24. Maïmoun L, Renard E, Humbert L, et al. Modification of bone mineral density, bone geometry and volumetric BMD in young women with obesity. *Bone*. 2021;150:116005. https://doi.org/10.1016/j.bone.2021.116005
25. Jain RK, Picazo ML, Humbert L, Dickens L, Vokes T. Bone structural parameters as measured by 3-dimensional dual-energy X-ray absorptiometry are superior in Black women and demonstrate unique associations with prior fracture versus White women. *Endocr Pract*. 2025;31(2):152–158. https://doi.org/10.1016/j.eprac.2024.10.015
26. Gracia-Marco L, Gonzalez-Salvatierra S, Garcia-Martin A, et al. 3D-DXA hip differences in patients with acromegaly or adult growth hormone deficiency. *J Clin Med*. 2021;10(4):657. https://doi.org/10.3390/jcm10040657
27. Toussirot E, Winzenrieth R, Aubin F, Wendling D, Vauchy C, Desmarets M. Areal bone mineral density, trabecular bone score and 3D-DXA analysis of proximal femur in psoriatic disease. *Front Med (Lausanne)*. 2024;11:1341077. https://doi.org/10.3389/fmed.2024.1341077
28. Ubago-Guisado E, Moratalla-Aranda E, González-Salvatierra S, et al. Do patients with type 2 diabetes have impaired hip bone microstructure? A study using 3D modeling of hip dual-energy X-ray absorptiometry. *Front Endocrinol (Lausanne)*. 2023;13:1069224. https://doi.org/10.3389/fendo.2022.1069224
29. Guerra FS, Palladino N, Winzenrieth R, Guglielmi G. Advanced 3D-DXA insights into bone density changes in hyperparathyroidism. *J Diabetes Metab Disord*. 2024;23(2):2191–2199. https://doi.org/10.1007/s40200-024-01487-3
30. Kužma M, Vaňuga P, Ságová I, et al. Non-invasive DXA-derived bone structure assessment of acromegaly patients: a cross-sectional study. *Eur J Endocrinol*. 2019;180(3):201–211. https://doi.org/10.1530/EJE-18-0881
31. Maïmoun L, Aouinti S, Puech M, et al. Modification of bone architecture following sleeve gastrectomy: a five-year follow-up. *J Bone Miner Res*. 2024;40(2):251–261.
32. Maïmoun L, Gelis A, Serrand C, et al. Alteration of volumetric bone mineral density parameters in men with spinal cord injury. *Calcif Tissue Int*. 2023;113(3):304–316. https://doi.org/10.1007/s00223-023-01110-2
33. Harding AT, Weeks BK, Lambert C, Watson SL, Weis LJ, Beck BR. Effects of supervised high-intensity resistance and impact training or machine-based isometric training on regional bone geometry and strength in middle-aged and older men with low bone mass: the LIFTMOR-M semi-randomised controlled trial. *Bone*. 2020;136:115362. https://doi.org/10.1016/j.bone.2020.115362
34. Gil-Cosano JJ, Ubago-Guisado E, Sánchez MJ, et al. The effect of an online exercise programme on bone health in paediatric cancer survivors (iBoneFIT): study protocol of a multi-centre randomized controlled trial. *BMC Public Health*. 2020;20(1):1520. https://doi.org/10.1186/s12889-020-09607-3
35. Humbert L, Whitmarsh T, De Craene M et al. 3D reconstruction of both shape and bone mineral density distribution of the femur from DXA images. In: *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE; 2010:456–459.
36. O'Rourke D, Beck BR, Harding AT, Watson SL, Pivonka P, Martelli S. Assessment of femoral neck strength and bone mineral density changes following exercise using 3D-DXA images. *J Biomech*. 2021;119:110315. https://doi.org/10.1016/j.jbiomech.2021.110315
37. Dudle A, Gugler Y, Pretterklieber M, Ferrari S, Lippuner K, Zysset P. 2D-3D reconstruction of the proximal femur from DXA scans: evaluation of the 3D-Shaper software. *Front Bioeng Biotechnol*. 2023;11:1111020. https://doi.org/10.3389/fbioe.2023.1111020
38. Clotet J, Martelli Y, Di Gregorio S, del Río Barquero LM, Humbert L. Structural parameters of the proximal femur by 3-dimensional dual-energy X-ray absorptiometry software: comparison with quantitative computed tomography. *J Clin Densitom*. 2018;21(4):550–562. https://doi.org/10.1016/j.jocd.2017.05.002
39. Ramamurthi K, Ahmad O, Engelke K, et al. An in vivo comparison of hip structure analysis (HSA) with measurements obtained by QCT. *Osteoporos Int*. 2011;23(2):543–551. https://doi.org/10.1007/s00198-011-1578-1
40. Whitmarsh T, Treece GM, Gee AH, Poole KES. Mapping bone changes at the proximal femoral cortex of postmenopausal women in response to alendronate and teriparatide alone, combined or sequentially. *J Bone Miner Res*. 2015;30(7):1309–1318. https://doi.org/10.1002/jbmr.2454
41. Whitmarsh T, Treece GM, Gee AH, Poole KES. The effects on the femoral cortex of a 24 month treatment compared to an 18 month treatment with teriparatide: a multi-trial retrospective analysis. *PLoS One*. 2016;11(2):e0147722. https://doi.org/10.1371/journal.pone.0147722
42. Huininga K, Koromani F, van Velsen E, Zillikens C, Rivadeneira F. From 2D to 3D Imaging: 3D shaper and EOS Edge applications in the Rotterdam study. *JBMR Plus*. 2024;8(Supplement_1):i1-i340. Published: 20 May 2024.
43. Huininga K, Koromani F, Zillikens MC, van Velsen E, Rivadeneira F. Use of 3d shaper analysis for the assessment of fracture risk

in a population-based setting. *J Bone Miner Res*. 2024;39(Supplement_1):i1–i406. https://doi.org/10.1093/jbmr/zjae183
44. Qasim M, Picazo ML, Wills CR, et al. 3D-DXA based finite element modelling for femur strength prediction: evaluation against QCT. *J Clin Densitom*. 2024;27(2):101471. https://doi.org/10.1016/j.jocd.2024.101471
45. Iki M, Winzenrieth R, Tamaki J, et al. Predictive ability of novel volumetric and geometric indices derived from dual-energy X-ray absorptiometric images of the proximal femur for hip fracture compared with conventional areal bone mineral density: the Japanese Population-based Osteoporosis (JPOS) cohort study. *Osteoporos Int*. 2021;32(11):2289–2299. https://doi.org/10.1007/s00198-021-06013-2
46. Grassi L, Väänänen SP, Ristinmaa M, Jurvelin JS, Isaksson H. Prediction of femoral strength using 3D finite element models reconstructed from DXA images: validation against experiments. *Biomech Model Mechanobiol*. 2016;16(3):989–1000. https://doi.org/10.1007/s10237-016-0866-2
47. Grassi L, Fleps I, Sahlstedt H, et al. Validation of 3D finite element models from simulated DXA images for biofidelic simulations of sideways fall impact to the hip. *Bone*. 2021;142:115678. https://doi.org/10.1016/j.bone.2020.115678
48. Grassi L, Väänänen SP, Jehpsson L, et al. 3D finite element models reconstructed from 2D dual-energy X-ray absorptiometry (DXA) images improve hip fracture prediction compared to areal BMD in osteoporotic fractures in men (MROS) Sweden cohort. *J Bone Miner Res*. 2023;38(9):1258–1267. https://doi.org/10.1002/jbmr.4878
49. Grassi L, Väänänen SP, Voss A, et al. DXA-based 3D finite element models predict hip fractures better than areal BMD in elderly women. *Bone*. 2025;195:117457. https://doi.org/10.1016/j.bone.2025.117457
50. Allison SJ, Poole KES, Treece GM, et al. The influence of high-impact exercise on cortical and trabecular bone mineral content and 3D distribution across the proximal femur in older men: a randomized controlled unilateral intervention. *J Bone Miner Res*. 2015;30(9):1709–1716. https://doi.org/10.1002/jbmr.2499
51. Whitmarsh T, Fritscher KD, Humbert L, Del-Rio-Barquero LM, Schubert R, Frangi AF. Hip fracture discrimination using 3D reconstructions from dual-energy X-ray absorptiometry. In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE; 2011:1189–1192.
52. Winzenrieth R, Ominsky MS, Wang Y, Humbert L, Weiss RJ. Differential effects of abaloparatide and teriparatide on hip cortical volumetric BMD by DXA-based 3D modeling. *Osteoporos Int*. 2021;32(3):575–583. https://doi.org/10.1007/s00198-020-05806-1
53. England JR, Cheng PM. Artificial intelligence for medical image analysis: a guide for authors and reviewers. *Am J Roentgenol*. 2019;212(3):513–519. https://doi.org/10.2214/AJR.18.20490
54. Michael Lewiecki E, Betah D, Humbert L, et al. 3D-modeling from hip DXA shows improved bone structure with romosozumab followed by denosumab or alendronate. *J Bone Miner Res*. 2024;39(4):473–483. https://doi.org/10.1093/jbmr/zjae028
55. Winzenrieth R, Kostenuik P, Boxberger J, Wang Y, Humbert L. Proximal femur responses to sequential therapy with abaloparatide followed by alendronate in postmenopausal women with osteoporosis by 3D modeling of hip dual-energy X-ray absorptiometry (DXA). *JBMR Plus*. 2022;6(4):e10612. https://doi.org/10.1002/jbm4.10612
56. Winzenrieth R, Humbert L, Boxberger JI, Weiss RJ, Wang Y, Kostenuik P. Abaloparatide effects on cortical volumetric BMD and estimated strength indices of hip subregions by 3D-DXA in women with postmenopausal osteoporosis. *J Clin Densitom*. 2022;25(3):392–400. https://doi.org/10.1016/j.jocd.2021.11.007
57. Bennett CM, Miller MB, Wolford GL. Neural correlates of interspecies perspective taking in the post-mortem Atlantic salmon: an argument for multiple comparisons correction. *NeuroImage*. 2009;47:S125. https://doi.org/10.1016/S1053-8119(09)71202-9
58. Poole KES, Treece GM, Gee AH, et al. Denosumab rapidly increases cortical bone in key locations of the femur: a 3D bone mapping study in women with osteoporosis. *J Bone Miner Res*. 2014;30(1):46–54.
59. Sato M, Westmore M, Ma YL, et al. Teriparatide [pth(1–34)] strengthens the proximal femur of ovariectomized nonhuman primates despite increasing porosity. *J Bone Miner Res*. 2004;19(4):623–629. https://doi.org/10.1359/JBMR.040112
60. Hadji P, Kamali L, Thomasius F, Horas K, Kurth A, Bock N. Real-world efficacy of a teriparatide biosimilar (rgb-10) compared with reference teriparatide on bone mineral density, trabecular bone score, and bone parameters assessed using quantitative ultrasound, 3D-Shaper® and high-resolution peripheral computer tomography in postmenopausal women with osteoporosis and very high fracture risk. *Osteoporos Int*. 2024;35(12):2107–2116. https://doi.org/10.1007/s00198-024-07208-z
61. Sone T, Ohnaru K, Sugai T, et al. The effects of abaloparatide on hip geometry and biomechanical properties in Japanese osteoporotic patients assessed using DXA-based hip structural analysis: results of the Japanese phase 3 ACTIVE-J trial. *Arch Osteoporos*. 2023;18(1):146. https://doi.org/10.1007/s11657-023-01344-5
62. Genant HK, Engelke K, Bolognese MA, et al. Effects of romosozumab compared with teriparatide on bone density and mass at the spine and hip in postmenopausal women with low bone mass. *J Bone Miner Res*. 2016;32(1):181–187. https://doi.org/10.1002/jbmr.2932
63. 3D-SHAPER MEDICAL. *Webinar - Use of 3d-Shaper in Clinical Research: State of the Art and Potential Applications* https://www.youtube.com/watch?v=xujeNrv0mlI. Chairman: Pr. David Kendler, University of Vancouver, British Columbia, Canada; Speakers: Dr. Ludovic Humbert and Dr. Renaud Winzenrieth, 3D-Shaper Medical; 2021.
64. Burgos EC, Di Gregorio S, Macías JG. et al. Datos de referencia de mediciones óseas en modelos 3D de femur proximal en población española cn DXA: Proyecto SEIOMM-3D-SHAPER. *Rev Osteoporos Metab Miner*. 2019;11(3):7–14.
65. Brance ML, Saravi FD, Henríquez MM. et al. Reference values of three-dimensional proximal femur parameters from bone densitometry images in healthy subjects from Argentina. World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (WCO-IOF-ESCEO 2020): Poster Abstracts. *Osteoporos Int*. 2020;31(Suppl 1):133–621. https://doi.org/10.1007/s00198-020-05696-3
66. Jennings I. *3d Shaper Technology: Revolutionizing Bone Health Analysis*. https://www.youtube.com/watch?v=te6a_Oz3QMI.
67. Al-Hashimi L, Klotsche J, Ohrndorf S, Gaber T, Hoff P. Trabecular bone score significantly influences treatment decisions in secondary osteoporosis. *J Clin Med*. 2023;12(12):4147. https://doi.org/10.3390/jcm12124147
68. U.S. Food and Drug Administration. *510(k) Summary: K220822*. 2022.
69. Broy SB, Cauley JA, Lewiecki ME, Schousboe JT, Shepherd JA, Leslie WD. Fracture risk prediction by non-BMD DXA measures: the 2015 ISCD official positions part 1: hip geometry. *J Clin Densitom*. 2015;18(3):287–308. https://doi.org/10.1016/j.jocd.2015.06.005
70. Whitmarsh T, Humbert L, De Craene M, Del Rio Barquero LM, Frangi AF. Reconstrucción de la forma del fémur y densidad mineral ósea en 3D Para el diagnóstico de osteoporosis a partir de dxa. XIV Congreso SEIOMM 2009. *Rev Mult Gerontol*. 2009;18(Supl 1).
71. Whitmarsh T, Humbert L, Del Río Barquero LM, Frangi AF. Volumetric bone mineral density estimation using a 3D reconstruction method from dual-energy X-ray absorptiometry. *J Bone Miner Res*. 2011;26(Supl 1):1S. https://doi.org/10.1002/jbmr.564.
72. Whitmarsh T, Fritscher KD, Humbert L, et al. *A Statistical Model of Shape and Bone Mineral Density Distribution of the Proximal Femur for Fracture Risk Assessment*. Springer; 2011:393, 400.
73. Whitmarsh T, Fritscher KD, Humbert L, et al. Hip fracture discrimination from dual-energy X-ray absorptiometry by

statistical model registration. *Bone*. 2012;51(5):896–901. https://doi.org/10.1016/j.bone.2012.08.114

74. Martelli Y, Whitmarsh T, Humbert L, et al. A software framework for 3D reconstruction and fracture risk assessment of the proximal femur from dual-energy X-ray absorptiometry. In: *Proceedings of VPH 2012: Virtual Physiological Human - Integrative approaches to computational biomedicine*. 2012. Presented at VPH 2012, 2012.
75. Slart RHJA, Punda M, Ali DS, et al. Updated practice guideline for dual-energy X-ray absorptiometry (DXA). *Eur J Nucl Med Mol Imaging*. 2025;52(2):539–563. https://doi.org/10.1007/s00259-024-06912-6

JBMR Plus, 2025, 9, ziaf075
https://doi.org/10.1093/jbmrpl/ziaf075
Review