This is a repository copy of *The relationship between treatment-related changes in total hip BMD measured after 12, 18, and 24 mo and fracture risk reduction in osteoporosis clinical trials: the FNIH-ASBMR-SABRE project*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/218263/

Version: Published Version

The relationship between treatment-related changes in total hip BMD measured after 12, 18, and 24 mo and fracture risk reduction in osteoporosis clinical trials: the FNIH-ASBMR-SABRE project

Tatiane Vilaca[1],* [iD], **Marian Schini[1], Li-Yung Lui[2], Susan K. Ewing[3], Austin R. Thompson[3]** [iD],
Eric Vittinghoff[3], Douglas C. Bauer[3,4], Richard Eastell[1] [iD]**, Dennis M. Black[3], Mary L. Bouxsein[5]**

[1]Division of Clinical Medicine, University of Sheffield, Sheffield S5 7AU, United Kingdom
[2]Research Institute, California Pacific Medical Center, San Francisco, CA 94158, United States
[3]Department of Epidemiology and Biostatistics, University of California at San Francisco, San Francisco, CA 94158, United States
[4]Department of Medicine, University of California at San Francisco, San Francisco, CA 94158, United States
[5]Center for Advanced Orthopedic Studies, Beth Israel Deaconess Medical Center and Department of Orthopedic Surgery, Harvard Medical School, Boston, MA 02215, United States

*Corresponding author: Tatiane Vilaca, Metabolic Bone Centre, Northern General Hospital, Herries Road, Sheffield S5 7AU, United Kingdom (t.vilaca@sheffield.ac.uk).

Abstract

There is a strong association between total hip bone mineral density (THBMD) changes after 24 mo of treatment and reduced fracture risk. We examined whether changes in THBMD after 12 and 18 mo of treatment are also associated with fracture risk reduction. We used individual patient data ($n = 122\,235$ participants) from 22 randomized, placebo-controlled, double-blind trials of osteoporosis medications. We calculated the difference in mean percent change in THBMD (active-placebo) at 12, 18, and 24 mo using data available for each trial. We determined the treatment-related fracture reductions for the entire follow-up period, using logistic regression for radiologic vertebral fractures and Cox regression for hip, non-vertebral, "all" (combination of non-vertebral, clinical vertebral, and radiologic vertebral) fractures and all clinical fractures (combination of non-vertebral and clinical vertebral). We performed meta-regression to estimate the study-level association (r^2 and 95% confidence interval) between treatment-related differences in THBMD changes for each BMD measurement interval and fracture risk reduction. The meta-regression revealed that for vertebral fractures, the r^2 (95% confidence interval) was 0.59 (0.19, 0.75), 0.69 (0.32, 0.82), and 0.73 (0.33, 0.84) for 12, 18, and 24 mo, respectively. Similar patterns were observed for hip: $r^2 = 0.27$ (0.00, 0.54), 0.39 (0.02, 0.63), and 0.41 (0.02, 0.65); non-vertebral: $r^2 = 0.27$ (0.01, 0.52), 0.49 (0.10, 0.69), and 0.53 (0.11, 0.72); all fractures: $r^2 = 0.44$ (0.10, 0.64), 0.63 (0.24, 0.77), and 0.66 (0.25, 0.80); and all clinical fractures: $r^2 = 0.46$ (0.11, 0.65), 0.64 (0.26, 0.78), and 0.71 (0.32, 0.83), for 12-, 18-, and 24-mo changes in THBMD, respectively. These findings demonstrate that treatment-related THBMD changes at 12, 18, and 24 mo are associated with fracture risk reductions across trials. We conclude that BMD measurement intervals as short as 12 mo could be used to assess fracture efficacy, but the association is stronger with longer BMD measurement intervals.

Keywords: Bone mineral density, osteoporosis medication, fracture risk reduction, surrogate, meta-regression, randomised controlled trial, vertebral fracture, non-vertebral fracture, clinical fracture, hip fracture

Lay Summary

In this study, we looked at how changes in hip bone density over time relate to the risk of fractures in people taking osteoporosis medications. We analysed data from over 122 000 participants across 22 different clinical trials. We found that the increase in bone density measured after 12, 18, and 24 mo of treatment was linked to the risk of fractures. Specifically, greater improvements in bone density were associated with fewer fractures in the spine, hips, and other bones. Using statistical methods, we calculated the strength of this association. We discovered that the later, we measured BMD in people taking the medication, the stronger the link between improved bone density and reduced fracture risk became. Our findings suggest that bone density measurements after 12 mo of treatment could help predict how well a medication will prevent fractures. However, the best predictions came from bone density changes measured over longer periods.

Graphical Abstract
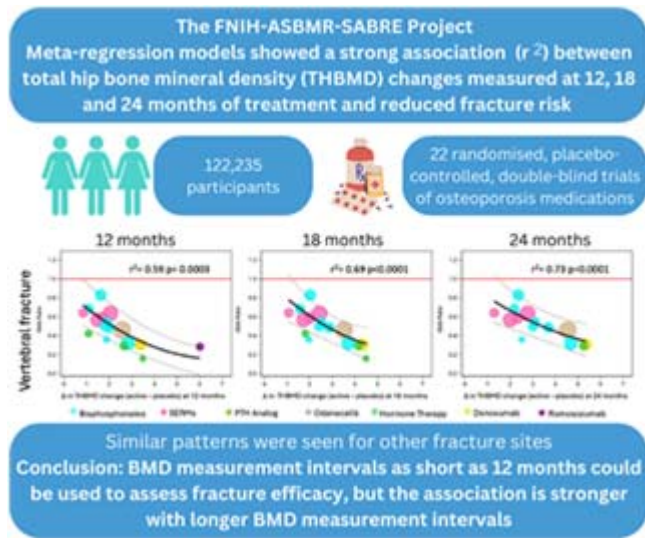
Introduction

We have shown that changes in total hip bone mineral density (THBMD) after 24 mo of treatment are strongly associated with fracture risk reductions across randomized trials of osteoporosis therapies, regardless of the mechanisms of action. Based on these findings, the treatment-related change in THBMD after 24 mo is an excellent surrogate endpoint for fractures in new clinical trials of osteoporosis therapies.[1,2] Several levels of evidence support BMD as a good surrogate marker. First, BMD is strongly associated with whole bone strength in laboratory-based studies of human cadaveric tissue.[3–5] Second, the treatment-related increase in BMD is plausibly associated with a decrease in the risk of fractures as population-based observational studies reveal a strong and consistent association between higher BMD and lower risk of fractures.[6–9] Third, meta-regressions using both published[10] and individual data[1] from randomized controlled trials of osteoporosis therapies have shown associations between improvements in BMD and reductions in the risk of fractures.

Our previous analyses evaluated the association between the change in THBMD after 24-mo of osteoporosis treatment and fracture risk reductions. The study investigated the efficacy of three different BMD measurement sites (total hip, femoral neck, and lumbar spine) as surrogate outcomes for fracture risk. Meta-regression analyses revealed similar performance among these sites within each fracture type, but changes in total hip and femoral neck BMD were more strongly associated with fracture reduction compared with lumbar spine BMD, which can be confounded by factors such as aortic calcification and degenerative changes, especially in older adults. At the hip, total hip measurements are generally more reproducible than those at the femoral neck, making total hip measurements preferable for future trials focusing on longitudinal changes in BMD.[1]

We also determined the surrogate threshold effect (STE), which corresponds to the magnitude of the treatment effect on the surrogate (ie THBMD change), that would predict a significant treatment effect on the final outcome (ie a reduction in the risk of fractures).[11,12] The use of THBMD change as a surrogate endpoint for fracture could reduce the size, duration, and ultimately the cost of trials for new osteoporosis therapies. Despite the clear benefits of reducing osteoporosis drug trials to 24 mo, some drugs might not be used for 24 mo due to mechanistic or safety reasons and shorter trial duration might be desirable. For example, some osteoporosis drugs are used only for 12 mo (eg romozosumab),[13] whereas other drugs, such as abaloparatide,[14] were approved based on 18-mo-long trials. However, the relationships between shorter treatment-related changes in BMD and fracture risk reduction are unknown.

Moreover, for the prior analyses of 24-mo changes in BMD, we included data from 81 497 participants enrolled in 16 RCT's. However, data from 6 trials that did not have 24-mo BMD information were not included.[13–18] Evaluation of shorter BMD measurement intervals would allow the inclusion of a greater number of trials in the analysis, thereby enhancing the generalizability of the findings.

Thus, the aim of the current analyses was to investigate the relationship between the treatment-related change in THBMD and the risk of fractures over shorter BMD measurement intervals. To do so, we conducted meta-regressions for BMD measurement intervals of 12, 18, and 24 mo. We hypothesized that longer BMD measurement intervals would result in stronger associations with anti-fracture efficacy.

Materials and methods
Search strategy and selection criteria

The methodology for the systematic review and meta-regression analysis has been previously described in detail.[1,2,10] Briefly, we aimed to identify and obtain individual patient data (IPD) from all randomized, placebo-controlled trials of osteoporosis medications with fracture outcomes. Using the IPD, we standardized the fracture definitions across all trials. We defined "all fractures" as the combination of non-vertebral, clinical vertebral, and radiographic vertebral fractures; and "all clinical fractures" as the combination of non-vertebral and clinical vertebral fractures. We excluded fractures due to major trauma (ie trauma sufficient to cause a fracture in a young, healthy individual) when reported. In one study, more than half of non-vertebral fractures were originally excluded due to trauma,[19] whereas we

included all non-vertebral fractures from that study. We used the individual study definitions for radiographic vertebral fractures, which were based on comparing the baseline with one or more follow-up lateral spine radiographs. The criteria for identifying radiographic vertebral fractures varied among trials as quantitative morphometry,[20] semi-quantitative assessment,[21] or a combination of these criteria were used.

We created standardized BMD (g/cm^2) values that were comparable across the different dual-energy X-ray absorptiometry devices used in the trials (Hologic, Bedford, MA; GE Lunar, Madison, WI; and Norland Corporation, Fort Atkinson, WI). Specifically, we converted Lunar and Norland BMD values to Hologic BMD values for the total hip, femoral neck, and lumbar spine using previously published methods.[22,23] For this analysis, we focused on changes in THBMD.

Data analysis

We used IPD from placebo-controlled trials for these analyses. When the study reported multiple doses, we combined the active treatment groups, regardless of the dose, except in two trials of risedronate.[24,25] In these two trials, patients randomized to 2.5 mg were excluded from follow-up early, and thus, only the 5 mg dose was included in the current analyses. When the study also included an active comparator, the active comparator group was excluded.[14,26] Furthermore, trials with 5 or less fractures of a specific type were not included in the meta-regression analysis for that specific fracture type.

We used IPD to estimate the difference (active-placebo) in the mean percent change in THBMD at 12, 18, and 24 mo for each study, when data were available. For 10 studies that did not measure BMD at 18 mo,[19,27−34] 18-mo BMD change was estimated as the average of the 12- and 24-mo BMD changes in the IPD dataset. We used percentage change rather than absolute change in total hip BMD because we previously reported that percent BMD changes were as informative as absolute changes.[1,35]

We also used IPD to estimate the log relative risk (RRs) for incident fractures for active vs. placebo for each study. When time to the fracture was known (hip, non-vertebral, all and all clinical fractures), we used Cox proportional hazard models to estimate the treatment effect on fracture risk reduction within each study, with results outputted as log hazard ratios (log HRs). In contrast, for radiographic vertebral fractures, where time to event was unknown, we used logistic regression to estimate treatment effect on fracture risk reduction, with results outputted as log odds ratios (log ORs). We determined the treatment-related fracture reductions for the entire follow-up period, except for the FRAME study. Is this trial, we only included data from the first 12 mo which is the placebo-controlled period. The association between treatment and fracture risk (eg OR or HR) that we calculated sometimes differed from the original published results due to different fracture definitions, the exclusion of traumatic fractures, or updates to the clinical trial dataset after publication.

For each study, we merged the treatment-related differences in mean percent change in THBMD at 12, 18, and 24 mo with the log RRs for incident fractures for active vs. placebo to create a study-level database. We used this database to perform the meta-regression analyses, using each trial as the unit of analysis. Each study was weighted by the inverse of its standard error of the log RR for the given fracture outcome: study weight = 1/standard error of log RR.

For each BMD measurement interval, we used a weighted linear regression model on the study-level dataset to estimate the association of the active-placebo difference in mean percentage change in THBMD during the given interval with the log HR or log OR for incident fracture for active vs. placebo, weighted by the inverse of its variance. This was implemented in SAS using the generalized linear model (GLM) procedure. The GLM model provided the r^2, with 95% CI, summarizing the variance explained by the association between treatment-related difference in mean percent change in THBMD and fracture risk reduction. To illustrate the results, we plotted the back-transformed ORs or HRs against the active-placebo difference in percentage change in THBMD for each BMD measurement interval. Each trial is represented by a circle of size proportional to the inverse of the variance of the log HR or OR for the given fracture outcome. To a first approximation, the size of each circle is proportional to the number of fractures in that trial. We added the back-transformed fitted regression line with 95% prediction limits to the corresponding plots, plotting the line from the smallest to largest treatment-related THBMD differences observed in these studies. We calculated the STE as previously described,[1,2] namely the active-placebo difference in mean THBMD percent change at which the upper 95% prediction limit from the meta-regression intersects an HR or OR = 1.0 (no treatment effect on fracture risk reduction).[11,12]

All analyses were by intention to treat, without regard to adherence to treatment, and were performed using SAS software (version 9.4, SAS Institute Inc., Cary, NC, USA) and Stata software (version 17, StataCorp. LLC, College Station, TX, USA).

Results

In the current analyses, we included data from 122 235 participants from 22 randomized, placebo-controlled trials for whom we had information on both THBMD change and incident fractures (Table S1). Data came from the following trials of osteoporosis medications: 10 bisphosphonate,[17,25,27−29,36−40] 1 odanacatib,[41] 2 hormone therapy— 1 conjugated equine oestrogen[18] and 1 conjugated equine oestrogen plus medroxyprogesterone acetate,[16] 3 parathyroid hormone -PTH receptor agonists[14,15,19], 1 denosumab,[31] 4 selective estrogen receptor modulator,[26,32−34] and 1 romosozumab.[13] The median study follow-up ranged from 12 to 100 mo. The number of trials used in each analysis ranged from 14 to 21, depending on the specific fracture type and BMD measurement interval combination (Table 1).

Depending on the fracture type, the 12-mo analyses included 80 235–122 017 participants enrolled in 17–21 trials, adding ~10 000 more participants to the vertebral fracture analysis and 40 000 more participants to the hip, non-vertebral, all, and all clinical analyses than the respective 24-mo analyses. The mean net differences in THBMD changes at 12 mo in the active vs. placebo treatment groups ranged from 0.84% to 6.02% across the studies.

The meta-regression revealed significant associations between the treatment-related difference (active—placebo) in THBMD change at 12 mo and reductions in risk for all fracture types, such that greater gains in THBMD were associated with larger fracture risk reductions (Table 2, Figure 1). The association between the treatment-related difference in THBMD change and fracture risk reduction

Table 1. Number of studies, participants, and fractures included in the meta-regression analysis comparing BMD measurement intervals of 12, 18, and 24 mo.

Fracture outcome		BMD measurement interval		
		12 mo	18 mo	24 mo
Vertebral	# of studies	17	16	14
	# of participants[a]	80 235	73 592	70 447
	# of fractures	4573	4497	4402
Hip	# of studies	18	15	14
	# of participants[a]	117 469	83 034	80 502
	# of fractures	1243	843	837
Non-vertebral	# of studies	21	17	15
	# of participants[a]	122 016	85 673	81 496
	# of fractures	10 210	6574	6440
All	# of studies	21	17	15
	# of participants[a]	122 017	85 674	81 497
	# of fractures	14 739	10 685	10 459
All clinical	# of studies	21	17	15
	# of participants[a]	122 016	85 673	81 496
	# of fractures	11 496	7495	7353

[a] Total number of participants used in calculating the OR or HR for each fracture type.

Table 2. r^2, 95% confidence intervals and p-values for the association between treatment-related differences in THBMD change and fracture risk reduction over 12, 18, and 24 mo.

Fracture outcome		BMD measurement interval		
		12 mo	18 mo	24 mo
Vertebral	r^2	0.59	0.69	0.73
	95% CI	(0.19, 0.75)	(0.32, 0.82)	(0.33, 0.84)
	p-value	.0003	<.0001	.0001
Hip	r^2	0.27	0.39	0.41
	95% CI	(0.00, 0.54)	(0.02, 0.63)	(0.02, 0.65)
	p-value	.03	.01	.01
Non-vertebral	r^2	0.27	0.49	0.53
	95% CI	(0.01, 0.52)	(0.10, 0.69)	(0.11, 0.72)
	p-value	.02	.002	.002
All	r^2	0.44	0.63	0.66
	95% CI	(0.10, 0.64)	(0.24, 0.77)	(0.25, 0.80)
	p-value	.001	.0002	.0002
All clinical	r^2	0.46	0.64	0.71
	95% CI	(0.11, 0.65)	(0.26, 0.78)	(0.32, 0.83)
	p-value	.0007	.0001	<.0001

tended to be stronger for vertebral fractures ($r^2 = 0.59$, $p=.0003$), than for hip ($r^2 = 0.27$, $p=.03$), non-vertebral ($r^2 = 0.27$, $p=.02$), all ($r^2 = 0.44$, $p=.001$), and all clinical fractures ($r^2 = 0.46$, $p=.0007$). Illustration of this association is shown in Figure 1 for vertebral and all clinical fractures, whereas graphs for the other fracture types are shown in Figures S1–S3. The STE, the minimum treatment-related increase in THBMD associated with a significant reduction in the risk of fracture, varied by fracture type (Table 3) and was lowest for vertebral fracture (0.84%) and highest for hip fracture (1.76%).

The difference (active—placebo) in THBMD change at 12 mo was less than 3.67% for all trials except for the FRAME trial of romosozumab, where the THBMD difference (active—placebo) was 6.02%. However, similar r^2 values were observed in the meta-regression when the FRAME trial was excluded from the analyses: $r^2 = 0.64$ ($p = .0002$) for vertebral fractures, $r^2 = 0.33$ ($p = .01$) for hip fractures, $r^2 = 0.36$ ($p = .005$) for non-vertebral fractures, $r^2 = 0.51$ ($p = .0004$) for all fractures, and $r^2 = 0.56$ ($p = .0002$) for all clinical fractures.

The analyses using 18-mo THBMD changes included 73 592−85 892 participants enrolled in 15−17 trials, depending on the fracture type (Table 1). Like the 12-mo analysis, the meta-regression also revealed significant associations between the treatment-related difference (active—placebo) in THBMD change at 18 mo and reductions in risk for all fracture types (Table 2, Figure 1). The r^2 values for the 18-mo analyses were highest for vertebral, all, and all clinical fractures ($r^2 = 0.63$–0.69). As expected, the STE values at 18 mo were higher than for the 12-mo analysis (Table 3).

The meta-regression results using 24-mo treatment-related differences in THBMD change, several of which were previously reported,[1] reveal moderate to strong associations with fracture risk reductions (Table 2, Figure 1). Specifically, the treatment-related difference in THBMD changes were associated with reductions in risk for vertebral ($r^2 = 0.73$, $p = .0001$), hip ($r^2 = 0.41$, $p = .01$), and non-vertebral fractures ($r^2 = 0.53$, $p = .002$). In the current analysis, we have also included all fractures ($r^2 = 0.66$, $p = .0002$) and all clinical fractures ($r^2 = 0.71$, $p < .0001$). The STE values at 24 mo
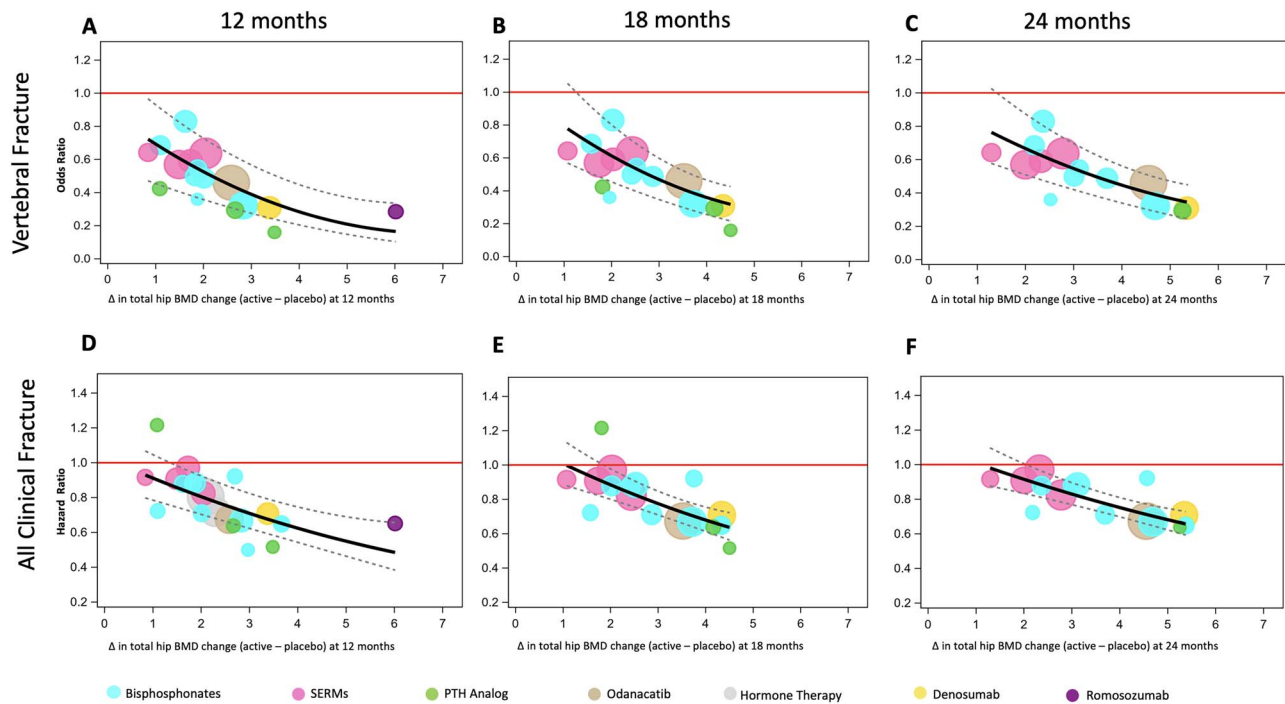
Figure 1. Meta-regression plots displaying the association of active-placebo differences in THBMD percent change and vertebral and all clinical fracture risk reduction for 12- (A, D), 18- (B, E), and 24-mo (C,F) BMD measurement intervals, respectively. Dashed lines represent 95% prediction intervals. Individual trials are represented by circles with areas that are approximately proportional to the number of fractures in the trial. Drugs of the same class are represented by symbols of the same color. The red horizontal line is the odds/hazard ratio of 1.0 (no treatment effect on fracture risk), and the STE is the point where the upper 95% prediction interval intersects this line.

Table 3. STE[a] for BMD measurement intervals of 12, 18, and 24 mo.

Fracture type	BMD measurement interval		
	12 mo	**18 mo**	**24 mo**
Vertebral (%)	0.84	1.27	1.43
Hip (%)	1.76	2.59	3.07
Non-vertebral (%)	1.16	1.91	2.13
All (%)	1.30	1.65	1.83
All clinical (%)	1.37	1.85	2.04

[a]STE is the minimum treatment-related difference (active-placebo) in total hip BMD change required to predict a significant reduction in fracture risk.

were higher than values for 12 and 18 mo: 1.43% for vertebral, 3.07% for hip, 2.13% for non-vertebral, 1.83% for all fractures, and 2.04% for all clinical fractures (Table 3).

Discussion

Previously, we showed that treatment-related differences in in total hip BMD change at 24 mo were associated with reductions in fracture risk. In the current analyses, we expanded this work, aiming to determine whether treatment-related differences in total hip BMD change over shorter intervals would also be associated with fracture risk reductions. Using meta-regressions based on IPD, we found significant associations between treatment-related differences in total hip BMD change at 12, 18, and 24 mo and reductions in vertebral, hip, non-vertebral, all, and all clinical fractures. We observed stronger associations, as measured by greater r^2 values, with longer BMD measurement intervals, though r^2 values were largely similar at 18 and 24 mo, and the 95% confidence interval for r^2 for the three timepoints overlapped.

Not surprisingly, the STEs were also smaller with shorter BMD intervals since the magnitude of treatment-related differences in BMD changes is lower with shorter measurement intervals. Proportional hazards in Cox regression entail a constant relative risk of an event over time. This implies that the hazard function for any two groups (placebo and active drug) remains proportional throughout the observation period. Conversely, treatment-related differences in BMD changes are not linear over time. For this reason, at the 12-mo time point, the STE was more than half that at 24 mo, as the increase in BMD in the second year is often less than in the first year.[27,36,40]

Whereas there is no established threshold to define the r^2 value from the meta-regression required to establish that a surrogate endpoint is valid, the literature suggests that an r^2 value >0.65 is an appropriate target for the acceptability of a surrogate endpoint.[11] Our results indicate that longer BMD measurement intervals increase the likelihood of the r^2 value reaching this threshold. For example, the r^2 value from the meta-regression was close to or exceeded this threshold at the

18- and 24-mo BMD intervals for vertebral, all, and all clinical fractures (r^2 = 0.63–0.71).

However, longer studies might not be feasible or desirable for drugs whose effects are optimized by or restricted to shorter periods. For example, romosozumab is licensed to be used only for 12 mo as that is when it showed anti-fracture efficacy. In addition, trials of PTH analogs were shorter than 24 mo due to safety concerns regarding osteosarcoma, based on animal studies. Furthermore, there are ethical concerns about longer duration placebo-controlled trials in individuals at moderate or high risk for fracture. The current analysis provides evidence that the association between THBMD change and fracture risk reduction due to treatment is present as early as 12 mo, although the r^2 values were close to the desirable threshold only for vertebral fracture (r^2 = 0.59).

Alternative approaches for trial design using BMD as a surrogate endpoint when it is desirable to only use the "test" drug for 18 mo or less could include sequential treatment. This approach has been used in several recent phase III trials.[13,42,43] For example, in the FRAME trial, romosozumab was compared with placebo for the first 12 mo, whereafter both groups received denosumab for an additional 12 mo. A longer BMD interval could then be used to estimate the comparative efficacy of the sequential treatment.

The current analyses included osteoporosis drugs with several mechanisms of action, showing that the use of THBMD change as a surrogate endpoint for fracture risk reduction applies across all categories of osteoporosis drugs. However, the optimum observation time might need to be adjusted for the particularities of each drug since the mechanisms of action are different. Some drugs might need to be used longer to positively affect BMD, and therefore, very short observation periods might not be adequate. As mentioned above, other drugs might have their use limited by safety concerns, making longer observation periods not desirable or possible. Finally, some drugs might have limited periods of action and require specific follow-up periods. Therefore, different trial durations and/or designs might be required in selected cases. Having flexibility to choose a BMD measurement interval that is most appropriate given the specific characteristics of the new drug treatment is a positive aspect of using THBMD change as a surrogate endpoint. We included the data from the initial 12 mo of the FRAME trial, when romosozumab was compared with placebo. Romosozumab showed an increase in THBMD 64% higher than the second-highest increase observed. Despite this greater increase, the romosozumab trial still fell within the 95% prediction limits of the meta-regression curve. Because extreme values can have an impact on regression analyses, we conducted a sensitivity analysis excluding the FRAME trial. The r^2 for vertebral, hip, and non-vertebral fractures was numerically greater with the inclusion of romosozumab, while for all, and all clinical fractures, the r^2 was numerically smaller, but overall, the results observed were similar with and without data from the FRAME trial providing evidence that the observed association was not unduly influenced by this study. This also illustrates how the inclusion of drugs with varying mechanisms of action might increase the variability of the results since the data are less homogeneous, but it also enhances the results' generalizability.

For each BMD measurement interval, there were different sets of studies due to the variable timing of BMD measurements and availability of fracture assessments in each of the studies included. It would be desirable to compare the

r^2 for 12, 18, and 24 mo in the same set of studies, to establish the most appropriate BMD measurement interval. However, this approach would limit the comparison to the set of studies from which we have data on 24-mo BMD and thereby exclude several trials of osteoanabolic therapies. The analyses at 12 and 18 mo include more studies, but the set of studies is different, and thus, direct comparisons of the r^2 values between the three BMD measurement intervals must be interpreted carefully.

This study has several strengths. We used a comprehensive database of IPD from RCTs following standardized criteria for the analysis of total hip BMD and fracture definition. We analysed the association between BMD change and fracture reduction for three different BMD measurement intervals, allowing for the particularities of the study design and drugs' mechanisms of action. The analysis of shorter BMD measurement intervals allowed the inclusion of additional studies that did not have BMD measurements at 24 mo. Because the power and precision of these regressions depend on the number studies included, the increase in the number of trials included makes the analyses more robust. Notably, only one osteoanabolic drug (teriparatide) was included in the 24-mo analyses, but there are four included in the 12-mo analysis, three PTH analogs and romosozumab, enhancing the generalizability to the findings. Moreover, the current analyses expanded our prior work by including a new fracture category, namely "all clinical" that may be useful for trial sponsors and regulatory agencies.

This study also has limitations. For several studies THBMD was not measured at 18 mo, and therefore, we estimated 18-mo BMD change based on the 12- and 24-mo measurements. The trials included in these analyses were all placebo-controlled mono-therapy trials, and further analyses are needed to demonstrate the association between BMD changes and fracture reduction in trials of combination/sequential treatment and in trials with an active comparator. Trials of anabolic drugs were smaller, had a relatively shorter follow-up, and were not powered to show an effect on hip fractures. Thus, most of the data on hip fractures come from bigger and longer trials of anti-resorptive drugs. Thus, hip fracture outcomes from longer trials of anabolic drugs would be helpful; however, these data are currently not available, as the trials have not been conducted. Few of the studies included men, and thus, sex-stratified analyses were not possible. However, BMD-bridging studies are routinely used to assess and approve osteoporosis drugs in men, as there is no indication that the association between the treatment-related change in BMD and reduction in fracture differs by sex.

Notably, these results cannot be applicable to individual patients since the analyses aim to assess the ability of changes in BMD to serve as a surrogate endpoint for fractures for future clinical trials, and thereby, they focus on relative risk reductions between the groups. This would not be appropriate to estimate absolute risk reductions in individuals. In addition, we used total hip BMD to assess treatment-related changes in BMD, but we acknowledge that osteoporosis diagnosis by DXA should be based on the lowest BMD site measured, and treatment decisions should be guided by fracture risk estimates.

It is important to note that additional guidance may be implemented by regulatory agencies to allow a new drug to advance to a registration trial using BMD as a surrogate endpoint for fracture. For example, the new drug will likely

need to be tested in preclinical studies, including demonstrating that drug treatment maintains the association between bone mass and bone strength. Clinical studies incorporating iliac crest biopsies may also be required to ensure normal bone histopathology following treatment. In addition, the STE defines the change in BMD that would predict any decrease in the risk of fractures with 95% certainty, but it is possible that regulatory agencies would require a given decrease in the risk of fracture and, consequently, greater increases in BMD. For example, for spine fractures, the STE for 24 mo was 1.4%, but a 30% decrease in the risk of spine fracture would demand a 3.0% increase in THBMD.[2]

In conclusion, we found significant associations between treatment-related differences in total hip BMD change and fracture risk reductions for BMD measurement intervals shorter than 24 mo. These observations, based on robust IPD from randomized placebo-controlled trials, suggest that, when needed, BMD measurement intervals shorter than 24-mo could be used as a surrogate endpoint for fracture outcomes in future clinical trials. This flexibility could be particularly helpful for new treatments that may be administered for a limited period due to efficacy and/or safety considerations. Furthermore, shorter trial durations could lessen the time it takes for new therapies to be available to patients. Ultimately, regulatory agencies will decide under what circumstances a shorter BMD measurement interval would be allowed.

Author contributions

Tatiane Vilaca (Conceptualization, Investigation, Visualization, Writing—original draft, Writing—review & editing), Marian Schini (Conceptualization, Investigation, Visualization, Writing—review & editing), Li-Yung Lui (Data curation, Formal analysis, Methodology, Writing—review & editing), Susan Ewing (Data curation, Formal analysis, Methodology, Writing—review & editing), Austin Thompson (Formal analysis, Project administration, Writing—review & editing), Eric Vittinghoff (Conceptualization, Formal analysis, Methodology, Writing—review & editing), Douglas Bauer (Conceptualization, Funding acquisition, Writing—review & editing), Richard Eastell (Conceptualization, Methodology, Supervision, Writing—original draft, Writing—review & editing), Dennis Black (Conceptualization, Funding acquisition, Methodology, Writing—original draft, Writing—review & editing), and Mary Bouxsein (Conceptualization, Funding acquisition, Methodology, Supervision, Writing—original draft, Writing—review & editing).

Supplementary material

Supplementary material is available at *Journal of Bone and Mineral Research* online.

Funding

Conflicts of interest

T.V. received consultancy funding from Pharmacosmos. R.E. receives consultancy funding from Immunodiagnostic Systems, Sandoz, Samsung, CL Bio, CureTeQ, Biocon, Takeda, UCB, meeting presentations for Pharmacosmos, Alexion, UCB, Amgen and Osteolabs and grant funding from Alexion. M.S. received consultancy from Kyowa Kirin International.
A.T. None.

E.V. None.
L.L. None.
S.K.E. None.
D.C.B. None.
M.L.B. Advisory Board/Consulting: Angitia, Beryl Therapeutics. Lecture honoraria: Alexion.
D.M.B. Denosumab membership for Eli Lilly.

Data availability

All study data were acquired by requesting IPD from study sponsors. An overarching data use agreement was created between all parties and individual data use agreements were created between individual study sponsors, FNIH, and University of California, San Francisco (UCSF). Per the data sharing agreements that we have with each sponsor, the data can be used for surrogate marker analyses, including any surrogate qualification processes with regulatory authorities. However, other uses of the data are restricted by this agreement, and UCSF is not allowed to share the data.

References

1. Black DM, Bauer DC, Vittinghoff E, et al. Treatment-related changes in bone mineral density as a surrogate biomarker for fracture risk reduction: meta-regression analyses of individual patient data from multiple randomised controlled trials. *Lancet Diabetes Endocrinol*. 2020;8(8):672–682. https://doi.org/10.1016/S2213-8587(20)30159-5

2. Eastell R, Vittinghoff E, Lui LY, et al. Validation of the surrogate threshold effect for change in bone mineral density as a surrogate endpoint for fracture outcomes: the FNIH-ASBMR SABRE project. *J Bone Miner Res*. 2022;37(1):29–35. https://doi.org/10.1002/jbmr.4433

3. Bouxsein ML, Zysset P, Glüer CC, et al. Perspectives on the non-invasive evaluation of femoral strength in the assessment of hip fracture risk. *Osteoporos Int*. 2020;31(3):393–408. https://doi.org/10.1007/s00198-019-05195-0

4. Moro M, Hecker AT, Bouxsein ML, Myers ER. Failure load of thoracic vertebrae correlates with lumbar bone mineral density measured by DXA. *Calcif Tissue Int*. 1995;56(3):206–209. https://doi.org/10.1007/BF00298611

5. Johannesdottir F, Thrall E, Muller J, et al. Comparison of non-invasive assessments of strength of the proximal femur. *Bone*. 2017;105:93–102. https://doi.org/10.1016/j.bone.2017.07.023

6. Black DM, Cauley JA, Wagman R, et al. The ability of a single BMD and fracture history assessment to predict fracture over 25 years in postmenopausal women: the study of osteoporotic fractures. *J Bone Miner Res*. 2018;33(3):389–395. https://doi.org/10.1002/jbmr.3194

7. Cummings SR, Black DM, Nevitt MC, et al. Bone density at various sites for prediction of hip fractures. The study of osteoporotic fractures research group. *Lancet*. 1993;341(8837):72–75. https://doi.org/10.1016/0140-6736(93)92555-8

8. Cummings SR, Cawthon PM, Ensrud KE, et al. BMD and risk of hip and nonvertebral fractures in older men: a prospective study and comparison with older women. *J Bone Miner Res*. 2006;21(10):1550–1556. https://doi.org/10.1359/jbmr.060708

9. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ*. 1996;312(7041):1254–1259. https://doi.org/10.1136/bmj.312.7041.1254

10. Bouxsein ML, Eastell R, Lui LY, et al. Change in bone density and reduction in fracture risk: a meta-regression of published trials. *J Bone Miner Res*. 2019;34(4):632–642. https://doi.org/10.1002/jbmr.3641

11. Ciani O, Buyse M, Drummond M, et al. Use of surrogate end points in healthcare policy: a proposal for adoption of a validation

framework. *Nat Rev Drug Discov*. 2016;15(7):516. https://doi.org/10.1038/nrd.2016.81

12. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat*. 2006;5(3):173–186. https://doi.org/10.1002/pst.207

13. Cosman F, Crittenden DB, Adachi JD, et al. Romosozumab treatment in postmenopausal women with osteoporosis. *N Engl J Med*. 2016;375(16):1532–1543. https://doi.org/10.1056/NEJMoa1607948

14. Miller PD, Hattersley G, Riis BJ, et al. Effect of abaloparatide vs placebo on new vertebral fractures in postmenopausal women with osteoporosis: a randomized clinical trial. *JAMA*. 2016;316(7):722–733. https://doi.org/10.1001/jama.2016.11136

15. Greenspan SL, Bone HG, Ettinger MP, et al. Effect of recombinant human parathyroid hormone (1-84) on vertebral fracture and bone mineral density in postmenopausal women with osteoporosis: a randomized trial. *Ann Intern Med*. 2007;146(5):326–339. https://doi.org/10.7326/0003-4819-146-5-200703060-00005

16. Cauley JA, Robbins J, Chen Z, et al. Effects of estrogen plus progestin on risk of fracture and bone mineral density: the Women's Health Initiative randomized trial. *JAMA*. 2003;290(13):1729–1738. https://doi.org/10.1001/jama.290.13.1729

17. Pols HA, Felsenberg D, Hanley DA, et al. Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Fosamax international trial study group. *Osteoporos Int*. 1999;9(5):461–468

18. Jackson RD, Wactawski-Wende J, AZ LC, et al. Effects of conjugated equine estrogen on risk of fractures and BMD in postmenopausal women with hysterectomy: results from the women's health initiative randomized trial. *J Bone Miner Res*. 2006;21(6):817–828. https://doi.org/10.1359/jbmr.060312

19. Neer RM, Arnaud CD, Zanchetta JR, et al. Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *N Engl J Med*. 2001;344(19):1434–1441. https://doi.org/10.1056/NEJM200105103441904

20. Eastell R, Cedel SL, Wahner HW, et al. Classification of vertebral fractures. *J Bone Miner Res*. 1991;6(3):207–215. https://doi.org/10.1002/jbmr.5650060302

21. Genant HK, Wu CY, van Kuijk C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res*. 1993;8(9):1137–1148. https://doi.org/10.1002/jbmr.5650080915

22. Hui SL, Gao S, Zhou XH, et al. Universal standardization of bone density measurements: a method with optimal properties for calibration among several instruments. *J Bone Miner Res*. 1997;12(9):1463–1470. https://doi.org/10.1359/jbmr.1997.12.9.1463

23. Lu Y, Fuerst T, Hui S, Genant HK. Standardization of bone mineral density at femoral neck, trochanter and Ward's triangle. *Osteoporos Int*. 2001;12(6):438–444. https://doi.org/10.1007/s001980170087

24. Reginster J, Minne HW, Sorensen OH, et al. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral efficacy with Risedronate therapy (VERT) study group. *Osteoporos Int*. 2000;11(1):83–91. https://doi.org/10.1007/s001980050010

25. Harris ST, Watts NB, Genant HK, et al. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. Vertebral efficacy with Risedronate therapy (VERT) study group. *JAMA*. 1999;282(14):1344–1352. https://doi.org/10.1001/jama.282.14.1344

26. Silverman SL, Christiansen C, Genant HK, et al. Efficacy of bazedoxifene in reducing new vertebral fracture risk in postmenopausal women with osteoporosis: results from a 3-year, randomized, placebo-, and active-controlled clinical trial. *J Bone Miner Res*. 2008;23(12):1923–1934. https://doi.org/10.1359/jbmr.080710

27. Black DM, Cummings SR, Karpf DB, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. *Fracture Intervention Trial Research Group Lancet*. 1996;348(9041):1535–1541. https://doi.org/10.1016/S0140-6736(96)07088-2

28. Cummings SR, Black DM, Thompson DE, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the fracture intervention trial. *JAMA*. 1998;280(24):2077–2082. https://doi.org/10.1001/jama.280.24.2077

29. Lyles KW, Colón-Emeric CS, Magaziner JS, et al. Zoledronic acid and clinical fractures and mortality after hip fracture. *N Engl J Med*. 2007;357(18):1799–1809. https://doi.org/10.1056/NEJMoa074941

30. Bone HG, Dempster DW, Eisman JA, et al. Odanacatib for the treatment of postmenopausal osteoporosis: development history and design and participant characteristics of LOFT, the long-term Odanacatib fracture trial. *Osteoporos Int*. 2015;26(2):699–712. https://doi.org/10.1007/s00198-014-2944-6

31. Cummings SR, San Martin J, McClung MR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med*. 2009;361(8):756–765. https://doi.org/10.1056/NEJMoa0809493

32. Cummings SR, McClung M, Reginster JY, et al. Arzoxifene for prevention of fractures and invasive breast cancer in postmenopausal women. *J Bone Miner Res*. 2011;26(2):397–404. https://doi.org/10.1002/jbmr.191

33. Cummings SR, Ensrud K, Delmas PD, et al. Lasofoxifene in postmenopausal women with osteoporosis. *N Engl J Med*. 2010;362(8):686–696. https://doi.org/10.1056/NEJMoa0808692

34. Ettinger B, Black DM, Mitlak BH, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple outcomes of Raloxifene evaluation (MORE) investigators. *JAMA*. 1999;282(7):637–645. https://doi.org/10.1001/jama.282.7.637

35. Schini M, Vilaca T, Lui LY, et al. Pre-treatment bone mineral density (BMD) and the benefit of pharmacologic treatment on fracture risk and BMD change: analysis from the FNIH-ASBMR SABRE project. *J Bone Miner Res*. 2024;39(5):544–550. https://doi.org/10.1093/jbmr/zjae040

36. Liberman UA, Weiss SR, Broll J, et al. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The alendronate phase III osteoporosis treatment study group. *N Engl J Med*. 1995;333(22):1437–1444. https://doi.org/10.1056/NEJM199511303332201

37. Orwoll E, Ettinger M, Weiss S, et al. Alendronate for the treatment of osteoporosis in men. *NEnglJMed*. 2000;343(9):604–610. https://doi.org/10.1056/NEJM200008313430902

38. Chesnut CH 3rd, Skag A, Christiansen C, et al. Effects of oral ibandronate administered daily or intermittently on fracture risk in postmenopausal osteoporosis. *J Bone Miner Res*. 2004;19(8):1241–1249. https://doi.org/10.1359/JBMR.040325

39. Recker R, Stakkestad JA, Chesnut CH 3rd, et al. Insufficiently dosed intravenous ibandronate injections are associated with suboptimal antifracture efficacy in postmenopausal osteoporosis. *Bone*. 2004;34(5):890–899. https://doi.org/10.1016/j.bone.2004.01.008

40. Black DM, Delmas PD, Eastell R, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *N Engl J Med*. 2007;356(18):1809–1822. https://doi.org/10.1056/NEJMoa067312

41. McClung MR, O'Donoghue ML, Papapoulos SE, et al. Odanacatib for the treatment of postmenopausal osteoporosis: results of the LOFT multicentre, randomised, double-blind, placebo-controlled trial and LOFT extension study. *Lancet Diabetes Endocrinol*. 2019;7(12):899–911. https://doi.org/10.1016/S2213-8587(19)30346-8

42. Saag KG, Petersen J, Brandi ML, et al. Romosozumab or alendronate for fracture prevention in women with osteoporosis. *N Engl J Med*. 2017;377(15):1417–1427. https://doi.org/10.1056/NEJMoa1708322

43. Bone HG, Cosman F, Miller PD, et al. ACTIVExtend: 24 months of alendronate after 18 months of Abaloparatide or placebo for postmenopausal osteoporosis. *J Clin Endocrinol Metab*. 2018;103(8):2949–2957. https://doi.org/10.1210/jc.2018-00163