HHS Public Access

Author manuscript

Curr Osteoporos Rep. Author manuscript; available in PMC 2024 March 15.

Published in final edited form as:

Curr Osteoporos Rep. 2022 October ; 20(5): 309–319. doi:10.1007/s11914-022-00743-w.

A review of CT-based fracture risk assessment with finite element modeling and machine learning

Ingmar Fleps[1,*], **Elise F. Morgan**[1]

[1]College of Mechanical Engineering, Boston University

Abstract

Purpose of Review—We reviewed advances over the past three years in assessment of fracture risk based on CT scans, considering methods that use finite element models, machine learning, or a combination of both.

Recent Findings—Several studies have demonstrated that CT-based assessment of fracture risk, using finite element modeling or biomarkers derived from machine learning, to be equivalent to currently used clinical tools. Phantomless calibration of CT scans for bone mineral density enables accurate measurements from routinely taken scans. This opportunistic use of CT scans for fracture risk assessment is facilitated by high quality automated segmentation with deep learning, enabling workflows that do not require user intervention. Modeling of more realistic and diverse loading conditions, as well as improved modeling of fracture mechanisms, has shown promise to enhance our understanding of fracture processes and improve the assessment of fracture risk beyond the performance of current clinical tools.

Summary—CT-based screening for fracture risk is effective and, by analyzing scans that were taken for other indications, could be used to expand the pool of people screened, therefore improving fracture prevention. Finite element modeling and machine learning both provide valuable tools for fracture risk assessment. Future approaches should focus on including more loading-related aspects of fracture risk.

Keywords

Fragility fractures; elderly; bone strength; finite element method; machine learning; computed tomography

Introduction

Osteoporosis is a strong risk factor for fragility fractures, however fewer than 50% of people with fragility fractures are classified as having osteoporosis (1,2), highlighting the need for

[*]Corresponding author: Ingmar Fleps, flepsi@bu.edu.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Human and Animal Rights and Informed Consent

This article does not contain any studies with human or animal subjects performed by any of the authors.

Ethical Approval

This review article does not present any previously unpublished original research, and ethical approval is therefore not applicable.

improved strategies for assessing fracture risk. Individuals with osteoporosis are identified as having an areal bone mineral density (aBMD), measured with dual energy X-ray absorption (DXA), that is at least 2.5 standard deviations lower than the average aBMD of young women (3). Alternative clinical tools for the assessment of fracture risk are composite scores like the fracture risk assessment tool (FRAX), which performs similar to aBMD in assessing future fracture risk (4).

Finite element (FE) models based on computed tomography (CT) scans offer the possibility to improve risk assessments, by simulating clinically relevant loading of the bone and estimating the force that is required to initiate fracture. Compared to aBMD, FE models based on CT scans have been shown to be more accurate at predicting bone strength *ex vivo* (5,6). On clinical cohort data, FE-derived bone strength has performed better than aBMD at identifying individuals who will suffer a vertebral fracture and at least equivalent at identifying those will suffer a hip fracture (7,8). However, clinical adoption of the FE-based approaches has been slow. Potential reasons are the higher radiation exposure of CT scans compared to DXA scans, more time-consuming model generation, and the lack of regulatory procedures for approval of such technologies. This review discusses the recent progress that has been made towards overcoming these challenges.

Studies aiming to expand our mechanistic understanding of fragility fractures are now combining bone strength with subject-specific estimates of loading severity, which encompasses the magnitude, spatial distribution, duration, and frequency of the applied loads (9–11). Such knowledge might improve the fidelity of *ex vivo* laboratory models and computational models and help us to better understand why such a large percentage of fragility fractures in the elderly occur in individuals who do not have low aBMD.

In contrast to CT-based FE models, machine learning (ML) methods aim to identify and exploit patterns in data to predict variables of interest without the need for a mechanistic description of which features in the data are predictive or how they are related to these variables. Medical image processing is an area that has rapidly evolved since the development of deep learning, a type of ML that allows the user to use images directly as input, rather than requiring the user to first extract features from the images (12). Recent developments in the amount of available data and the frameworks for deep learning have provided powerful tools to process the spatial information encoded in CT scans for high quality image segmentation and biomarker extraction. This review discusses current ML tools to extract biomarkers and predict fracture risk, some of which are being implemented into web services and clinical tools with approval for medical reimbursement.

The purpose of this review is to give an overview of advances in fracture risk assessment based on CT scans that were made in the last 3 years, focusing on approaches that use finite element models, machine learning, or a combination of both. This review highlights key studies and concepts and is not intended as a systematic review.

Clinical use of FE-based assessment of fracture risk

The use of CT-based computational modeling for the diagnosis of osteoporosis and screening for fracture risk has gained increased traction due to the idea of opportunistic use of CT scans, which proposes the use of CT scans performed for other indications, such as colon cancer screening or urolithiasis (13). Because already existing scans are used, no additional radiation exposure is required for the assessment of fracture risk, making these scans an ethical and cost-effective alternative that can be used complementary to FRAX scores and aBMD.

Studies on clinical cohorts have shown promising results regarding the use of FE-derived femoral strength and volumetric bone mineral density (vBMD), both obtained from CT scans. A study that used opportunistic abdominal CT-scans (FOCUS cohort, n =3938) found that Hip fracture risk assessment with vBMD and FE-derived femoral strength was at least equivalent to DXA-derived aBMD (8). CT-based predictors also remained predictive of fracture risk even when controlling for aBMD. These finding are consistent with a study (AGES-RS study cohort, n = 601) that found FE-derived femoral strength to be more predictive of hip fracture than was QCT-derived aBMD (14). In this study, models with a non-linear material implementation were more predictive than linear models. Contrary to previous studies (15,16) predictions of fracture risk did not improve when considering multiple loading scenarios. Overall, these studies show that fracture risk assessment with FE-derived femoral strength is effective at predicting hip fracture risk with an accuracy that is at least comparable to that provided by aBMD (Table 1).

Similarly, vertebral CT scans have been demonstrated to be useful for the assessment of vertebral fracture risk (7). A study using a case-control cohort (AGES-RS cohort, 401 subjects) demonstrated that vBMD and FE-derived vertebral strength from the thoracic and lumbar spine can both be used effectively for fracture risk prediction (17). This allows for the extension of the pool of CT scans that can be used for opportunistic assessment of fracture risk. Recent findings in smaller cohorts further support the usefulness of vBMD and FE-derived vertebral strength for the assessment of vertebral fracture risk (18,19). These findings were limited to CT scans without contrast agents, which are frequently administered to increase the contrast of soft tissues in the image. To allow for the use of scan with and without contrast agents, a separate analysis is required to investigate the usefulness of scans with contrast agents.

The first guidelines for reporting of computational modeling studies were released by the FDA in 2016 (20) followed by a standard for verification, validation, and credibility assessment of computational models for the application to medical devices (21). The first *in silico* tools (E.g. VirtuOst, O.N. Diagnostics, Berkeley, USA) for the assessment of fracture risk based on CT scans have now been approved by the FDA and are eligible for medical reimbursement. These tools provide BMD measurements and bone strength assessment based on pre-existing CT scans. Challenges for the cost effectiveness and applicability of these tools are the generation of high-quality segmentation masks and the need for BMD calibration, given that clinical scans often don’t include a calibration phantom. Advances in these areas will be discussed in the following two sections.

Automatic segmentation

Recent advances using deep learning for image segmentation have demonstrated promise for fully automated segmentation of tissues from medical imaging. Deep learning frameworks learn segmentation models from medical images without the need for the separate development of methods to extract features and subsequently use these models to predict segmentation masks (12). Deep learning architectures for image segmentation rely on ground truth data during training to provide feedback to the model. Therefore, pre-labelled data are required to train the model for the desired task. Datasets that contain the variability of data that is encountered in clinical practices are essential for training. Further, some of the pre-labelled data should be reserved as a test set, to evaluate the performance of the trained model.

In the last couple of years, deep learning frameworks for high quality, fully automated segmentation have been made publicly available (22,23). A fully convolutional neural network structure has been used to segment, to label vertebrae, and to judge if the complete vertebra is visible in the scan with the same ML framework. The study used data from four different CT and MRI datasets that were non-uniform with respect to the vertebral levels and number of vertebrae included in the field of view. A dice coefficient, a measure of the number of pixels that both segmentations have in common compared to the average number of pixels in individual segmentations, of 0.931–0.965 was achieved, outperforming previous automated segmentation methods. The top performers in a 2020 competition among approaches for automatic segmentation of vertebrae from CT scans (24) achieved an average dice coefficient of 0.888, a clear improvement over the winning models from the competition just one year prior (dice coefficient of 0.789). The public availability of the training and test data of this challenge (https://github.com/anjany/verse) provides a means for further development of vertebral segmentation algorithms and continue the rapid development that has been achieved in this area.

Fully automated segmentation of proximal femur has received less attention, which is likely because its less complex shape is easier to extract from CT scans. Automated methods have been published using a combination of thresholding and watershed algorithm (25) or by multiple thresholding and morphing of a template shape (26). These methods resulted in highly accurate segmentation (dice coefficients of 0.97 and 0.93); however, both methods were only tested on a small number of femora. Segmentation models based on deep learning for the bones in the pelvic region (27) and femoral head (28) used larger datasets of 200 CT scan (20 test scans) and 136 CT scans (27 test scans) and performed similar to other automated methods with respect to dice coefficient. A recent study performed training on 60 femora and tested the model on 200 femur achieving dice coefficients ranging from 0.96–0.99 (29). The quality of automated segmentation algorithms has reached comparable levels to inter-operator differences between manual segmentations (30) providing a solid foundation for the use of these algorithms. Further work should involve testing on larger datasets, including data acquired with different scanners, as well as for different fields of view, reconstruction kernels and scanner parameters.

Phantomless calibration

Calibration of CT scans for the calculation of BMD based on the image greyscale values is required for generation of subject-specific FE models and for extraction of imaging biomarkers of fracture risk, such as bone mineral density, cortical bone mass and cortical thickness. In research, this calibration is typically done with an in-line phantom that contains volumes of known mineral density that are used to extract a linear relationship between image grey scale values and BMD. Alternative calibration methods are asynchronous calibration, which uses a single scan of a calibration phantom to calibrate all scans taken with a scanner within a predetermined time frame, and internal material calibration, also called phantomless calibration, which uses the grey scale values of tissues within the scan to derive a calibration equation without the need for a phantom. Phantomless calibration allows for opportunistic assessment of BMD and the generation of FE models from CT scans regardless of the presence of a calibration phantom.

Several procedures for phantomless calibration have been proposed (31–33) using different anatomical regions. Reference density values for different tissues can be extracted by calculating them from calibrated CT scans (32) or by using standardized reference values from the National Institute for Standards and Technology (31,33,34). Excellent agreement for BMD and bone strength predicted using phantomless vs. in-line phantom calibration has been achieved for the vertebra (coefficients of determination (R^2) between 0.94–0.99) (32,34,35). These studies used air as one of the reference volumes and a second reference volume with a large difference in the mean HU value, such as aortic blood, adipose tissue, muscle, or bone. For the calibration of femoral BMD and FE-computed strength predictions, coefficients of determination ranged from 0.86–0.99 (32,34) showing a large spread. Potential sources for error for phantomess calibrations are fluctuation in the scanner performance and natural inter-subject variation in the tissue composition. A study on 121 CT scans of the lumbar spine region of elderly subjects found the theoretical error due to inter-subject tissue composition differences to be <10% (33). They reported promising results for combinations of air with aortic blood or muscle tissue. However, the ideal choice of tissue combination might be dependent on the population of interest or the individual's pathologies. According to this study, fatty muscle infiltration, which is common in the elderly, could lead to BMD errors > 30% due to differences in muscle composition. Careful evaluation of phantomless calibration procedures for subjects with such biases is therefore required.

Reference volumes used in studies that investigated phantomless calibration were segmented manually or in semi-automated fashion, which could lead to loss of repeatability of the predictions. However, inter-operator reanalysis precision has been found to be excellent by three different studies (31,32,34), using 2–3 operators per study, suggesting that phantomless calibration can be reliably implemented for clinical workflows. Automation of the segmentation of the tissues of interest will allow for a comparison of in-line and phantomless calibration on large clinical cohorts to develop robust algorithms.

Advances in the prediction of fracture mechanisms with finite element modeling

Although bone strength is typically the primary outcome of FE modeling of the vertebra and femur, other outcomes, such as fracture location and progression, if predicted with high accuracy, might provide additional information regarding fracture prevention and treatment. While the prediction of fracture locations in the femur during stance can be estimated with high accuracy (36,37), prediction under sideways fall loading conditions is more challenging (38,39). The extended finite element method (XFEM) (40), and a combination of FE method and phase field method (41), allow for improved modeling of discontinuities that appear during fracture processes and enable tracking of fracture paths (42,43), which might lead to more accurate prediction of fracture locations.

XFEM uses local enrichment functions to describe discontinuities that arise due to crack growth. As such, local damage can be considered, and crack paths can be tracked without the need of local mesh refinement between solution steps, which greatly increases computational efficiency. First implementations into image-based FE models of the vertebra (44) and the femur (43) used strain-based criteria for crack initiation, because the yield strain of bone is not very sensitive to differences in BMD. Both studies provided a proof-of-concept implementation into a small number of specimens. Applications to larger data set and other clinically relevant loading scenarios will be needed to further evaluate the potential of XFEM for improved prediction of fracture locations.

The combination of phase-field method and FE method originated from crystallography simulations but has recently been applied to micromechanical simulations of bone (45) and macroscopical structures of brittle biomaterials, such as artificial femora (46). This method uses a diffusive phase field in combination with the constitutive material model of the FE simulation. Cracks are modelled by the value of the phase field, which reduces the stiffness of the elements in similar manner to a damage parameter in continuum damage theory. However, the diffusive zone of the phase field provides a smooth transition from fully cracked to intact material. A study on image-based humerus models considered the fracture energy to be a function of the Young’s modulus, indirectly linking the fracture energy to BMD (42). The methodology was applied to a single specimen that had been tested experimentally, showing good agreement regarding the fracture location. Recently published open-source implementations of element formulations for combined FE and phase-field modeling in Abaqus (47,48) may simplify the adoption of this method into validation studies for fracture predictions to evaluate its potential to produce more accurate predictions of fracture initiation and propagation.

Variations in cortical thickness are another aspect related to bone fragility that is challenging to assess based on clinical CT scans, because the thickness of thin cortices is frequently smaller than the image resolution. Methods for cortical thickness measurement from clinical CT scans have been developed (49,50), but the integration of these measurements into FE models is not trivial. A recent study developed a framework that combined statistical shape modeling with a ML model to estimate cortical thickness for the generation of FE models of the femur with an accurate representation of the cortical shell (51). The study

analyzed the robustness of the framework based on a dataset of 72 femora, demonstrating that a high-quality finite element discretization can be achieved while providing comparable accuracy for cortical thickness estimates to state of the art methods.

Improvements regarding the fidelity of laboratory loading conditions

Models that consider the subject-specific loading environment of a bone in combination with FE-computed values of bone strength in this environment aim to better understand fracture processes and improve fracture risk assessment.

Most hip fractures in the elderly are the result of a falls from standing height with impact to the lateral aspect of the proximal thigh. Laboratory experiment of such impacts with volunteers are limited to low loading severity due to the risk of injury. *Ex vivo* experiments offer an alternative which allows simulation of impacts that can lead to injury. Modeling 11 subject-specific sideways fall impacts with cadaveric tissue of bones, cartilage and ligaments of the pelvic region embedded in a surrogate material for the soft tissues in the hip region, resulted in impacts with femoral fracture, pelvic fracture and no fractures (9,52). These experiments in combination with complimentary FE models (53,54) demonstrated that low severity impact underestimated the stiffness of the pelvic region by a factor of 3–5 compared to an impact due to a fall from standing. The experimentally observed clinically relevant fractures allowed for the validation of FE models for the differentiation of fracture outcomes. This modeling approach was used to investigate its potential for the assessment of fracture risk using a cohort of 254 female subjects and found its performance to be comparable to CT-derived aBMD (55). However, when the cohort was limited to self-reported fallers (122 subjects), a larger and statistically significant improvement in the AUC was found for the fall models compared to aBMD, indicating that reliable information for the risk of falling might be required as additional input for estimates of the load applied to the femur and femoral strength. Due to the detailed description of the fracture event, such models could also offer an *in silico* framework for subject-specific assessment of preventive treatments such as hip protectors (56), pharmaceutical treatment or surgical augmentation (57), but extensive validation for this type of assessment is still required. In a study on a case-control cohort of 98 CT-scans of postmenopausal women taken post fracture, a complex multi-scale model resulted in an improvement in hip fracture risk assessment when compared to FE-derived femoral strength alone (AUC = 0.85 and 0.82, respectively) (58). The model considered different fall scenarios, body anthropometrics, and assessments of femoral strength specific to the loading alignment. Further validation of these models on larger prospective cohorts could establish the predictive ability of such models.

In contrast to hip fracture, the loading conditions leading to vertebral fractures are less well known. Vertebral fractures can be the result of overloading but have also been reported to slowly progress over time with often unknown initiation. To investigate the potential for vertebral overloading, strength estimates based on CT scans have been combined with subject-specific multi-body dynamics simulations of the vertebral column that were informed by lateral scout views and CT scans, allowing for the calculation of load-to-strength ratios across 250 individual and 109 static postures of daily living (11). Highest load-to-strength ratios were most often found in the lower thoracic and lumbar spine and

task involving loading with a flexed spine, but varied subject-specifically with respect to the posture and vertebral level. A limitation of the study was that load was directly compared to BMD-based vertebral strength estimates without considering activity and (intervertebral disc) IVD specific loading patterns.

Modeling the vertebral loading through the adjacent IVDs could result in more realistic loading of the vertebral endplate, which might improve predictions of bone strength (59) and failure locations (60). Including the IVD in image-based FE model improved the prediction of fracture forces compared to a model that assumed uniform endplate displacements in a study on ten embalmed lumbar vertebral column segments that included two IVDs and three vertebrae without posterior elements (59). These results are in line with a study which compared experimentally measured vertebral deformations with FE models for lumbar spine segments, including three vertebrae and two IVDs, in axial compression and concluded that inaccuracies in the assumed vertebral loading conditions limit the accuracy of vertebral FE models regarding the prediction of vertebral deformation (60).

Recent studies have made important steps towards more realistic load application and load estimates that will help to improve our understanding of bone fragility. Further development of these models could lead to more sophisticated methods for fracture risk assessment and targeted selection of available preventive treatments –such as hip protectors, muscle strengthening, and posture— based on subject-specific considerations of loading and bone strength.

Machine learning approaches for vertebral fracture detection

Vertebral fractures have been shown to be predictive of future osteoporotic fractures but are currently underreported in clinical practice (61). Clinically, identification of vertebral compression fractures relies on vertebral fracture assessment (VFA) based on DXA scans or radiological assessment on CT scan according to the Genant grade (62), which classifies fractures based on a reduced anterior, medial, or posterior vertebral body height into four categories. Automated vertebral fracture detection, such as by ML, on medical images would allow for a more widespread analysis of images to identify fractures that currently go unnoticed.

An algorithm based on neural networks detected vertebral fracture from CT scans with an accuracy of 89.2%, a sensitivity of 85.2%, and a specificity of 95.8%, which was comparable to the performance of the original medical records of the study when compared to the evaluation by a radiological expert (63). The study used central sagittal slices of 1432 CT scans, split into 80% training data, 10% validation data, and 10% test data. The images had already been labelled for “compression deformity” or “compression fracture” in the database. A study that used a previously trained commercial ML tool (Zebra Medical vision) on a dataset of 1696 lumbar CT scans of patient older than 50 reported a sensitivity for identifying vertebral fractures of 54% and a specificity of 92% when considering vertebral fractures for all Genant grades (64). Limiting the analysis to the identification of fractures with a Genant grade of 2–3 increased the sensitivity to 65% while keeping the specificity at 92%. 64% more vertebral fractures were identified in the dataset than were originally

documented in the medical records, demonstrating the usefulness of the tool for automated clinical identification of vertebral fractures.

Although the sensitivity of ML methods for vertebral fracture detection can likely be improved, its performance is comparable to current manual vertebral fracture detection. Consequently, the use of these automated algorithms could reduce manual effort in clinical practice and allow for a more widespread analysis of CT scans for vertebral fractures.

Machine learning approaches for bone mineral density assessment and bone strength

ML methods have been developed that can predict DXA-equivalent BMD from CT scans through automated segmentation processes and processing of grey scale values, thus eliminating operator variability and increasing throughput. A combination of ML-based segmentation of the lumbar vertebrae in combination with a trained linear regression model predicted T-score equivalent to those derived from DXA-based aBMD from CT scans with a correlation coefficient of 0.85 (65). The study used 1693 pairs of CT and DXA scans for the development and assessment of the model and concluded that the CT-based assessment would have been 82% accurate in the identification of individuals with osteoporosis. These results are in agreement with a study that trained a convolutional neural network for the prediction of DXA-equivalent BMD from CT scans and found them to be well correlated (66). The study used 183 paired CT and DXA scans to train the network and evaluated its performance in a test set of 45 subjects from the same medical center ($r=0.852$) and in a test set of 50 patient from a second medical center that did not contribute data to the training ($r=0.840$). A limitation of this approach is the use of aBMD based on DXA as a gold standard instead of experimental BMD metrics such as ash density, which is a measure of bone mineral mass divided by the bone volume.

Predictions of bone strength based on ML could reduce the effort during pre-processing and result in even faster predictions. General regression neural networks and support vector regression models have been recently explored for the prediction of vertebral strength (67). Bone strength predicted by CT-derived FE models of 80 vertebral bodies were used to train models based on regional grey scale related parameters of the vertebral body extracted from CT scans. 54 of the vertebrae were used for training and 26 were used to test the trained models. Both ML techniques were able to predict vertebral strength with high accuracy (mean absolute error < 8%), providing computationally efficient estimates of bone strength. Although these results provide a promising new direction for the prediction of vertebral strength, this study used manual preprocessing steps for the generation of input features for the ML models.

Machine learning approaches for fracture risk assessment

Current ML approaches using CT scans focus on the extraction of biomarkers, which provide a classification or regression target for training. Consequently, fracture risk assessment using ML tools has been focused on combining ML-based biomarkers for fracture risk assessment, rather than training algorithms directly for fracture outcome, which

includes a stronger probabilistic component. Using commercial ML tools (Zebra Medical vision) a recent study derived aBMD equivalent BMD, vBMD, and prevalent vertebral fractures from CT scans of 48,277 individuals that had undergone abdominal or chest CT scans (68). The follow-up period was five years. The ML tools were developed in separate studies on different data sets and were applied to the data of the study without additional training. The use of these CT-based biomarkers resulted in equivalent AUC values for hip fracture risk assessment compared to FRAX scores without aBMD (AUC of 0.76 and 0.751, respectively) and slightly higher AUC values for the prediction of major osteoporotic fracture (AUC of 0.709 and 0.69, respectively). A combination of FRAX scores with CT-based biomarkers resulted in an increase in the AUC, demonstrating that automated opportunistic screening based on CT scans can be effective at identifying individuals at high risk of fracture and can be combined with FRAX scores. A limitation of the methodology is that only 83.6% of CT datasets could be processed successfully. These findings are in agreement with a study that analyzed lumbar vertebral attenuation, muscle attenuation, and visceral-to-subcutaneous fat ratio extracted with ML from CT-scans and compared their ability to identify individual that will suffer major osteoporotic fractures to fracture risk assessment with FRAX (69). CT-based biomarkers for 9223 scans were extracted with a deep learning tool for automatic segmentation in combination with an automated feature extraction algorithm, both of which had been developed and validated in previous studies on separate cohorts. Biomarkers from CT for trabecular bone attenuation and muscle attenuation performed comparable to FRAX without BMD (AUC 0.71–0.72 compared to 0.72) for predicting individuals with major osteoporotic fractures. Interestingly, including BMD into FRAX reduced its ability to identify individuals that will suffer fragility fractures. For the assessment of hip fracture risk, CT-based muscle attenuation performed better than FRAX (AUC 0.75 and 0.71, respectively). The automated CT analysis in this study was highly robust, failing for less than 1.4% of scans.

Although current CT-based biomarkers perform only comparably to FRAX regarding fracture risk assessment, both studies highlight that fracture risk screening with biomarkers from CT can be used to complement FRAX scores and DXA-based aBMD to increase the overall screening coverage and enable improved prevention of fractures (Table 1).

Summary and future directions

CT-based fracture risk assessment using FE models and ML has been found to offer comparable or slightly improved performance to currently used tools, making them a valid option for opportunistic fracture risk screening (Table 1). The high level of automation enabled by ML makes ML methods attractive for clinical use and for the analysis of large datasets for research purposes. A small number of such tools has now been approved clinically, allowing their integration into clinical workflows.

Despite these exciting developments for the application of CT-based biomarkers for the assessment of fracture risk, improvements regarding the sensitivity and specificity have been small. A reason might be the focus of clinical studies on risk factors that purely relate to BMD or bone strength. Interestingly, CT attenuation of muscle on lumbar CT scans in Pickardt et al (69) was as good as CT attenuation of bone at predicting individuals at risk

of fracture. Combining information of bone strength with parameters of physical fitness, lifestyle, mechanical loading, and risk of falling, might therefore enable improved fracture risk assessment. A challenge related to the development of such models is the availability of data in clinical cohorts, which rarely contain comprehensive data related to physical activity, motion analysis, and falls risk. New clinical cohorts that include this type of information will be needed to develop mechanistic and data-driven methods for fracture risk assessment.

CT-based biomarkers extracted with ML are currently developed by comparing them to other image-based biomarkers such as aBMD or FE-derived bone strength. This approach limits their ability to predict future fractures to the predictive ability of the used gold standard. ML approaches that aim to overcome this limitation should be trained against experimentally measured bone strength or for fracture outcome directly. However, the generation of a large enough experimental dataset is time consuming and costly, but likely necessary step forward.

Phantomless calibration and automatic segmentation based on deep learning will likely become standard in processing of CT scans. Future approaches could combine both methods to automatically extract the tissue volumes for phantomless calibration. Diverse and unbiased datasets are essential for the development of such methods. Data collections such as the VerSe 2020 dataset for the spine (70) would be a great asset for other anatomical regions to develop and compare algorithms for segmentation, scan calibration, and biomarker extraction. Extending these methods with ML algorithms for automatic identification of pathological tissues and biases due to contrast agents could further improve the robustness of phantomless calibration.

The development of more "*in vivo* like" laboratory experiments and computational models will lead to a better understanding of fracture risk and fracture mechanisms. For the development of these models, comprehensive validation of the prediction of fracture forces and fracture location is needed. The use of vertebral column models that include the adjacent IVDs and posterior elements could allow for more realistic loading of the vertebra and consequently might result in more clinically relevant estimates of vertebral fracture risk and failure behavior. Current modeling approaches for multiple vertebrae and adjacent IVDs consider simplified material descriptions for the nucleus pulposus and annulus fibrosus, that neither consider the biphasic nature of the material nor material property changes due to disc degeneration (71). Combining such IVD models with vertebral bone models so simulate the behavior of the spine under immediate and creep loading might provide further insights into fracture mechanisms. Explicit modeling of the endplate tissue might result in more accurate prediction of vertebral bone loading and fracture initiation (72). Full field methods, such as digital volume correlation, allow for measurements of 3D deformation fields of the vertebra which can be used for the validation of the prediction of vertebral mechanisms with FE models (73,74). A combination of FE models with models for spinal kinematics and loading could then provide a comprehensive understanding of the fracture mechanics of the spine across tasks of daily living and subject-specific spine anatomies. Ultimately, ML models that are trained based on FE models or other mechanical models could be used to predict bone loading and bone strength, providing fast assessment of fracture risk across a large number of loading conditions.

Laboratory models, and the resulting fracture risk assessment tool, currently neglect the probabilistic nature of fracture events. Combining *ex vivo* models and FE models with real-life observational data regarding loading conditions and their frequency (75) could improve our understanding of individuals with high fracture risk. ML algorithms for automatic posture detection and tracking could help manage the large amounts of data that would be required to extract such information from video data. FE models could be used to replicate loading events that were recorded on video to test the ability of models to predict real-life fracture events. Moreover, the analysis of video data and wearable sensors will improve our understanding of the frequency of loading events, which could be used to combine mechanical models with probabilistic considerations for improved fracture risk assessment.

Conclusion

Robust and effective methods using finite element and machine learning are available for clinical assessment of fracture risk from CT scans. These methods have a comparable accuracy to tools currently used in clinical settings to identify individuals at elevated risk of fracture. The high degree of automation in these methods and the possibility to use scans that were taken for other indications will enable an increase in the number of people that undergo screening for fracture risk and consequently improve fracture prevention. Promising new research approaches for improved fracture risk assessment investigate the interaction between subject-specific bone loading and bone strength to develop mechanistic and data-based approaches that go beyond the assessment of bone fragility to better understand fracture mechanisms and fracture risk.

Acknowledgements

This work was supported by grant AR054620 from the National Institutes of Health

References

Papers of particular interest, published recently, have been highlighted as:

* Of importance

** Of major importance

1. Schuit SCE, Van der Klift M, Weel A, De Laet C, Burger H, Seeman E, et al. Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. Bone. 2004;34(1):195–202. [PubMed: 14751578]
2. Wainwright SA, Marshall LM, Ensrud KE, Cauley JA, Black DM, Hillier TA, et al. Hip fracture in women without osteoporosis. J Clin Endocrinol Metab. 2005;90(5):2787–93. [PubMed: 15728213]
3. Kanis JA, McCloskey EV, Johansson H, Oden A, Melton III LJ, Khaltaev N. A reference standard for the description of osteoporosis. Bone. 2008;42(3):467–75. [PubMed: 18180210]
4. Kanis JA, Cooper C, Rizzoli R, Reginster JY. Executive summary of the European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Calcif Tissue Int. 2019;104(3):235–8. [PubMed: 30796490]
5. Cody DD, Gross GJ, Hou FJ, Spencer HJ, Goldstein SA, Fyhrie DP. Femoral strength is better predicted by finite element models than QCT and DXA. J Biomech [Internet]. 1999;32(10):1013–20. Available from: http://ac.els-cdn.com/S0021929099000998/1-s2.0-S0021929099000998-main.pdf?_tid=33b5b734-

ae9a-11e5-9548-00000aacb35f&acdnat=1451441482_79796d7ca1594b74b3e1d6a2e4a2539b [PubMed: 10476839]

6. Crawford RP, Cann CE, Keaveny TM. Finite element models predict in vitro vertebral body compressive strength better than quantitative computed tomography. Bone. 2003;33(4):744–50. [PubMed: 14555280]
7. Kopperdahl DL, Aspelund T, Hoffmann PF, Sigurdsson S, Siggeirsdottir K, Harris TB, et al. Assessment of incident spine and hip fractures in women and men using finite element analysis of CT scans. J Bone Miner Res [Internet]. 2014;29(3):570–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23956027 [PubMed: 23956027]

8 **. Adams AL, Fischer H, Kopperdahl DL, Lee DC, Black DM, Bouxsein ML, et al. Osteoporosis and hip fracture risk from routine computed tomography scans: the fracture, osteoporosis, and CT utilization study (FOCUS). J Bone Miner Res. 2018;33(7):1291–301. [PubMed: 29665068] Largest cohort study for fracture risk assessment with finite element-derived bone strength.

9. Fleps I, Fung A, Guy P, Ferguson SJ, Helgason B, Cripton PA. Subject-specific ex vivo simulations for hip fracture risk assessment in sideways falls. Bone. 2019;125:36–45. [PubMed: 31071479]
10. Hayes WC, Myers ER, Robinovitch SN, Van Den Kroonenberg A, Courtney a C, McMahon T a. Etiology and prevention of age-related hip fractures. Bone. 1996;18(1 Suppl):77S–86S. [PubMed: 8717551]
11. Mokhtarzadeh H, Anderson DE, Allaire BT, Bouxsein ML. Patterns of Load-to-Strength Ratios Along the Spine in a Population-Based Cohort to Evaluate the Contribution of Spinal Loading to Vertebral Fractures. J Bone Miner Res. 2021;36(4):704–11. [PubMed: 33253414]
12. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. [PubMed: 26017442]
13. Chao Y-S, Sinclair A, Morrison A, Hafizi D, Pyke L. The Canadian Medical Imaging Inventory 2019–2020. 2021;

14 *. Fleps I, Pálsson H, Baker A, Enns-Bray W, Bahaloo H, Danner M, et al. Finite element derived femoral strength is a better predictor of hip fracture risk than aBMD in the AGES Reykjavik study cohort. Bone. 2022;154:116219. [PubMed: 34571206] Large cohort study that found improved fracture risk assessment compared to CT-based aBMD and compares different material implementations.

15. Qasim M, Farinella G, Zhang J, Li X, Yang L, Eastell R, et al. Patient-specific finite element estimated femur strength as a predictor of the risk of hip fracture: the effect of methodological determinants. Osteoporos Int. 2016;27(9):2815–22. [PubMed: 27108118]
16. Falcinelli C, Schileo E, Balistreri L, Baruffaldi F, Bordini B, Viceconti M, et al. Multiple loading conditions analysis can improve the association between finite element bone strength estimates and proximal femur fractures: A preliminary study in elderly women. Bone [Internet]. 2014;67:71–80. Available from: http://linkinghub.elsevier.com/retrieve/pii/S8756328214002476 [PubMed: 25014885]

17 **. Johannesdottir F, Allaire B, Kopperdahl DL, Keaveny TM, Sigurdsson S, Bredella MA, et al. Bone density and strength from thoracic and lumbar CT scans both predict incident vertebral fractures independently of fracture location. Osteoporos Int. 2021;32(2):261–9. [PubMed: 32748310] Large cohort study on assessement of vertebral fracture risk based on CT-derived finite element models.

18. Dieckmeyer M, Rayudu NM, Yeung LY, Löffler M, Sekuboyina A, Burian E, et al. Prediction of incident vertebral fractures in routine MDCT: Comparison of global texture features, 3D finite element parameters and volumetric BMD. Eur J Radiol. 2021;141:109827. [PubMed: 34225250]
19. Allaire BT, Lu D, Johannesdottir F, Kopperdahl D, Keaveny TM, Jarraya M, et al. Prediction of incident vertebral fracture using CT-based finite element analysis. Osteoporos Int. 2019;30(2):323–31. [PubMed: 30306225]
20. Services USD of H and H, Administration F and D. Reporting of Computational Modeling Studies in Medical Device Submissions [Internet]. 2016. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/reporting-computational-modeling-studies-medical-device-submissions

21. Center for Devices and Radiological. Assessing Credibility of Computational Modeling through Verification and Validation: Application to Medical Devices [Internet]. 2018. Available from: https://www.asme.org/codes-standards/find-codes-standards/v-v-40-assessing-credibility-computational-modeling-verification-validation-application-medical-devices

22. Lessmann N, Van Ginneken B, De Jong PA, Išgum I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. Med Image Anal. 2019;53:142–55. [PubMed: 30771712]

23. Müller D, Kramer F. MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning. BMC Med Imaging. 2021;21(1):1–11. [PubMed: 33407190]

24**. Sekuboyina A, Husseini ME, Bayat A, Löffler M, Liebl H, Li H, et al. VerSe: a vertebrae labelling and segmentation benchmark for multi-detector CT images. Med Image Anal. 2021;73:102166. [PubMed: 34340104] Comparison of state-of-the-art deep learning frameworks for image segmetation and labelling on a diverse dataset of the spine.

25. Kim JJ, Nam J, Jang IG. Fully automated segmentation of a hip joint using the patient-specific optimal thresholding and watershed algorithm. Comput Methods Programs Biomed. 2018;154:161–71. [PubMed: 29249340]

26. Väänänen SP, Grassi L, Venäläinen MS, Matikka H, Zheng Y, Jurvelin JS, et al. Automated segmentation of cortical and trabecular bone to generate finite element models for femoral bone mechanics. Med Eng Phys. 2019;70:19–28. [PubMed: 31280927]

27 *. Hemke R, Buckless CG, Tsao A, Wang B, Torriani M. Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment. Skeletal Radiol. 2020;49(3):387–95. [PubMed: 31396667] Study demonstrating muti-label segmentation with high accuracy for tissues in the pelvic region.

28. Balagopal A, Kazemifar S, Nguyen D, Lin M-H, Hannan R, Owrangi A, et al. Fully automated organ segmentation in male pelvic CT images. Phys Med Biol. 2018;63(24):245015. [PubMed: 30523973]

29. Bjornsson PA, Helgason B, Palsson H, Sigurdsson S, Gudnason V, Ellingsen LM. Automated femur segmentation from computed tomography images using a deep neural network. In: Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging. International Society for Optics and Photonics; 2021. p. 116001E.

30. Pauchard Y, Fitze T, Browarnik D, Eskandari A, Pauchard I, Enns-Bray W, et al. Interactive graph-cut segmentation for fast creation of finite element models from clinical ct data for hip fracture prediction. Comput Methods Biomech Biomed Engin. 2016;19(16):1693–703. [PubMed: 27161828]

31 *. Winsor C, Li X, Qasim M, Henak CR, Pickhardt PJ, Ploeg H, et al. Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses. Bone. 2021;143:115759. [PubMed: 33212317] Study investigating the influence of different tissue combinations for phantomless calibration applied to hip BMD and finite element derived bone strength.

32*. Lee DC, Hoffmann PF, Kopperdahl DL, Keaveny TM. Phantomless calibration of CT scans for measurement of BMD and bone strength—inter-operator reanalysis precision. Bone. 2017;103:325–33. [PubMed: 28778598] Study investating the inter-operator variability of phantomless calibration at the spine and hip

33 *. Bartenschlager S, Dankerl P, Chaudry O, Uder M, Engelke K. BMD accuracy errors specific to phantomless calibration of CT scans of the lumbar spine. Bone. 2022;157:116304. [PubMed: 34973497] Study analysing the expected influence of tissue variability on BMD accuracty with phantomless calibration.

34 *. Michalski AS, Besler BA, Michalak GJ, Boyd SK. CT-based internal density calibration for opportunistic skeletal assessment using abdominal CT scans. Med Eng Phys. 2020;78:55–63. [PubMed: 32059948] Study comparing in-line, asynchronous and phantomless calibration for different applied to hip BMD and finite element derived bone strength.

35 *. Prado M, Khosla S, Chaput C, Giambini H. Opportunistic application of phantom-less calibration methods for fracture risk prediction using QCT/FEA. Eur Radiol. 2021;31(12):9428–35. [PubMed: 34047849] Study investigating the influence of different tissue combinations for phantomless calibration applied to vertebral BMD and finite element derived bone strength.

36. Duchemin L, Mitton D, Jolivet E, Bousson V, Laredo JD, Skalli W. An anatomical subject-specific FE-model for hip fracture load prediction. Comput Methods Biomech Biomed Engin. 2008;11(2):105–11. [PubMed: 17891675]

37. Bessho M, Ohnishi I, Matsuyama J, Matsumoto T, Imai K, Nakamura K. Prediction of strength and strain of the proximal femur by a CT-based finite element method. J Biomech [Internet]. 2006/10/13. 2007;40(8):1745–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17034798 [PubMed: 17034798]

38. Ariza O, Gilchrist S, Widmer RP, Guy P, Ferguson SJ, Cripton PA, et al. Comparison of explicit finite element and mechanical simulation of the proximal femur during dynamic drop-tower testing. J Biomech [Internet]. 2015;48(2):224–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25527888 [PubMed: 25527888]

39. Dragomir-Daescu D, Op Den Buijs J, McEligot S, Dai Y, Entwistle RC, Salas C, et al. Robust QCT/FEA models of proximal femur stiffness and fracture load during a sideways fall on the hip. Ann Biomed Eng. 2011;39(2):742–55. [PubMed: 21052839]

40. Melenk JM, Babuška I. The partition of unity finite element method: basic theory and applications. Comput Methods Appl Mech Eng. 1996;139(1–4):289–314.

41. Miehe C, Welschinger F, Hofacker M. Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field FE implementations. Int J Numer Methods Eng. 2010;83(10):1273–311.

42 *. Shen R, Waisman H, Yosibash Z, Dahan G. A novel phase field method for modeling the fracture of long bones. Int j numer method biomed eng. 2019;35(8):e3211. [PubMed: 31062516] First implementation of fracture predictions with the phase field methods in CT-based heterogeneous whole bone simulations.

43 *. Gustafsson A, Tognini M, Bengtsson F, Gasser TC, Isaksson H, Grassi L. Subject-specific FE models of the human femur predict fracture path and bone strength under single-leg-stance loading. J Mech Behav Biomed Mater. 2021;113:104118. [PubMed: 33125949] Implementation of XFEM into CT-based heteogeneous whole bone simulations with promising results for stance loading.

44. Giambini H, Qin X, Dragomir-Daescu D, An K-N, Nassr A. Specimen-specific vertebral fracture modeling: a feasibility study using the extended finite element method. Med Biol Eng Comput. 2016;54(4):583–93. [PubMed: 26239163]

45. Maghami E, Josephson TO, Moore JP, Rezaee T, Freeman TA, Karim L, et al. Fracture behavior of human cortical bone: Role of advanced glycation end-products and microstructural features. J Biomech. 2021;125:110600. [PubMed: 34246065]

46. Wu C, Fang J, Zhang Z, Entezari A, Sun G, Swain MV, et al. Fracture modeling of brittle biomaterials by the phase-field method. Eng Fract Mech. 2020;224:106752.

47. Navidtehrani Y, Betegón C, Martínez-Pañeda E. A simple and robust Abaqus implementation of the phase field fracture method. Appl Eng Sci. 2021;6:100050.

48 *. Molnár G, Gravouil A. 2D and 3D Abaqus implementation of a robust staggered phase-field solution for modeling brittle fracture. Finite Elem Anal Des. 2017;130:27–38.Study describing an open source Abaqus implementation of the phase field methods in a user defined element formuation.

49. Yosibash Z, Trabelsi N, Hellmich C. Subject-Specific p-FE Analysis of the Proximal Femur Utilizing Micromechanics-Based Material Properties. Int J Multiscale Comput Eng. 2008;6(5):483–98.

50. Treece GM, Gee AH. Independent measurement of femoral cortical thickness and cortical bone density using clinical CT. Med Image Anal. 2015;20(1):249–64. [PubMed: 25541355]

51. Chandran V, Maquer G, Gerig T, Zysset P, Reyes M. Supervised learning for bone shape and cortical thickness estimation from CT images for finite element analysis. Med Image Anal. 2019;52:42–55. [PubMed: 30471462]

52. Fleps I, Vuille M, Melnyk A, Ferguson SJ, Guy P, Helgason B, et al. A novel sideways fall simulator to study hip fractures ex vivo. PLoS One. 2018;13(7):e0201096. [PubMed: 30040858]

53 **. Fleps I, Guy P, Ferguson SJ, Cripton PA, Helgason B. Explicit finite element models accurately predict subject-specific and velocity-dependent kinetics of sideways fall impact. J Bone Miner

Res. 2019;34(10):1837–50. [PubMed: 31163090] Validation study for finite element models that combine loading due to a fall with fracture predictions.

54. Fleps I, Enns-Bray WS, Guy P, Ferguson SJ, Cripton PA, Helgason B. On the internal reaction forces, energy absorption, and fracture in the hip during simulated sideways fall impact. PLoS One. 2018;13(8).

55 **. Enns-Bray WS, Bahaloo H, Fleps I, Pauchard Y, Taghizadeh E, Sigurdsson S, et al. Biofidelic finite element models for accurately classifying hip fracture in a retrospective clinical study of elderly women from the AGES Reykjavik cohort. Bone. 2019;120.Study that applies an impact model of the body for a sideways fall to a clinical cohort showing that fracture risk assessment for fallers could potentially be improved by incluing loading estimates into the finite element methodology.

56. Galliker ES, Laing AC, Ferguson SJ, Helgason B, Fleps I. The influence of fall direction and hip protector on fracture risk: FE model predictions driven by experimental data. Ann Biomed Eng. 2022;1–13.

57. Fung A, Fleps I, Cripton PA, Guy P, Ferguson SJ, Helgason B. Prophylactic augmentation implants in the proximal femur for hip fracture prevention: An in silico investigation of simulated sideways fall impacts. J Mech Behav Biomed Mater. 2022;126:104957. [PubMed: 34861519]

58. Bhattacharya P, Altai Z, Qasim M, Viceconti M. A multiscale model to predict current absolute risk of femoral fracture in a postmenopausal population. Biomech Model Mechanobiol. 2018;1–18.

59 *. Anitha DP, Baum T, Kirschke JS, Subburaj K. Effect of the intervertebral disc on vertebral bone strength prediction: A finite-element study. Spine J. 2020;20(4):665–71. [PubMed: 31841703] Study that demonstrates improved vertebral strength prediction when considering loading through the intervertebral disc compared to uniform displacement loading.

60 *. Hussein AI, Louzeiro DT, Unnikrishnan GU, Morgan EF. Differences in trabecular microarchitecture and simplified boundary conditions limit the accuracy of quantitative computed tomography-based finite element models of vertebral failure. J Biomech Eng. 2018;140(2).The study highlight the imprtance of loading assumptions for the prediction of vertebral deformations during fracture with finite element models.

61. Carberry GA, Pooler BD, Binkley N, Lauder TB, Bruce RJ, Pickhardt PJ. Unreported vertebral body compression fractures at abdominal multidetector CT. Radiology. 2013;268(1):120–6. [PubMed: 23449956]

62. Genant HK, Wu CY, Van Kuijk C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. J bone Miner Res. 1993;8(9):1137–48. [PubMed: 8237484]

63 *. Tomita N, Cheung YY, Hassanpour S. Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. Comput Biol Med. 2018 Jul 1;98:8–15. [PubMed: 29758455] Automated vertebral fracture detection from CT scans with implications for its usefulness in clinical practice.

64 *. Kolanu N, Silverstone EJ, Ho BH, Pham H, Hansen A, Pauley E, et al. Clinical utility of computer-aided diagnosis of vertebral fractures from computed tomography images. J Bone Miner Res. 2020;35(12):2307–12. [PubMed: 32749735] Development and validation of the automated vertebral fracture detection model used in the commercial tool from Zebra Medical vision.

65 *. Krishnaraj A, Barrett S, Bregman-Amitai O, Cohen-Sfady M, Bar A, Chettrit D, et al. Simulating dual-energy X-ray absorptiometry in CT using deep-learning segmentation cascade. J Am Coll Radiol. 2019;16(10):1473–9. [PubMed: 30982683] Development and validation of the machine learning model for predicting DXA equivalent aBMD used in the commercial tool from Zebra Medical vision.

66 *. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network. Eur Radiol. 2020;30(6):3549–57. [PubMed: 32060712] Development and validation of the machine learning model for predicting DXA equivalent aBMD.

67. Zhang M, Gong H, Zhang K. Prediction of lumbar vertebral strength of elderly men based on quantitative computed tomography images using machine learning. Osteoporos Int. 2019;30(11):2271–82. [PubMed: 31401661]

68 **. Dagan N, Elnekave E, Barda N, Bregman-Amitai O, Bar A, Orlovsky M, et al. Automated opportunistic osteoporotic fracture risk assessment using computed tomography scans to aid in FRAX underutilization. Nat Med [Internet]. 2020;26(1):77–82. Available from: 10.1038/s41591-019-0720-z [PubMed: 31932801] Application of machine learning based biomarkers (prevalent vertebreal fracture, DXA equivalent aBMD, vBMD) for retrospective fracture risk assessment in a large clincal cohort.

69 **. Pickhardt PJ, Graffy PM, Zea R, Lee SJ, Liu J, Sandfort V, et al. Automated abdominal CT imaging biomarkers for opportunistic prediction of future major osteoporotic fractures in asymptomatic adults. Radiology. 2020;297(1):64–72. [PubMed: 32780005] Application of machine learning based biomarkers (CT attenuation for bone, muscle, and viseral to subcutaneous fat ratio) for retrospective fracture risk assessment in a large clincal cohort.

70. Liebl H, Schinz D, Sekuboyina A, Malagutti L, Löffler MT, Bayat A, et al. A computed tomography vertebral segmentation dataset with anatomical variations and multi-vendor scanner data. Sci data. 2021;8(1):1–7. [PubMed: 33414438]

71. Newman HR, DeLucca JF, Peloquin JM, Vresilovic EJ, Elliott DM. Multiaxial validation of a finite element model of the intervertebral disc with multigenerational fibers to establish residual strain. JOR Spine. 2021;4(2):1–16.

72. Wu Y, Loaiza J, Banerji R, Blouin O, Morgan E. Structure-function relationships of the human vertebral endplate. JOR spine. 2021;4(3):e1170. [PubMed: 34611592]

73. Jackman TM, DelMonaco AM, Morgan EF. Accuracy of finite element analyses of CT scans in predictions of vertebral failure patterns under axial compression and anterior flexion. J Biomech. 2016;49(2):267–75. [PubMed: 26792288]

74. Costa MC, Tozzi G, Cristofolini L, Danesi V, Viceconti M, Dall'Ara E. Micro finite element models of the vertebral body: validation of local displacement predictions. PLoS One. 2017;12(7):e0180151. [PubMed: 28700618]

75 *. Yang Y, Komisar V, Shishov N, Lo B, Korall AMB, Feldman F, et al. The Effect of Fall Biomechanics on Risk for Hip Fracture in Older Adults: A Cohort Study of Video-Captured Falls in Long-Term Care. J Bone Miner Res. 2020;35(10):1914–22. [PubMed: 32402136] Study reporting on the circumstances of falls in the elderly and how they relate to hip fractures.

Table 1:

Summary of fracture risk assessment performance in studies using finite element (FE) model or machine learning (ML)

Study	Fracture location	Method	Sex	AUC	Sensitivity	Specificity	N	Follow up [years]
Adams, 2018 [8]	hip	FE	Women Men	0.73 0.75	0.66 0.56	0.66 0.76	850 465	5
		vBMD		0.72 0.71	0.52 0.45	0.77 0.82		
		DXA-aBMD		0.72 0.73	0.56 0.43	0.77 0.83		
Enns-Bray, 2019 [51]	hip	FE + loading	Women	0.73	-	-	254	5
		CT-aBMD		0.70	-	-		
Fleps, 2022 [14]	hip	FE	Women Men	0.74 0.78	-	-	362 239	5
		CT-aBMD		0.69 0.72	-	-		
Bhattacharya, 2018 [54]	hip	FE + loading	Women	0.85	0.78	0.82	98	Post-fracture imaging
		FE		0.82	-	-		
		aBMD		0.75	-	-		
Allaire,2019 [19]	Vertebra	FE	Pooled	0.80	0.46	0.87	88	5
		vBMD		0.82	0.27	0.81		
Johannesdottir, 2021 [17]	Vertebra	FE	Pooled	-	0.64	0.60	401	6
		vBMD		-	0.79	0.49		
Dieckmeyer, 2021 [18]	Vertebra	FE	Pooled	0.71	-	-	32	1
		Texture		0.75	-	-		
Dagan, 2020 [65]	Major osteoporotic	ML tool	Pooled	0.71	0.67	0.65	48’227	5
	Hip			0.76	0.93	0.37		
	Major osteoporotic	FRAX noBMD		0.69	0.64	0.64		
	Hip			0.76	0.91	0.37		
Pickardt, 2020 [66]	Major osteoporotic	ML tool	Pooled	0.68	0.22	0.90	9’223	10
	Hip			0.68	0.34	0.90		
	Major osteoporotic	FRAX noBMD		0.65	0.04	0.99		
	Hip			0.62	0.16	0.95		