Calcified Tissue International (2023) 112:631–646
https://doi.org/10.1007/s00223-023-01078-z

REVIEW ARTICLE

The Clinical Effectiveness of Denosumab (Prolia®) for the Treatment of Osteoporosis in Postmenopausal Women, Compared to Bisphosphonates, Selective Estrogen Receptor Modulators (SERM), and Placebo: A Systematic Review and Network Meta-Analysis

Magdalena Ruth Moshi[1] · Konstance Nicolopoulos[1] · Danielle Stringer[1] · Ning Ma[1] · Mathias Jenal[2] · Thomas Vreugdenburg[1]

Received: 26 September 2022 / Accepted: 24 December 2022 / Published online: 5 April 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

To assess the effectiveness and safety of denosumab (Prolia®) compared to bisphosphonates (alendronate, ibandronate, risedronate, zoledronate), selective estrogen receptor modulators (SERMs; bazedoxifene, raloxifene) or placebo, for the treatment of osteoporosis in postmenopausal women (PMW). Systematic searches were run in PubMed, Embase & Cochrane Library on 27-April-2022. Randomized controlled trials (RCTs) that included osteoporotic PMW allocated to denosumab, SERMs, bisphosphonates, or placebo were eligible for inclusion. RCTs were appraised using Cochrane Risk of Bias 2.0. Bayesian network and/or pairwise meta-analyses were conducted on predetermined outcomes (i.e. vertebral/nonvertebral fractures, bone mineral density [BMD], mortality, adverse events [AEs], serious AEs (SAEs), withdrawals due to AEs, AEs caused by denosumab discontinuation). A total of 12 RCTs ($k = 22$ publications; $n = 25{,}879$ participants) were included in the analyses. Denosumab, reported a statistically significant increase in lumbar spine (LS) and total hip (TH) BMD, compared to placebo. Similarly, denosumab also resulted in a statistically significant increase in TH BMD compared to the raloxifene and bazedoxifene. However, relative to denosumab, alendronate, ibandronate and risedronate resulted in significant improvements in both femoral neck (FN) and LS BMD. With regards to vertebral fractures and all safety outcomes, there were no statistically significant differences between denosumab and any of the comparator. Relative to placebo, denosumab was associated with significant benefits in both LS and TH BMD. Additionally, denosumab (compared to placebo) was not associated with reductions in vertebral and nonvertebral fractures. Finally, denosumab was not associated with improvement in safety outcomes, compared to placebo. These findings should be interpreted with caution as some analyses suffered from statistical imprecision.

Keywords Network meta-analysis · Systematic review · Denosumab · Bisphosphonates · Selective estrogen receptor modulators · Postmenopausal women

Introduction

Osteoporosis is a bone disorder characterized by low bone mass and density, resulting in skeletal fragility and an increased fracture risk in the spine, hip, wrist, pelvis and humerus [1–3]. Bone properties associated with predicting fracture risk include bone mineral density (BMD) and bone turnover markers [4–7].

Globally, it is estimated that over 200 million people currently have osteoporosis [8], with one in three women and one in five men age 50 or older presenting with osteoporotic fractures [9]. Postmenopausal women are at an increased

✉ Magdalena Ruth Moshi
college.asernip@surgeons.org

[1] Research & Evaluation incorporating ASERNIP–S, Royal Australasian College of Surgeons, 24 King William Street, Kent Town, South Australia 5067, Australia

[2] Health Technology Assessment Section, Health Insurance Benefits Division, Health and Accident Insurance Directorate, Federal Office of Public Health (FOPH), Bern, Switzerland

risk of developing primary osteoporosis due to reduced estrogen levels [10, 11].

Osteoporotic fractures, also called fragility fractures, are bone injuries that occur from low-energy trauma such as falls from standing height [12–14]. These types of fractures result in reduced health-related quality of life (HRQoL), increased disability and increased mortality [11, 15, 16]. Vertebral fractures are specifically associated with an increased mortality rate of 10–20% in people with osteoporosis [3, 16].

Antiresorptive agents are often used to prevent osteoporosis and osteoporotic fractures by increasing bone mass and density [17]. Antiresorptive treatments include bisphosphonates such as alendronate, ibandronate, risedronate and zoledronate; selective estrogen receptor modulators (SERMs), such as bazedoxifene and raloxifene; and monoclonal antibody against receptor activator of nuclear factor kappa-B ligand (RANK-L), denosumab.

Denosumab (Prolia®) is regularly used in clinical practice across developed countries to treat osteoporosis in postmenopausal women and is often reimbursed through health insurance [18–21]. In 2017, pharmacovigilance reports warned that discontinuation of denosumab therapy in patients with osteoporosis can lead to increased rates of bone turnover, significant bone mineral loss (in some cases below baseline levels) and increased vertebral fracture risk [22]. Such complications have not been observed after the discontinuation of other osteoporosis therapies (i.e. bisphosphonates, SERMS) due to differences in their mode of action.

The aim of this study was to evaluate the clinical effectiveness and safety (using randomized control trials [RCT]) of denosumab, compared to bisphosphonates and SERMs, for the treatment of postmenopausal women with osteoporosis in developed countries.

This was achieved through conducting a systematic review and network meta-analyses of all available evidence on denosumab, bisphosphonate and SERMs in postmenopausal women with osteoporosis.

Method

This systematic review followed an a priori protocol and is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement for network meta-analysis [23]. A completed PRISMA network meta-analysis checklist is detailed in the *Supplementary document*. Similarly, the study protocol is available from the authors upon request.

Search Strategy

Two systematic literature searches were conducted in three databases (PubMed, Embase, The Cochrane Library) from inception to 27 April 2022 (*Supplementary Table S1 & S2*). The first section of the search was sensitive and identified the literature relevant to denosumab in people with osteoporosis. The second section of the search was more specific and was combined with a methodological filter to limit the identified literature associated with osteoporotic patients on bisphosphonates or SERMs to randomized controlled trials (RCTs). In order to capture any additional evidence that may have otherwise not been identified, grey literature searches of specialty websites were conducted (*Supplementary Tables S4 & S5*), clinical trial registries were reviewed (*Supplementary Table S3*), and the reference lists of included publications were pearled by authors. In addition, clinical experts (i.e. oncologist, rheumatologists, endocrinologists) were consulted.

Study Selection

Studies were considered for inclusion if they were RCTs and met the predetermined eligibility criteria (complete criteria is available in *Table 1* & *Supplementary Table S6*). For the purposes of this review, developed countries were defined as World Health Organization (WHO)-Mortality Stratum A countries. This was implemented to ensure that the included RCTs had comparable causes of death and burden of disease [24].

The search results were screened by title and abstract against predetermined inclusion criteria by three reviewers. Inter-rater reliability was checked via discordance among reviewers on a test sample ($k = 200$); the calculated Fleiss kappa score was high (kappa = 0.872) [25]. All articles deemed potentially relevant were then reviewed by full text by two reviewers, independently. Conflicts between reviewers on study inclusion were settled via consensus. If consensus could not be reached, a third reviewer decided whether to include or exclude the citation.

Data Extraction

One reviewer independently extracted data into a standardized template, which was then checked against the original study record by a second reviewer. Disagreements were settled by discussion or utilization of a third reviewer. Data of interest included trial information, demographic information, intervention and comparator, outcomes of interest and any other noteworthy features. Both intention-to-treat (preferentially utilized) and per-protocol information, as well

Table 1 Study selection criteria

	Inclusion criteria	Exclusion criteria
Population	Postmenopausal women with osteoporosis (with a reduction of more than 2.5 SD in osteodensitometry or in case of a fracture)	Men with osteoporosis (primary or secondary) Patients with hormone-sensitive cancer receiving hormone therapy Patients with multiple myeloma Patients with bone metastases (from solid tumors) Patients with giant-cell tumors Patients with hypercalcemia of malignancy refractory to bisphosphonate treatment Patients with kidney disease
Intervention	Denosumab 60 mg (Prolia®)	Denosumab 120 mg (Xgeva®)
Comparator	All bisphosphonates: Alendronate: 70 mg tablet once weekly Ibandronate: 150 mg tablet once monthly, 3 mg/3 mL infusion trimonthly Risedronate: 35 mg tablet once weekly Zoledronate: 4 mg/5 mL or 4 mg/100 mL (i.e. post-dilution) infusion monthly/trimonthly/annually, 5 mg/100 mL infusion annually All SERMs: Bazedoxifene: 20 mg tablet once daily Raloxifene: 60 mg tablet once daily Placebo	All other interventions
Outcomes	Efficacy/effectiveness: *Primary* Vertebral fractures Nonvertebral fractures Health-related quality of life (HRQoL) *Secondary* Bone mineral density (BMD) – measured at femoral neck (FN), lumbar spine (LS), total hip (TH) and/or trochanter (TRO) Fracture risk (i.e. FRAX®) Safety: Mortality Treatment-related adverse events (AE) Serious adverse events (SAE) Withdrawal due to treatment-related AE AE upon discontinuation of denosumab (e.g. rebound effect)[a]	
Study design/ publication type	Randomized controlled trials (RCT)	Case reports Conference abstracts Editorial Expert opinion Letter to the editors Narrative review articles Non-randomized studies of interventions (NRSI) Single-arm studies (i.e. case-series)
Language	English, German, French, Italian	All other languages
Country	WHO-Mortality Stratum A countries: Andorra, Australia, Belgium, Brunei, Canada, Croatia, Cuba, Cyprus, Czech Republic (Czechia), Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Luxembourg, Malta, Monaco, The Netherlands, New Zealand, Norway, Portugal, San Marino, Singapore, Slovenia, Spain, Sweden, Switzerland, United Kingdom (UK), and United States of America (USA)	
Length of follow	Longest point of follow-up	None
Date limitation	None	None

AE adverse events; *BMD* bone mineral density; *FN* femoral neck; *FRAX* ® fracture risk assessment tool; *HRQoL* health-related quality of life; *LS* lumbar spine; *mg* milligrams; *mL* milliliters; *NRSI* non-randomized studies of interventions; *RCT* randomized controlled trial; *SAE* serious adverse events; *SD* standard deviations; *SERM* selective estrogen receptor modulators; *TH* total hip; *TRO* trochanter; *UK* United Kingdom; *USA* United States of America; *WHO* World Health Organization

[a]Upon discontinuing denosumab treatment, the patients BMD losses may increase above baseline levels and/or the patient may experience an increased rate of vertebral fractures [26, 27]

as adjusted and unadjusted (preferentially utilized) results were extracted for data analysis. For studies that reported outcomes graphically, *WebPlotDigitizer* was used to estimate numerical values [28].

Assessment of Quality of Evidence

The quality of RCTs was evaluated using Cochrane Risk of Bias 2.0 (RoB 2.0) [29]. The appraisal was performed by one reviewer and checked by a second reviewer. Any differences were settled via consensus. If consensus could not be reached, a third reviewer was consulted. Risk of bias (RoB) was assessed on a per-outcome basis for clinical effectiveness and safety outcomes. However, for the ROB2 domains of randomization and blinding of personal/participants, the effectiveness and safety outcomes were evaluated together.

Data Analysis

Network Meta-Analyses

A network meta-analysis was performed if data was available for three treatment arms across a minimum of two trials. The network meta-analyses were performed using a Bayesian inference. The analysis was performed under the assumption of a consistency model using a random-effects model. Random-effects models were used in the analyses in order to account for any variations in the possible effect modifiers (i.e. compliance, age, baseline fractures etc.) as well as to account for discrepancies in how the intervention and comparators were delivered in the included trials. The referent comparator for each model was placebo, as most of the available direct evidence was reporting treatment effect relative to placebo. Default and non-informative priors with standard normal distribution and sufficiently wide standard deviations (SD) were used to compute the posterior distribution data, as it was computationally feasible [30, 31].

The duration of follow-up extracted and used in the network meta-analysis for the dichotomous outcomes of vertebral fractures, nonvertebral fractures, mortality, adverse events (AEs), serious AEs (SAEs) and withdrawals due to AEs was the total number of events at the last reported timepoint. Risk ratios (RR) and accompanying 95% credible interval (95% CrI) were calculated using link function *logit* and a binomial family distribution. With regards to continuous outcomes, mean percentage change for BMD was extracted for each reported timepoint. In situations where the mean percentage change was not provided (e.g. T-score, g/cm^2, nmol) it was imputed from the measurements at baseline and timepoints. The mean follow-up period (± 1 SD of the follow-up period) from all included trials was calculated, and results at this timepoint were used in the analyses. Mean difference (MD) and accompanying 95% CrI were calculated using a link function *identity* and a normal family distribution [32]. Missing values were imputed using methods and formulae detailed in the Cochrane Handbook [33–35].

Modelling, Convergence and Output The Bayesian network meta-analysis was performed in RStudio using the BUGSnet (Bayesian inference Using Gibbs Sampling to conduct network meta-analysis) package [32, 36, 37].

For dichotomous outcomes, a burn-in of 70,000 iterations of Markov chain Monte Carlo (MCMC) simulation was used where results were discarded. A burn of 700,000 iterations of MCMC simulations was run to estimate parameters. To assess the consistency assumption, an inconsistency model using a random-effects model was run. A burn-in of 5,000 iterations of MCMC simulation and an additional 25,000-iteration of MCMC simulations were run to compare the parameter for the assessment of consistency.

For continuous outcomes, a burn-in of 10,000 iterations of MCMC simulation was used where results were discarded. A burn of 100,000 iterations of MCMC simulations was run to estimate parameters. 3For the purpose of assessing the consistency assumption, an inconsistency model using a random-effects model was run. A burn-in of 5,000 iterations of MCMC simulation and an additional 25,000-iteration of MCMC simulations were run to compare the parameter for the assessment of consistency.

The Gelman-Rubin statistic as defined in Brooks and Gelman (i.e. potential scale reduction factor [PSRF] between 1 and 1.05) was used to assess whether convergence had been met in both the consistency and inconsistency models [31, 38, 39].

The results were presented in forest plots. These forest plots included the pooled treatment effects of each intervention relative to placebo. The plots also presented the treatment ranking within the network as well as the equivalent surface under the cumulative ranking curve (SUCRA).

The league tables that present treatment effects and 95% CrI for all possible pairs of interventions in each network are available in the *Supplementary document*.

Network diagrams were drawn to illustrate the geometry of the treatment network in each analysis. The size of each network node is proportional to the sample size of that node, and the thickness of the lines connecting the nodes is proportional to the number of included trials (i.e. direct evidence).

Assessment of Heterogeneity and Inconsistency. Cochrane's Q-statistic was used to derive the conventional I^2 values to characterize statistical heterogeneity for both continuous and dichotomous outcomes [40, 41]. The results of the heterogeneity assessment are available in the *Supplementary document.*

Inconsistency at the global level was assessed by reviewing the fit of consistency and inconsistency models using

leverage plots, as well as comparing the deviance information criterion (DIC) score for both consistency and inconsistency models [32]. A difference in DIC scores of 0 to 5 between models was considered minimal, a difference of 5 to 10 was substantial and, finally, a difference that was greater than 10 was significant and eliminated the validity of the results of the model with the higher DIC [42]. The presence of local inconsistency was evaluated by a plot that compared the posterior MD of each data point produced by the consistency and inconsistency models [32]. In situations where networks do not have closed loops, a DIC score could not be calculated [43]. Global and local inconsistency tables are available in the *Supplementary document.*

Meta-Regressions

Meta-regressions were conducted to evaluate whether there was an association between the included antiresorptive treatments and the age of postmenopausal women. The meta-regressions were only conducted if 10 or more trials were included [40].

Sensitivity Analysis

In addition to the main analyses, sensitivity analyses were conducted to review the impact that high and moderate RoB had on the various analyses. This was achieved by rerunning the respective analyses and only including trials that had a low risk of attrition bias, selection bias and reporting bias. The analyses only focused on these three domains, as the risks of performance bias and detection bias in the included trials were low.

Assessment of Publication Bias

Publication bias was assessed using comparison-adjusted funnel plots (*Supplementary document*) [44]. This method requires a minimum of 10 trials per outcome [45].

Results

The results of the literature search are summarized in Fig. 1. The searches identified 22,979 articles. A total of 4,753 duplicate citations were removed and 18,226 items were reviewed by title and abstract. In total, 1,227 articles were reviewed by full text. A total of 12 RCTs ($k = 22$ publications) met the study selection criteria (Table 1) [46–67]. Table 2 details the characteristics of the 12 included trials. There was no available RCT evidence that met the predetermined selection criteria (Table 1) to investigate the effect of denosumab on HRQoL, trochanteric (TRO) BMD and fracture risk assessment (FRAX®) in postmenopausal women with osteoporosis (Table 2).

Risk of Bias (RoB)

The RoB 2.0 graph and summary are reported in Fig. 2. Across the five domains, all included RCTs ($n = 12$), presented a high RoB overall. The specific domain scores for each included trial are described in *Supplementary Table S7*.

Effectiveness

Vertebral Fractures

Vertebral fracture data were available from nine RCTs that had a combined sample size of 19,710 (Fig. 3a) [51, 54–57, 60, 61, 63–65]. Neither denosumab nor the other active treatments were statistically significant compared to placebo after 12 to 84 months of treatment (Fig. 4). Similarly, none of the pairwise comparisons were statistically significant (Supplementary Table S8). Of these treatments, denosumab had the highest probability (SUCRA = 79.56) of being the most effective at preventing vertebral fractures, whereas bazedoxifene had the lowest probability (SUCRA = 26.09) and was ranked as the least effective active treatment. The network did not show any significant evidence of statistical heterogeneity or local inconsistency (Supplementary Figure S2 & Table S17). Global inconsistency could not be estimated as a DIC score could not be generated (Supplementary Table S18).

Nonvertebral Fractures

Nonvertebral fracture data were available from seven RCTs that had a combined sample size of 21,873 (Fig. 3b) [47, 51, 55, 56, 60, 64, 67]. Compared to all treatments, denosumab was not associated with statistically significant changes. In addition, only the pairwise comparisons between alendronate and placebo were statistically significant (Supplementary Table S9). Denosumab had the lowest probability (SUCRA = 29.51) and was ranked as the least effective treatment in the network (Fig. 4). Risedronate had the highest probability (SUCRA = 95.69) of being the most effective treatment in the network. The network did not show any significant evidence of statistical heterogeneity or inconsistency (Supplementary Figure S3, Table S19 & Table S20). In addition, sensitivity analysis suggested that results were not impacted by reporting bias (Supplementary Table S35).

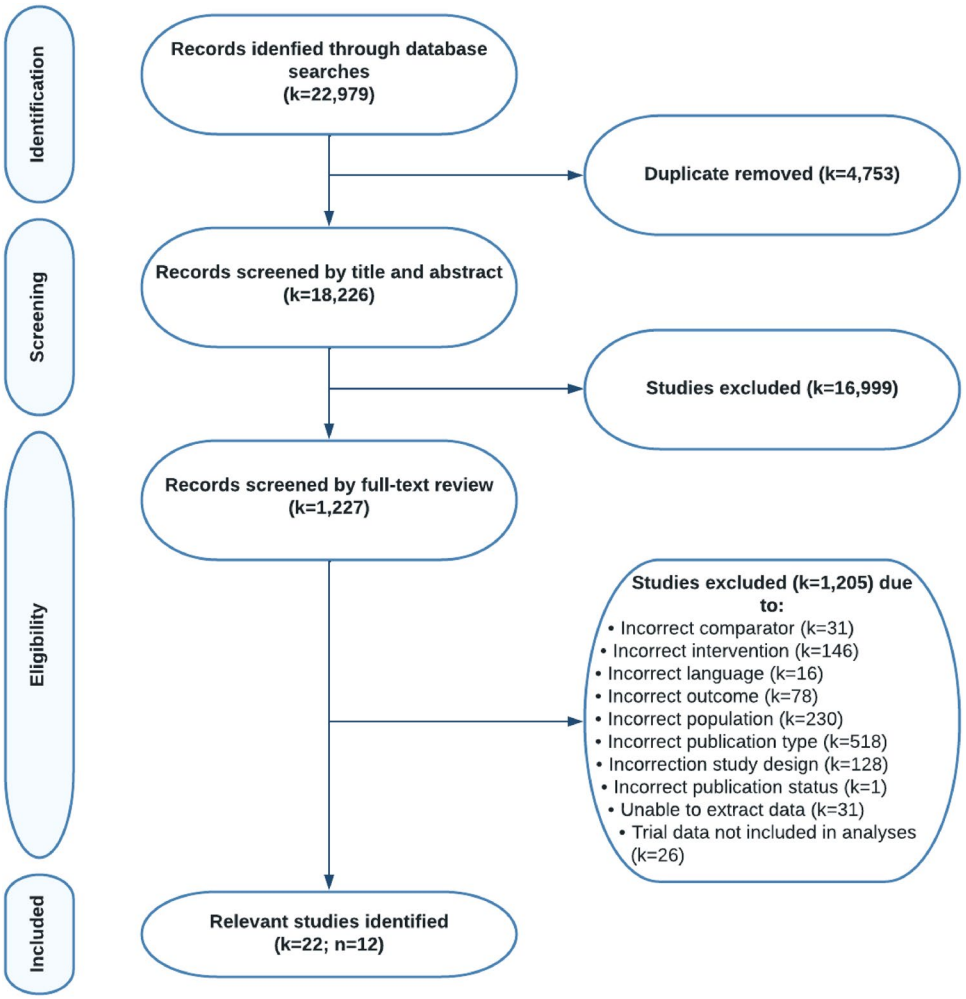

Fig. 1 PRISMA flow diagram. *PRISMA* preferred reporting items for systematic reviews and meta-analyses, *RCTs* randomized controlled trials. *K* number of individual publications. *n* number of RCTs—an RCT can be included in multiple publications

Femoral Neck (FN) BMD

Data on BMD measured at the femoral neck (FN) were available from eight RCTs that had a combined sample size of 12,128 (Fig. 3c) [47, 54–56, 58, 59, 62, 64]. Relative to denosumab, at 19 (± 1 SD) months, alendronate, ibandronate and risedronate resulted in significant improvements in FN BMD of 11.47% (MD 11.47; 95% Crl 1.39, 21.96), 11.02% (MD 11.02; 95% Crl 0.82, 21.37) and 9.67% (MD 9.67; 95% Crl 0.88, 18.72), respectively (Supplementary Table S10). Of these treatments, alendronate had the highest probability (SUCRA = 94.05) of being the most effective treatment and denosumab had a low probability (SUCRA = 52.31) and was ranked as the fourth most effective treatment (Fig. 5). Overall, there was low total heterogeneity within the network (Supplementary S22). There was no evidence of local or global inconsistency in the network (Supplementary Figure S4 & Table S21).

Lumbar Spine (LS) BMD

Data on BMD measured at the lumbar spine (LS) were available from nine RCTs that had a combined sample size of 10,092 (Fig. 3d) [47, 51, 54–56, 62, 64, 67]. Figure 5 indicates that denosumab (compared to placebo) can improve LS BMD by 7.67% (MD 7.67; 95% Crl 3.11, 12.22) at 20 (± 1 SD) months. Relative to denosumab, alendronate, ibandronate and risedronate resulted in significant improvements in LS BMD of 13.32% (MD 13.32; 95% Crl 4.90, 21.64), 13.16% (MD 13.16; 95% Crl 4.75, 21.50) and 9.52% (MD 9.52; 95% Crl 2.66, 16.37), respectively (Supplementary Table S11). Alendronate had the highest probability (SUCRA = 94.57) of being the most effective treatment in the network, and denosumab (SUCRA = 55.34) was ranked as the fourth most effective treatment. The entire network showed substantial to considerable total heterogeneity (Supplementary Table S24). However, the network arm that compared placebo to denosumab presented low heterogeneity.

Table 2 Characteristics of included RCTs assessing clinical effectiveness and safety

Study/trial, date	Country	Age (yo)[a] *mean ± s.d*	Follow-up duration (months)	Intervention		Comparator(s)		Outcome(s)												
								Effectiveness									*Safety*			
				Treatment	Sample size	Treatment	Sample size	Vertebral fracture	Non-vertebral fracture	HRQoL	BMD				FRAX®	Mortality	AE	SAE	Withdrawals due to AE	AE upon discontinuation
											FN	LS	TH	TRO						
FREEDOM, 2009 [46, 48, 50, 51] NCT00089791	32 countries[b]	72.3 ± 5.2	84	DEN (60 mg/6mo, subq)	3,902	PLB	3,906	✓	✓	✗	✗	✓	✓	✗	✗	✓	✓	✓	✓	✓
Nakamura et al., 2012 [61][c]	Japan	64.8 ± 6.6	12	DEN (60 mg/6mo, subq)	54	PLB	55	✓	✗	✗	✗	✗	✓	✗	✗	✗	✓	✓	✗	✗
Miller et al., 2016 [59] NCT01732770	7 countries[d]	67.3 ± 7.9	12	DEN (60 mg/6mo, subq)	321	ZOL (5 mg/y, IV)	322	✗	✗	✗	✓	✗	✓	✗	✗	✗	✓	✓	✓	✗
Morii et al., 2003 [60][e]	Japan	64.7 ± 6.3	12	RLX (60 mg/d, p.o.)	92	PLB	97	✓	✓	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✗
MORE, 1999 [53, 54][e]	25 countries[f]	67 ± NR	48	RLX (60 mg/d, p.o.)	2,557	PLB	2,576	✓	✗	✗	✓	✓	✗	✗	✗	✓	✗	✗	✗	✗
Silverman et al., 2008 [49, 52, 63, 66, 67] NCT00205777	28 countries[g]	66.5 ± 6.7	84	BAZ (20 mg/d, p.o.)	1,886	RLX (60 mg/d, p.o.)	1,849	✓	✓	✗	✗	✓	✓	✗	✗	✓	✓	✓	✓	✗
						PLB	1,885													
Itabashi et al., 2011 [56][h] NCT00238745	Japan	63.42 ± 6.3	24	BAZ (20 mg/d, p.o.)	143	PLB	142	✓	✓	✗	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗
MOTION, 2008 [58] MM17385	NR	65.6 ± NR	12	IBN (150 mg/mo, p.o.)	887	ALN (70 mg/w, p.o.)	873	✗	✗	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗
TRIO, 2014 [62] NCT00666627	UK	67.1 ± 7.2	24	IBN (150 mg/mo, p.o.)	57	ALN (70 mg/w, p.o.)	57	✗	✗	✗	✓	✓	✗	✗	✗	✗	✓	✓	✗	✗
						RIS (35 mg/w, p.o.)	58													
ZEST, 2015 [55] NCT00558012	USA	85.4 ± 0.5	24	ZOL (5 mg/y, IV)	89	PLB	92	✓	✓	✗	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗
Palomba et al., 2005 [64, 65]	Italy	51.9 ± 3.1	36	RIS (35 mg/w, p.o.)	75	PLB	45	✓	✓	✗	✓	✓	✗	✗	✗	✗	✓	✗	✗	✗
HORIZON-PFT, 2007 [47, 57] NCT00049829	27 countries[i]	73.1 ± 5.4	36	ZOL (5 mg/y, IV)	3,889	PLB	3,876	✓	✓	✗	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗

Table 2 Continued

AE adverse events; *ALN* alendronate; *BAZ* bazedoxifene; *BMD* bone mineral density; *d* day/s; *DEN* denosumab; *FN* femoral neck; *FRAX®* fracture risk assessment tool; *FREEDOM* fracture reduction evaluation of denosumab in osteoporosis every 6 months; *HORIZON-PFT* health outcomes and reduced incidence with zoledronic acid once yearly—pivotal fracture trial; *HRQoL* health-related quality of life; *IBN* ibandronate; *IV* intravenous; *LS* lumbar spine; *mo* month/s; *MOTION* monthly oral therapy with ibandronate for osteoporosis intervention; mg: milligrams; NR: not reported; PLB: placebo; p.o.: orally; RLX: raloxifene; RIS: risedronate; s.d.: standard deviations; *SAE* serious adverse events; *subq* subcutaneous injection; *TH* total hip; *TRO* trochanter; *UK* United Kingdom; *USA* United States of America; *w* week/s; *y* year/s; *yo* years old; *ZEST* zoledronic acid in frail elders to strengthen bone; *ZOL* zoledronate

[a]In situations where only treatment-specific age (mean ± s.d.) was provided, the total population age (mean ± s.d.) was combined using the formulae detailed in *Cochrane Handbook for Systematic Reviews of Interventions (version 6.1)* [40]

[b]Countries: USA, Canada, Argentina, Brazil, Mexico, Australia, New Zealand, Australia, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, UK, Czech Republic, Estonia, Greece, Hungary, Latvia, Lithuania, Malta, Poland, Slovak Republic, Romania, Bulgaria, Serbia [51]

[c]Findings on denosumab (14 mg/6mo) and (100 mg/6mo) were excluded as these dosages of denosumab are not eligible for reimbursement when they are indicated to treat osteoporosis with postmenopausal women [61]

[d]Countries: Australia, Canada, Belgium, Denmark, Poland, Spain, USA [59]

[e]Findings on raloxifene (120 mg/d) were excluded as this dosage is not eligible for reimbursement when it is indicated to treat osteoporosis with postmenopausal women [53, 54, 60, 68–70]

[f]Countries: Argentina, Australia, Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Hungary, Israel, Italy, Mexico, the Netherlands, New Zealand, Poland, Singapore, Slovak Republic, Slovenia, Spain, Sweden, UK, USA, Canada [54]

[g]Countries: Argentina, Australia, Belgium, Brazil, Bulgaria, Canada, Chile, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Lithuania, Mexico, The Netherlands, New Zealand, Norway, Poland, Romania, Russia, Slovakia, South Africa, Spain, USA [67]

[h]Findings on bazedoxifene (40 mg/d) were excluded as this dosage is not eligible for reimbursement when it is indicated to treat osteoporosis with postmenopausal women [56]

[i]Countries: Argentina, Australia, Austria, Belgium, Brazil, Canada, China, Columbia, Finland, France, Germany, Hong Kong, Hungary, Israel, Italy, Korea, Mexico, New Zealand, Norway, Poland, Russia, Sweden, Switzerland, Taiwan, Thailand, UK, USA [47]

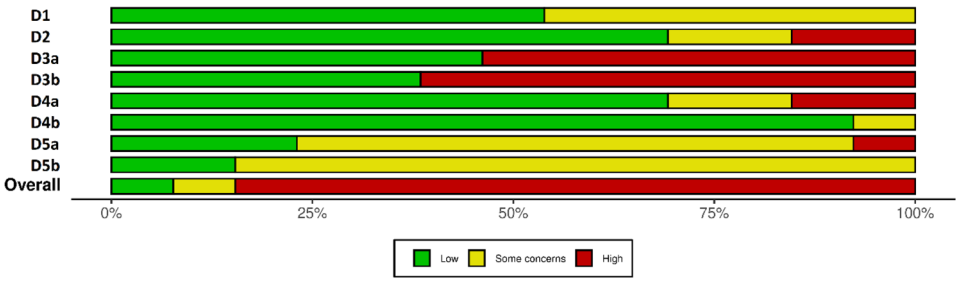

Fig. 2 Risk of bias graph for RCTs assessing clinical effectiveness and safety outcomes combined. *D1* randomization process; *D2* blinding of participants/personnel; *D3a* missing outcome data (fracture/safety); *D3b* missing outcome data (BMD); *D4a* measurement of the outcome (fracture/safety); *D4b* measurement of the outcome (BMD); *D5a* selective reporting (fracture/safety); *D5b* selective reporting (BMD). *BMD* bone mineral density; *RCTs* randomized controlled trials

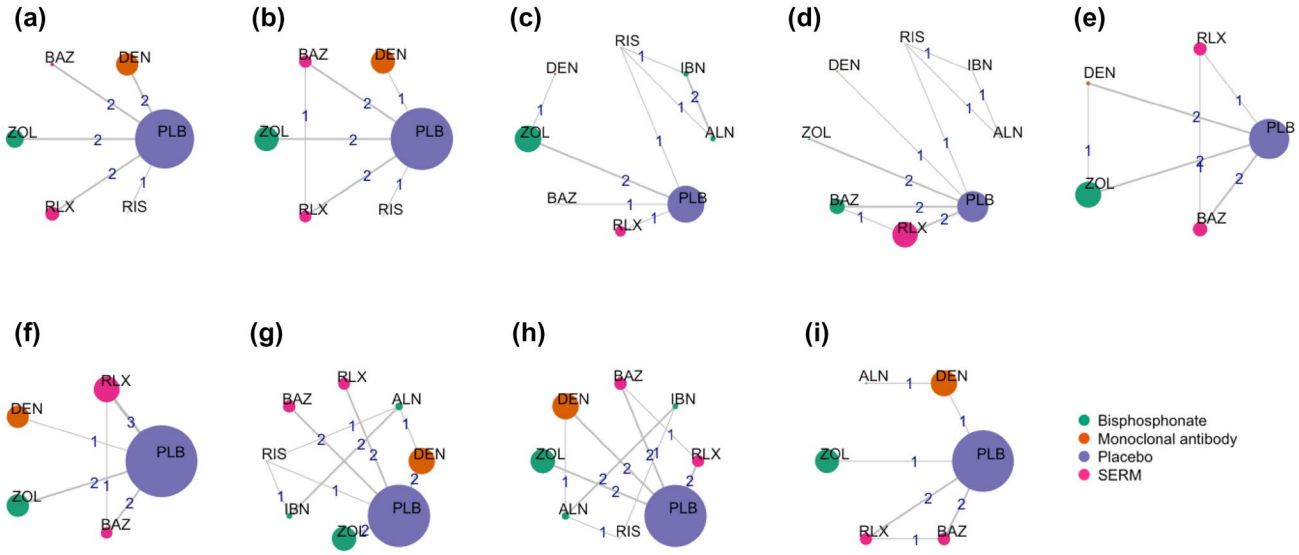

Fig. 3 Network diagrams. (**a**) vertebral fractures; (**b**) nonvertebral fractures; (**c**) FN BMD; (**d**) LS BMD; (**e**) TH BMD; (**f**) mortality; (**g**) AEs; (**h**) SAEs; (**i**) withdrawal due to AEs. *AE* adverse events; *ALN* alendronate; *BAZ* bazedoxifene; *BMD* bone mineral density; *DEN* denosumab; *FN* femoral neck; *IBN* ibandronate; *LS* lumbar spine; *PLB* placebo; *RIS* risedronate; *RLX* raloxifene; *SAE* serious adverse events; *SERM* selective estrogen receptor modulators; *TH* total hip; *ZOL* zoledronate

There was no evidence of inconsistency in the network (Supplementary Figure S5 & Table 23).

Total Hip (TH) BMD

Data on total hip (TH) BMD were available from seven RCTs that had a combined sample size of 13,666 (Fig. 3e) [47, 51, 55, 56, 59, 61, 67]. Denosumab (MD 4.55; 95% Crl 3.08, 6.05) significantly improved TH BMD after 19 (± 1 SD) months of treatment, relative to placebo (Fig. 5). Furthermore, denosumab resulted in a 2.67% (MD 2.67; 95% Crl 0.05, 5.23) and 3.07% (MD 3.07; 95% Crl 0.75, 5.21) improvement in TH BMD compared to the raloxifene and bazedoxifene, respectively (Supplementary Table S12). Denosumab had the highest probability (SUCRA = 89.01) of being the most effective treatment at increasing TH BMD and bazedoxifene had the lowest probability (SUCRA = 31.46) and was ranked as least effective active treatment. There was moderate total heterogeneity within the

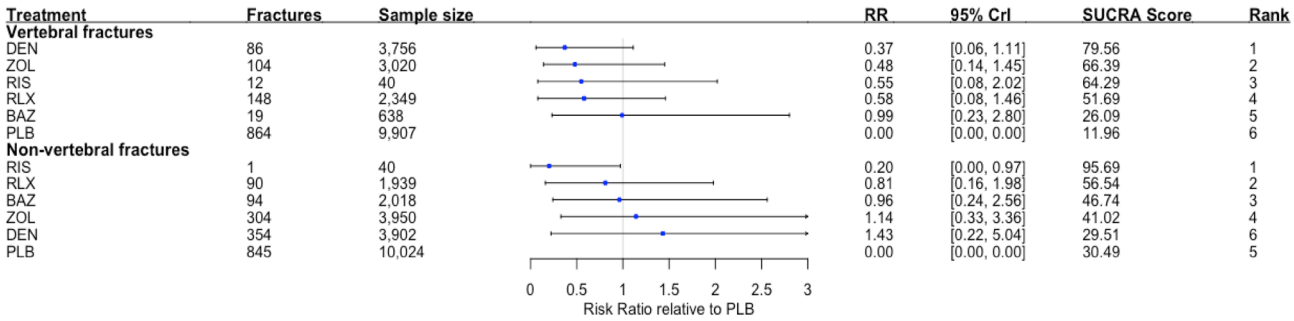

Fig. 4 Forest plot indicating the RR of vertebral fractures (relative to placebo). *BAZ* bazedoxifene; *CrI* credible interval; *DEN* denosumab; *PLB* placebo; *RIS* risedronate; *RLX* raloxifene; *RR* risk ratio; *SUCRA* surface under the cumulative ranking curve; *ZOL* zoledronate. *Credible interval (CrI)* interval within which RR values will fall with a specific probability; can be interpreted as a confidence interval [71]. Surface under the cumulative ranking curve (SUCRA): probability that a specific treatment is among the most effective options (i.e. 'best') in the network. A SUCRA value of 100% suggests that the treatment is the most effective treatment included in the network; a value of 0% suggests that the included treatment is the least effective treatment in the network [72]. Rank: position of treatment hierarchy within the network based on the SUCRA score, with 1 representing the most effective treatment

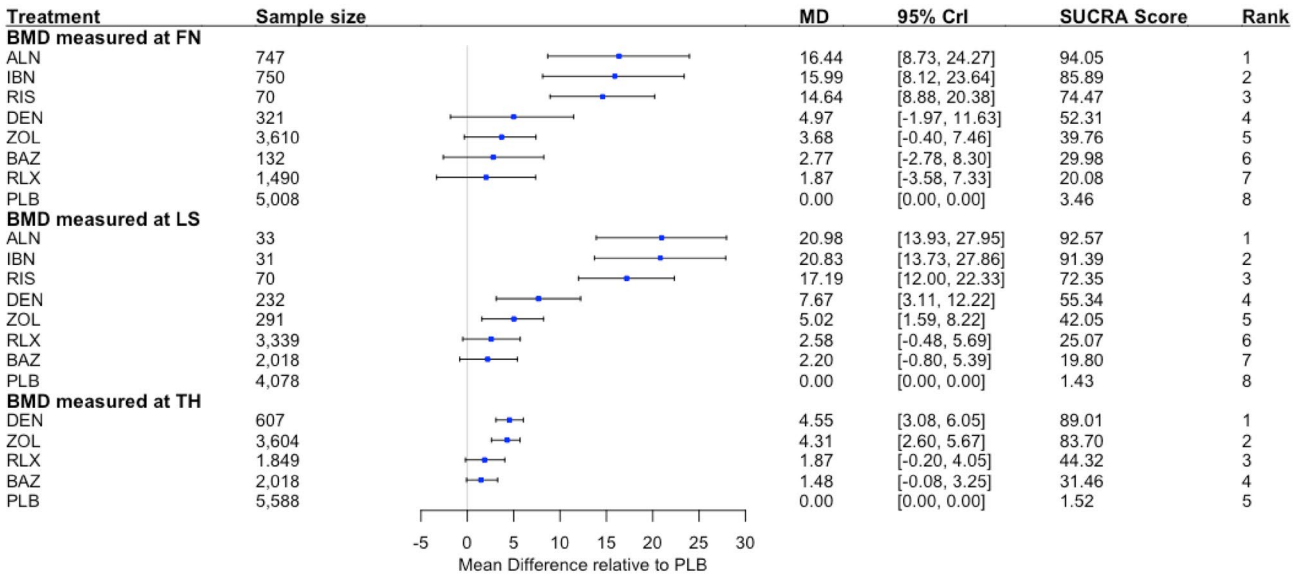

Fig. 5 Forest plot indicating the mean percentage difference in BMD (relative to placebo). *ALN* alendronate; *BAZ* bazedoxifene; *BMD* bone mineral density; *CrI* credible interval; *DEN* denosumab; *FN* femoral neck; *IBN* ibandronate; *LS* lumbar spine; *MD* mean difference; *PLB* placebo; *RIS* risedronate; *RLX* raloxifene; *SD* standard deviation; *SUCRA* surface under the cumulative ranking curve; *TH* total hip; *ZOL* zoledronate. *Credible interval (CrI)* interval within which RR values will fall with a specific probability. A credible interval can be interpreted as a confidence interval [71]. Surface under the cumulative ranking curve (SUCRA): probability that a specific treatment is among the most effective options (i.e. 'best') in the network. A SUCRA value of 100% suggests that the treatment is the most effective treatment included in the network; a SUCRA value of 0% suggests that the included treatment is the least effective treatment in the network [72]. Rank: position of treatment hierarchy within the network based on the SUCRA score, with 1 representing the most effective treatment. It is difficult to determine whether any of the statistically significant results are also clinically significant, as there is no verified scale that associates an increase in BMD with a decrease in the risk of vertebral or nonvertebral fractures [73–75]

network (Supplementary Table S26). Moreover, there was no evidence of local and global inconsistency (Supplementary Figure S6 & Table S25).

Safety

Mortality

Mortality data were available from seven RCTs that had a combined sample size of 26,882 (Fig. 3f) [47, 49, 51–53,

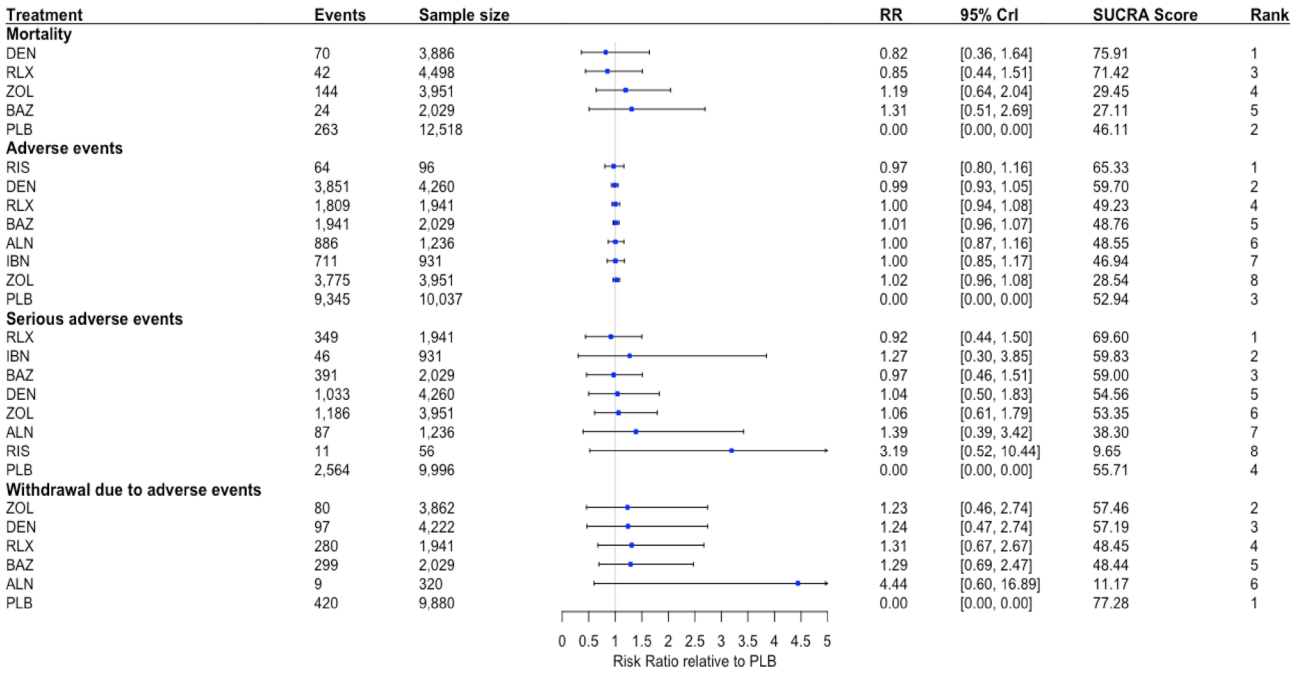

Fig. 6 Forest plot indicating the RR of safety outcomes (relative to placebo). *ALN* alendronate; *BAZ* bazedoxifene; *CrI* credible interval; *DEN* denosumab; *IBN* ibandronate; *PLB* placebo; *RIS* risedronate; *RLX* raloxifene; *RR* risk ratio; *SUCRA* surface under the cumulative ranking curve; *ZOL* zoledronate. *Credible interval (CrI)*: interval within which RR values will fall with a specific probability. A credible interval can be interpreted as a confidence interval [71]. Surface under the cumulative ranking curve (SUCRA): probability that a specific treatment is among the most effective options (i.e. 'best') in the network. A SUCRA value of 100% suggests that the treatment is the most effective treatment included in the network; a SUCRA value of 0% suggests that the included treatment is the least effective treatment in the network [72]. Rank: position of treatment hierarchy within the network based on the SUCRA score, with 1 representing the most effective treatment.

55, 56, 60, 66]. Relative to all included treatments, denosumab was not associated with statistically significant changes in mortality after 12 to 60 months. Similarly, none of the pairwise comparisons were statistically significant (Supplementary Table S13). Denosumab had the highest probability (SUCRA = 75.91) of having a favorable mortality profile, noting the difference in ranking between denosumab and raloxifene was negligible (Fig. 6). Overall, the network presented moderate evidence of statistical heterogeneity (Supplementary Table S28). There was no evidence of local inconsistency in the network (Supplementary Figure S7). Global inconsistency could not be estimated (Supplementary Table S27).

AE

AE data were available from 12 RCTs that had a combined sample size of 24,481 (Fig. 3g) [47, 49, 51, 55, 56, 58–62, 64]. Compared to all treatments, denosumab was not associated with statistically significant improvements within 12 to 36 months of treatment. Furthermore, none of the pairwise comparisons were statistically significant (Supplementary Table S14). Risedronate had the highest probability (SUCRA = 65.33) of having a favorable AE profile, with denosumab ranked as having the second highest probability (SUCRA = 59.70), noting that there were no differences in the reported relative effects across interventions (Fig. 6). The network did not show any significant evidence of statistical heterogeneity (Supplementary Table S30) or inconsistency (Supplementary Figure S8 & Table S29).

SAE

SAE data were available from 11 RCTs that had a combined sample size of 24,400 (Fig. 3h) [47, 49, 51, 55, 56, 58–62]. After 12 to 36 months of treatment, compared to all treatments, denosumab was not associated with statistically significant improvements. None of the pairwise comparisons were statistically significant (Supplementary Table S15). Raloxifene had the highest probability (SUCRA = 69.60) of having a favorable SAE profile (Fig. 6). Meanwhile, denosumab was ranked fifth (SUCRA = 54.56). The network did not show any significant evidence of statistical heterogeneity or inconsistency (Supplementary Figure S9, Table S31 & Table S32). Furthermore, sensitivity analysis suggested that

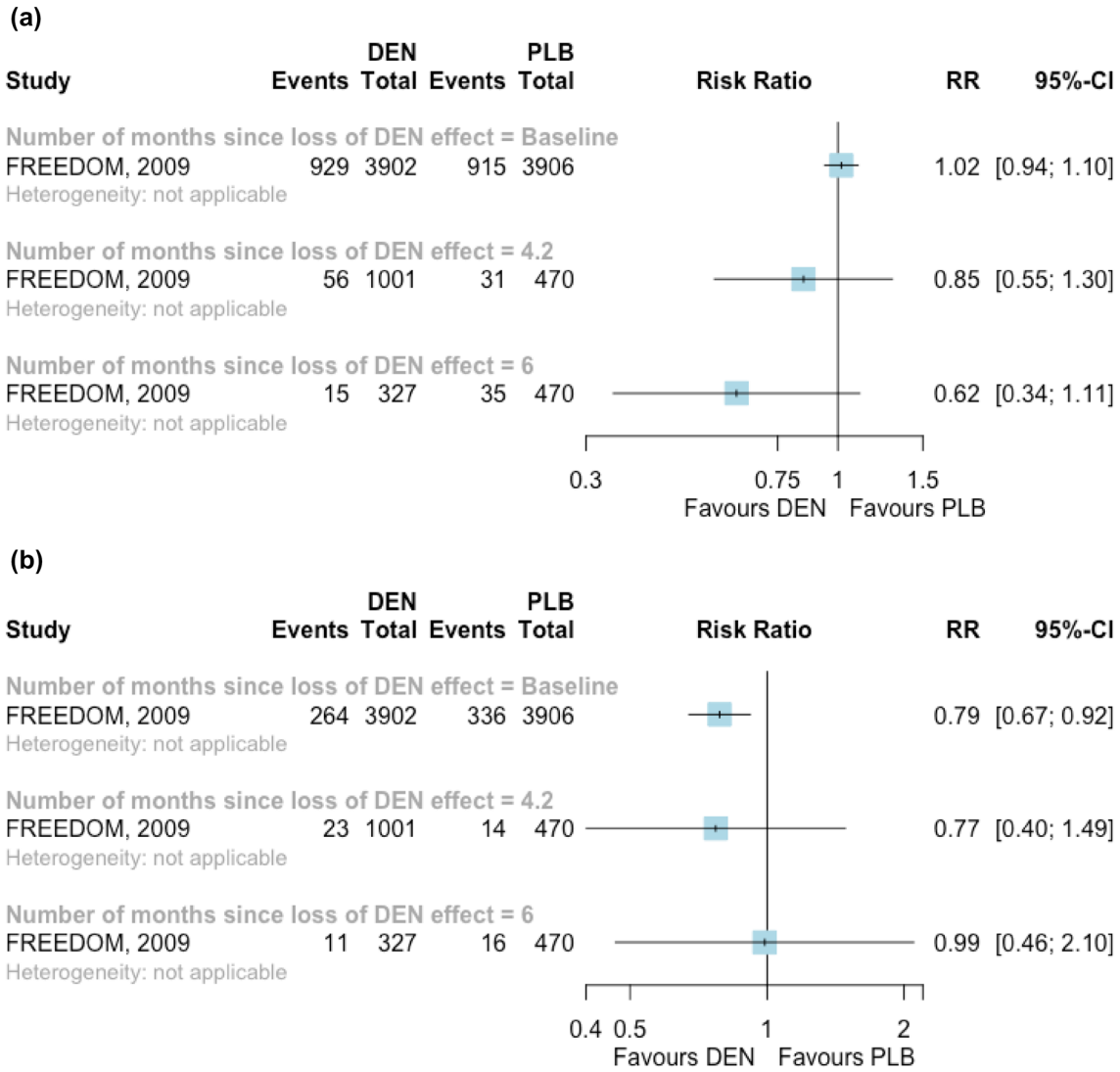

Fig. 7 Forest plots indicating RR of fracture outcomes after denosumab discontinuation (relative to placebo). (**a**) vertebral fractures; (**b**) nonvertebral fractures. *CI* confidence interval; *DEN* denosumab; *PLB* placebo; *RR* risk ratio. Loss of effect defined as 6 months post-last dose of denosumab. An additional 1-month study visit window was also reported [48, 50]

results were not impacted by reporting bias (Supplementary Table S35).

Study Withdrawal Due to Treatment-Related AEs

Data on study withdrawals due to treatment-related AEs were available from six RCTs that had a combined sample size of 22,254 (Fig. 3i) [46, 47, 49, 56, 59, 60]. After a treatment regimen of 12 to 36 months, compared to all treatments, denosumab was not associated with statistically significant improvements (Fig. 6). None of the pairwise comparisons were statistically significant (Supplementary Table S16). Zoledronate had the highest probability (SUCRA = 57.46) of having the least number of withdrawals due to treatment-related AEs (Fig. 6), with denosumab closely ranked as second with a comparably high probability (SUCRA = 57.19). The network did not show any significant evidence of statistical heterogeneity or inconsistency (Supplementary Figure S10, Table S33 & Table S34). Moreover, the sensitivity analysis suggested that results were not impacted by reporting bias (Supplementary Table S35).

AE Upon Denosumab Discontinuation

Data on AEs upon denosumab discontinuation were only available from a single RCT that compared denosumab to placebo for vertebral and nonvertebral fractures [48, 50]. The sample size was 7,808 at baseline, 1,471 at 4.2 months after a loss of denosumab treatment effect and 797 at 6 months

(median) [48, 50]. Loss of the denosumab treatment effect occurs approximately 6 months after the last dose.

Overall, there was likely no significant change in vertebral (Fig. 7a) and nonvertebral fracture (Fig. 7b) rates between denosumab and placebo at 4.2 and 6 months after the loss of the denosumab treatment effect, noting the evidence for this outcome is highly uncertain.

Discussion

A comprehensive literature search identified 12 RCTs ($k=22$ publications) comparing denosumab to bisphosphonates (alendronate, ibandronate, risedronate, zoledronate), SERMs (bazedoxifene, raloxifene) and placebo in postmenopausal women.

Vertebral fracture, mortality, AEs, SAEs and withdrawal due to treatment-related AEs reported no significant differences for any intervention compared to placebo. Contrastingly, risedronate was found to be statistically significant at preventing nonvertebral fractures. Denosumab was statistically significant at improving BMD compared to placebo when measured at the LS and TH. It is important to note that it is difficult to determine whether any of the statistically significant BMD results are also clinically significant, as there is no verified scale that associates an increase in BMD with a decrease in the risk of vertebral or nonvertebral fractures [73–75].

Network meta-analyses could not be conducted on the published data available on AEs upon denosumab discontinuation. There appears to be no significant change in vertebral and nonvertebral fracture rates between baseline and 4.2 to 6 months after denosumab discontinuation (loss of effect). It is important to note, however, that there were significant losses to follow-up at both timepoints in each treatment, and the results are from a single study. As such, the results presented are subject to considerable uncertainty and should be interpreted with caution.

The sensitivity analyses conducted to evaluate the impact of reporting bias on the results were consistent with the main analyses for fractures (vertebral and nonvertebral), mortality, AEs, SAEs and withdrawal due to treatment-related AEs. The impact of reporting bias on the main analysis could not be assessed for BMD at LS and FN, as none of the included trials posed a low RoB in this category. Similarly, the effects of attrition bias and selection bias on the main analyses could not be assessed for the included outcomes as none of the trials presented a low bias in either of these categories. Sensitivity analyses could not be conducted to determine the impact of bias on AEs associated with denosumab discontinuation, as each pairwise meta-analyses only included a single trial.

There was low to moderate heterogeneity for all but one of the network meta-analyses conducted. The network meta-analysis conducted on LS BMD showed substantial to considerable heterogeneity. There was no evidence of strong inconsistency in any of the meta-analyses conducted.

A single meta-regression found that there was a slight association between the effectiveness of denosumab and postmenopausal age. The regression indicated that denosumab was slightly less effective at preventing SAEs in older postmenopausal women (aged 75 to 85). Additional meta-regressions on the remaining outcomes could not be conducted as each network included fewer than 10 trials.

No publication bias was identified in the network meta-analyses that were conducted on AEs and SAEs (Supplementary Fig. 11 & 12). Publication bias could only be assessed in these two outcomes as they were the only network meta-analyses that included the minimum 10 trials required.

In general, the findings of this review complement a previously reported network meta-analysis by Simpson et al. 2020 [76]. The direction of treatment effect when comparing denosumab to placebo for vertebral fractures and BMD measured at the FN are generally in accordance with the findings of Simpson et al. 2020 [76]. Contrastingly, this network meta-analysis utilized RR and a random-effects model to assess dichotomous outcomes, while Simpson et al. 2020 utilized hazards ratios [76]. This resulted in the analysis by Simpson et al. 2020 not being able to account for the timing of treatment benefits post-fracture [76].

There are several key limitations of this review. First, many of the meta-analyses suffer from statistical imprecision due to small sample sizes in certain treatment arms. Low samples sizes can result in wide uncertainty margins that do not accurately reflect the *true effect* of the treatment (compared to placebo). Finally, it is still unclear how heterogeneity and inconsistency affect the findings of a network meta-analysis. Mills et al. 2013 reported that it is unclear how moderate to considerable levels of heterogeneity and inconsistency may impact the reliability of results (generated from both direct and indirect evidence) [41].

It will be of great interest to conduct further research to address the effects of denosumab on HRQoL, FRAX®, TRO BMD and AEs upon discontinuation in postmenopausal women with osteoporosis within WHO-Mortality Stratum A countries. Due to the lack of applicable evidence for these outcomes, an evidence-based decision cannot be made to inform policy decisions in developed countries.

Conclusion

In conclusion, denosumab had varying impact on clinical outcomes in post-menopausal women with osteoporosis. Relative to placebo, denosumab was associated with significant improvements in BMD measured at both the LS and TH. Denosumab also resulted in an improvement in TH BMD compared to the raloxifene and bazedoxifene. However, relative to denosumab, alendronate, ibandronate and risedronate resulted in significant improvements in LS BMD. Moreover, compared to all treatment, denosumab was not associated with statistically significant improvement in fractures (vertebral and nonvertebral). Finally, regarding safety outcomes, relative to all included treatments (i.e. SERMs, placebo, bisphosphonates), denosumab was not associated with statistically significant changes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00223-023-01078-z.

Acknowledgements The authors would like to Virginie Gaget for her assistance with collating the background information.

Author contributions MM, KN, MJ, and TV: designed the study. MM and KN: prepared the manuscript. MM, KN, DS, and TV: contributed to the data collection and analysis. MM and NM: were responsible for statistical analysis of the data. All authors revised the paper critically for intellectual content and approved the final version. All authors agree to be accountable for the work and to ensure that any questions relating to the accuracy and integrity of the paper are investigated and properly resolved.

Funding The research project was funded by the Swiss Federal Office of Public Health (FOPH).

Data availability All data generated or analyzed during this study are included in this published article or in the data repositories listed in the references.

Declarations

Conflict of interest Magdalena Moshi, Konstance Nicolopoulos, Danielle Stringer, Mathias Jenal, Ning Ma, Thomas Vreugdenburg has no conflicts to disclose.

informed consent No informed consent was not required as this was secondary research.

References

1. Heaney RP (1992) The natural history of vertebral osteoporosis is low bone mass an epiphenomenon? Bone 13(2):S23-26
2. Kanis JA (1990) Osteoporosis and osteopenia. J bone mineral res 5:209–211
3. Riggs BL, Melton Iii L (1995) The worldwide problem of osteoporosis: insights afforded by epidemiology. Bone 17:S505–S511
4. Bonnick SL, Shulman L (2006) Monitoring osteoporosis therapy: bone mineral density, bone turnover markers, or both? Am J Med 119:S25–S31
5. Lane NE (2006) Epidemiology, etiology, and diagnosis of osteoporosis. Am J Obstet Gynecol 194:S3-11
6. Naylor K, Eastell R (2012) Bone turnover markers: use in osteoporosis. Nat Rev Rheumatol 8:379–389
7. Stone KL, Seeley DG, Lui LY, Cauley JA, Ensrud K, Browner WS, Nevitt MC, Cummings SRJJOB, Research M (2003) BMD at multiple sites and risk of fracture of multiple types: long-term results from the study of osteoporotic fractures. J bone and mineral res 18:1947–1954
8. Cooper C, Campion G, Melton LJ 3rd (1992) Hip fractures in the elderly: a world-wide projection. Osteoporos Int 2:285–289
9. Natural Therapy Pages (2019) World Osteoporosis Day 2019. In: World Osteoporosis Day 2019. Natural Therapy Pages
10. Khosla S, Melton LJ 3rd, Riggs BL (2002) Clinical review 144: estrogen and the male skeleton. J Clin Endocrinol Metab 87:1443–1450
11. Poole KE, Compston JE (2006) Osteoporosis and its management. Bmj 333:1251–1256
12. Barry P, Aspray T, Briers K (2012) Osteoporosis: assessing the risk of fragility fracture (NICE Clinical Guidance 146. National Institute for Health and Care Excellence
13. Kanis JA, Oden A, Johnell O, Jonsson B, de Laet C, Dawson A (2001) The burden of osteoporotic fractures: a method for setting intervention thresholds. Osteoporos Int 12:417–427
14. Organization WH (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: report of a WHO study group [meeting held in Rome from 22 to 25 June 1992]. World Health Organization
15. Darbà J, Kaskens L, Pérez-Álvarez N, Palacios S, Neyro JL, Rejas J (2015) Disability-adjusted-life-years losses in postmenopausal women with osteoporosis: a burden of illness study. BMC Public Health 15:324
16. Ioannidis G, Papaioannou A, Hopman WM, Akhtar-Danesh N, Anastassiades T, Pickard L, Kennedy CC, Prior JC, Olszynski WP, Davison KS (2009) Relation between fractures and mortality: results from the Canadian multicentre osteoporosis study. CMAJ 181:265–271
17. Sözen T, Özışık L, Başaran NÇ (2017) An overview and management of osteoporosis. Eur J Rheumatol 4:46–56
18. Born R, Zwahlen M (2013) Disparities in bone density measurement history and osteoporosis medication utilisation in Swiss women: results from the Swiss Health Survey 2007. BMC Musculoskelet Disord 14:10
19. Bundesamt für Gesundheit (2021) Spezialitätenliste. http://www.spezialitaetenliste.ch/Default.aspx. Accessed 24 May 2021
20. United Nations Department of Economic and Social Affairs (UN DESA), United Nations Conference on Trade and Development (UNCTAD), Economic Commission for Africa (UNECA), Economic Commission for Europe (UNECE), Economic Commission for Latin America and the Caribbean (UNECLAC) EaS-CfAatPU, Economic and Social Commission for Western Asia (UNESCWA), report. UNWTOUactt (2022) World Economic Situation and Prospects 2022
21. Lewiecki EM (2018) New and emerging concepts in the use of denosumab for the treatment of osteoporosis. Ther Adv Musculoskelet Dis 10:209–223
22. Livio F (2018) Pharmacovigilance update. Rev Med Suisse 14:81–84
23. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JPA, Straus S, Thorlund K, Jansen JP, Mulrow C, Catalá-López F, Gøtzsche PC, Dickersin K, Boutron I, Altman DG, Moher D (2015) The PRISMA extension statement for reporting of systematic reviews incorporating network

meta-analyses of health care interventions: checklist and explanations. Ann Intern Med 162:777–784
24. World Health Organization (2020) List of Member States by WHO Region and Mortality Stratum. In: Cost effectiveness and strategic planning (WHO-CHOICE). World Health Organization
25. Fleiss JL (1971) Measuring nominal scale agreement among many raters. Psychol Bull 76:378
26. Tripto-Shkolnik L, Rouach V, Marcus Y, Rotman-Pikielny P, Benbassat C, Vered I (2018) Vertebral fractures following denosumab discontinuation in patients with prolonged exposure to bisphosphonates. Calcif Tissue Int 103:44–49
27. Bone HG, Bolognese MA, Yuen CK, Kendler DL, Miller PD, Yang Y-C, Grazette L, San Martin J, Gallagher JC (2011) Effects of denosumab treatment and discontinuation on bone mineral density and bone turnover markers in postmenopausal women with low bone mass. J Clin Endocrinol Metab 96:972–980
28. Ankit Rohatgi (2020) WebPlotDigitizer. https://automeris.io/WebPlotDigitizer/. Accessed 8 Feb 2020
29. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernan MA, Hopewell S, Hrobjartsson A, Junqueira DR, Juni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 366:l4898. https://doi.org/10.1136/bmj.l4898
30. Thorlund K, Thabane L, Mills EJ (2013) Modelling heterogeneity variances in multiple treatment comparison meta-analysis—are informative priors the better solution? BMC Med Res Methodol 13:2
31. van Valkenhoef G, Lu G, de Brock B, Hillege H, Ades AE, Welton NJ (2012) Automating network meta-analysis. Res Synth Methods 3:285–299
32. Béliveau A, Boyne DJ, Slater J, et al (2019) BUGSnet: an R package to facilitate the conduct and reporting of Bayesian-network Meta-analyses. BMC Med Res Methodol 19(1):196. https://doi.org/10.1186/s12874-019-0829-2
33. Higgins JPT, J.; Chandler, J.; Cumpston, M.; Li, T; Page, M.J.; Welch, V.A. (editors) (2020) Imputing standard deviations for changes from baseline. https://handbook-5-1.cochrane.org/chapter_16/16_1_3_2_imputing_standard_deviations_for_changes_from_baseline.htm
34. Bracken M (1992) Statistical methods for analysis of effects of treatment in overviews of randomized trials. Effective care of the newborn infant. 1992:13–20
35. Lajeunesse MJ (2016) Facilitating systematic reviews, data extraction and meta-analysis with the metagear package for r. Methods Ecol Evol 7:323–330
36. R Core Team (2020) R: A language and environment for statistical computing. In: Foundation for Statistical Computing, Vienna, Austria
37. RStudio Team (2020) RStudio: Integrated Development for R. In:RStudio, Boston, USA
38. Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 7:434–455
39. Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472
40. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V (2020) Cochrane Handbook for Systematic Reviews of Interventions version 6.1. Cochrane. https://training.cochrane.org/handbook. Accessed Sep 2020
41. Mills E, Thorlund K, Ioannidis J (2013) Demystifying trial networks and network meta-analysis. BMJ (Clinical research ed) 346:f2914
42. University of Cambridge, MRC Biostatistics Unit (n.d.) DIC: Deviance Information Criteria. http://www.mrcbsu.cam.ac.uk/software/bugs/the-bugs-project-dic/. Accessed Oct 2021
43. Krahn U, Binder H, König J (2013) A graphical tool for locating inconsistency in network meta-analyses. BMC Med Res Methodol 13:35
44. Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G (2013) Graphical tools for network meta-analysis in STATA. PLoS ONE 8:e76654
45. Chaimani A, Salanti G (2012) Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. Res Synth Methods 3:161–176
46. Adachi JD, Bone HG, Daizadeh NS, Dakin P, Papapoulos S, Hadji P, Recknor C, Bolognese MA, Wang A, Lin CJF, Wagman RB, Ferrari S (2017) Influence of subject discontinuation on long-term nonvertebral fracture rate in the denosumab FREEDOM Extension study. BMC Musculoskelet Disord 18:174
47. Black DM, Delmas PD, Eastell R, Reid IR, Boonen S, Cauley JA, Cosman F, Lakatos P, Leung PC, Man Z, Mautalen C, Mesenbrink P, Hu H, Caminis J, Tong K, Rosario-Jansen T, Krasnow J, Hue TF, Sellmeyer D, Eriksen EF, Cummings SR, Trial HPF (2007) Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. N Engl J Med 356:1809–1822
48. Brown JP, Roux C, Torring O, Ho PR, Beck Jensen JE, Gilchrist N, Recknor C, Austin M, Wang A, Grauer A, Wagman RB (2013) Discontinuation of denosumab and associated fracture incidence: analysis from the fracture reduction evaluation of denosumab in osteoporosis every 6 months (FREEDOM) trial. J bone mineral res 28:746–752
49. Christiansen C, Chesnut CH 3rd, Adachi JD, Brown JP, Fernandes CE, Kung AW, Palacios S, Levine AB, Chines AA, Constantine GD (2010) Safety of bazedoxifene in a randomized, double-blind, placebo- and active-controlled Phase 3 study of postmenopausal women with osteoporosis. BMC Musculoskelet Disord 11:130
50. Cummings SR, Ferrari S, Eastell R, Gilchrist N, Jensen JB, McClung M, Roux C, Torring O, Valter I, Wang AT, Brown JP (2018) Vertebral fractures after discontinuation of denosumab: a post hoc analysis of the randomized placebo-controlled FREEDOM trial and its extension. J bone mineral res 33:190–198
51. Cummings SR, San Martin J, McClung MR, Siris ES, Eastell R, Reid IR, Delmas P, Zoog HB, Austin M, Wang A, Kutilek S, Adami S, Zanchetta J, Libanati C, Siddhanti S, Christiansen C, Trial F (2009) Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med 361:756–765
52. de Villiers TJ, Chines AA, Palacios S, Lips P, Sawicki AZ, Levine AB, Codreanu C, Kelepouris N, Brown JP (2011) Safety and tolerability of bazedoxifene in postmenopausal women with osteoporosis: results of a 5-year, randomized, placebo-controlled phase 3 trial. Osteoporos Int 22:567–576
53. Delmas PD, Ensrud KE, Adachi JD, Harper KD, Sarkar S, Gennari C, Reginster JY, Pols HA, Recker RR, Harris ST, Wu W, Genant HK, Black DM, Eastell R, Mulitple Outcomes of Raloxifene Evaluation I (2002) Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. J Clin Endocrinol Metab 87:3609–3617
54. Ettinger B, Black DM, Mitlak BH, Knickerbocker RK, Nickelsen T, Genant HK, Christiansen C, Delmas PD, Zanchetta JR, Stakkestad J, Gluer CC, Krueger K, Cohen FJ, Eckert S, Ensrud KE, Avioli LV, Lips P, Cummings SR (1999) 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 282:637–645
55. Greenspan SL, Perera S, Ferchak MA, Nace DA, Resnick NM (2015) Efficacy and safety of single-dose zoledronic acid for

osteoporosis in frail elderly women: a randomized clinical trial. JAMA Intern Med 175:913–921
56. Itabashi A, Yoh K, Chines AA, Miki T, Takada M, Sato H, Gorai I, Sugimoto T, Mizunuma H, Ochi H, Constantine GD, Ohta H (2011) Effects of bazedoxifene on bone mineral density, bone turnover, and safety in postmenopausal Japanese women with osteoporosis. J bone mineral res 26:519–529
57. Jacques RM, Boonen S, Cosman F, Reid IR, Bauer DC, Black DM, Eastell R (2012) Relationship of changes in total hip bone mineral density to vertebral and nonvertebral fracture risk in women with postmenopausal osteoporosis treated with once-yearly zoledronic acid 5 mg: the HORIZON-Pivotal fracture trial (PFT). J J Bone Miner Res 27:1627–1634
58. Miller PD, Epstein S, Sedarati F, Reginster JY (2008) Once-monthly oral ibandronate compared with weekly oral alendronate in postmenopausal osteoporosis: results from the head-to-head MOTION study. Curr Med Res Opin 24:207–213
59. Miller PD, Pannacciulli N, Brown JP, Czerwinski E, Nedergaard BS, Bolognese MA, Malouf J, Bone HG, Reginster JY, Singer A, Wang C, Wagman RB, Cummings SR (2016) Denosumab or zoledronic acid in postmenopausal women with osteoporosis previously treated with oral bisphosphonates. J Clin Endocrinol Metab 101:3163–3170
60. Morii H, Ohashi Y, Taketani Y, Fukunaga M, Nakamura T, Itabashi A, Sarkar S, Harper K (2003) Effect of raloxifene on bone mineral density and biochemical markers of bone turnover in Japanese postmenopausal women with osteoporosis: results from a randomized placebo-controlled trial. Osteoporos Int 14:793–800
61. Nakamura T, Matsumoto T, Sugimoto T, Shiraki M (2012) Dose-response study of denosumab on bone mineral density and bone turnover markers in Japanese postmenopausal women with osteoporosis. Osteoporos Int 23:1131–1140
62. Paggiosi MA, Peel N, McCloskey E, Walsh JS, Eastell R (2014) Comparison of the effects of three oral bisphosphonate therapies on the peripheral skeleton in postmenopausal osteoporosis: the TRIO study. Osteoporos Int 25:2729–2741
63. Palacios S, Silverman SL, de Villiers TJ, Levine AB, Goemaere S, Brown JP, De Cicco NF, Williams R, Hines TL, Mirkin S, Chines AA, Bazedoxifene Study G (2015) A 7-year randomized, placebo-controlled trial assessing the long-term efficacy and safety of bazedoxifene in postmenopausal women with osteoporosis: effects on bone density and fracture. Menopause 22:806–813
64. Palomba S, Manguso F, Orio F Jr, Russo T, Oppedisano R, Sacchinelli A, Falbo A, Tolino A, Zullo F, Mastrantonio P (2008) Effectiveness of risedronate in osteoporotic postmenopausal women with inflammatory bowel disease: a prospective, parallel, open-label, two-year extension study. Menopause 15:730–736
65. Palomba S, Orio F Jr, Manguso F, Falbo A, Russo T, Tolino A, Tauchmanova L, Colao A, Doldo P, Mastrantonio P, Zullo F (2005) Efficacy of risedronate administration in osteoporotic postmenopausal women affected by inflammatory bowel disease. Osteoporos Int 16:1141–1149
66. Silverman SL, Chines AA, Kendler DL, Kung AW, Teglbjaerg CS, Felsenberg D, Mairon N, Constantine GD, Adachi JD, Bazedoxifene Study G (2012) Sustained efficacy and safety of bazedoxifene in preventing fractures in postmenopausal women with osteoporosis: results of a 5-year, randomized, placebo-controlled study. Osteoporos Int 23:351–363
67. Silverman SL, Christiansen C, Genant HK, Vukicevic S, Zanchetta JR, de Villiers TJ, Constantine GD, Chines AA (2008) 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 mineral res: official J American Soc Bone Mineral Res 23:1923–1934
68. Cauley JA, Norton L, Lippman ME, Eckert S, Krueger KA, Purdie DW, Farrerons J, Karasik A, Mellstrom D, Ng KW, Stepan JJ, Powles TJ, Morrow M, Costa A, Silfen SL, Walls EL, Schmitt H, Muchmore DB, Jordan VC, Ste-Marie LG (2001) Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Multiple outcomes of raloxifene evaluation. Breast Cancer Res Treat 65:125–134
69. Siris E, Adachi JD, Lu Y, Fuerst T, Crans GG, Wong M, Harper KD, Genant HK (2002) Effects of raloxifene on fracture severity in postmenopausal women with osteoporosis: results from the MORE study. multiple outcomes of raloxifene evaluation. Osteoporos Int 13:907–913
70. Uusi-Rasi K, Beck TJ, Semanick LM, Daphtary MM, Crans GG, Desaiah D, Harper KD (2006) Structural effects of raloxifene on the proximal femur: results from the multiple outcomes of raloxifene evaluation trial. Osteoporos Int 17:575–586
71. Dias S, Caldwell DM (2019) Network meta-analysis explained. Arch Dis Childhood - Fetal Neonatal Edition 104:F8
72. Rouse B, Chaimani A, Li T (2017) Network meta-analysis: an introduction for clinicians. Intern Emerg Med 12:103–111
73. European Medicines Agency (2006) Committee for Medicinal Products for Human Use (CHMP): Guideline on the evaluation of medicinal products in the treatment of primary osteoporosis. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatmentprimary-osteoporosis_en.pdf. Accessed 21 July 2021
74. Cranney A, Welch V, Wells G, Adachi J, Shea B, Simon L, Tugwell P (2001) Discrimination of changes in osteoporosis outcomes. J Rheumatol 28:413–421
75. Guralnik J, Bandeen-Roche K, Bhasin SAR, Eremenco S, Landi F, Muscedere J, Perera S, Reginster JY, Woodhouse L, Vellas B (2020) Clinically meaningful change for physical performance: perspectives of the ICFSR task force. J Frailty Aging 9:9–13
76. Simpson EL, Martyn-St James M, Hamilton J, Wong R, Gittoes N, Selby P, Davis S (2020) Clinical effectiveness of denosumab, raloxifene, romosozumab, and teriparatide for the prevention of osteoporotic fragility fractures: a systematic review and network meta-analysis. Bone 130:115081

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.