Q J Med 2005; **98**:403–413
doi:10.1093/qjmed/hci070
Advance Access publication 6 May 2005

Review

QJM

Osteoporosis therapy: an example of putting evidence-based medicine into clinical practice

D.J. HOSKING[1], P. GEUSENS[2] and R. RIZZOLI[3]

From the [1]Nottingham City Hospital, Nottingham, United Kingdom, [2]Biomedical Research Institute—LUC, Diepenbeek, Belgium, and [3]Division of Bone Diseases, University Hospital, Geneva, Switzerland

Summary

A major aim of evidence-based medicine is to assist clinical decision-making by providing the most current and reliable medical information. Systematic reviews and meta-analyses are important tools in this process. Systematic reviews identify and compile relevant evidence, while meta-analyses summarize and quantify the results of such reviews. Results from meta-analyses are, at present, the main source of summary evidence for the efficacy of treatments for a specific condition. They are important tools for helping to choose among treatment options, although they cannot be used to directly compare the magnitude of the effect of various therapies. However, the methods used and the consequent clinical value of the results, may be poorly understood by clinicians, who may therefore not take full advantage of the evidence.

Recently, a panel of experts in osteoporosis and evidence-based medicine applied rigorous, validated, scientific standards to produce a systematic review and meta-analysis of randomized controlled trials of anti-resorptive agents used to treat osteoporosis. They found that, although several agents reduced the risk of vertebral fracture, only two, alendronate and risedronate, demonstrated convincing evidence for both non-vertebral and vertebral fracture risk reductions. The clinical implication of these results is that there are important differences in anti-fracture efficacy among currently available agents. In the absence of evidence from head-to-head clinical trials and because of the systematic nature and methodological rigor of the analyses, these data provide important information for clinical decision-making.

Introduction

In recent years, emphasis on 'evidence-based' clinical decision making has increased, but some confusion about the principles and their use remains. The practice of evidence-based medicine involves the application, in the management of individual patients, of the best available scientific information. Large, well-conducted clinical trials or meta-analyses of these trials are considered to provide the best quality evidence. Systematic reviews performed according to standard guidelines provide quantitative, objective summaries of the available evidence and thereby provide important information to the clinician.[1]

Address correspondence to Dr D.J. Hosking, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB. e-mail: dhosking@ncht.trent.nhs.uk

Osteoporosis is widely prevalent in the older population, and frequently results in fragility fractures, which are associated with morbidity, mortality, considerable health care costs, and diminished quality of life.[2,3] Both men and women experience bone loss with age that can lead to osteoporosis, but postmenopausal women are particularly at risk because of the substantial decline in bone mass and changes in bone architecture that are associated with oestrogen deficiency. In the past decade, well-designed clinical trials have shown that several treatments decrease fracture risk in postmenopausal women with low bone mass, and clinicians now must evaluate the relative merits of these options in order to make informed decisions about therapy.

Recently, a team of experts in evidence-based medicine methodology and osteoporosis research (The Osteoporosis Methodology Group, OMG, and The Osteoporosis Research Advisory Group, ORAG) conducted a series of systematic reviews and meta-analyses of antiresorptive osteoporosis therapies.[4–12] The OMG/ORAG meta-analyses collectively represent a comprehensive summary of the efficacy of treatments for osteoporosis. The objectives of the current article are to describe why systematic reviews and meta-analyses provide the highest level of evidence, and use this information to illustrate its application to the current treatments for osteoporosis.

Hierarchy of clinical evidence

Scientific evidence varies in strength and reliability, and can be represented as a hierarchy of relative quality, as presented in Figure 1.[13] This hierarchy is arranged according to sound statistical principles, and evidence used in decision-making should come from as high in the hierarchy as possible. Large, well-conducted randomized controlled trials (RCTs) and meta-analyses of results of multiple large RCTs occupy the top of the hierarchy and, ideally, medical decisions should be based on the results of such studies. In contrast, personal experience, even that of experts or respected authorities, provides the poorest quality evidence, because it is the most subjective and prone to bias. Although RCTs have certain limitations, they are considered to provide the best evidence of the effectiveness of a particular therapy. Randomization helps ensure that patients in the active treatment and control groups have similar characteristics at baseline, while blinded evaluation of outcome measures according to uniform criteria reduces the chance of bias in the analysis and interpretation of the trial results. As a consequence, differences observed between treatment and control groups can be attributed confidently to the effect of the treatment rather than to other factors such as the characteristics of the participants or to incidental confounders, which should be equally distributed among groups.[14]

Even among RCTs, the strength of reported evidence can vary. Large trials are more able to identify treatment effects than smaller trials, and provide more precise estimates of the magnitude of any such effect. Smaller RCTs are more commonly available, but the results may be conflicting or inconclusive because the statistical power may be inadequate to permit precise estimates of effect. The strongest evidence derived from a single study comes from analyses of primary and secondary outcomes that were specified before the study began. '*Post hoc*' and subgroup analyses have less value, because the chance of identifying spurious differences (coincidences) becomes greater when multiple analyses are performed, and because randomization was not based on such subgroups.[15] Observational studies, including case reports, case series, cross-sectional surveys, case-control studies, and prospective cohort studies are subject to a variety of possible biases. Fundamentally, the bias reflects the fact that study groups are not similar in all respects except for the treatment.[16,17]

Meta-analysis combines results from multiple trials, some of which may have had limited ability to demonstrate a treatment effect because of their small size. It increases the effective size of the study population and thereby the power to detect and quantify a treatment effect.

Systematic review and meta-analyses of osteoporosis therapies

Inclusion criteria and assessment of methodological quality

The OMG/ORAG investigators defined, and strictly adhered to, detailed pre-specified methods for the systematic review, to help ensure that the reviews would be comprehensive with reduced chance of bias. Only RCTs, the highest level of evidence, were included and the methodological quality of each trial was determined using formal objective criteria. Studies were included in the review only if deemed eligible by multiple independent reviewers, and four aspects of trial design and conduct that affect validity were evaluated for each.

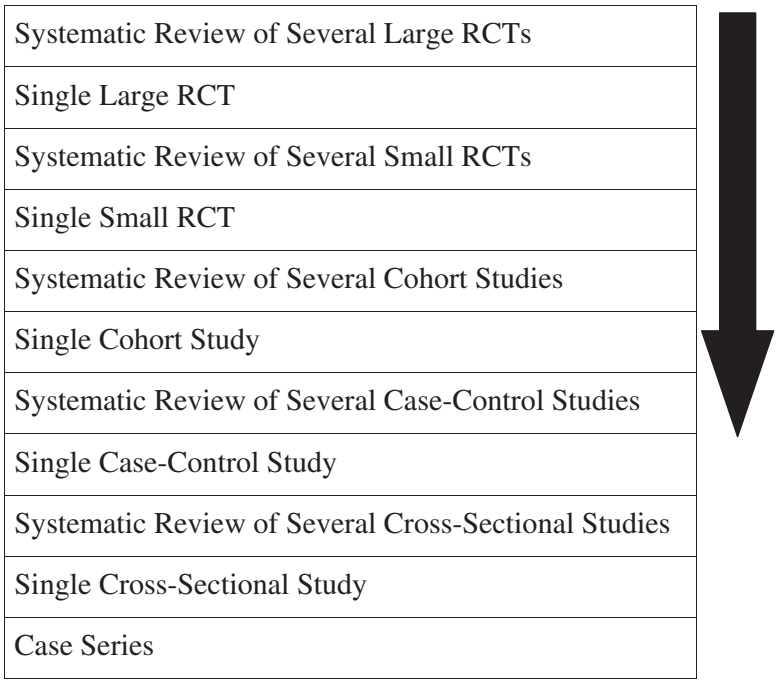

Figure 1. Hierarchy of clinical evidence.

(i) Concealment of randomization

Individuals responsible for determining eligibility should not know which treatment participants will be receiving if they enter the trial. Trials that inadequately conceal patient allocation tend to overestimate treatment effects.[18]

(ii) Blinding to treatment

Trials are more likely to yield unbiased results if participants and investigators, including those who assess outcomes, are unaware of treatment allocation.[19]

(iii) Completeness of follow-up

Trials with large losses to follow-up (defined as >20% for these meta-analyses) are at risk of losing the balance of baseline characteristics initially achieved by randomization. Although the investigators did not explain their choice of threshold, losses >20% are typically considered to jeopardize the validity of trial results.[20] If participants with fractures preferentially drop out of the placebo group in osteoporosis trials, the observed results may underestimate the treatment effect, whereas the converse will be the case if fracture drop-outs are concentrated in the treatment group. The effects of this imbalance cannot be calculated or corrected for by any statistical method, and results from such studies must be interpreted with great caution.[21]

(iv) Analysis according to intention-to-treat (ITT)

The ITT principle requires that all participants be analysed in the group to which they were randomized, irrespective of whether they received the intended treatment or completed the trial. Although this may be counter-intuitive, it is a conservative approach because it tends to underestimate the treatment effect. Failure to use ITT (for example, including only participants who took all study medication in the primary analysis) can introduce bias, because it can destroy the original randomization by selecting subgroups, and lead to imbalance in baseline characteristics.[21,22]

Scrupulous attention to methodology is essential to ensure the validity of a meta-analysis. RCTs of osteoporosis treatment often use multiple doses of drugs and different durations of treatment. Pooling of studies simplifies the systematic review and increases the statistical power of the meta-analysis, but is considered appropriate only if the same magnitude of effect can be demonstrated across a range of patients, interventions, and outcomes. This is evaluated statistically as heterogeneity, by comparing a model that includes a variable for dose or year of treatment with a corresponding model that does not consider dose or duration. According to the OMG/ORAG protocol, study results were pooled only when these two models did not differ significantly, so that heterogeneity was absent.

Table 1 Criteria used by OMG/ORAG to distinguish osteoporosis prevention and treatment studies

Criterion	Prevention	Treatment
Population	Postmenopausal women	Postmenopausal women
Study classification	Primary prevention	Secondary prevention
T-score[a]	Within 2 SD of mean peak bone mass	At least 2 SD below peak bone mass
Prevalence of vertebral fracture at baseline[b]	⩽20% of population	>20% of population
Mean age[c]	⩽62 years	>62 years

[a]Because of controversy about a threshold of 2 SD, analyses for alendronate and risedronate were repeated using a threshold of 2.5 SD below young adult normal. [b]This criterion applied if no BMD results were provided. [c]This criterion applied if neither BMD results nor vertebral fracture prevalence was provided.

In order to examine any observed heterogeneity in greater detail, the OMG/ORAG investigators pre-defined factors that might be related to this effect, including differences in (i) prevention vs. treatment populations, (ii) prevalent vertebral fracture(s), (iii) losses to follow-up during the study, (iv) levels of calcium intake during trials, and (v) concomitant use of vitamin D supplements. For example, the ORAG/OMG investigators hypothesized that treatment effects could differ in women with moderately low bone density (prevention studies) and women with established osteoporosis (treatment studies). They therefore differentiated prevention from treatment trials using a predefined set of criteria, including BMD, history of previous fracture and age (Table 1). Statistically significant heterogeneity was defined as $p \leqslant 0.10$. If the heterogeneity p value for any factor was lower than 0.10, chance alone was considered to be an unlikely explanation for the difference, and that factor was considered responsible for the heterogeneity. If heterogeneity was encountered, those groups were reported separately, because pooling of results was not considered statistically sound. For example, if prior fracture was associated with heterogeneity of treatment effect, then groups with and without such a history would be reported separately. In the case of alendronate, as will be explained in greater detail in the results, such heterogeneity was identified when evaluating the effects of doses $\geqslant$10 mg (vs. doses <10 mg) on non-vertebral fracture incidence; these doses were thus reported separately.

Two general mathematical models with differing underlying principles are used in meta-analysis: fixed effects and random effects.[23] Functionally, the random effects model usually produces wider confidence intervals, and is considered to be the more conservative option. In the fixed-effects model, the studies included are considered to comprise the entire universe of information from which the most precise estimate is to be derived. Variability, therefore, derives mostly from study size, and larger studies are weighted more heavily. In contrast, the random effects model assumes that the results included represent only a random sample of the possible estimates of treatment effect. In this model, variability includes both within and between study components, resulting in wider confidence intervals. The OMG/ORAG investigators chose to use the more conservative random effects model; thus, statistically significant treatment effects were relatively more difficult to demonstrate than if a fixed-effects model had been used.

Outcomes

BMD, vertebral fracture, and non-vertebral fracture were the outcomes of interest in the OMG/ORAG meta-analyses, and separate analyses were conducted for each of these outcomes for each of the agents. Fracture analyses were based on the proportion of patients with new fractures rather than the number of new fractures, because analysing the number of fractures overestimates the treatment effect unless special statistical techniques are used (which was not possible in these meta-analyses). This is of particular importance for osteoporosis, since the occurrence of a fracture represents a major risk for experiencing a subsequent fracture.[24] Adverse event profiles and discontinuation rates were also evaluated when available.

Studies with positive results may be more likely to be published than studies with negative results, and this can overestimate the actual treatment effect when data are combined. Moreover, trial results may be published more than once, which may also compromise the validity of a systematic review. To identify and limit this effect (known as publication bias) as much as possible, data

Table 2 Methodological quality of studies included in the meta-analyses

Intervention	No. of studies	Blinding	Concealed treatment allocation	ITT analysis	Loss to follow-up		
					<5%	5–20%	>20%
Alendronate	11	11	11	10	2	6	3
Risedronate	8	8	6	8	–	1	7
Hormone therapy	57	31	5	Incomplete	Incomplete data		
Vitamin D	25	18	10	9	–	8	13
Calcitonin	30	16	15	4	6	13	9
Raloxifene (pooled)	7	7	7	7	–	3	3
Etidronate (400 mg, vertebral)	13	6	9	12	1	5	7
Calcium	15	13	13	1	–	13	2

From Cranney A, *et al.* IX: Summary of meta-analyses of therapies for postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:570–8. Reproduced with permission of The Endocrine Society, Copyright 2002.

from unpublished studies are included in the meta-analyses, duplicate publications are identified and omitted, and statistical tests are performed to detect such bias. A funnel plot, of study sample size or precision (which depends on sample size) on the horizontal axis and treatment effect size on the vertical axis, is generally used in this process.[23] If publication bias exists, such that null or unfavourable results are not published, the lower left corner of the plot, which should show small studies with null or negative results, will be empty.

Meta-analysis results

Methodological rigor

The literature review process identified eight antiresorptive agents for treating postmenopausal osteoporosis, for which systematic reviews could be performed: alendronate, calcitonin, calcium, etidronate, hormone replacement therapy (because results of the Women's Health Initiative had not been published at the time of the review, they were not included in this meta-analysis), raloxifene, risedronate, and vitamin D (including pharmacological doses and nutritional supplements). The reviewers found that studies of alendronate and raloxifene best fulfilled the four predefined criteria for study validity (concealment of allocation, blinding, completeness of follow-up, and intention-to-treat analysis), thus allowing greater confidence in results of these studies. (Table 2) Studies of etidronate, risedronate, and calcitonin were considered to be compromised by relatively large losses to follow-up, and studies of calcitonin and calcium generally failed to use intention-to-treat analysis, so that less confidence could be placed in these results. In addition, calcitonin trials frequently failed to conceal treatment allocation, and reported effects on BMD were larger in calcitonin trials with inadequate concealment, supporting the inference that this biased the study results.

Vertebral fracture

Statistically significant reductions in the risk of vertebral fracture, which ranged from 36% to 48%, were found in the meta-analyses for alendronate, risedronate, etidronate, raloxifene, and vitamin D (Table 3). Treatment effects for alendronate, risedronate, etidronate, and hormone replacement therapy were very consistent across studies (*homo*geneity of effect), regardless of dose or duration.

Non-vertebral fracture

Only alendronate and risedronate significantly reduced the risk of non-vertebral fracture (Table 3). Alendronate, in doses <10 mg, decreased the risk of non-vertebral fractures by 49% (95%CI 31–62%), and there were sufficient data to conclude that the risk reduction was similar at all non-vertebral sites, including the hip. Moreover, these risk reductions were similar for fracture sites that had not previously been considered to be related to osteoporosis. Alendronate demonstrated heterogeneity for non-vertebral fracture incidence, such that doses <10 mg had a significantly lesser effect (13% reduction; 95%CI 2% increase, 27% reduction). As a consequence, these results could not be pooled, and were reported separately. Furthermore, in one large trial of alendronate

Table 3 Magnitude of therapy effect on fracture risk reduction

Intervention	Non-vertebral fractures				Vertebral fractures			
	Relative risk reduction (95%CI)	Trials (patients)	Relative risk	*p*	Relative risk reduction (95%CI)	Trials (patients)	Relative risk	*p*
Agents with statistically significant reductions in non-vertebral fracture, with or without reduction in vertebral fractures								
Alendronate >10 mg/day (non-vertebral)	49% (31–62%)	6 (3723)	0.51	<0.01				
Alendronate >5 mg/day (vertebral)					48% (35–56%)	8 (9360)	0.52	<0.01
Risedronate ⩾2.5 mg/day (non-vertebral)	27% (13–39%)	7 (12 958)	0.73	<0.01				
Risedronate ⩾2.5 mg/day (vertebral)					36% (23–46%)	5 (2604)	0.64	0.01
Agents with statistically significant reductions in vertebral fracture only								
Vitamin D	23% (−4 to 43%)	6 (6187)	0.77†	0.09	37% (12–55%)	8 (1130)	0.63	<0.01
Calcitonin	20% (−9 to 41%)	1 (1245)	0.80	0.16	21% (0–38%)	1 (1108)	0.79	0.05
Raloxifene (pooled)	9% (−6 to 21%)	2 (6961)	0.91	0.24	40% (30–50%)	1 (6828)	0.60	0.01
Etidronate (400 mg, vertebral)	1% (−42 to 31%)	7 (867)	0.99	0.97	37% (8–56%)	9 (1076)	0.63	0.02
Agents without statistically significant reduction in fracture								
Calcium	14% (−72 to 57%)	2 (222)	0.86	0.66	23% (−9 to 46%)	5 (576)	0.77	0.14
Hormone therapy	13% (−8 to 29%)	6 (3986)	0.87	0.10	34% (−7 to 59%)	5 (3117)	0.66	0.12

Agents are presented in order by magnitude of nonvertebral fracture risk reduction within each category. Only the two newer bisphosphonates, alendronate and risedronate, significantly reduced both vertebral and non-vertebral fractures. †Significant heterogeneity between studies, based on $p \leqslant 0.10$. From Cranney A, *et al.* IX: Summary of meta-analyses of therapies for postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:570–8. Reproduced with permission of The Endocrine Society, Copyright 2002.

(Fracture Intervention Trial), 5 mg was given for 2 years followed by 10 mg for the remainder of the study,[25] and heterogeneity of dose effect rendered this study ineligible for inclusion in the meta-analysis of non-vertebral fracture.

Risedronate (2.5 and 5 mg doses combined, because they exhibited homogeneity of effect) was the only other antiresorptive agent to show a significant reduction in non-vertebral fracture risk (27%, 95%CI 13–39%), but data were insufficient to comment on site-specific effects, although hip fracture risk reduction was shown in a single trial.[26] Neither calcium nor hormone replacement therapy significantly reduced either vertebral or non-vertebral fracture risk.

Bone mineral density (BMD)

The biggest gains in lumbar spine BMD, ranging from 4.5% to 7.5%, were seen with alendronate, risedronate, and hormone therapy (Figure 2). More modest BMD increases, in the range of 2.4% to 4.1%, were seen with calcitonin, etidronate, and raloxifene. Neither calcium nor vitamin D had a significant effect on spine BMD. The relative increases in BMD for these eight agents were similar at the hip. In general, the magnitude of effect on BMD correlated with non-vertebral fracture risk reduction, so that the agents that produced the greatest increase in BMD (alendronate and risedronate) demonstrated the greatest anti-fracture efficacy, especially at non-vertebral sites. The magnitude of vertebral fracture reduction was less closely associated with BMD change: alendronate, raloxifene, and risedronate all reduced vertebral fracture incidence, despite differing increases in BMD. Higher doses of alendronate, hormone replacement therapy, and risedronate were associated with greater increases in BMD, but no such dose-response relationship was shown for calcitonin or calcium. Longer duration of treatment produced larger increases in BMD with alendronate, hormone therapy, raloxifene, and risedronate, but not with other agents. Tests for heterogeneity did not detect differences in the magnitude of effects on lumbar spine BMD between prevention and treatment populations, nor were significant differences identified based on years

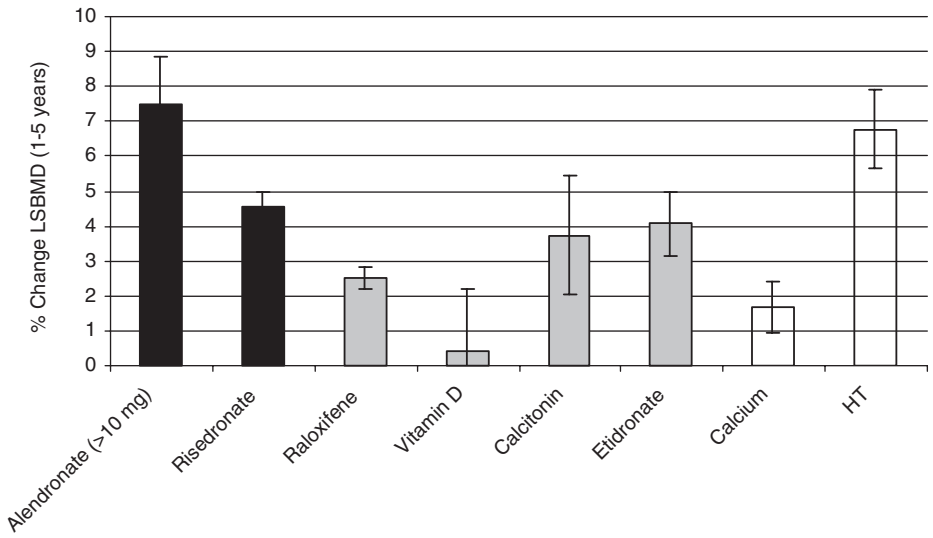

Figure 2. Mean percent change in BMD at the lumbar spine for agents in OMG/ORAG meta-analyses. Magnitude of increase in spine BMD after 1–5 years on treatment versus placebo. Longer treatment durations resulted in larger BMD effects with alendronate, risedronate, raloxifene and HT. This was not the case for vitamin D, calcitonin, or calcium. Greater increases in BMD were reported in agents that significantly reduced the risk of non-vertebral fractures. Black bars indicate agents with confirmed non-vertebral and vertebral fracture risk reduction. Gray bars indicate agents with vertebral fracture risk reduction only. White bar indicates agents without significant fracture risk reductions.

of follow-up. This is shown in Figure 2, which presents the mean percent change in BMD for each of the agents, based on pooled data, generally using data from the last year of a multi-year trial.

Discontinuation due to adverse events

Tolerability for all of the agents examined was good and only vitamin D (RR 1.37, 95%CI 1.01–1.88) and, barely, raloxifene (RR 1.15, 95%CI 1.00–1.33), had discontinuation rates that were significantly different from placebo (Table 4).

Table 4 Relative risks of discontinuing therapy for adverse events

Intervention	Relative risk (95%CI)
Alendronate	1.15 (0.93–1.42)
Risedronate	0.94 (0.80–1.10)
Hormone therapy	NA
Vitamin D	1.37 (1.01–1.88)
Calcitonin	Similar in treatment and control groups
Etidronate	0.93 (0.70–1.23)
Raloxifene	1.15 (1.00–1.33)
Calcium	NA

NA, not available.

Discussion

Because meta-analysis pools data from as many studies as possible, the estimates of treatment effect are more precise than those available from any individual trial. However, the validity of the meta-analysis depends on the methodological rigor with which it is conducted. In the OMG/ORAG analyses, the proportion of participants who incurred fractures in each group was used to estimate the treatment effects, which provides consistency and comparability among studies and analyses. As a consequence, these estimates of effect using proportional analysis may differ from the results reported in the original publications in which time-to-event, or survival, analysis was used. This is especially relevant in the case of vertebral fractures, where in the presence of relatively large losses to follow-up (even if the losses are random), time-to-event analysis results in a larger treatment effect than does proportional analysis.[27]

Reducing the risk of non-vertebral fracture is a critical objective of osteoporosis treatment, because these fractures have important effects on health and quality of life. There is at present convincing evidence of non-vertebral fracture risk reduction for only alendronate and risedronate. Alendronate (>10 mg/day) reduced the risk of non-vertebral fractures by 49% (95%CI 31–62%), with consistent risk reductions at all fracture sites, including those not traditionally considered to be related to osteoporosis. Risedronate (>2.5 mg/day)

reduced the risk of non-vertebral fractures by 27% (95%CI 13–39%), but because of power limitations, the analysis could not conclude that this reduction pertained at individual fracture sites,[6,12] except for hip fractures in women 70–79 years of age with low BMD.[26]

Reduction of vertebral fracture risk should be another therapeutic goal, given the literature on the impact of fractures on quality of life, morbidity and mortality.[2,28] Based on the methodological quality, the magnitude of the treatment effect, the narrowness of the confidence intervals and the consistency of results among studies, evidence of vertebral fracture risk reduction was strongest for alendronate and risedronate, but raloxifene, vitamin D, etidronate, and calcitonin also reduced this risk.

The type of analysis performed in individual trials may have important implications for reported efficacy. In situations of considerable loss to follow-up early in the trial, Kaplan-Meier or Cox proportional hazards time-to-event analysis can yield biased estimates of the true reduction in vertebral fracture risk.[27] Time-to-event analysis censors individuals at the time they are lost to follow-up or have the clinical event of interest, so they are considered as part of the denominator only up to this point. When numbers of drop-outs are large, especially early in the study, the estimates of cumulative incidence at these time points carry a disproportionate weight in calculating the final overall cumulative incidence, since the Kaplan-Meier estimator weights early events more heavily than later ones. This generally has little effect on non-vertebral fracture outcomes. Furthermore, because time to event for each trial participant is not available to meta-analysts, neither Kaplan-Meier nor Cox proportional hazards methods can be used to derive summary estimates. As noted by the ORAG/OMG authors, additional research is necessary to demonstrate whether agents that offer only vertebral fracture protection can improve quality of life.[12]

The choice among treatment options is usually determined by a combination of the clinician's perception of the quality of the clinical trial evidence for a particular agent and the patient's personal priorities/preferences. In the case of osteoporosis, these considerations may include non-skeletal factors such as wanted or un-wanted effects on breast, urogenital tract or cardiovascular system. As noted by the OMG/ORAG investigators, the magnitude of the summary treatment effects of two or more agents cannot be directly compared to one another, because the estimates were derived from different sets of populations that were not randomized with respect to one another. As a consequence, even if the confidence intervals for two estimates do not overlap, a statistically significant difference between these agents cannot be inferred.

Head-to-head trials with fracture endpoints would provide the most reliable and valid comparison between the effectiveness of agents, but such trials would be very large and expensive and are, consequently, unlikely to be performed. For example, in order to demonstrate equivalence between two treatments that reduce fracture incidence by 50% relative to placebo in a population with a 5% fracture rate, the sample size per group would need to be $>30\,000$.[29] Several head-to-head trials have been conducted to compare the relative effects of antiresorptive therapies on BMD and bone turnover,[30–35] but none was designed to assess effects on fracture incidence. However, among antiresorptive agents, the magnitude of the BMD increase is positively associated with non-vertebral fracture reduction,[36,37] although the extent to which it may explain this effect is a particularly contentious issue.[38]

Although the systematic review provides powerful evidence as to the consistency of effect on BMD and fracture risk reduction, it does not specifically address issues of comparison between agents. However, statistical techniques for indirectly comparing agents to a common comparator (such as placebo) generally show agreement with results of meta-analyses and head-to-head RCTs comparing the same agents.[39–41] In effect, by mathematically combining placebo groups for the active agents, strata are constructed that include participants from each of the strata included in each of the treatment groups. As a consequence, comparisons between the agents can be mediated through this combined placebo group. This procedure adds to variability of estimates of treatment effects, so that adjusted indirect comparisons tend to be more conservative, with wider confidence intervals and less likelihood of demonstrating statistical significance.[39] When these techniques were applied to the OMG/ORAG results, alendronate had significantly greater efficacy in preventing non-vertebral fracture than any of the other agents.[42]

Systematic reviews and meta-analyses do have limitations that largely depend on the methodological quality of the original studies. If the original studies contain one or more biases, or if negative studies are unpublished, the resultant meta-analysis may be flawed. Analyses that neglect important secondary outcomes and side-effects are incomplete.[4] It must also be recognized that a systematic review is helpful in making therapeutic decisions

only if relevant patient populations and clinical outcomes are evaluated.[43] The OMG/ORAG systematic reviews focused on treatment effects in postmenopausal women with known low bone density, the population at highest risk of fracture, so the findings are useful for the clinical management of such patients.

Concurrent with the publication of these meta-analyses, data from the Women's Health Initiative (WHI), a large randomized controlled trial evaluating the effects of oestrogen plus progestin (Prempro) in post-menopausal women, were published. The WHI investigators reported that this regimen of hormone therapy reduced vertebral fractures by 34% and non-vertebral fractures by 23%.[44] The magnitude of fracture risk reduction reported in WHI was similar to the meta-analysis estimates, and if WHI data are added to the OMG/ORAG meta-analyses, hormone treatment significantly reduced the risk of both vertebral (RR 0.70, 95%CI 0.52–0.94) and non-vertebral (RR 0.78, 95%CI 0.64–0.96) fractures. (Cranney, personal communication) Even more recently, results of the oestrogen-only WHI trial have shown a similar reduction in fracture incidence.[45] However, because of the risks associated with long-term use of hormone therapy, it is no longer recommended solely for prevention or treatment of osteoporosis.[46,47]

In summary, the results from the OMG/ORAG meta-analyses provide a useful insight into the strength of the evidence for effectiveness of different osteoporosis therapies. Results from systematic reviews and meta-analyses, considered together with factors such as extra-skeletal benefits, risks, cost, and patient preferences, can help clinicians to identify the most effective agents for treating postmenopausal osteoporosis.

Acknowledgements

The authors acknowledge the useful comments on earlier versions of this paper by Dr Lois E. Wehren, Merck & Co., Inc., Whitehouse Station, NJ, USA, formerly at the University of Maryland School of Medicine, Baltimore, MD, USA

References

1. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? *JAMA* 1994; **271**:59–63.
2. Cauley JA, Thompson DE, Ensrud KC, Scott JC, Black D. Risk of mortality following clinical fractures. *Osteoporos Int* 2000; **11**:556–61.
3. Ross PD. Osteoporosis: frequency, consequences, and risk factors. *Arch Intern Med* 1996; **156**:1399–411.
4. The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. Meta-analyses of therapies for postmenopausal osteoporosis. I. Systematic reviews of randomized trials in osteoporosis: introduction and methodology. *Endocr Rev* 2002; **23**:496–507.
5. Cranney A, Wells G, Willan A, Griffith L, Zytaruk N, Robinson V, Black D, Adachi J, Shea B, Tugwell P, Guyatt G, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. II. Meta-analysis of alendronate for the treatment of postmenopausal women. *Endocr Rev* 2002; **23**:508–16.
6. Cranney A, Tugwell P, Adachi J, Weaver B, Zytaruk N, Papaioannou A, Robinson V, Shea B, Wells G, Guyatt G, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. III. Meta-analysis of risedronate for the treatment of postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:517–23.
7. Cranney A, Tugwell P, Zytaruk N, Robinson V, Weaver B, Adachi J, Wells G, Shea B, Guyatt G, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. IV. Meta-analysis of raloxifene for the prevention and treatment of postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:524–8.
8. Wells G, Tugwell P, Shea B, Guyatt G, Peterson J, Zytaruk N, Robinson V, Henry D, O'Connell D, Cranney A, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. V. Meta-analysis of the efficacy of hormone replacement therapy in treating and preventing osteoporosis in postmenopausal women. *Endocr Rev* 2002; **23**:529–39.
9. Cranney A, Tugwell P, Zytaruk N, Robinson V, Weaver B, Shea B, Wells G, Adachi J, Waldegger L, Guyatt G, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. VI. Meta-analysis of calcitonin for the treatment of postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:540–51.
10. Shea B, Wells G, Cranney A, Zytaruk N, Robinson V, Griffith L, Ortiz Z, Peterson J, Adachi J, Tugwell P, Guyatt G, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. VII. Meta-analysis of calcium supplementation for the prevention of postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:552–9.
11. Papadimitropoulos E, Wells G, Shea B, Gillespie W, Weaver B, Zytaruk N, Cranney A, Adachi J, Tugwell P, Josse R, Greenwood C, Guyatt G, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. VIII. Meta-analysis of the efficacy of vitamin D treatment in preventing osteoporosis in postmenopausal women. *Endocr Rev* 2002; **23**:560–9.
12. Cranney A, Guyatt G, Griffith L, Wells G, Tugwell P, Rosen C, The Osteoporosis Methodology Group, The Osteoporosis Research Advisory Group. IX. Summary of meta-analysis of therapies for postmenopausal osteoporosis. *Endocr Rev* 2002; **23**:570–8.
13. Devereaux PJ, Yusuf S. The evolution of the randomized controlled trial and its role in evidence-based decision making. *J Intern Med* 2003; **254**:105–13.
14. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet* 2002; **359**:57–61.
15. Oxman AD, Guyatt GH. A consumer's guide to subgroup analysis. *Ann Intern Med* 1992; **116**:78–84.

16. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002; **359**:248–52.
17. Herrington DM, Howard TD. From presumed benefit to potential harm – hormone therapy and heart disease. *N Engl J Med* 2003; **349**:519–21.
18. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet* 2002; **359**:614–18.
19. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002; **359**:696–700.
20. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine*, 2nd edn. Boston, Little, Brown and Company, 1991.
21. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; **359**:781–5.
22. Schulz KF. Randomised trials, human nature, and reporting guidelines. *Lancet* 1996; **348**:596–8.
23. Petitti DB. Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine. New York, Oxford University Press, 1994.
24. Klotzbuecher CM, Ross PD, Landsman PB, Abbott TA, Berger M. Patients with prior fractures have an increased risk of future fractures: a summary of the literature and statistical synthesis. *J Bone Miner Res* 2000; **15**:721–39.
25. Black DM, Thompson DE, Bauer DC, Ensrud K, Musliner T, Hochberg MC, Nevitt MC, Suryawanshi S, Cummings SR, for the FIT Research Group. Fracture risk reduction with alendronate in women with osteoporosis: the Fracture Intervention Trial. *J Clin Endocrinol Metab* 2000; **85**:4118–24.
26. McClung MR, Geusens P, Miller PD, Zippel H, Bensen WG, Roux C, Adami S, Fogelman I, Diamond T, Eastell R, Meunier PJ, Reginster J-Y. Effect of risedronate on the risk of hip fracture in elderly women. *N Engl J Med* 2001; **344**:333–40.
27. Black DM, Thompson DE, Teutsch C, de Papp AE. Patterns in loss to follow-up can influence estimates of risk reduction of fracture (abstract). *J Bone Miner Res* 2003; **18(suppl. 2)**:S249.
28. Ensrud KE, Thompson DE, Cauley JA, Nevitt MC, Kado DM, Hochberg MC, Santora AC, Black DM. Prevalent vertebral deformities predict mortality and hospitalization in older women with low bone mass. *J Am Geriatr Soc* 2000; **48**:241–9.
29. Kanis JA, Oden A, Johnell O, Caulin F, Bone HG, Alexandre J-M, Abadie E, Lekkerkerker F. Uncertain future of trials in osteoporosis. *Osteoporos Int* 2002; **13**:443–9.
30. Adami S, Passeri M, Ortolani S, Broggini M, Carratelli L, Caruso I, Gandolini G, Gnessi L, Laurenzi M, Lombardi A, Norbiato G, Pryor-Tillotson S, Reda C, Romanini L, Subrizi D, Wei L, Yates AJ. Effects of oral alendronate and intranasal salmon calcitonin on bone mass and biochemical markers of bone turnover in postmenopausal women with osteoporosis. *Bone* 1995; **17**:383–90.
31. Hosking D, Adami S, Felsenberg D, Andia JC, Valimaki M, Benhamou L, Reginster J-Y, Yacik C, Ryback-Feglin A, Petruschke RA, Zaru L, Santora AC. Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opinion* 2003; **19**:383–94.
32. Rosen CJ, Hochberg M, Bonnick S, McClung M, Miller P, Broy S, Kagan R, Chen E, Petruschke RA, Thompson DE, de Papp AE, for the Fosamax Actonel Comparison Trial (FACT) investigators. Treatment with once-weekly alendronate 70 mg compared to once-weekly risedronate 35 mg in women with postmenopausal osteoporosis: a randomized, double-blind study. *J Bone Miner Res* 2005; **20**:141–51.
33. Greenspan SL, Emkey RD, Bone HGr, Weiss SR, Bell NH, Downs RWJ, McKeever C, Miller SS, Davidson M, Bolognese MA, Mulloy AL, Heyden N, Wu M, Kaur A, Lombardi A. Significant differential effects of alendronate, estrogen, or combination therapy on the rate of bone loss after discontinuation of treatment of postmenopausal osteoporosis: a randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 2002; **137**:875–83.
34. Ravn P, Bidstrup M, Wasnich RD, Davis JW, McClung MR, Balske A, Coupland C, Sahota O, Kaur A, Daley M, Cizza G. Alendronate and estrogen-progestin in the long-term prevention of bone loss: four-year results from the Early Postmenopausal Intervention Cohort Study. *Ann Intern Med* 1999; **131**:935–42.
35. Johnell O, Scheele WH, Lu Y, Reginster J-Y, Need AG, Seeman E. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. *J Clin Endocrinol Metab* 2002; **87**:985–92.
36. Wasnich RD, Miller PD. Antifracture efficacy of antiresorptive agents are related to changes in bone density. *J Clin Endocrinol Metab* 2000; **85**:231–6.
37. Hochberg M, Ross PD, Black DM, Cummings SR, Genant HK, Nevitt M, Barrett-Connor E, Musliner T, Thompson D. Larger increases in bone mineral density during alendronate therapy are associated with a lower risk of new vertebral fractures in women with postmenopausal osteoporosis. *Arthritis Rheum* 1999; **42**:1246–54.
38. Delmas PD, Seeman E. Changes in bone mineral density explain little of the reduction in vertebral or nonvertebral fracture risk with anti-resorptive therapy. *Bone* 2004; **34**: 599–604.
39. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; **50**:683–91.
40. Song F, Glenny A-M, Altman DG. Indirect comparison in evaluating relative efficacy illustrated by antimicrobial prophylaxis in colorectal surgery. *Control Clin Trials* 2000; **21**:488–97.
41. Song F, Altman DG, Glenny A-M, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *Br Med J* 2003; **326**:472–6.
42. Wehren LE, Hosking D, Hochberg MC. Putting evidence-based medicine into clinical practice: comparing anti-resorptive agents for the treatment of osteoporosis. *Curr Med Res Opinion* 2004; **20**:525–31.
43. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. *JAMA* 1994; **272**:1367–71.
44. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy

postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002; **288**:321–33.

45. Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA* 2004; **291**:1701–12.
46. Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* 2003; **362**:419–27.
47. US Preventive Services Task Force. Postmenopausal hormone replacement therapy for primary prevention of chronic conditions: recommendations and rationale. *Ann Intern Med* 2002; **137**:834–9.