



In Silico Clinical Trial for Osteoporosis Treatments to Prevent Hip Fractures: Simulation of the Placebo Arm

Giacomo Savelli¹ · Sara Oliviero¹ · Antonino A. La Mattina² · Marco Viceconti^{1,2}

Received: 17 May 2024 / Accepted: 14 October 2024 / Published online: 22 November 2024
© The Author(s) 2024

Abstract

Osteoporosis represents a major healthcare concern. The development of novel treatments presents challenges due to the limited cost-effectiveness of clinical trials and ethical concerns associated with placebo-controlled trials. Computational models for the design and assessment of biomedical products (*In Silico* Trials) are emerging as a promising alternative. In this study, a novel *In Silico* Trial technology (*BoneStrength*) was applied to replicate the placebo arms of two concluded clinical trials and its accuracy in predicting hip fracture incidence was evaluated. Two virtual cohorts (N = 1238 and 1226, respectively) were generated by sampling a statistical anatomy atlas based on CT scans of proximal femurs. Baseline characteristics were equivalent to those reported for the clinical cohorts. Fall events were sampled from a Poisson distribution. A multiscale stochastic model was implemented to estimate the impact force associated to each fall. Finite Element models were used to predict femur strength. Fracture incidence in 3 years follow-up was computed with a Markov chain approach; a patient was considered fractured if the impact force associated with a fall exceeded femur strength. Ten realizations of the stochastic process were run to reach convergence. Each realization required approximately 2500 FE simulations, solved using High-Performance Computing infrastructures. Predicted number of fractures was 12 ± 2 and 18 ± 4 for the two cohorts, respectively. The predicted incidence range consistently included the reported clinical data, although on average fracture incidence was overestimated. These findings highlight the potential of *BoneStrength* for future applications in drug development and assessment.

Keywords Osteoporosis · Hip fracture · *In silico* clinical trial · Computational biomechanics · Retrospective validation · Multiscale model

Introduction

Osteoporosis (OP) represents a significant health concern worldwide, particularly among the ageing population. As life expectancy increases, especially in developed countries, the social and economic burden of osteoporosis is becoming more pronounced [1]. In osteoporosis (OP), the progressive reduction in bone mineral density and architectural

deterioration leads to an increased risk of fragility fractures. The majority of hip fractures (up to 95%) happen as consequence of a fall [2] and are particularly severe, as they are associated with a 20% mortality rate within a year and high management costs. Therefore, preventing hip fractures and developing effective osteoporosis treatments is of paramount importance [1, 3–5].

Existing interventions to prevent hip fractures primarily include lifestyle interventions and pharmacological treatments. Lifestyle changes such as weight-bearing exercise are recommended to strengthen bones and muscles [6], while fall prevention strategies such as balance training and home safety assessments may reduce the risk of falls [7]. Calcium and vitamin D supplements are commonly prescribed to support bone health [8]. Pharmacological treatments include bisphosphonates, such as alendronate and risedronate [9–11], parathyroid hormone analogues, such as teriparatide [12], RANKL inhibitor (denosumab) [13], sclerostin

Associate Editor Sean S. Kohles oversaw the review of this article.

✉ Giacomo Savelli
giacomo.savelli4@unibo.it

¹ Department of Industrial Engineering, Alma Mater Studiorum - University of Bologna, Bologna, Italy

² Medical Technology Lab, IRCSS - Istituto Ortopedico Rizzoli, Via di Barbiano 1/10, 40136 Bologna, Italy

inhibitors such as romosozumab [14], calcitonin analogues [15], selective oestrogen receptor modulators (SERMs) such as raloxifene [16, 17].

Currently, randomised clinical trials (RCTs) are the gold standard for evaluating the efficacy and safety of new treatments [18–20]. Placebo-controlled clinical trials are methodologically preferable, due to the presence of an internal control and the relatively smaller sample size required to demonstrate treatment effectiveness [21]. On the other hand, placebo-controlled clinical trials present ethical concerns, when effective treatments are available, especially in vulnerable elderly individuals at a high risk of fracture, who are deprived of potentially beneficial interventions [21]. Possible alternatives include the recruitment of individuals at low risk or the use of active control in place of the placebo group. For both cases, it has been reported that the required sample sizes would be prohibitive, due to the small difference in fracture incidence between the trial arms [21]. Therefore, there is a pressing need to develop alternative approaches that can improve drug development by complementing and/or replacing traditional clinical trials. As an example, among these the SABRE (Study to Advance BMD as a Regulatory Endpoint) project aims to reduce the length and cost of clinical trials by qualifying changes in bone mineral density (BMD) as a surrogate non-invasive endpoint for fractures in osteoporosis drug trials [22].

Computational tools for the design and assessment of biomedical products, i.e. *In Silico* Clinical Trials (ISCTs), are also emerging as promising alternatives [23]. ISCTs leverage computational models to simulate an intervention of interest in a virtual cohort, representative of the population of interest, enabling the testing of various interventions and hypotheses. Potential advantages associated with the use of ISCTs include the opportunity to test multiple interventions in the same cohort, to improve clinical trial design, to test treatments without subjecting patients to potential risks associated with clinical trials, to increase sample size (*In Silico*-augmented clinical trials), and to reduce the required time and costs [23–25].

Subject-specific Finite Element (FE) models of the human femur based on Computed Tomography (CT) scans have been extensively validated [26–28], and excellent accuracy in predicting failure load in a wide range of loading conditions has been reported (Standard Error of Estimate, SEE = 15% for failure load, SEE = 7% for strains, $R^2 = 0.89$ between FE predictions and experimental measurements) [29–31]. Additionally, in a recent study [32] a multiscale model of side fall has been developed and applied to predict subject-specific current absolute risk of fracture (ARF0). Various stochastic parameters are used to simulate a wide range of falling scenarios, and fracture probability is obtained by comparing the impact force associated with multiple fall events with an FE-based failure load map. This

probability, combined with the fall rate, is used to calculate ARF0. In a retrospective cohort, ARF0 had superior stratification accuracy compared to areal bone mineral density (aBMD) to stratify fractured vs. non-fractured patients [32]. ARF0 could potentially be used as a surrogate endpoint for osteoporosis phase I or phase II clinical trials in place of aBMD, after the required prospective clinical validation and qualification [33].

Currently, in phase III clinical trials regulatory bodies require fracture incidence as primary endpoint to evaluate and demonstrate the efficacy of OP treatments, as a more clinically relevant outcome [34, 35]. In our group, we are developing an ISCT technology (*BoneStrength*), based on a Markov Chain algorithm, to simulate the equivalent of a phase III clinical trial *In Silico*, able to predict fracture incidence as the final endpoint. La Mattina et al., (2023) developed a method to generate a virtual cohort of proximal femurs from a statistical atlas and demonstrated that the virtual cohort was representative of the physical cohort used to build the atlas, in terms of baseline aBMD and fracture risk (ARF0) [36]. In this study, the method was used to generate two virtual cohorts and replicate the placebo arms of two concluded clinical trials using *BoneStrength*. Fall events were stochastically sampled from a Poisson distribution, the multiscale fall model was used to estimate the impact load, and FE models to assess femur strength. Predicted fracture incidence was compared with the clinical data. This retrospective validation represents a mandatory step in the development of an ISCT technology to assess the efficacy of OP treatments, using fracture incidence as the predicted endpoint.

The aim of this study was to evaluate the accuracy of *BoneStrength* in predicting fracture incidence in two virtual populations of postmenopausal women, whose baseline characteristics replicated those reported for the placebo arms of two concluded clinical trials, by comparing the model predictions against the clinical outcome.

Materials and Methods

In this study, we replicated the placebo arms of two clinical trials reported in the literature using the *In Silico* trial methodology developed in our group (*BoneStrength*). We generated virtual cohorts of Finite Element models of proximal femurs, by matching the baseline characteristics of the populations recruited in the studies. Subsequently, fracture incidence was predicted for both cohorts using a Markov Chain process. A Poisson distribution was used to model the probability of fall events, while a multiscale model was applied to simulate side falls and predict impact load. Fractures were predicted by comparing impact load with failure load associated with each fall event, obtained with

subject-specific FE models. Three years of follow-up were simulated for both cohorts to estimate hip fracture incidence and compare the results with the clinical data.

Clinical Data for Retrospective Validation

Two large concluded clinical trials from the literature were selected: the Long-term Intervention on Fractures with Tibolone (LIFT) and Fracture Reduction Evaluation of Denosumab in Osteoporosis Every 6 Months (FREEDOM) [37, 38]. Selection criteria were based on the population of interest (postmenopausal women, T-score < -1, average age > 55 years), the presence of a placebo arm, and reported hip fracture incidence. Details of the baseline characteristics of the cohorts recruited in the selected trials and fracture incidence are reported in Table 1.

We generated two virtual cohorts of proximal femurs to match the baseline characteristics of the placebo arms in terms of Total Hip aBMD.

Virtual Cohort

The procedure applied to generate the virtual cohorts has been described in detail in La Mattina et al., 2023 [36] and is shortly summarised here. Briefly, a dataset of 94 CT-based FE-models of human femurs with high quality 10-node tetrahedral meshes (ANSYS ICEM CFD 2019R2, ANSYS Inc., USA) was used to generate a statistical anatomy atlas through principal component analysis (PCA). The original cohort was recruited in the context of a retrospective study, described in detail in Yang et al., 2014 [39]. The statistical atlas included both geometry (coordinates of nodes and reference system points) and density (density-based element-specific Young's moduli) features. By sampling the inverse cumulative functions associated with each principal component, 93 stochastic variables following the same distributions were obtained, which were combined to generate both geometry and material property distribution for virtual femurs. To estimate the knee rotation centre position, proximal femur models were fitted with an anatomy atlas [40].

Table 1 Baseline characteristics of the populations recruited in the reference studies (placebo groups)

	Cummings et al., 2008 (LIFT)	Cummings et al., 2009 (FREEDOM)
Total Hip aBMD [g/cm ²]	0.722 ± 0.096	0.709 ± 0.099
Age [yrs]	68.2 ± 5.2	72.3 ± 5.2
BMI [kg/m ²]	25.7 ± 3.4	26.0 ± 4.2
Number of participants	2257	3906
Follow-up duration [yrs]	3	3
Number of hip fractures	14	43

The procedure was implemented in MATLAB (MATLAB 2022a, MathWorks Inc., USA).

We generated 1270 and 1249 synthetic femurs to match the cohorts recruited in the LIFT and the FREEDOM trials, respectively. Virtual patients were progressively sampled from the distributions of principal components and total hip aBMD was calculated for each. Subsequently, patients were included in the virtual cohort or discarded until the population with the target numerosity and statistical distribution of total hip aBMD was obtained. From the two virtual cohorts, 32 and 24 models were respectively excluded due to convergence issues or excessive element distortion. Subject-specific height and weight were determined for each virtual patient based on their correlation with DXA data obtained using a linear regression [36], with a Pearson's correlation coefficient of 0.4 was between height, weight, and femoral neck aBMD. The number of virtual patients was chosen to reproduce the typical sample size of phase III clinical trials, in the order of more than 1000 patients [18, 21, 41] while reducing the computational resources required for simulation.

Fall Model

The impact force associated with each fall was determined with a multiscale model, as described in Bhattacharya et al. 2019 [32]. Briefly, the model estimates the force transmitted to the femur during a side fall by considering both subject-specific and stochastic parameters, which are sampled from probability distributions appropriately defined for the population of interest. Three models are included: a mechanistic body-floor impact model to compute peak impact force; a ground-skeleton force transfer model; a subject-specific FE model. The ground-skeleton force transfer model is applied to attenuate the impact force, to account for the attenuation provided by the presence of soft tissues around the hip (subject-specific parameter), different flooring materials (e.g. carpets), and wearable devices (such as hip protectors) [32]. Subject-specific soft tissue thickness was derived from individual BMI using a linear relationship tailored to the reference population [42]. The first two models were implemented using standard Python libraries NumPy and SciPy. Lastly, the FE model (paragraph 2.4) is used to calculate the femur strength in each loading condition.

In this work, we have assumed that muscle contraction and subsequent stiffening of the dynamic system had a negligible effect on the impact force transferred from the ground to the bone, as it is very unlikely that elderly individuals would be able to significantly contract their muscles before the impact. Falls, which occur unexpectedly and unintentionally during daily activities, result in impacts within a time frame of 300–750 milliseconds [43–46]. Even though subjects can intervene in the dynamics of falls, it has been

estimated that the initial 300–450 milliseconds are dedicated to recognizing and attempting to prevent the fall [43, 44, 46]. Additionally, in elderly individuals, there is a general increase in reaction time, especially in choice reaction tasks that involve selecting an appropriate response strategy [47–53]. Comparative studies have estimated that this difference is in the order of hundreds of milliseconds, particularly in experimental setups that involve secondary mental tasks and considering the delay in EMG activation of the lower limbs [52, 54, 55]. Therefore, in this study it was assumed that no or negligible muscle contraction could occur in the time frame of the fall event.

The occurrence of falls in the community-dwelling elderly population was modelled as the occurrence of discrete events following a Poisson distribution, which assumes:

- Equiprobability of the occurrence of an event for all sub-intervals.
- Possibility for an event to occur multiple times within the same sub-interval.
- Independence of events occurring in different sub-intervals.

Under these assumptions, the probability that an event with an average probability of occurrence λ will occur n times is expressed by the equation:

$$P_{\lambda}(n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

We determined the number of falls per year for each patient by sampling the distribution (MATLAB 2022a, MathWorks Inc., USA). The value of $\lambda = 0.65$ falls/year was based on data from the literature for the population of interest, in terms of falls per person per year [56], obtained from concluded clinical trials where fall-assessment techniques of proven accuracy were applied [57–59]. Implicitly, it was assumed the strong hypothesis of non-variability of fall risk within the population, among different time intervals, and within the intervals themselves.

FE Models

Subject-specific FE models were used to predict the biomechanical failure load associated with each fall event. Each FE model had a quadratic tetrahedral mesh with maximum element size of 3 mm and heterogeneous isotropic linear elastic material properties [60]. The impact force was applied at the centre of the femoral head. For each fall, the impact direction between the greater trochanter and the ground was stochastic and defined by two angles, namely the intra-extra rotation (α) and abduction-adduction (β) angles of the femur. The intervals $\alpha = [-30^\circ, 30^\circ]$ and $\beta = [0^\circ, 30^\circ]$ were discretized

with a step of 1° , resulting in 1891 possible force directions. It was assumed:

- Equiprobability among the possible values of α and β .
- Reciprocal independence of α and β values.
- Independence of α and β from time, subject-specific factors, and the number of falls.

For each fall, a FE simulation was run by applying a 1000 N load in the centre of the femoral head with the associated orientation (ANSYS Mechanical APDL 2022a, ANSYS Inc., USA) [31]. This approach was used to predict direction-specific failure load, as in previous studies it has been reported that femur strength is strongly dependent on the impact angle [31]. Simulating multiple side fall directions also had a significantly better stratification accuracy compared with single-load strength in a retrospective cohort [31]. The distal diaphysis was coupled with a node in the centre of the knee joint with multi-point constraints, therefore the femur was allowed to rotate around an axis perpendicular to load direction through the knee centre. A rigid plane was positioned at the furthest point of the greater trochanter along the load direction and a non-linear frictionless contact between the femur surface and the rigid plane was applied [31].

Failure load was determined by calculating the load intensity required to generate a principal tensile strain of 0.73% or a principal compressive strain of 1.04% [61] (whichever occurred first) on the surface of the femur. Nodal principal strains were averaged over a sphere with a radius of 3 mm (NumPy and SciPy libraries, Python) [62]. This fragile fracture criterion has been previously validated, showing excellent prediction accuracy compared to experimental results obtained with cadaveric bones [61].

Fracture Incidence

The Markov Chain algorithm used to predict fracture incidence in the virtual cohort is reported in Figure 1. For each patient, fall events were sampled from the Poisson distribution, as described in paragraph 2.3. In the event of a fall, direction and impact load were stochastically sampled from the corresponding distributions, and the FE simulation was run to obtain failure load. Virtual patients were classified as fractured when the impact force exceeded the failure load in the impact direction and excluded from subsequent simulations. This, besides saving computational time, simulates the role of hip fracture as a clinical endpoint. Three years of follow-up were simulated with a time step of 1 year by iteratively sampling fall events, comparing the impact force with the corresponding failure load, and excluding fractured patients. The computed incidence of hip fractures over three years was compared with the clinical data.

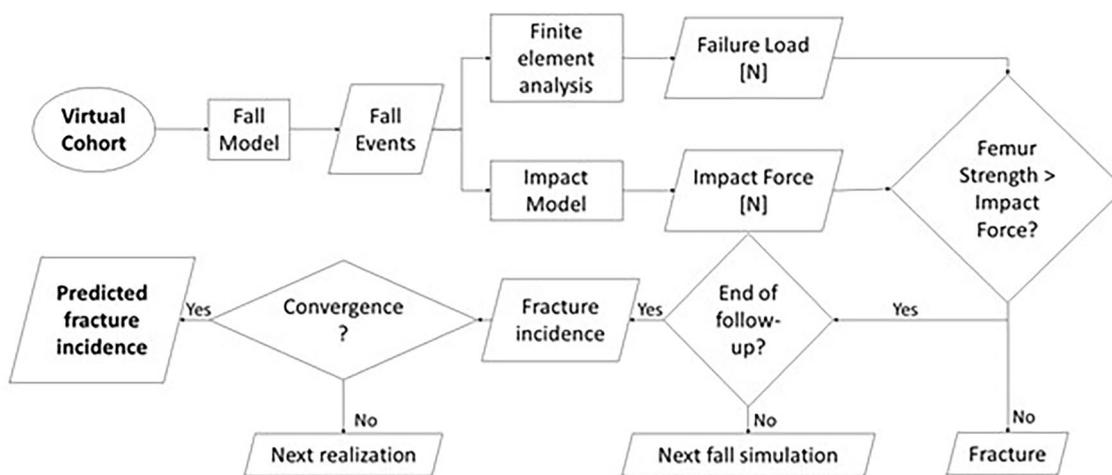


Fig. 1 Markov-Chain algorithm architecture

For each virtual cohort, 10 realizations of the Markov Chain process were run, to reach convergence of the estimated fracture incidence. Each realization required approximately 2500 FE simulations (4 cores and 28 GB of RAM per simulation).

This pipeline is currently running on High-Performance Computing Infrastructure Ares (Cyfronet, Poland) through the Model Execution Environment (MEE, <https://mee.cyfronet.pl/>).

Statistical Analysis

A two-sample t-test (ttest2 function, MATLAB 2022a, MathWorks Inc., USA) was used to verify that the distributions of Total Hip aBMD and BMI (which are strongly correlated with hip fracture risk [63]) in the virtual and reference cohorts were not significantly different.

Convergence of the predicted fracture incidence obtained with the Markov Chain algorithm was assessed by evaluating the mean number of fractures as a function of the number of realizations. Bootstrapping was employed to determine the relative difference in the mean obtained by including an additional realization. A relative difference of 1% was used as threshold to define convergence (NumPy and SciPy libraries, Python).

Results

The distributions of Total Hip aBMD in the two virtual cohorts were not significantly different compared to the cohorts recruited in the LIFT study (t-test, $p=0.31$) and the FREEDOM study ($p=0.38$). BMI distributions were significantly different between virtual cohorts and clinical

data for both studies ($p<0.001$). A comparison between the distributions of Total Hip aBMD and BMI in the two virtual cohorts and the reference populations is reported in Figure 2.

Fracture incidence predicted by the model converged after 10 realizations (relative difference of cumulative mean = 0.36% for LIFT, and 0.32% for FREEDOM).

For the virtual cohort based on the LIFT study ($N=1238$ virtual patients), 12 ± 2 fractures in three years were predicted. For the virtual cohort based on the FREEDOM study ($N=1225$) 18 ± 4 hip fractures were obtained over a simulated follow-up of three years. Minimum and maximum number of fractures predicted across all realisations are reported in Table 2.

Discussion

In this study, the *In Silico* trial methodology *BoneStrength* was applied to predict the hip fractures in two simulated cohorts of postmenopausal women, designed to replicate the baseline characteristics of participants recruited in the placebo arms of two concluded clinical trials. Predicted fracture incidence was compared to the clinical data reported in the studies to assess the model predictive ability.

The range of hip fracture incidence predicted by the model was consistent with the clinical data reported in both studies, highlighting the approach was able to adequately model the phenomenon of interest; nevertheless, on average the number of fractures was consistently overestimated compared with the clinical data. This could be linked to the simplified approach applied to model the frequency and dynamics of falls, as discussed in the following paragraphs. The occurrence of hip fractures in a given cohort depends on several factors such as the frequency of fall events, which

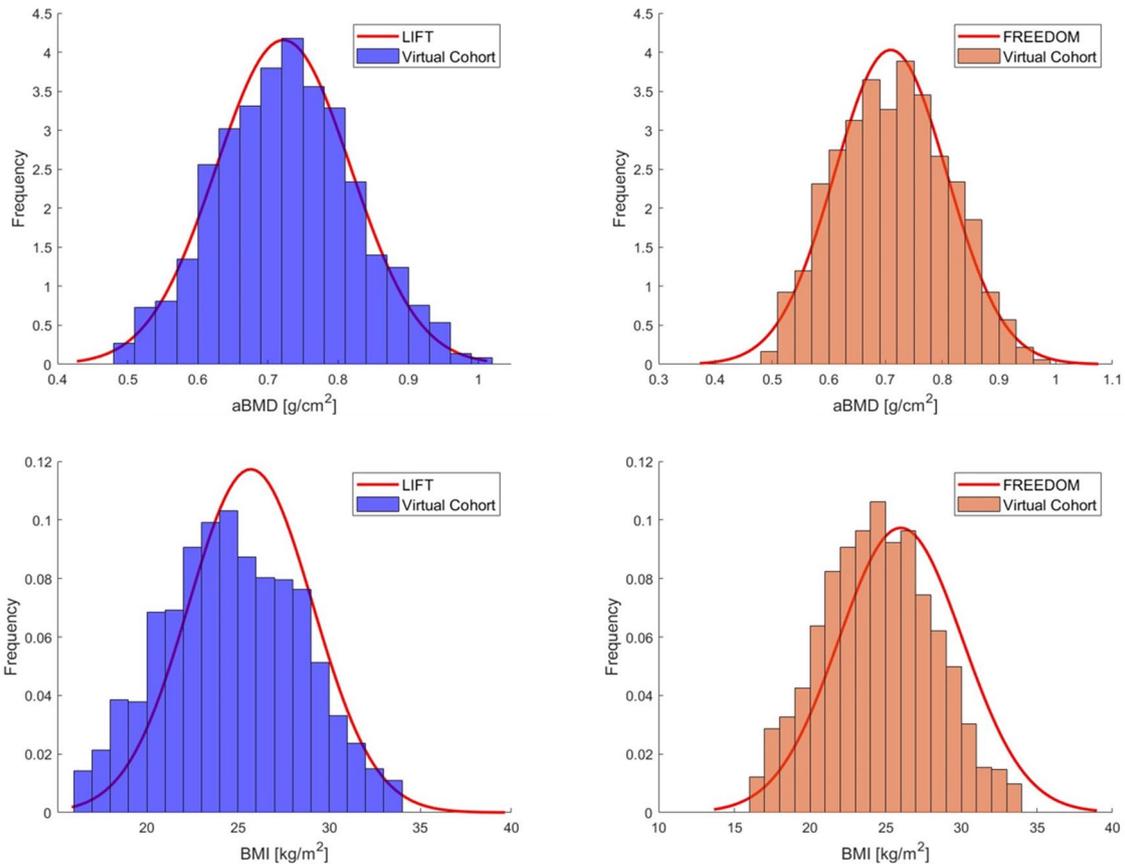


Fig. 2 Distributions of Total Hip aBMD and BMI in the virtual cohorts compared with clinical data reported in the reference studies

Table 2 Number of fractures predicted for the two virtual cohorts in three years follow up: average and SD, minimum and maximum across all realisations

	Predicted fracture incidence (Mean ± SD)	Min predicted fracture incidence	Max predicted fracture incidence	Reported fracture incidence [37, 38]
LIFT (N = 1238)	12 ± 2	8	15	7.7 (14 in 2257 patients)
FREEDOM (N = 1225)	18 ± 4	13	25	13.5 (43 in 3906 patients)

may be influenced by external factors (e.g. presence of fall hazards), the ability of the patient to prevent or attenuate the fall, the characteristics of the impact surface, and more. In our model these are represented with a probability distribution associated with each phenomenon; in this perspective, a clinical trial can be considered as one observation (or realisation) of the process outcome, while more than one trial would be needed to collect data about its variability.

The envisioned application for this *In Silico* trial technology is to support the assessment of treatment efficacy. In this context, a key strength is the possibility to predict fracture incidence, which is the clinical endpoint required in phase III by regulatory authorities, in a representative population

of interest. The stochastic approach also allows to model the complexity and variability of the fall and fracture phenomena. Nevertheless, limitations linked to the generation of virtual cohorts and to the fall model are discussed in the following paragraphs.

In this study, the approach used to model propensity to fall in the population was simplified; it was assumed that the underlying probability mass function associated with *n* fall events was equal for all virtual patients. Nevertheless, propensity to fall is not homogeneous among subjects, with approximately 30% of subjects experiencing one or more falls per year and 70% of non-fallers [37]. Additionally, fall rate is strongly correlated with older age, general frailty (e.g.

dynapenia, presence of comorbidities, equilibrium and postural impairment) and lower education level [64]. Similarly, a previous fall is a significant risk factor for further falls and is associated with fear of falling in the elderly [65]; therefore, assuming the independence of subsequent events also represents a limitation of the current approach. Future developments will be focused on including inter-subject variability in fall rate, e.g. by integrating rule-based predictors based on subject-specific risk factors [66]. Lastly, it is important to highlight that the value of λ (average fall rate) used in this study was based on data from a systematic review reported for the population of interest, in terms of falls per person per year [56]. In this systematic review, concluded clinical trials that performed fall assessment using techniques of proven efficacy were selected [57–59]. The median fall rate was 0.65 falls/year, which differs deeply from the fall rate reported in the clinical trials replicated in this study (0.03–0.05 falls per person per year) [37, 38]. This difference may be due to the fact that falls were recorded as adverse events by clinicians; it has been shown that this approach may lead to an underestimation of recorded falls, given the tendency to record injury-related events, while potentially overlooking non-injurious falls [57, 58]. In this study, fall rate reported in Gillespie et al., (2012) [56] was considered more reliable and applied to model fall incidence.

Likely, a correlation also exists between fall frequency and ability to attenuate a fall, as ageing and general frailty may affect not only the frequency but also the dynamics of falls, although their dependence is not trivial given the lack of a clear relationship [32, 65, 67–69]. For example, incorrect weight shifting has been suggested as the main cause for falling in the elderly [45, 70], which likely produces a significant population-specific difference in fall dynamics, and therefore in impact velocity and force, correlated with age and frailty. In this study, clinical trials cohorts were replicated by matching the baseline aBMD distributions. Implicitly, we assumed that age does not affect postural parameters, although this is a simplification. In fact, predicted fracture incidence was less overestimated for the virtual cohort replicating the FREEDOM cohort, which was on average older compared to the one enrolled in the LIFT study; 73.6% of subjects were older than 70 years of age and 31.6% older than 75; on the other hand, in the LIFT study only 40% of subjects were older than 70 [37, 38]. This discrepancy suggests that the model could be improved by including the effect of frailty, for example on postural reflex, considering that in older cohorts a cognitive delay would result in lower attenuation of the impact force upon falling.

The anatomy atlas used to generate virtual cohorts was based on data from a retrospective study [39] to obtain geometry and BMD distribution. Other subject-specific parameters (height, weight, soft tissue thickness) were derived from either internal correlations with other parameters or

from literature data reported for similar populations. While this approach allows to introduce the population variability in the model to some extent, it is dependent on the amount and quality of data available; additionally, depending on the application, it might be necessary to adapt some of the parameter distributions to model different populations (e.g. ethnicity).

As reported previously, the Body Mass Index (BMI) distributions associated with the virtual cohorts were significantly different from those reported in the reference studies. Our model generates height and weight values for each virtual model based on the relative correlations with BMD obtained from the physical cohort [36]. In both virtual cohorts, the individual correlations between bone mineral density and BMI align with existing literature [45, 71]. More importantly, no significant variation in hip fracture risk has been reported for the corresponding range of BMI [63]. Therefore, this difference is not expected to have an impact on predicted fracture incidence and was considered acceptable.

Another limitation of the study is that no osteoporosis progression was modelled, by assuming that FE material properties were constant over the simulated follow-up years. While bone loss may increase fracture incidence due to structural weakening, nevertheless our preliminary simulations indicate that, in the relatively short follow-up period of the reference studies (three years) [37, 38], a decrease in BMD had limited effect on predicted fracture incidence. Therefore, in first approximation this parameter was kept constant and will be included in the model in future developments, while this work was focused on the algorithm to predict hip fractures in the population of interest and replication of concluded clinical trials to assess its accuracy.

Lastly, the aim of this model was the prediction of hip fractures only, while fractures at other anatomical sites (e.g. pelvic) may occur. Despite this limitation, hip fractures are the second most common clinical endpoint for evaluating the efficacy of osteoporosis treatments, following vertebral fractures [21].

A future development we are exploring is the possibility to use a machine learning-based surrogate model for predicting failure load in multiple fall directions, which could potentially replace the FE simulations. If appropriately validated, this model would allow to significantly reduce the computational cost of the *In Silico* trial simulation.

BoneStrength represents the first *In Silico* Clinical Trial methodology for the development and assessment of interventions to reduce hip fractures. This technology can potentially be used in the future to improve the time- and cost-effectiveness of drug development in osteoporosis and to produce evidence on treatments' efficacy for preventing hip fractures in postmenopausal women. Potential applications include the improvement of clinical trial design and

de-risking, e.g. by reducing the required sample size and/or partially replacing human experimentation, which could also potentially overcome the ethical concerns around the inclusion of high-risk patients in placebo-controlled trials [19, 21].

In conclusion, the *In Silico* trial technology *BoneStrength* was applied for the replication of the placebo arms of two clinical trials on OP treatments, showing its potential for predicting hip fracture incidence. The model refinement and validation will be carried out for application in the drug development process and assessment.

Acknowledgements We gratefully acknowledge Polish high-performance computing infrastructure PLGrid (HPC Centres: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2023/016227.

Funding This study was supported by the European Commission through the H2020 project “In Silico World: Lowering barriers to ubiquitous adoption of In Silico Trials” (topic SC1-DTH-06-2020, grant ID 101016503).

Data Availability Model data are available in the University of Bologna institutional repository under terms of CC BY 4.0 International license at the link: <https://doi.org/10.6092/unibo/amsacta/7934>.

Declarations

Conflict of interest The authors declare that they do not have any financial or personal relationships with other people or organisations that could have inappropriately influenced this study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Salari, N., et al. The global prevalence of osteoporosis in the world: a comprehensive systematic review and meta-analysis. *J. Orthop. Surg.* 16(1):609, 2021. <https://doi.org/10.1186/s13018-021-02772-0>.
- Veronese, N., H. Kolk, and S. Maggi. Epidemiology of Fragility fractures and social impact. In: *Orthogeriatrics: The Management of Older Patients with Fragility Fractures*, 2nd ed., edited by P. Falaschi, and D. Marsh. Cham: Springer, 2021.
- S. WHO Scientific Group on the Prevention and Management of Osteoporosis (2000): Geneva, Prevention and management of osteoporosis: report of a WHO scientific group, World Health Organization, 2003. <https://apps.who.int/iris/handle/10665/42841>. Accessed: Aug. 01, 2023.
- Borgström, F., et al. Fragility fractures in Europe: burden, management and opportunities. *Arch. Osteoporos.* 15(1):59, 2020. <https://doi.org/10.1007/s11657-020-0706-y>.
- Lin, J. T., and J. M. Lane. Osteoporosis: a review. *Clin. Orthop. Relat. Res.* 425:126, 2004. <https://doi.org/10.1097/01.blo.0000132404.30139.f2>.
- Benedetti, M. G., G. Furlini, A. Zati, and G. Letizia Mauro. The effectiveness of physical exercise on bone density in osteoporotic patients. *BioMed Res. Int.* 2018:4840531, 2018. <https://doi.org/10.1155/2018/4840531>.
- Musculoskeletal rehabilitation after hip fracture: a review. *Archives of Osteoporosis*. <https://link.springer.com/article/https://doi.org/10.1007/s11657-010-0050-8>. Accessed: Jul. 25, 2024.
- Tabatabai, L. S., and D. E. Sellmeyer. Nutritional supplements and skeletal health. *Curr. Osteoporos. Rep.* 19(1):23–33, 2021. <https://doi.org/10.1007/s11914-020-00651-x>.
- Tsuda, T., Y. Hashimoto, Y. Okamoto, W. Ando, and K. Ebina. Meta-analysis for the efficacy of bisphosphonates on hip fracture prevention. *J. Bone Miner. Metab.* 38(5):678–686, 2020. <https://doi.org/10.1007/s00774-020-01096-z>.
- Cummings, S. R., L.-Y. Lui, R. Eastell, and I. E. Allen. Association between drug treatments for patients with osteoporosis and overall mortality rates: a meta-analysis. *JAMA Intern. Med.* 179(11):1491–1500, 2019. <https://doi.org/10.1001/jamainternmed.2019.2779>.
- Cardozo, B., E. Karatza, and V. Karalis. Osteoporosis treatment with risedronate: a population pharmacokinetic model for the description of its absorption and low plasma levels. *Osteoporos. Int.* 32(11):2313–2321, 2021. <https://doi.org/10.1007/s00198-021-05944-0>.
- Gallagher, J. C., H. K. Genant, G. G. Crans, S. J. Vargas, and J. H. Krege. Teriparatide reduces the fracture risk associated with increasing number and severity of osteoporotic fractures. *J. Clin. Endocrinol. Metab.* 90(3):1583–1587, 2005. <https://doi.org/10.1210/jc.2004-0826>.
- Matsumoto, T., and I. Endo. RANKL as a target for the treatment of osteoporosis. *J. Bone Miner. Metab.* 39(1):91–105, 2021. <https://doi.org/10.1007/s00774-020-01153-7>.
- Händel, M. N., et al. Fracture risk reduction and safety by osteoporosis treatment compared with placebo or active comparator in postmenopausal women: systematic review, network meta-analysis, and meta-regression analysis of randomised clinical trials. *BMJ.* 381:e068033, 2023. <https://doi.org/10.1136/bmj-2021-068033>.
- Muñoz-Torres, M., G. Alonso, and P. Mezquita Raya. Calcitonin therapy in osteoporosis. *Treat. Endocrinol.* 3(2):117–132, 2004. <https://doi.org/10.2165/00024677-200403020-00006>.
- Johnell, O., W. H. Scheele, Y. Lu, J.-Y. Reginster, A. G. Need, and E. Seeman. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. *J. Clin. Endocrinol. Metab.* 87(3):985–992, 2002. <https://doi.org/10.1210/jcem.87.3.8325>.
- Gennari, L., D. Merlotti, F. Valleggi, G. Martini, and R. Nuti. Selective estrogen receptor modulators for postmenopausal osteoporosis. *Drugs Aging.* 24(5):361–379, 2007. <https://doi.org/10.2165/00002512-200724050-00002>.
- Boag, J. W., J. L. Haybittle, J. F. Fowler, and E. W. Emery. The number of patients required in a clinical trial. *Br. J. Radiol.* 44(518):122–125, 1971. <https://doi.org/10.1259/0007-1285-44-518-122>.
- Adami, G., K. G. Saag, and M. I. Danila. Pragmatic clinical trials in osteoporosis. *Curr. Osteoporos. Rep.* 17(6):521–526, 2019. <https://doi.org/10.1007/s11914-019-00551-9>.
- Bentley, C., S. Cressman, K. van der Hoek, K. Arts, J. Dancy, and S. Peacock. Conducting clinical trials—costs, impacts, and

- the value of clinical trials networks: a scoping review. *Clin. Trials*. 16(2):183–193, 2019. <https://doi.org/10.1177/1740774518820060>.
21. Kanis, J. A., et al. Uncertain future of trials in osteoporosis. *Osteoporos. Int. J. Establ. Result. Coop. Eur. Found. Osteoporos. Natl. Osteoporos. Found. USA*. 13(6):443–449, 2002. <https://doi.org/10.1007/s001980200052>.
 22. ASBMR SABRE Team Submits Full Qualification Plan to FDA for Approval of Bone Mineral Density as a Surrogate Endpoint for Future Trials of New Anti-Osteoporosis Drugs. American Society for Bone and Mineral Research. <https://www.asbmr.org/about/news-release-detail/asbmr-sabre-team-submits-full-qualification-plan-t>. Accessed: Jul. 31, 2024.
 23. Pappalardo, F., G. Russo, F. M. Tshinanu, and M. Viceconti. In silico clinical trials: concepts and early adoptions. *Brief. Bioinform.* 20(5):1699–1708, 2019. <https://doi.org/10.1093/bib/bby043>.
 24. Viceconti, M., A. Henney, and E. Morley-Fletcher. In silico clinical trials: how computer simulation will transform the biomedical industry. *Int. J. Clin. Trials*. 3(2):37–46, 2016. <https://doi.org/10.18203/2349-3259.ijct20161408>.
 25. Viceconti, M., and E. Dall'Ara. From bed to bench: how in silico medicine can help ageing research. *Mech. Ageing Dev.* 177:103–108, 2019. <https://doi.org/10.1016/j.mad.2018.07.001>.
 26. Miura, M., et al. Prediction of fracture load and stiffness of the proximal femur by CT-based specimen specific finite element analysis: cadaveric validation study. *BMC Musculoskelet. Disord.* 18(1):536, 2017. <https://doi.org/10.1186/s12891-017-1898-1>.
 27. Gustafsson, A., M. Tognini, F. Bengtsson, T. C. Gasser, H. Isaksson, and L. Grassi. Subject-specific FE models of the human femur predict fracture path and bone strength under single-leg-stance loading. *J. Mech. Behav. Biomed. Mater.* 113:104118, 2021. <https://doi.org/10.1016/j.jmbm.2020.104118>.
 28. Keaveny, T. M. Biomechanical computed tomography—noninvasive bone strength analysis using clinical computed tomography scans. *Ann. N. Y. Acad. Sci.* 1192(1):57–65, 2010. <https://doi.org/10.1111/j.1749-6632.2009.05348.x>.
 29. Schileo, E., L. Balistreri, L. Grassi, L. Cristofolini, and F. Taddei. To what extent can linear finite element models of human femora predict failure under stance and fall loading configurations? *J. Biomech.* 47(14):3531–3538, 2014. <https://doi.org/10.1016/j.jbiomech.2014.08.024>.
 30. Falcinelli, C., 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*. 67:71–80, 2014. <https://doi.org/10.1016/j.bone.2014.06.038>.
 31. Altai, Z., M. Qasim, X. Li, and M. Viceconti. The effect of boundary and loading conditions on patient classification using finite element predicted risk of fracture. *Clin. Biomech.* 68:137–143, 2019. <https://doi.org/10.1016/j.clinbiomech.2019.06.004>.
 32. Bhattacharya, P., Z. Altai, M. Qasim, and M. Viceconti. A multiscale model to predict current absolute risk of femoral fracture in a postmenopausal population. *Biomech. Model. Mechanobiol.* 18(2):301–318, 2019. <https://doi.org/10.1007/s10237-018-1081-0>.
 33. Aldieri, A., C. Curreli, J. A. Szyszko, A. A. La Mattina, and M. Viceconti. Credibility assessment of computational models according to ASME V&V40: application to the Bologna biomechanical computed tomography solution. *Comput. Methods Programs Biomed.* 240:107727, 2023. <https://doi.org/10.1016/j.cmpb.2023.107727>.
 34. Bouxsein, M. L., and P. D. Delmas. Considerations for development of surrogate endpoints for antifracture efficacy of new treatments in osteoporosis: a perspective. *J. Bone Miner. Res.* 23(8):1155–1167, 2008. <https://doi.org/10.1359/jbmr.080301>.
 35. Li, Z., A. A. Chines, and M. P. Meredith. Statistical validation of surrogate endpoints: is bone density a valid surrogate for fracture? *J. Musculoskelet. Neuronal Interact.* 4(1):64–74, 2004.
 36. La Mattina, A. A., F. Baruffaldi, M. Taylor, and M. Viceconti. Statistical properties of a virtual cohort for in silico trials generated with a statistical anatomy atlas. *Ann. Biomed. Eng.* 51(1):117–124, 2023. <https://doi.org/10.1007/s10439-022-03050-8>.
 37. Cummings, S. R., et al. The effects of tibolone in older postmenopausal women. *N. Engl. J. Med.* 2008. <https://doi.org/10.1097/01.ogx.0000338097.79940.0c>.
 38. Cummings, S. R., et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N. Engl. J. Med.* 361(8):756–765, 2009. <https://doi.org/10.1056/NEJMoa0809493>.
 39. Yang, L., W. J. M. Udall, E. V. McCloskey, and R. Eastell. Distribution of bone density and cortical thickness in the proximal femur and their association with hip fracture in postmenopausal women: a quantitative computed tomography study. *Osteoporos. Int.* 25(1):251–263, 2014. <https://doi.org/10.1007/s00198-013-2401-y>.
 40. Qasim, M., 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. J. Establ. Result. Coop. Eur. Found. Osteoporos. Natl. Osteoporos. Found. USA*. 27(9):2815–2822, 2016. <https://doi.org/10.1007/s00198-016-3597-4>.
 41. Kanis, J., et al. Study design in osteoporosis: a European perspective. *J. Bone Miner. Res.* 18(6):1133–1138, 2003. <https://doi.org/10.1359/jbmr.2003.18.6.1133>.
 42. Schacter, I., and W. D. Leslie. Estimation of trochanteric soft tissue thickness from dual-energy X-ray absorptiometry. *J. Clin. Densitom. Off. J. Int. Soc. Clin. Densitom.* 17(1):54–59, 2014. <https://doi.org/10.1016/j.jocd.2013.01.007>.
 43. Hsiao, E. T., and S. N. Robinovitch. Common protective movements govern unexpected falls from standing height. *J. Biomech.* 31(1):1–9, 1997. [https://doi.org/10.1016/S0021-9290\(97\)00114-0](https://doi.org/10.1016/S0021-9290(97)00114-0).
 44. Le Goic, M., et al. An initial passive phase that limits the time to recover and emphasizes the role of proprioceptive information. *Front. Neurol.* 2008. <https://doi.org/10.3389/fneur.2018.00986>.
 45. Robinovitch, S. N., et al. Video capture of the circumstances of falls in elderly people residing in long-term care: an observational study. *Lancet Lond. Engl.* 381(9860):47–54, 2013. [https://doi.org/10.1016/S0140-6736\(12\)61263-X](https://doi.org/10.1016/S0140-6736(12)61263-X).
 46. Oladi, S., L. Uiga, K. Hebert-Losier, and R. S. W. Masters. Testing the efficacy of a motor analogy designed to promote safe landing by older adults who fall accidentally: a study protocol for a randomised control study. *BMJ Open*. 12(8):e060144, 2022. <https://doi.org/10.1136/bmjopen-2021-060144>.
 47. Fozard, J. L., M. Verccryssen, S. L. Reynolds, P. A. Hancock, and R. E. Quilter. Age differences and changes in reaction time: the Baltimore longitudinal study of aging. *J. Gerontol.* 49(4):P179–P189, 1994. <https://doi.org/10.1093/geronj/49.4.p179>.
 48. Wilkinson, R. T., and S. Allison. Age and simple reaction time: decade differences for 5,325 subjects. *J. Gerontol.* 44(2):P29–P35, 1989. <https://doi.org/10.1093/geronj/44.2.p29>.
 49. Gottsdanker, R. Age and simple reaction time. *J. Gerontol.* 37(3):342–348, 1982. <https://doi.org/10.1093/geronj/37.3.342>.
 50. Woods, D. L., J. M. Wyma, E. W. Yund, T. J. Herron, and B. Reed. Age-related slowing of response selection and production in a visual choice reaction time task. *Front. Hum. Neurosci.* 2015. <https://doi.org/10.3389/fnhum.2015.00193>.
 51. Mendelson, D. N., M. S. Redfern, R. D. Nebes, and J. R. Jennings. Inhibitory processes relate differently to balance/reaction time dual tasks in young and older adults. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 17(1):1–18, 2010. <https://doi.org/10.1080/13825580902914040>.

52. Deary, I. J., and G. Der. Reaction time, age, and cognitive ability: longitudinal findings from age 16 to 63 years in representative population samples. *Aging Neuropsychol. Cogn.* 12(2):187–215, 2005. <https://doi.org/10.1080/138255805990969235>.
53. Vincent, A. S., T. Roebuck-Spencer, K. Gilliland, and R. Schlegel. Automated neuropsychological assessment metrics (v4) traumatic brain injury battery: military normative data. *Mil. Med.* 177(3):256–269, 2012. <https://doi.org/10.7205/MILMED-D-11-00289>.
54. Pijnappels, M., K. Delbaere, D. L. Sturnieks, and S. R. Lord. The association between choice stepping reaction time and falls in older adults—a path analysis model. *Age Ageing.* 39(1):99–104, 2010. <https://doi.org/10.1093/ageing/afp200>.
55. Choice Stepping Response and Transfer Times: Effects of Age, Fall Risk, and Secondary Tasks. *The Journals of Gerontology: Series A*, Oxford Academic. <https://academic.oup.com/biomedgerontology/article/62/5/537/541102>. Accessed 09 May 2023.
56. Gillespie, L. D., et al. Interventions for preventing falls in older people living in the community. *Cochrane Database Syst. Rev.* 2012. <https://doi.org/10.1002/14651858.CD007146.pub3>.
57. Zecevic, A. A., A. W. Salmoni, M. Speechley, and A. A. Vandervoort. Defining a fall and reasons for falling: comparisons among the views of seniors, health care providers, and the research literature. *Gerontol.* 46(3):367–376, 2006. <https://doi.org/10.1093/geron/46.3.367>.
58. Pannurat, N., S. Thiemjarus, and E. Nantajeewarawat. Automatic fall monitoring: a review. *Sensors.* 2014. <https://doi.org/10.3390/s140712900>.
59. Kärkkäinen, M. K., et al. Does daily vitamin D 800 IU and calcium 1000 mg supplementation decrease the risk of falling in ambulatory women aged 65–71 years? A 3-year randomized population-based trial (OSTPRE-FPS). *Maturitas.* 65(4):359–365, 2010. <https://doi.org/10.1016/j.maturitas.2009.12.018>.
60. Grassi, L., N. Hraiech, E. Schileo, M. Ansaloni, M. Rochette, and M. Viceconti. Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur. *Med. Eng. Phys.* 33(1):112–120, 2011. <https://doi.org/10.1016/j.medengphys.2010.09.014>.
61. Schileo, E., F. Taddei, L. Cristofolini, and M. Viceconti. Subject-specific finite element models implementing a maximum principal strain criterion are able to estimate failure risk and fracture location on human femurs tested in vitro. *J. Biomech.* 41(2):356–367, 2008. <https://doi.org/10.1016/j.jbiomech.2007.09.009>.
62. Bayraktar, H. H., E. F. Morgan, G. L. Niebur, G. E. Morris, E. K. Wong, and T. M. Keaveny. Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. *J. Biomech.* 37(1):27–35, 2004. [https://doi.org/10.1016/S0021-9290\(03\)00257-4](https://doi.org/10.1016/S0021-9290(03)00257-4).
63. De Laet, C., et al. Body mass index as a predictor of fracture risk: a meta-analysis. *Osteoporos. Int. J. Establ. Result Coop. Eur. Found. Osteoporos. Natl. Osteoporos. Found. USA.* 16(11):1330–1338, 2005. <https://doi.org/10.1007/s00198-005-1863-y>.
64. Xu, Q., X. Ou, and J. Li. The risk of falls among the aging population: a systematic review and meta-analysis. *Front. Public Health.* 10:902599, 2022. <https://doi.org/10.3389/fpubh.2022.902599>.
65. Vellas, B. J., S. J. Wayne, L. J. Romero, R. N. Baumgartner, and P. J. Garry. Fear of falling and restriction of mobility in elderly fallers. *Age Ageing.* 26(3):189–193, 1997. <https://doi.org/10.1093/ageing/26.3.189>.
66. Cattelani, L., et al. FRAT-up, a Web-based fall-risk assessment tool for elderly people living in the community. *J. Med. Internet Res.* 17(2):e41, 2015. <https://doi.org/10.2196/jmir.4064>.
67. Weijer, R. H. A., M. J. M. Hoozemans, O. G. Meijer, J. H. van Dieën, and M. Pijnappels. The short- and long-term temporal relation between falls and concern about falling in older adults without a recent history of falling. *PLoS ONE.* 16(7):e0253374, 2021. <https://doi.org/10.1371/journal.pone.0253374>.
68. Asai, T., K. Oshima, Y. Fukumoto, Y. Yonezawa, A. Matsuo, and S. Misu. The association between fear of falling and occurrence of falls: a one-year cohort study. *BMC Geriatr.* 22(1):393, 2022. <https://doi.org/10.1186/s12877-022-03018-2>.
69. Liu, M., et al. Fear of falling is as important as multiple previous falls in terms of limiting daily activities: a longitudinal study. *BMC Geriatr.* 21(1):350, 2021. <https://doi.org/10.1186/s12877-021-02305-8>.
70. Luo, Y. H., M. N. Sarvi, P. D. Sun, and J. Ouyang. A subject-specific dynamics model for predicting impact force in elderly lateral fall. *Appl. Mech. Mater.* 446–447:339–343, 2014. <https://doi.org/10.4028/www.scientific.net/AMM.446-447.339>.
71. Rinonapoli, G., et al. Obesity and bone: a complex relationship. *Int. J. Mol. Sci.* 2021. <https://doi.org/10.3390/ijms222413662>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.