Analyses of the cost-effectiveness of pooled alendronate and risedronate, compared with strontium ranelate, raloxifene and teriparatide. Response of the Osteoporosis Guideline Development Group (GDG) to the assessment report

Summary

The GDG has considered the assessment report carefully, and offers the following comments and points for discussion at the Appraisal Committee's meeting on September 6th, 2006.

The assessment report is a summary of the cost effectiveness analyses and informs the recommendations that will be made by the appraisal committee. We recognise that the purpose of the report is to provide the evidence that sources recommendations, however, there are some phrases in the CE threshold tables that can be interpreted as a recommendation. We are concerned as this can and has led to confusion and may have generated comments in other submissions that will be inappropriate following the committee's interpretation of the data. The GDG expects that the committee will interpret the cost effectiveness analyses in the light of what is clinically reasonable, an approach that is consistent with that taken for TA87.

One important example of this is the management of women in the older age group (over 75 years). If the cost effectiveness thresholds in the report were applied without interpretation, all women would need to have a BMD scan before treatment. This would produce significantly increased NHS costs by demanding additional DXA scans in the elderly (who, in general, do not currently have DXA scans) and would be seen as clinically unreasonable. We note however, the normal distribution of T-scores in this age group suggests that the majority of self identifying women would already have the threshold T-scores, so that BMD scanning is unnecessary. The clinically reasonable interpretation would be a recommendation to treat women over 75 without BMD scanning.

The primary and secondary prevention categories of the previous ACDs and TA87 have been replaced in the report by opportunistic detection and self identification. From the description in the report, these two sets of population groups are not at all identical, although we note that full definitions of the new groups are yet to be produced. The GDG has made suggestions for these definitions.

We note a lack of information on which clinical risk factors should be used for risk assessment, and include suggestions for these which are consistent with those that inform the WHO model. We have also identified certain groups of patients who should be excluded from the appraisal, but will be covered in the guideline as part of its wider scope. This is felt to be crucial by the GDG. Cross reference to the guideline would be appropriate for these groups.

The GDG requests that absolute 10 year hip fracture probabilities are included in parentheses alongside the thresholds by T score and number of risk factors. This was agreed at the remodelling discussion meeting in January, as a means of facilitating the insertion of the appraisal recommendations into the guideline. We assume that this omission was an oversight.

We note that the cost-effectiveness of identification/treatment only approaches the threshold of £20,000 per QALY for the women close to the upper T-score threshold. The majority of women eligible for testing or treatment fall well below this limit, and this inevitably produces a more conservative result. We expect that this will be balanced by appropriate selection of sensitivity analyses.

The GDG generally favours the base case, with the exception of the cost of bisphosphonate, although we would prefer the base case to use the efficacy data of alendronate alone, together with up-to-date fracture costs. We were pleased to note that the committee has requested additional analyses to take account of the significant reduction in price of generic alendronate. We recommend that this is adopted as the base case, and suggest further modifications should include using the efficacy data for alendronate alone and the fracture costs of Stevenson et al.

We expect that if there is a requirement for more than one parameter to be changed at a time, the analysis will be re-run, rather than making extrapolations from two or more single-parameter sensitivity analyses.

Detailed discussion

No	Page	Subject	Comments
1	Gen	General	Generally, the document is clear and the analyses and summaries well presented and easy to follow. However, we have experienced some confusion because the threshold statements appear to be making recommendations, such as 'do not BMD' or 'BMD and treat where T-score<-2.5'. It would be preferable for the report to state that BMD + treatment is not cost effective for this age/CRF group and to explain that this takes into account the T-score distribution at that age. In the second example, a statement such as 'BMD + treatment where T-score<-2.5' would be more helpful
2	Gen	General	The GDG notes that a number of sensitivity analyses have been carried out that alter single variables, but at this stage no combinations of variables have been included. The GDG anticipates that if the preferred case is to adopt two or more of these sensitivity analyses, further runs of the model will be carried out before recommendations are made. We understand that this would be fairly straightforward to produce.
3	Gen	General	The GDG notes that significant restructuring of the patient groups has taken place in the current report, which was unexpected with little explanation given. We refer to the change from the original division into primary and secondary prevention to the new groups of opportunistic assessment and self-identification.
4	Gen	General	It is not clear if these new groups are intended to replace the populations used in the primary and secondary prevention appraisals respectively, or if the two groups will be combined in a single new appraisal. If two appraisals are required, we believe the original titles will have to be revised, and further consultation will be needed (see point 5 below).
5	Gen	General	The new groupings do not overlap with the original primary and secondary prevention populations. The self ID group includes women with an 'acute' fracture, whereas TA87 refers to 'postmenopausal women who have already had an osteoporosis-related fracture'. Although, the term 'acute fracture' has not been defined in the current report, it is unlikely to include all prior fractures (see also point 10 below). In addition, the inclusion of the other categories in the self ID group (rheumatoid arthritis and women who have

			taken high doses of glucocorticoids) has further diminished the overlap of this group with those defined in the secondary prevention appraisal – TA87 states that *'long-term corticosteroid use, as a principal risk factor, requires separate consideration and is not covered in this guidance'.*
6	Gen	General	The treatment thresholds are more conservative for the base case than for TA87. This is likely to be a consequence of changes in the maximum acceptable incremental cost effectiveness ratio (MAICER) from £30k to £20k.
7	Gen	General	Both points 5 and 6, if they are reproduced in the appraisal recommendations, will have far reaching consequences for those Trusts that are already implementing TA87.
8	Gen	General	The GDG notes that the analysis has been carried out only for women who are 50 years or older. Younger women at risk of osteoporosis are rare and will be covered by the guideline. We anticipate that the appraisal will cross refer to the guideline for these and other patient groups excluded (see also point 10 below).
9	3	Change of groups	The GDG understands that the change of groups into self-ID and opportunistic assessment was necessary for the cost effectiveness analysis in order to include added costs for GP time for ascertaining risks in the latter group. However, there is a need to define the two groups more effectively.
10	3	Self ID group	For the self ID group, there is a need to define the qualifying population, and the GDG makes the following recommendations: 1) The term 'acute fracture' should be changed to 'presenting with a fracture', which is defined as: a patient who reports a personal history of fragility fracture at any time – this would be in line with TA87: '*This guidance specifically applies to women who present with clinically apparent fractures identified directly by symptoms or indirectly during routine consultations for other purposes. The use of screening for asymptomatic fractures is not covered by this guidance*'. The GDG considers this to be the most clinically relevant definition. We note that older people undergoing multifactorial falls risk assessment and who have a fracture would then be included in the self identification group (NICE guideline 21) and women with asymptomatic fractures would be included in the opportunistic group. 2) The category 'rheumatoid arthritis' should refer only to the specific disorder. It is not an exemplar.

		Exclusions from the appraisals	3) There is a range of other conditions/treatments that put a patient at risk of a low BMD. This includes women who have had an untreated premature menopause, patients with COPD, cystic fibrosis and IBD, patients with prolonged immobilisation, women treated with aromatase inhibitors for breast cancer, etc. These are rare, but all are relevant in the development of osteoporosis. We expect them to be excluded from the appraisal, and will address these within the wider scope of the guideline. A complete list of these conditions is included in Appendix A. We suggest the appraisals cross refer to the guideline. 4) Similarly, patients who are on high doses of glucocorticoids should be excluded from the appraisal (as they were in TA87) and this group will also be covered by the guideline.
11	3	Opportunistic group	The GDG suggests defining this group as those patients who attend a GP surgery for reasons other than osteoporosis and who are screened in some way. The GDG expect to advise GPs to routinely assess for osteoporosis in relevant patient populations.
12	4	Modelling methodology	We note that the cost-effectiveness of identification/treatment only approaches the threshold of £20,000 per QALY for the women close to the upper T-score threshold. The majority of women eligible for testing or treatment fall well below this limit. For example, under the base case analysis the overall cost-effectiveness of the strategy for 70-74 year old women identified through the opportunistic screening strategy is only £14,257. This approach inevitably makes the base case results more conservative. Our view is that this should be taken into consideration when building in sensitivity analyses that make the model more conservative.
13	4	Pooling of alendronate and risedronate	The GDG has reservations about the assumption that alendronate and risedronate are sufficiently similar to be pooled. Alendronate and risedronate, by virtue of their different chemical compositions (figure 1) have differences in their potency in enzyme inhibition, pharmacokinetics, bone affinity, and osteoclast uptake. Both drugs are amino-bisphosphonates and are distinct from the simpler etidronate. We would prefer to see analysis by individual drug. This is more important when there is a large price differential between alendronate and risedronate.
14	4	Pooling of alendronate and	Further, we are uncertain if the appraisal will make recommendations for alendronate and risedronate together, and then etidronate separately, or the bisphosphonates as a whole. The latter would have consequences for the guideline, which needs to consider ibandronate in addition to the other bisphosphonates. The former would raise concerns because etidronate is considerably more cost effective than alendronate or risedronate, but with

		risedronate	unproven efficacy for non-vertebral fractures. In the absence of additional evidence for etidronate, the GDG suggests that the committee do not recommend the use of etidronate as a first line treatment, despite its higher cost effectiveness, and gives separate recommendations for alendronate and risedronate.
15	8	Persistence	The GDG accepts the base case value for persistence of 50%, despite the UK prescription event monitoring study indicating 75%.
16	8	FIT trial	The GDG does not recommend the adoption of the post-hoc subgroup analyses in the FIT trial. Post-hoc analyses should be treated with the utmost caution: the women were not stratified by severity and then randomised to treatments, therefore the groups are essentially non-randomised comparisons. There is no evidence about baseline characteristics within the subgroups to give any idea about the comparability of the intervention and placebo groups, but even if there were, there may be differences in other characteristics not considered (or reported). The problem with subgroup analyses within trials is that the more trial data is split post hoc, the more likely results will be found by chance. This is a widely acknowledged, with the correct way to examine the effect of, say, age would be to use regression methods on the full dataset. The post-hoc subgroup analysis of the hip fracture data in the FIT trial is a good illustration of the potentially flawed nature of this type of analysis. Splitting gives a relative risk of 0.44 (95%CI 0.19, 1.02) for the osteoporosis group (T<-2.5), and 1.84 (95%CI 0.66, 5.09) for the osteopenia group. The latter effect has a wide confidence interval and its results could be found by chance. In addition, what is the clinical meaning of alendronate doubling the relative risk of hip fracture in patients with a T-score of, say, -2.3, whilst the T-score is halved in patients with a T-score of, say, -2.6.
17	8	FIT trial and osteopenia	It is appropriate to mention at this point the retrospective regression analysis recently published[1]. This analysed the incidence of vertebral fracture, in terms of binary risk factors, in patients receiving risedronate or placebo, and who had very similar baseline characteristics across the two intervention groups. Amongst other things, the analysis concluded that, for BMD above and below T-score -2.5, the treatment-by-BMD interaction was of borderline significance for the femoral neck (p=0.100) and for the lumbar spine (p=0.052). The relative risk for the osteopenia group was 0.69 and still statistically significant.

18	8	Fracture efficacy by risk factor	The GDG does not support the use of sensitivity analyses 4 and 5, i.e. the suggestion that the efficacy of bisphosphonates is reduced or zero in patients who have risk factors other than low BMD and a prior fracture. For the risk factor, use of glucocorticoids, there are three alendronate or risedronate studies, in men and women. Meta-analysis for all doses shows the effect to be statistically significant for vertebral fracture (RR=0.44 (95%CI 0.23, 0.84)), but not for non-vertebral fracture, although only a relatively small number of patients were included in the meta-analysis (figure 2)
19	8	Fracture efficacy by risk factor	In addition, the published regression analysis of risedronate vs placebo data[1] shows that all treatment-by-non-BMD subgroup interactions were not significant ($p>0.2$) – these included BMI (above/below 26 kg/m^2) and smoking status (current/previous or none). Neither of these factors affected the efficacy of risedronate. On the basis of limited evidence, there is little rationale for assuming the efficacy of bisphosphonates should be set to 50% or 0% for risk factors other than BMD and fracture status. The GDG advises the use of the base case value of 100%.
20	8	Fracture costs	The GDG recommends that up-to-date costs are used, preferably those of Stevenson et al. We are not clear why costs dating from 2000 are preferred to those in more recent studies.
21	9	Utility multiplier vertebral fracture	The GDG's view is that vertebral and hip fractures should not be given the same value. The values in the base case were taken from a single, internally consistent study, which showed that vertebral fractures were more painful than hip fractures, hence the higher utility value for the former. There is no basis for the arbitrary change proposed by this sensitivity analysis.
22	9	Side effects	The GDG accepts the base case value for side effects of a disutility of 0.91. The 10 x base case disutility assumption is arbitrary and implausible.
23	9	Alendronate cost	The GDG recommends that this sensitivity analysis is adopted, but in a modified form using the efficacy of fracture for alendronate alone.

[1] Kanis JA, Barton IP, Johnell O. Risedronate decreases fracture risk in patients selected solely on the basis of prior vertebral fracture. Osteoporosis International 2005; 16:475

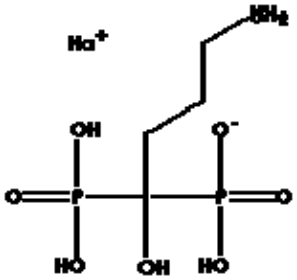
			We consider it highly likely that the cost of generic alendronate will be significantly reduced (and listed in the BNF) before the committee meets to discuss the recommendations. Indeed, we note that the committee has recently requested revised analyses based on the August price, but we are not clear what efficacy values will be used. The GDG suggests that the analyses should be re-run using the correct cost, and it should use the efficacy data for alendronate alone. Separate recommendations should be given for risedronate.
24	11	CPQ threshold	We note that the MAICER has been set at £20,000 for this report, which is at variance with that used in the appraisal for secondary prevention (TA87). This has led to substantial changes in the treatment thresholds.
25	11-end	Clinical risk factors	The eligible clinical risk factors should be defined. The GDG recommends the following risk factors for each of the groups (self ID and opportunistic): previous fracture, glucocorticoid therapy, family history of hip fracture, current smoking, alcohol ≥3 units/day, rheumatoid arthritis and is consistent with those used in the WHO model. The rationale for including some of these risk factors is given in Appendix B.
26	17-end	Absolute hip fracture risks	The GDG requests that absolute 10 year hip fracture probabilities are reported as the treatment thresholds by age, and that these are equated with the T-scores, being given in parentheses alongside the T-scores. This was agreed at the meeting in January, and will allow the guideline to include the WHO risk assessment calculator when it is available.
27	17	Recommend BMD first	The GDG recognises that the assessment report does not contain recommendations, and supports the approach used in TA87, which takes account of the decrease in T scores with age. For example, in the over 75 age group, TA87 correctly stated that DXA scanning was unnecessary in this population because '*it was considered very likely that women who have sustained a fragility fracture will have a low BMD (T score of -2.5 SD or below)*'. The GDG recommends that the same logic is applied to the over 75 age group in the new guidance - that the patient is treated without the need for a BMD scan. The GDG is concerned that a recommendation to carry out DXA scans in this age group would not only be impractical but also detrimental to health and quality of life (e.g. DXA scan of an elderly patient with an acute hip fracture). Finally, an important consequence of a requirement for BMD scans before treatment is a significant increase in NHS costs in order to provide additional DXA services for the elderly (who, generally, do not currently have

			DXA scans).
28	43-48	2nd line treatments	The GDG notes that strontium ranelate is not cost effective, even as a second line therapy, for some patient groups that will be treated with alendronate/risedronate as first line treatment. We suggest that careful consideration is given to how these women are managed. A recommendation that women who are intolerant of bisphosphonates will not be eligible for other treatment, may be difficult to apply in practice.

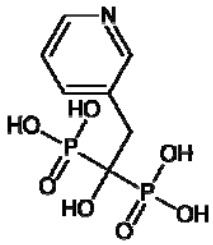
Figure 1: chemical structure of alendronate, risedronate and etidronate

Amino-bisphosphonates

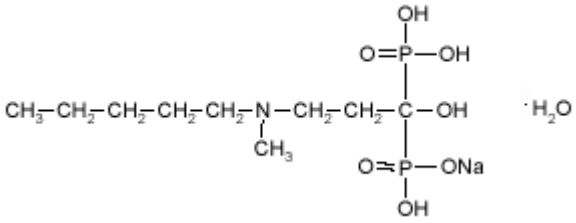
Alendronate
(aliphatic primary amine in side chain)



Risedronate
(pyridine – aromatic amine -group in side chain)



Ibandronate
(aliphatic tertiary amine in side chain)



Non-amino-bisphosphonates

Etidronate
(side chain methyl group)

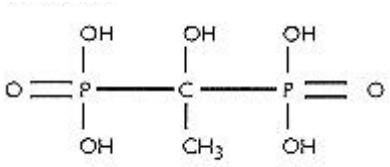
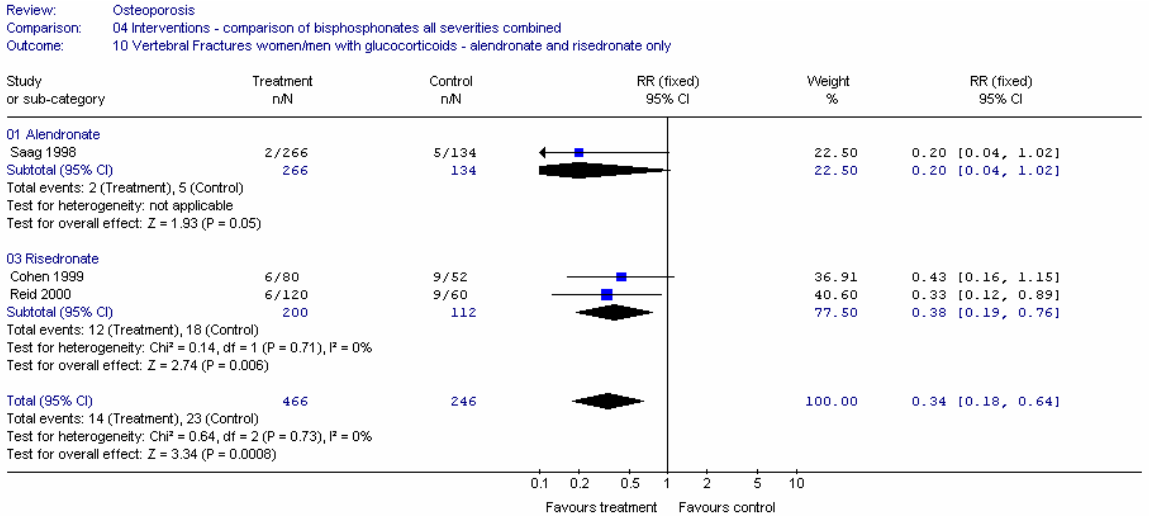
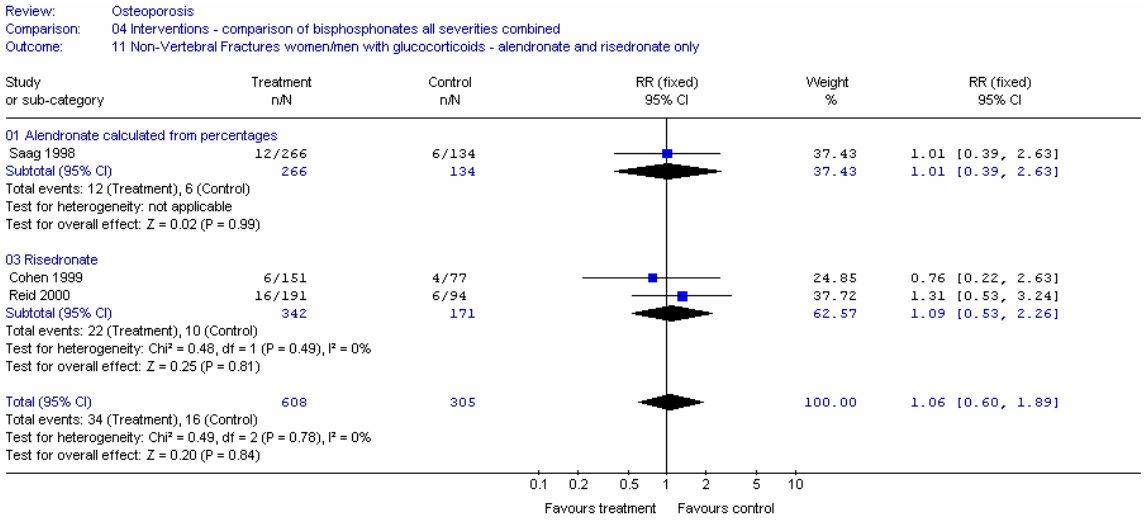


Figure 2: Alendronate or risedronate versus placebo for treatment of patients (women and men) receiving oral corticosteroids

a) Vertebral fractures



b) Non-vertebral fractures



Appendix A: Examples of patient groups that should be excluded from the appraisals, but covered by the guideline

- Premenopausal women
- Post-menopausal women less than 50 years
- Men
- Patients using oral glucocorticoids
- anorexia nervosa
- inflammatory bowel disease (such as Crohn's disease or ulcerative colitis)
- untreated hypogonadism in both sexes
- women with breast cancer treated with aromatase inhibitors

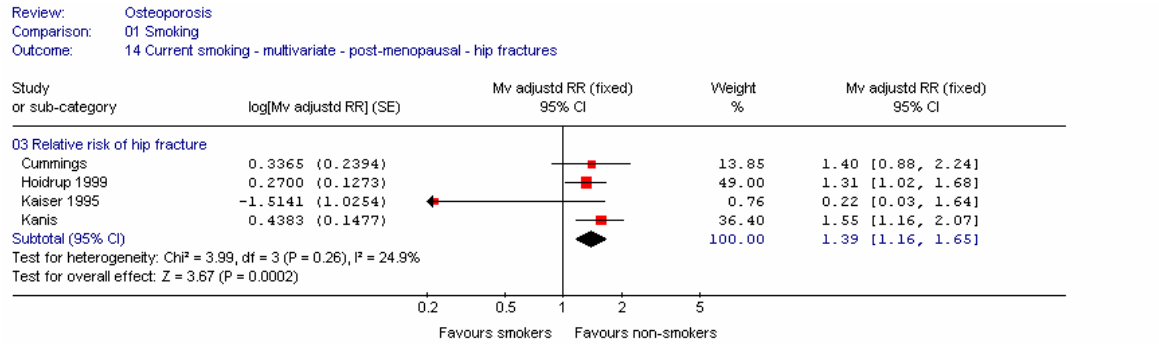
Appendix B: Risk factors smoking and alcohol

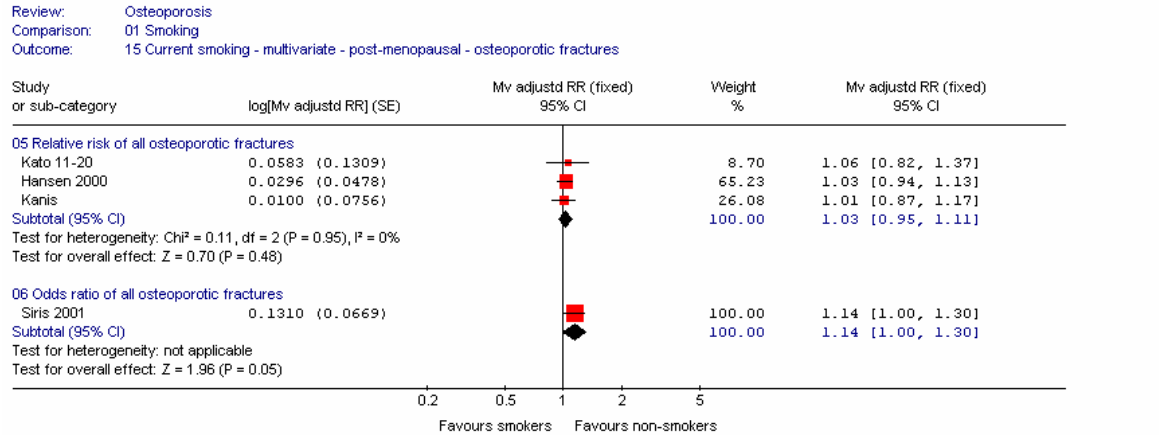
We present a brief meta-analysis of the relative risk of fracture from cohort studies that analysed their data with adjustment for multiple variables. The studies were exclusively or predominantly in post-menopausal women.

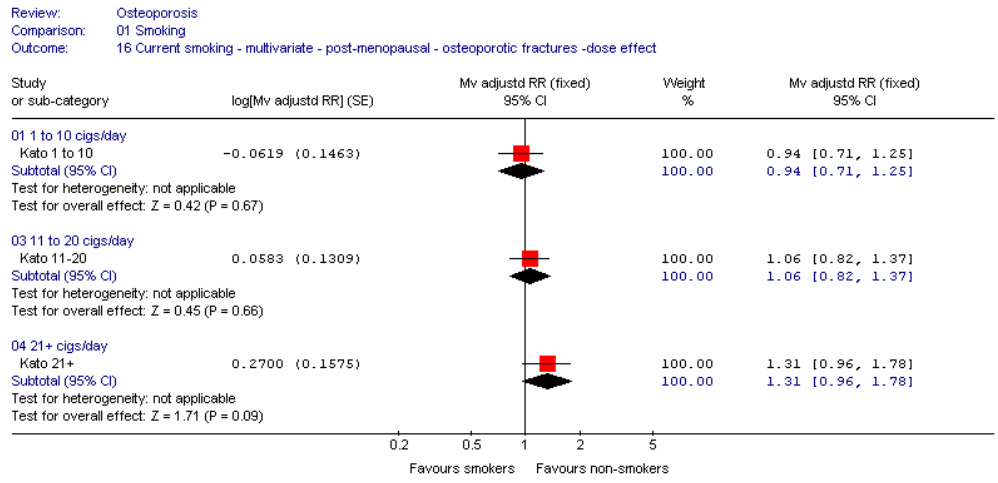
Smoking

The two figures below show a statistically significant effect of smoking on the rate of hip fracture and little difference for all osteoporotic fractures (although the odds ratio in the large Siris study (200,000) was of borderline significance for the latter analysis). There is also a clear dose effect of smoking, illustrated in the third figure for all osteoporotic fractures.

The GDG recommends current smoking as a risk factor for osteoporosis.



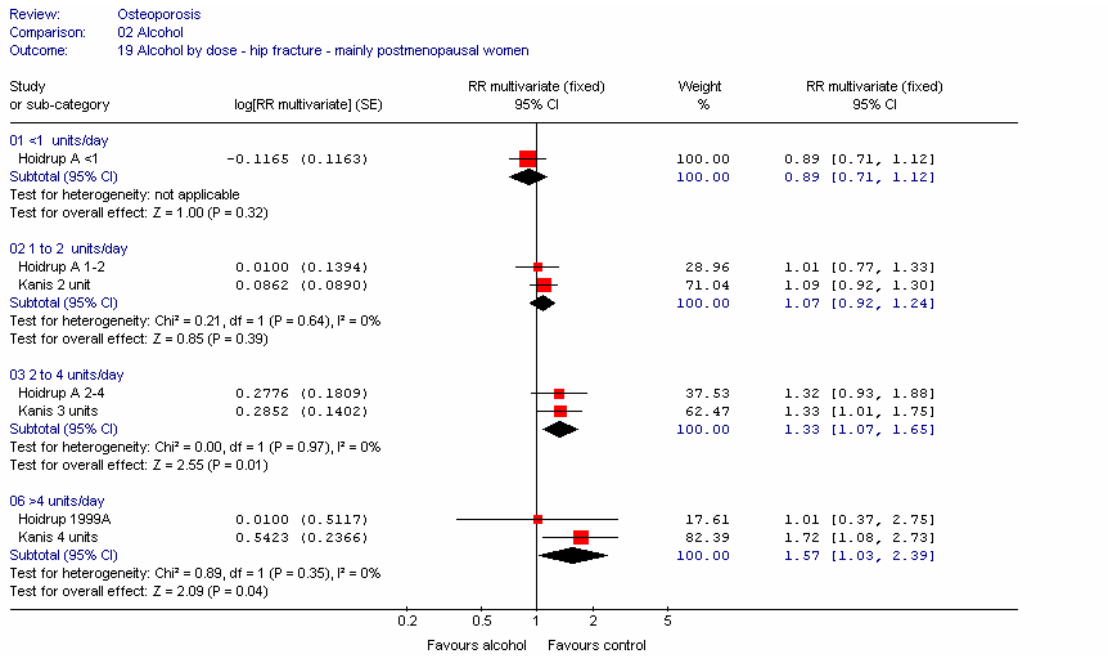


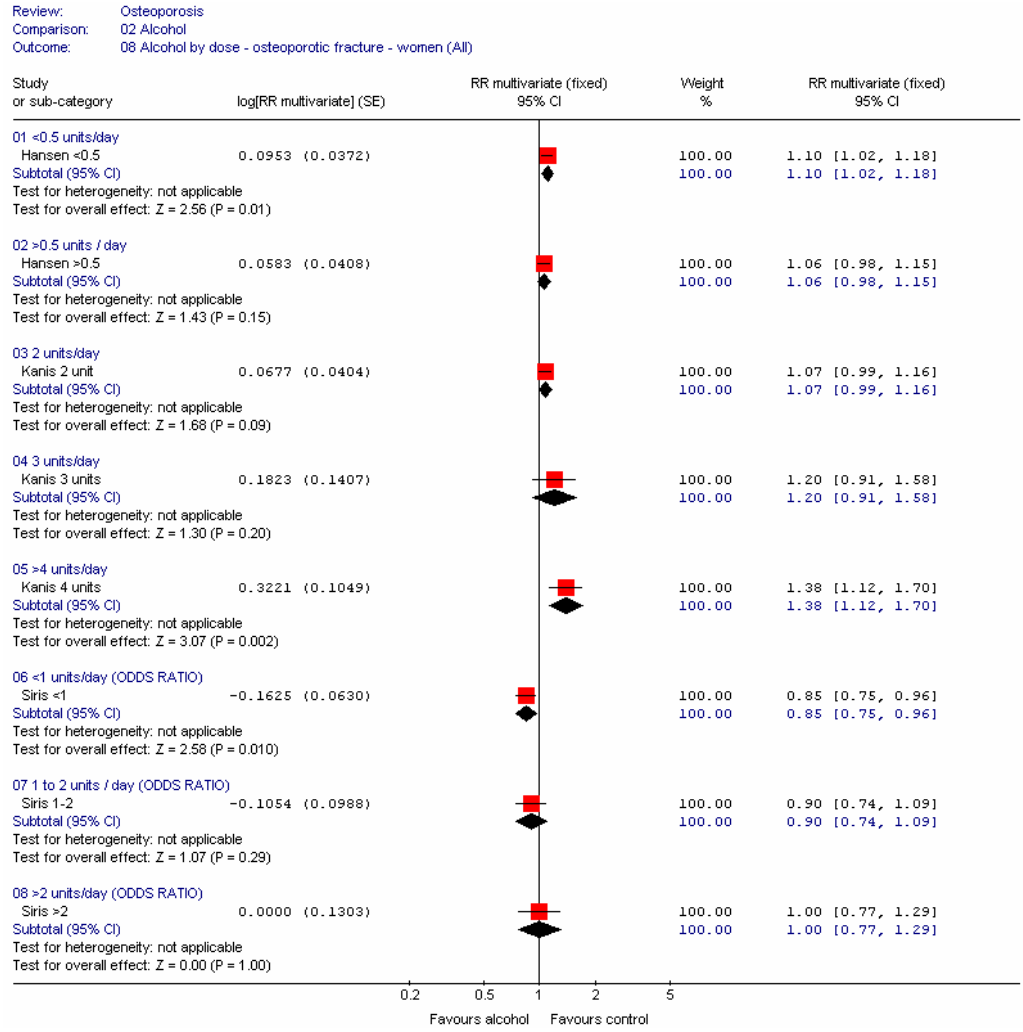


Alcohol

The Kanis meta-analysis and three studies (Hoidrup 1999A – Copenhagen; Hansen – Iowa Women's Health study; and Siris 2001 - NORA) looked at the effect of the amount of alcohol. The Hansen and Siris studies were in post-menopausal women only, and the other studies were mainly in post-menopausal women. The Kanis meta-analysis had a comparator cohort of 1 unit/day and the others had zero or <1 unit/week. Where doses were comparable, meta-analysis was carried out (despite the different comparators). The two figures below show the effects on hip fracture and any osteoporotic fracture.

There is evidently a dose effect of alcohol on the rate of fracture. On the grounds of hip fracture alone, **the GDG recommends alcohol as a risk factor for those who take 3 or more units per day of alcohol.**




Review: Osteoporosis
Comparison: 02 Alcohol
Outcome: 08 Alcohol by dose - osteoporotic fracture - women (All)
Study or sub-category
log[RR multivariate] (SE)
RR multivariate (fixed) 95% CI
Weight %
RR multivariate (fixed) 95% CI
01 <0.5 units/day
Hansen <0.5 0.0953 (0.0372) 100.00 1.10 [1.02, 1.18]
Subtotal (95% CI) 100.00 1.10 [1.02, 1.18]
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.56 (P = 0.01)
02 >0.5 units / day
Hansen >0.5 0.0583 (0.0408) 100.00 1.06 [0.98, 1.15]
Subtotal (95% CI) 100.00 1.06 [0.98, 1.15]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.43 (P = 0.15)
03 2 units/day
Kanis 2 unit 0.0677 (0.0404) 100.00 1.07 [0.99, 1.16]
Subtotal (95% CI) 100.00 1.07 [0.99, 1.16]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.68 (P = 0.09)
04 3 units/day
Kanis 3 units 0.1823 (0.1407) 100.00 1.20 [0.91, 1.58]
Subtotal (95% CI) 100.00 1.20 [0.91, 1.58]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.30 (P = 0.20)
05 >4 units/day
Kanis 4 units 0.3221 (0.1049) 100.00 1.38 [1.12, 1.70]
Subtotal (95% CI) 100.00 1.38 [1.12, 1.70]
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.07 (P = 0.002)
06 <1 units/day (ODDS RATIO)
Siris <1 -0.1625 (0.0630) 100.00 0.85 [0.75, 0.96]
Subtotal (95% CI) 100.00 0.85 [0.75, 0.96]
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.58 (P = 0.010)
07 1 to 2 units / day (ODDS RATIO)
Siris 1-2 -0.1054 (0.0988) 100.00 0.90 [0.74, 1.09]
Subtotal (95% CI) 100.00 0.90 [0.74, 1.09]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.07 (P = 0.29)
08 >2 units/day (ODDS RATIO)
Siris >2 0.0000 (0.1303) 100.00 1.00 [0.77, 1.29]
Subtotal (95% CI) 100.00 1.00 [0.77, 1.29]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.00 (P = 1.00)
0.2
0.5
1
2
5
Favours alcohol
Favours control