

Screening for Osteoporosis in Postmenopausal Women

A report
for the UK National Screening Committee

Lisa Peto
Dr Martin Allaby
January 2013

This report has been compiled by

Lisa Peto, Specialty Registrar in Public Health

Dr Martin Allaby, Consultant in Public Health Medicine

This work was undertaken by Solutions for Public Health (SPH) who received funding from the Department of Health via the National Screening Committee. Views expressed in this publication are those of the authors and not necessarily of the Department of Health.

Solutions for Public Health
4150 Chancellor Court
Oxford Business Park South
Oxford
OX4 2GX

Tel: +44 (0)1865 334700
www.sph.nhs.uk

Contents

Introduction.....	2
The Condition.....	3
The Test.....	7
The Treatment.....	10
The Programme	16
Implications for Policy.....	21
Implications for Research	25
Appendix A.....	26
Appendix B.....	29
Appendix C.....	30
References	31

Introduction

1. This paper reviews screening for osteoporosis in postmenopausal women against the UK National Screening Committee (UKNSC) criteria for appraising the viability, effectiveness and appropriateness of a screening programme (UKNSC 2003). It is based on a literature search conducted by the UKNSC in June 2011. Full details of the search strategy are set out in Appendix A.
2. The scope of this paper is limited to the identification and treatment of postmenopausal women who have osteoporosis, but who have not yet had a clinically apparent osteoporotic fragility fracture. Approximately half of all hip fractures will occur among this group, so it is important to assess whether population-based screening could prevent some of them. People who have already suffered a clinically apparent osteoporotic fragility fracture require good clinical management to reduce their risk of future fractures, but their management is outside the scope of this paper. Wherever possible, this paper draws on evidence that is directly relevant to women who have not yet suffered an osteoporotic fragility fracture, rather than trying to apply to a screen-detected population evidence obtained from women who have already experienced an osteoporotic fragility fracture. Unless stated otherwise, all comments in this paper regarding treatments for osteoporosis refer specifically to women who have osteoporosis but have not yet had a clinically apparent osteoporotic fragility fracture.
3. Osteoporosis is described by the World Health Organisation as a progressive systemic skeletal disease characterised by low bone mass and micro-architectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture (NOGG 2010).
4. Osteoporosis is defined operationally by the level of bone mass, measured as bone mineral density (BMD). Osteoporosis is defined by the World Health Organisation as having a BMD that is 2.5 standard deviations (SDs) or more below the young adult mean value for women (T-score equal to or less than -2.5 SD).¹ Severe or established osteoporosis denotes osteoporosis as defined above in the presence of one or more documented fragility fractures (NOGG 2010). Low bone density or mass (sometimes referred to as osteopenia) is diagnosed when BMD is between 1.0–2.5 SD below the reference mean (Nelson et al 2010).
5. This paper makes frequent reference to two recently published documents that concern testing and treatment for osteoporosis in the UK. In July 2010, the National Osteoporosis Guideline Group (NOGG)² published an updated version of its clinical guideline for the prevention and treatment of osteoporosis (NOGG 2010). In January 2011, NICE published an amended version of its appraisal of the use of alendronate, etidronate, risedronate, raloxifene and strontium ranelate for the primary prevention of osteoporotic fragility fractures in

¹ T-scores for people of all ages are based on the difference between the individual's BMD and the mean value for young adults, rather than the mean value for their own age group, because mean BMD falls with age; comparisons based on this gradually falling mean BMD, rather than the young adult mean, would underestimate the prevalence of osteoporosis. A different measure (the Z-score) is used to define the difference between an individual's BMD and the mean for their age group. As with the T-score, it expresses the difference in terms of multiples of the standard deviation of the BMD. The WHO definition of osteoporosis of a T score of -2.5 only applies to BMD measured by DXA at the proximal femur, lumbar spine and distal 1/3 radius. The definition does not apply to other methods of bone densitometry (quantitative CT or quantitative ultrasonography) or to other anatomical sites (calcaneus).

² NOGG represents the Bone Research Society, British Geriatrics Society, British Orthopaedic Association, British Society of Rheumatology, National Osteoporosis Society, Osteoporosis 2000, Osteoporosis Dorset, Primary Care Rheumatology Society, Royal College of Physicians and Society for Endocrinology

postmenopausal women (NICE 2011a). NOGG and NICE differ substantially in the treatment thresholds they recommend for the primary prevention of osteoporotic fractures. These differences are described in further detail in the relevant sections of this review.

6. NOGG does not advocate population screening to identify patients with osteoporosis (NOGG 2010 paras 21, 32, 47, 62). NOGG merely recommends 'selective case finding' for the prevention of osteoporosis (NOGG 2010 para 32).
7. In 2002, the US Preventive Services Task Force (USPSTF) recommended bone density screening for women 65 years or older and for women aged 60 to 64 years at increased risk for osteoporotic fractures (USPSTF 2002). In March 2011, the USPSTF widened the population recommended for osteoporosis screening in the USA to women aged 65 years or older and younger women whose fracture risk is equal to or greater than that of a 65-year-old white woman who has no additional risk factors (USPSTF 2011).
8. The USPSTF recommendations are based on a systematic review commissioned to update the 2002 USPSTF recommendations on osteoporosis screening (Nelson et al 2010). The review found no trials evaluating the effectiveness of a screening programme, but found sufficient evidence around the test and treatment for the USPSTF to conclude that there is moderate certainty that the net benefit of screening for osteoporosis by using DXA is at least moderate. The systematic review was well conducted and relatively up-to-date (searches up to the end of 2009) and hence has been used as a main source of evidence in this report. Since the UKNSC uses more stringent criteria than the USPSTF for appraising potential screening programmes, application of the UKNSC criteria to the same evidence base may result in policy recommendations that differ from those of the USPSTF.

The Condition

The condition should be an important health problem

9. Osteoporosis is a common disease which affects mainly postmenopausal women. It is estimated that more than two million women have osteoporosis in England and Wales. After the menopause, prevalence rises with age from approximately 2% at 50 years to more than 25% at 80 years (NICE 2011a). The main consequence of osteoporosis is an increased risk of fractures, most commonly fractures of the vertebral bodies, distal radius, hip and the proximal humerus (NOGG 2010). More than one-third of adult women and one in five men will sustain one or more osteoporotic fractures in their lifetime.³ It is estimated that annually there are 180,000 osteoporosis-related symptomatic fractures in England and Wales. Of these, 70,000 are hip fractures, 25,000 are clinical vertebral fractures, and 41,000 are wrist fractures (NICE 2011a). NOGG predicts that if changes are not made to present practice, osteoporotic fractures will double over the next 50 years (NOGG 2010).

³ Using data from the GP Research Database (GPRD), van Staa et al (2001) report a higher estimate (53%) for women's lifetime risk of any fracture from age 50 years.

10. Approximately 16% of women who are aged over 50 have suffered a prior fracture, likely to be a fragility fracture, since the age of 50. Conversely, 84% of the postmenopausal population has not suffered a fracture. Given that 50% of hip fracture sufferers have fractured before, 16% of the postmenopausal population account for 50% of future hip fracture cases (Mitchell 2011). Population-based screening is concerned with women who have not yet suffered a fracture, but who might be able to reduce their risk of fracture through screening if it were viable, effective and appropriate. Population-based screening is not relevant to women who have already suffered a clinically apparent fracture. Their needs should be addressed through good clinical management following their initial fracture.
11. Fractures are associated with chronic pain and disability, loss of independence, decreased quality of life, and increased mortality (USPSTF 2011). Fifty per cent of patients suffering a hip fracture can no longer live independently and 20% die within 12 months of the fracture (NOGG 2010). Vertebral fractures can be associated with curvature of the spine and loss of height and can result in pain, breathing difficulties, gastrointestinal problems and difficulties in performing daily living activities. It is thought that about 80% of vertebral fractures in women do not come to clinical attention (Strom et al 2011). UK-specific data indicate a 4.4-fold increase in mortality related to vertebral fracture (NICE 2011a).
12. The cost of osteoporosis to the National Health Service (NHS) is substantial. Fractures related to osteoporosis cost the NHS over £1.73 billion each year (NOGG 2010) and this is expected to rise to over £2.1 billion per year in 2020 (Gauthier et al 2011). Fractures in patients over 60 years account for more than 2 million hospital bed days in England. This exceeds the bed occupancy attributable to diabetes, ischaemic heart disease, heart failure or chronic obstructive pulmonary disease (NOGG 2010). Hip fractures account for the majority of the health service costs. The admission rate for hip fractures has increased in England by 2.1% per year since 1999 and hospital bed days have increased by 5.9% per year (NOGG 2010).

The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker, latent period or early symptomatic stage

13. The resorption and formation of bone is a continuous process throughout life. Bone formation exceeds bone resorption in youth, but by the fourth decade of life there is a gradual loss of bone mass. Osteoporosis is therefore usually an age-related disease. Women are at greater risk because the decrease in oestrogen production after the menopause accelerates bone loss (NICE 2011a). BMD is usually maximal by age 35 years and declines in women after age 40 years when the rate of new bone formation no longer equals the rate of bone resorption. The rate of decline in BMD is most rapid in women within two years of menopause and averages 2% to 4% a year during the first seven years after menopause. BMD may decline by 25% to 33% during this period. Afterward, loss continues, but at a slower rate of 1% to 2% a year (Slovik 2009).
14. BMD is the most important risk factor for fractures. Prospective studies have shown that the risk of fracture increases progressively with decreasing BMD, with the risk of fracture increasing approximately two-fold for each standard deviation decrease in BMD (NOGG

2010). In women aged 50 years, the lifetime risk of a vertebral fracture is estimated to be one in three, and that of hip fracture one in five (NICE 2011a).⁴

15. Increase in age is also an important risk factor due to its close relationship with BMD and the increased tendency to fall in older age (Nelson et al 2002). Other risk factors that are at least partly independent of BMD include parental history of hip fracture, alcohol, prior fracture, ethnicity, long-term systemic use of glucocorticoids and rheumatoid arthritis. Risk factors that are known to be indicators of low BMD and therefore additional risk factors for fracture include low body mass index (BMI), smoking and medical conditions such as ankylosing spondylitis, Crohn's disease, conditions that result in prolonged immobility, and untreated premature menopause (NOGG 2010 & NICE 2011a). There is a wide variation in the incidence of osteoporosis-related fractures worldwide, with a higher incidence seen in more developed countries (Strom et al 2011).

All the cost-effective primary prevention interventions should have been implemented as far as practicable

16. Primary prevention interventions aimed at decreasing fracture risk at a population level include those that aim to increase the level of physical activity at all ages, reduce the prevalence of smoking and increase dietary calcium intake. NOGG makes no recommendations on such interventions as they state that not all the modulating factors are necessarily causally related to osteoporosis and there is little evidence on their effect on fracture risk. This statement is based on a systematic review that was prepared for the clinical guidelines published by the Royal College of Physicians (RCP 1999, RCP 2000). Table 1, which is reproduced from NOGG (2010), summarises the findings of that systematic review. In addition, NOGG states that the uptake and compliance of such interventions have not been adequately assessed (NOGG 2010 paras 30-31).

Intervention	Effect on outcomes		
	BMD	Spine fracture	Hip fracture
Exercise	A	B	B
Calcium (+/- vitamin D) supplements	A	B	B
Dietary calcium	B	B	B
Smoking cessation	B	B	B
Reduced alcohol consumption	C	C	B
Fall prevention programmes		C	C
Hip protectors			B

Table 1 - Grade of recommendations for global strategies for the prevention of osteoporosis (NOGG 2010)⁵

⁴ Using data from the GPRD, van Staa et al (2001) reported much lower lifetime risks than NICE (2011a) for a 50 year old female of hip fracture (11.4%) and vertebral fracture (3.1%). The GPRD data probably exclude the majority of vertebral fractures (because they did not reach clinical attention) and this may account for the discrepant figures for vertebral fractures. However, the GPRD data are thought to be valid for hip fracture, and it is unclear why they give a lower risk of hip fracture than that stated by NICE (2011a).

⁵ Grade A = evidence levels Ia and Ib (from meta-analysis of RCTs or from at least one RCT); Grade B = evidence levels IIa, IIb and III (from at least one well designed controlled study without randomisation, from at least one other type of well designed quasi-experimental study or from well designed non-experimental descriptive studies); Grade C = evidence level IV (from expert committee reports or opinions and/or clinical experience of Authorities).

17. A Cochrane review assessing the effectiveness of exercise on reducing bone loss and preventing fractures in postmenopausal women was published after NOGG's guidelines (Howe et al 2011). The review found that people who engaged in combinations of exercise types had on average 3.2% less BMD loss at the spine than those who did not exercise and people who exercised by strength training had on average 1.03% less BMD loss at the neck of femur. No statistically significant effect was found on fractures.

If the carriers of a mutation are identified as a result of screening the natural history of people with this status should be understood, including the psychological implications.

18. Not relevant to screening for osteoporosis.

The Test

There should be a simple, safe, precise and validated screening test. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed

19. The World Health Organisation and the International Osteoporosis Foundation recommend that the reference technology for the diagnosis of osteoporosis is dual energy X-ray absorptiometry (DXA) applied to the femoral neck (NOGG 2010). Another test used to measure BMD is quantitative ultrasonography (QUS) of the heel. QUS is less expensive and more portable than DXA and does not expose patients to ionizing radiation.
20. The systematic review behind the USPSTF guidelines found that QUS of the heel can predict fractures as effectively as DXA (Nelson et al 2010). However USPSTF does not recommend using QUS for screening as the measurements obtained from QUS are not interchangeable with those obtained from DXA and it is not clear how QUS can be used to select people for medications that were proven to be efficacious on the basis of DXA criteria (USPSTF 2011). NOGG does not recommend the use of QUS for diagnosing osteoporosis as it has not been as well validated as DXA. However they state that this does not preclude the use of QUS in risk assessment (NOGG 2010 para 18).
21. The use of BMD alone to assess fracture risk has a high specificity but low sensitivity, so many osteoporotic fractures will occur in women who do not have osteoporosis as defined by a T-score ≤ 2.5 . The performance characteristics of the test can be improved by the concurrent consideration of risk factors that operate independently of BMD (NOGG 2010). Many risk assessment instruments have been developed to predict risk of low BMD and fractures, some of which can be used without a DXA scan.
22. The systematic review for the USPSTF identified 14 externally-validated risk assessment instruments that predict low bone density and 11 instruments that predict fractures (Nelson et al 2010). The authors concluded that risk assessment instruments are modest predictors of low bone density (area under the curve, 0.13 to 0.87) and fractures (area under the curve, 0.48 to 0.89). They concluded that simple and complex instruments perform similarly, but this statement seems to overlook the superior performance of the QFracture instrument for predicting fractures (area under the curve, 0.86 to 0.89). This instrument is relatively complex (17 items), but does not use a DXA scan and is based on risk factors that should be readily available in patients' health records. It was by far the largest study identified by Nelson (2010) and was developed from QRESEARCH, a UK database of anonymized general practice health records.
23. The performance of the WHO's FRAX tool, which is favoured by NOGG, was reported by Nelson et al (2010) as follows: area under the curve for osteoporotic fracture, 0.54–0.78; area under the curve for hip fracture, 0.65–0.81. However, these data cannot be compared with data given for QFracture in paragraph 22 above, because the area under the curve for osteoporotic fracture is derived from the fracture risk score and not from fracture probability. The importance of this distinction has recently been reviewed by Kanis et al (2012). The FRAX tool uses up to 12 items of clinical data to estimate risk (country, age, sex, weight, height,

previous fracture, parental history of hip fracture, smoking status, glucocorticoid use, history of rheumatoid arthritis, secondary osteoporosis and alcohol consumption ≥ 3 units per day). It can be used with or without a DXA scan.

24. An additional six studies evaluating externally-validated risk assessment instruments published after Nelson et al (2010) were found. The results of the studies all fell within the area under the curve range reported by Nelson et al (2010). The most noteworthy study, Collins et al (2011), was an independent evaluation of QFracture on 2.2 million adults registered with a general practice. It reported area under the curve characteristics of 0.89 for hip fractures and 0.82 for osteoporotic fractures in women. It is worth noting that although the tool is designed to be based on risk factors that are readily available and recorded in patients' health records, the study found that only 44.5% of women had complete data on smoking status, smoking category, alcohol consumption and body mass index. The authors sought to compare QFracture to FRAX, but the details for the calculation of an individual's risk using FRAX has not been published and the authors' request for an independent head-to-head comparison was not taken up by the developers of FRAX (Collins et al 2011).

25. In August 2012 NICE published Clinical Guideline 146 on assessing the risk of fragility fracture in osteoporosis (NICE 2012). The membership of the Guideline Development Group included the Chair of NOGG. The guidance states that two tools, FRAX and QFracture, are available for use in the UK and that it is not clear whether these tools are equally accurate and whether choice of tool should depend on circumstances.' Recommendations that are relevant to population-based screening in postmenopausal women include the following:
 - 1.1 Consider assessment of fracture risk in all women aged 65 years and over
 - 1.3 Estimate absolute risk when assessing risk of fracture (for example, the predicted risk of major osteoporotic or hip fracture over 10 years, expressed as a percentage).
 - 1.4 Use either FRAX (without a BMD value if a DXA scan has not previously been undertaken) or QFracture, within their allowed age ranges, to estimate 10-year predicted absolute fracture risk when assessing risk of fracture. Above the upper age limits defined by the tools, consider people to be at high risk.
 - 1.5 Interpret the estimated absolute risk of fracture in people aged over 80 years with caution, because predicted 10-year fracture risk may underestimate their short-term fracture risk.
 - 1.6 Do not routinely measure BMD to assess fracture risk without prior assessment using FRAX (without a BMD value) or QFracture.
 - 1.7 Following risk assessment with FRAX (without a BMD value) or QFracture, consider measuring BMD with DXA in people whose fracture risk is in the region of an intervention threshold for a proposed treatment, and recalculate absolute risk using FRAX with the BMD value.
 - 1.10 Consider recalculating fracture risk in the future if the original calculated risk was in the region of the intervention threshold for a proposed treatment and only after a minimum of 2 years.

Take into account that risk assessment tools may underestimate fracture risk in certain circumstances, for example if a person:

- has a high alcohol intake
- is taking high-dose oral or high-dose systemic glucocorticoids (more than 7.5 mg prednisolone or equivalent per day for 3 months or longer)
- has other causes of secondary osteoporosis.

26. The scope of NICE Clinical Guideline 146 does not include treatment intervention thresholds. It merely advises that healthcare professionals should follow local protocols or other national guidelines for advice on intervention thresholds (NICE 2012). As discussed below, NICE (2011a) and NOGG (2010) present very different perspectives on intervention thresholds. NICE will not be able to make further recommendations on drug interventions until its current technology appraisals dealing with drug interventions for reduced BMD are updated (Rabar et al 2012). As pointed out by Kanis et al (2012), the essential purpose of a risk assessment tool is to differentiate between those who should be offered treatment and those who should not; this means that until there is agreement on intervention thresholds it will not be possible to determine the relative merits of QFracture and FRAX for the purposes of a potential screening programme.

The test should be acceptable to the population

27. No studies were found that directly assess the acceptability of osteoporosis tests. A DXA scan is quick and painless and although it uses X-rays it is still likely to be considered safe by the population, as it is equivalent to only 1-2 day's exposure to natural background radiation and carries no significant fatal cancer risk (Damilakis et al 2010). A Scottish randomised controlled trial (RCT) assessing the effectiveness of screening for osteoporosis by DXA scan reported that 74% of women aged 45-54 years who were invited for screening in 1993 attended, which is a good indication that a DXA scan is an acceptable test (Barr et al 2010). Some risk assessment instruments only need information already recorded in GP records so these tests are likely to be acceptable to the population, but the acceptability of an osteoporosis screening programme as a whole would still need to be established.

There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals

28. The Royal College of Physicians recommends a set of blood tests (including standard tests for bone, liver, kidney and thyroid function) for men and women aged over 45 years who have or are at risk of osteoporosis (RCP 2000). These guidelines are endorsed in NOGG's recommendations (NOGG 2010 para 29).

If the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not being tested, should be clearly set out.

29. Not relevant to screening for osteoporosis.

The Treatment

There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment

Supportive treatments

30. NOGG makes recommendations on supportive treatment for high risk individuals which include promoting regular weight bearing exercise, modifying factors important in preventing falls (e.g. decreased visual acuity, consumption of medication that alters alertness and balance and home environmental hazards) and daily vitamin D and calcium supplementation in the housebound elderly and those in residential care and nursing homes (NOGG 2010 paras 32-36).

Drug treatments

31. The major interventions approved for the prevention and treatment of osteoporosis related fractures are bisphosphonates (alendronate, etidronate, ibandronate, risedronate and zoledronate), raloxifene, strontium, parathyroid hormone peptides, teriparatide and denosumab. All of these treatments have been shown to reduce the risk of vertebral fracture when given with calcium and vitamin D supplements (Strom et al 2011). There have been no trials to date assessing the effect on fracture-related morbidity and mortality (Nelson et al 2010). There have been no head-to-head trials with sufficient power to detect differences between interventions in terms of reduced risk of fractures.
32. Meta-analyses of the effects of treatments for osteoporosis that is detected through population-based screening should ideally include only women who have both osteoporosis and either no fractures, or only subclinical vertebral fractures. The literature search for this review did not find any meta-analyses that meet these criteria. The meta-analyses conducted by Nelson et al (2010) included RCTs that met one of the following criteria:
- Trial excluded individuals with previous vertebral or other presumably osteoporotic fractures.
 - Trial permitted individuals with previous osteoporotic fractures, but the overall proportion of participants with fractures was <20%, or the trial reported results separately for participants with and without previous fractures.
 - Trial did not report the proportion of participants with previous osteoporotic fractures, but inclusion criteria did not select individuals on the basis of presence of a previous fracture, and mean BMD T-scores were ≥ -3.0 .

These criteria allow inclusion of women who have osteopenia rather than osteoporosis, and may exclude some women with subclinical vertebral fractures. The meta-analyses used in NICE technology appraisal guidance 160 (NICE 2011a) included RCTs regardless of the baseline BMD and fracture status of the participants in the studies; these criteria allow inclusion of women who have osteopenia rather than osteoporosis, and inclusion of women with clinical fractures. Although neither the Nelson et al (2010) nor the NICE(2011a) meta-analyses are ideal for assessing the effects of treatments for screen-detected osteoporosis, they appear to be the best currently available. Table 2 shows that they reached broadly similar

conclusions, with the exception of whether alendronate specifically, and bisphosphonates as a group, produce a statistically significant reduction in hip fractures. Denosumab was not included in the NICE (2011a) TA160 guidance, but was looked at separately in guidance on denosumab for prevention of osteoporotic fractures developed through the single technology appraisal process (NICE 2010).

	RR (95% CI) for prevention of fractures			Reference
	Vertebral	Non-vertebral ⁶	Hip	
Alendronate	0.60 (0.44-0.83) 0.56 (0.46-0.68)	0.88 (0.55-1.4)	0.78 (0.4-1.38) 0.62 (0.4-0.98)	Nelson 2010 NICE 2011a
Bisphosphonates (any) ⁷	0.66 (0.50-0.89) 0.58 (0.51-0.67)	0.83 (0.64-1.08)	0.70 (0.44-1.11) 0.71 (0.58-0.87)	Nelson 2010 NICE 2011a
Parathyroid hormone peptides	0.32 (0.14-0.75)	0.97 (0.71-1.33)	No evidence	Nelson 2010
Raloxifene	0.61 (0.54-0.69) 0.65 (0.53-0.79)	0.97 (0.87-1.09)	0.97 (0.62-1.52) 1.13 (0.66-1.96)	Nelson 2010 NICE 2011a
Strontium	0.62 (0.55-0.71)	0.86 (0.74-0.99)	0.85 *(0.61-1.19)	NICE 2011a
Denosumab	0.32 (0.21-0.48)	0.81 (0.69-0.96)	0.61 (0.37-1.0)	NICE 2010
Teriparatide	0.36 (0.23-0.57)	0.49 (0.27-0.87)	0.25 (0.03-2.24)	NICE 2011a

Table 2 - Relative risks (RR) of the major pharmacological interventions for prevention of fractures

Results in bold are statistically significant.

The results for each intervention in Table 2 should not be directly compared due to the differences in populations involved. Across the interventions, the reduction in vertebral fracture rate ranges from 40-70%. A smaller reduction is seen for non-vertebral fractures, with the majority of interventions ranging from 5-20%.

33. Reduction in fracture risk has been shown to occur within one year of starting treatment for bisphosphonates, strontium and denosumab (Strom et al 2011). For most osteoporosis interventions the maximum length of treatment is not specified.
34. In response to post-marketing reports of rare but serious adverse events associated with bisphosphonates, such as atypical femur fractures, osteonecrosis of the jaw, and oesophageal cancer, the Food and Drug Administration performed a systematic review of long-term bisphosphonate efficacy. Their findings are as follows (Whitaker et al 2012):

The available data on long-term efficacy do not clearly identify subgroups of patients who are more likely to benefit from drug therapy beyond 3 to 5 years. Nevertheless, the emergence of safety concerns warrants consideration of new treatment algorithms for patients with osteoporosis. The available data do suggest that bisphosphonates may be safely discontinued in some patients without compromising therapeutic gains, but no adequate clinical trials have yet delineated how long the drugs' benefits are maintained after cessation ... Further investigation into the benefits and risks of long-term therapy, as well as surveillance of fracture risk after

⁶ Non-vertebral fractures are defined as fractures of the hip, leg, pelvis, wrist, hand, foot, rib or humerus, though this definition is not consistently applied to phase 3 studies (personal communication from NOGG).

⁷ The meta-analysis for bisphosphonates in NICE 2011a includes alendronate and risedronate only.

discontinuation of bisphosphonate therapy, will be crucial for determining the best regimen of treatment for individual patients with osteoporosis.

35. In a clinical commentary on this FDA report, Black et al (2012) wrote:

For clinicians, we believe that the current evidence base supports the following conclusions. Patients with low bone mineral density at the femoral neck (T score below -2.5) after 3 to 5 years of treatment are at the highest risk for vertebral fractures and therefore appear to benefit most from continuation of bisphosphonates. Patients with an existing vertebral fracture who have a somewhat higher (although not higher than -2.0) T score for bone mineral density may also benefit from continued therapy. Patients with a femoral neck T score above -2.0 have a low risk of vertebral fracture and are unlikely to benefit from continued treatment. We recognize that these conclusions, which are based on reductions in vertebral fractures, might change as additional data about long-term risks of bisphosphonate therapy become available.

Other adverse effects

36. Raloxifene has been shown to increase thromboembolic events (Nelson et al 2010), as has strontium ranelate (British National Formulary). Osteonecrosis of the jaw has been reported for denosumab (British National Formulary).
37. Gastrointestinal side effects are common with oral bisphosphonates. This, coupled with the complex instructions for administration, mean that compliance and adherence with treatment for osteoporosis is poor. Prescription-event monitoring studies in England indicate that 24% discontinued alendronate treatment within 1 year and 30% discontinued risedronate within 6 months (NICE 2011a). Poor adherence has been shown to be associated with reduced anti-fracture efficacy (Imaz et al 2010). However this effect may be overestimated since patients who fail to comply with placebo have poorer health outcomes than compliant patients (Strom et al 2011). For patients in whom they are recommended, annual injections with zoledronic acid or six monthly injections with denosumab provide an option to overcome the compliance issues related to weekly tablets.

Drugs recommended by NOGG and NICE

38. NOGG makes no distinction between prevention and treatment for pharmacological interventions. They recommend that alendronate be used as first-line treatment in the majority of cases due to its broad spectrum of anti-fracture efficacy and low cost. In women who are intolerant of alendronate or in whom it is contraindicated they recommend other bisphosphonates, denosumab, strontium or raloxifene. They only recommend using parathyroid hormone peptides in those at very high risk due to its high cost (NOGG 2010 para 39).
39. NICE have published separate guidance on primary and secondary prevention of osteoporotic fractures in postmenopausal women who have osteoporosis (NICE 2011a, NICE 2011b). The primary prevention guidance assessed the use of alendronate, etidronate, risedronate, raloxifene and strontium (NICE 2011a). Denosumab was assessed separately for primary and secondary prevention in guidance produced through the single technology assessment process (NICE 2010).

40. Essentially NICE recommends using alendronate for first-line treatment for the primary prevention of osteoporotic fractures, followed by etidronate and risedronate for second-line treatment and strontium for third-line treatment (NICE 2011a). NICE recommends denosumab for women who are unable to take alendronate, risedronate or etidronate (NICE 2010). Unlike NOGG, NICE does not recommend raloxifene for primary prevention. NICE has not made recommendations for the newer bisphosphonates, ibandronate and zoledronate. NICE has strict criteria on which women are eligible for treatment, and these are discussed further in the section below (NICE 2010 & 2011a).
41. The NICE criteria reflect assessments of cost-effectiveness, which are obviously sensitive to changes in the cost of the drugs. Since the cost-effectiveness analyses underpinning NICE TA160 were performed, both alendronate and risedronate have become available generically, and ibandronate and zoledronic acid will do so in the near future. Raloxifene has patent protection until 2014. The fall in costs of all these drugs will probably result in a re-evaluation by NICE in due course.

Treatments for osteopenia

42. The systematic review by Nelson et al (2010) identified only one RCT of treatment for reduced bone density among women without previous fragility fractures that reported results stratified according to baseline BMD (Cummings et al 1998). In women with osteopenia (T-scores between -1.6 and -2.0 or -2.0 and -2.5), there was a non-statistically significant trend towards decreased risk of vertebral fracture (RR, 0.82 [95% CI, 0.33–2.07] and RR, 0.54 [95% CI, 0.28–1.04], respectively), but no effect on any clinical fracture (RR, 1.14 [95% CI, 0.82–1.60] and RR, 1.03 [95% CI, 0.77–1.39], respectively). The literature search for the current paper did not identify any systematic reviews of treatment for osteopenia in women without previous clinically apparent fragility fractures. However, NOGG has highlighted a number of RCTs that include women with osteopenia. Having reviewed these RCTs we believe there is some evidence that some treatments for osteoporosis may also prevent fractures in some women with osteopenia. It is unclear to what extent this evidence is applicable to the full spectrum of women who would be found to have osteopenia if they were offered a population-based screening programme, and this question should be addressed through a systematic review.

There should be agreed evidence based policies covering which individuals should be offered treatment and the appropriate treatment to be offered

43. NICE and NOGG generally agree on which treatments should be offered for primary prevention of fractures. However, in two important respects they differ on which individuals should be offered treatment. The first major difference between NOGG and NICE is that, for women under 75, NICE only recommends primary preventative treatment for those who have osteoporosis confirmed by a DXA scan (NICE 2011a para 1.1). NOGG only requires women to have a DXA scan if their 10-year fracture probability is between a defined lower and upper assessment threshold. Women above the upper threshold can have treatment without the need for a DXA scan (NOGG 2010 para 50).
44. The second major difference between NOGG and NICE concerns thresholds for initiating treatment. NOGG recommends primary preventative treatment for women whose fracture risk is equivalent to the fracture risk of women of the same age who have already had a fracture.

Since the latter risk rises with age, the threshold for initiating primary preventative treatment also rises with age (NOGG 2010 para 51). NICE's treatment thresholds are fundamentally different, because they are based on identifying groups of women for whom treatment is expected to cost no more than £20,000 per QALY gained.⁸ NICE gives multiple treatment thresholds depending on the intervention, age and number of independent clinical risk factors for fracture. For each intervention recommended, the thresholds decrease with increasing age and number of risk factors (NICE 2011a paras 1.1-1.3). Table 3 shows the T-scores (SD) at (or below) which NICE recommends alendronate is taken.

Age	Number of indicators of low BMD and independent clinical risk factors for fractures		
	0	1	2
<65	nr	nr	-2.5 SD
65-69	nr	-2.5 SD	-2.5 SD
70-74	nr	-2.5 SD	-2.5 SD
75 or older	nr	-2.5 SD	any

Table 3 - T-scores at (or below) which NICE recommends alendronate is taken (NICE 2011a)

nr = not recommended

45. The literature search for this review did not identify any published comparisons of the 10-year probability of fracture at the age-specific interventions thresholds recommended by NOGG, versus the age-specific interventions thresholds recommended by NICE. However, NOGG have stated that the NICE thresholds are variously higher or lower than those provided by NOGG (personal communication from NOGG, November 2012).
46. NICE's treatment recommendations have been criticised on the grounds that many women who are eligible for alendronate, but unable to tolerate it (about one quarter of the total) would have to wait without any treatment till their BMD deteriorates sufficiently to become eligible for second-line treatment, which is more expensive (Dennison and Cooper 2011). The faq page of the NOGG website (www.shef.ac.uk/NOGG/faq.html) responds to criticisms that their treatment recommendations discriminate against the elderly, with an 80 year old woman having to have four times the risk of fracture compared to a 50 year old in order to access treatment.
47. The literature search for this review did not find any studies that compare the NICE guidelines against the NOGG recommendations in terms of their implications for the numbers of women who would be recommended for BMD measurement or drug treatment.

⁸ Kanis et al (2010) have argued that NICE's cost-effectiveness calculations are opaque, and neglect the impact of clinical risk factors on death hazards.

Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme

48. In 2007, the Information Centre commissioned a national evaluation of standards of care for osteoporosis and falls in primary care using the QRESEARCH general practice database (Hippisley-Cox et al 2007). They found that standards were highest in prescribing appropriate drugs, with almost three quarters of older women with diagnosed osteoporosis and a previous fragility fracture receiving the appropriate drugs. Standards for other aspects of care were lower, with only one in ten older women with a previous fragility fracture being referred for bone density assessment and less than one in fifty older people at high risk of falling being referred to a falls service or exercise programme.
49. NICE published clinical guidelines on the management of hip fractures in adults in June 2011 (NICE 2011c). The guidelines highlight a substantial variation in the extent, timing, manner and organisation of the necessary collaborative and multidisciplinary elements of effective hip fracture management, including the timely achievement of rehabilitation after surgery and delays in surgery. In addition to recommendations around analgesia, surgery and multidisciplinary rehabilitation, the guidelines make recommendations on effective imaging strategies to ensure hip fractures are not missed.
50. The National Hip Fracture Database (NHFD) was launched in 2007 with the aim of improving the care and secondary prevention of hip fracture in the UK. The 2011 NHFD National Report provides details on the casemix, care and outcomes of 53,443 cases of hip fracture from 176 hospitals (Currie et al 2011). The audit found that 66% of patients with hip fractures are discharged on bone protection medication (up from 57%) and 81% received a falls assessment prior to discharge (up from 63%). Since April 2010 the Department of Health Best Practice Tariff for hip fracture care has used financial incentives to drive adherence with the six core standards benchmarked by the NHFD, which include an assessment of bone health and risk of falling. In the two years following introduction of the tariff, the proportion of patients with fragility hip fracture for whom all six standards were met has risen from 24% to 55% (National Hip Fracture Database 2012).

The Programme

There should be evidence from high quality Randomised Controlled Trials that the screening programme is effective in reducing mortality or morbidity

51. No randomised controlled trials (RCTs) were found that assessed the effectiveness of an osteoporosis screening programme on reducing mortality. Two RCTs were identified that assessed the effectiveness of osteoporosis screening on reducing fracture incidence in Scotland (Barr et al 2005, Barr et al 2010). However these trials were conducted well over a decade ago and the results from the RCTs are no longer applicable, mainly due to changes in the recommended treatment.
52. In 1993, Barr et al randomised 4,800 women aged between 45 and 54 years of age to either receive a screening invitation letter to have a DXA scan at a screening unit no more than 32km away from their address, or to no invitation (Barr et al 2010). The results of the screened women were sent to their GP and the GP was advised to give them lifestyle advice and offer HRT to women “at risk” (defined as BMD in the lowest quartile of the first 1,000 women screened) when the women reached menopause, providing there were no contraindications. Nine years after randomisation, women were mailed a questionnaire that included questions on fractures and falls. Self reported fractures were verified through consultation of x-ray reports. Three quarters of those invited for screening attended and around two thirds of women were followed-up (60% in screening group and 57% in control group). The authors concluded that screening for osteoporosis reduces the incidence of fractures, but these conclusions were based on a per protocol analysis. A more appropriate intention-to-screen analysis, which includes all women randomised to screening, was also reported and showed a statistically non-significant reduction in risk of fractures of 20.9% (HR=0.79; 95%CI 0.6-1.04) and a small absolute risk reduction of 0.6%. These results are not a good reflection of the likely effectiveness of any current osteoporosis screening programme, because the trial population was younger than those likely to be eligible for screening today, the test did not include any form of fracture risk assessment, and HRT is no longer the first-line treatment for osteoporosis.
53. Another RCT (Barr et al 2005), which included 2,515 women aged over 70 years old, assessed the effectiveness of a screening visit that comprised a lifestyle questionnaire and a quantitative ultrasound heel scan. GPs of women in the active group who were found to be in the lowest quartile of broadband ultrasound attenuation and/or who had two or more risk factors for hip fracture were advised to prescribe a calcium and vitamin D supplement. One to three years after randomisation, women were mailed a questionnaire which included questions on falls and fractures. 68% of women randomised to screening attended and 83% of those screened and 79% in the control group were followed-up. A major flaw of the trial is that the women randomised to screening who did not attend were not followed up, so only a per protocol analysis can be carried out, which introduces selection bias. Excluding the non-attenders is likely to exaggerate the screening effect, because the baseline data show that the screening non-attender group was older and had poorer health on average than the screening attender group. Bearing in mind that the following estimate of effectiveness must be biased in favour of screening, the per protocol analysis shows a statistically significant 51% reduction in

fractures when adjusted for age and weight (OR=0.49; 95%CI, 0.30–0.81). In addition to the problems associated with using a per protocol analysis, this estimate is unlikely to be a good reflection of the effectiveness of a current screening programme, because women at high risk of fracture would now be offered bisphosphonates in addition to a calcium and vitamin D supplement, and quantitative ultrasound scan is unlikely to be used as a test.

54. A meta-analysis of RCTs of treatment (not screening) found an 11% reduction in mortality in the treated group (Bolland and Grey 2010). However, all the patients in one of the included studies (Lyles et al 2007) already had a hip fracture at the start of the trial, so the results are not directly relevant to population-based screening for osteoporosis. When the authors excluded this trial from the primary analysis, the overall result was no longer significant (RR of mortality = 0.94; 95% CI, 0.84–1.06).

There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically, socially and ethically acceptable to health professionals and the public

55. NOGG does not advocate population screening to identify patients with osteoporosis (NOGG 2010 paras 21, 47, 62; NICE 2011a para 1.8). NOGG recommends 'selective case finding' for the prevention of osteoporosis (NOGG 2010 para 32). The NICE appraisal refers to 'opportunistic identification, during visits to a healthcare professional for any reason, of postmenopausal women who are at risk of osteoporotic fragility fractures and who could benefit from drug treatment' (NICE 2011a para 1.8).
56. The high attendance rates reported in the screening RCTs conducted by Barr et al (68% & 74%) indicate that the public is likely to find a osteoporosis screening programme acceptable (Barr et al 2005, Barr et al 2010).

The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment)

57. Long-term evidence, ideally with a 10-20 year follow-up, would be needed to clarify whether the benefits of an osteoporosis screening programme would outweigh the harms. This ideal has to be balanced against the likelihood that the specific screening programme being evaluated will become dated as new tests or treatments emerge during the period of follow-up.

The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (ie. value for money).

58. Assessment against this criterion should be informed by an assessment of the cost of implementing a screening programme in relation to the expected benefits. This assessment should incorporate:
- the costs of operating a population-based call / recall system
 - the likely levels of uptake of the invitation to be screened and compliance with offers of DXA testing
 - the costs and benefits of the full range of treatments that are likely to be recommended as a result of screening, using up to date prices.

59. The health economic approach used to inform the treatment recommendations in NICE (2011a) TA160 did not incorporate any of the above. Instead, the approach involved identifying groups of women for whom the process of risk assessment using clinical criteria, followed by DXA scanning, followed by treatment with alendronate priced at £53.56 per year, would give an incremental cost-effectiveness ratio (ICER) of <£20,000 per QALY (NICE 2011a, para 4.2.14).

60. The health economic approach used to inform NICE (2011a) TA160 did not take account of incomplete compliance with offers of DXA testing, and it did not include the full range of treatments that are likely to be recommended as a result of screening. The guidance states that (NICE 2011a para 4.3.22):

The Committee considered an approach where the higher costs of risedronate and strontium ranelate were incorporated into the analysis by combining costs based on the estimated use of alendronate, risedronate and strontium ranelate. However, the overall cost effectiveness of such a combined approach for fracture prevention would be less favourable than that of alendronate. As a consequence, some women who would be eligible for treatment with alendronate as recommended in section 1.1 would not be offered treatment using such a combined approach. For this reason, the Committee did not consider the combined approach to be appropriate.

61. For these reasons, the health economic analyses that were used to inform NICE (2011a) TA160 cannot be used to imply that screening for osteoporosis is cost-effective, even if the groups of women offered screening were to match exactly the groups of women for whom treatment is recommended by NICE (2011a) TA160.

62. NOGG's recommendations were based on a cost-effectiveness analysis of alendronate in the management of osteoporosis conducted by Kanis et al (2008). Like NICE, the authors do not incorporate the criteria set out in paragraph 55, nor do they allow for incomplete compliance with offers of DXA testing and the full range of treatments that are likely to be recommended as a result of screening. In addition, both NICE (2011a) and Kanis et al (2008) assume that a drug intervention will be used for only five years, and that after stopping treatment the effect of treatment will reduce to zero in a linear manner over a five year period.

63. Kanis et al (2008) found that alendronate was cost effective at a willingness to pay of £20,000 per QALY at all ages in women with osteoporosis (T-score = -2.5 SD). They also looked at the cost-effectiveness of alendronate in women with clinical risk factors and found that, in the absence of BMD information, the combination of any two risk factors (prior fracture, family history, glucocorticoids, rheumatoid arthritis, alcohol > 3 units daily and current smoking) gives an incremental cost-effectiveness ratio (ICER) of less than £30,000 from the age of 50 years and less than £20,000 from the age of 65 years.

64. The main differences between the NICE (2011a) and Kanis et al (2008) cost-effectiveness analyses are set out in Table 5. Kanis et al (2008) conducted sensitivity analyses to model the effect of these differences and found that choice of time horizon had the most impact on the ICERs; they more than doubled when a 10-year rather than a lifetime horizon was used. Kanis et al (2008) point out that using a 10-year time horizon captures all the cost of treatment, but loses a component of the benefit and that NICE recommends using a lifetime horizon for chronic diseases. However, the NICE analysis does include lifetime mortality

effects, and they justify using a 10-year horizon for morbidity effects on the grounds of the uncertainty around health effects over a longer period of time and the older age groups involved (NICE 2011a). Indeed, it is not clear in the Kanis et al (2008) paper how the long term morbidity effects were modelled.

	NICE	Kanis et al (2008)
Discounting rates	6% for costs and 1.5% for benefits	3.5% for costs and benefits
Time horizons	Lifetime for mortality and 10-years for morbidity effects	Lifetime for mortality and morbidity effects
Source of evidence for effect size	Meta-analysis of alendronate and risedronate results conducted for the NICE guidelines (NICE 2008)	Meta-analysis of alendronate results conducted for the NICE guidelines (NICE 2008)
Costs for identifying women	Included	Only for those meeting the treatment threshold
Cost for alendronate	£53.56 per year ⁹	£95 per year
Source of evidence for costs of fracture	HRGs	Paper on the hospitalisation costs and outpatient costs of fragility fractures (Stevenson 2006)
Costs for side-effects	Included	Not included

Table 4 - Comparison of inputs for NICE (2011a) and Kanis et al (2008) cost-effectiveness analyses for alendronate in the management of osteoporosis

65. There is a need for a cost-effectiveness analysis of the screening programme as a whole, ideally based on the results of a RCT that assesses the effectiveness of a screening programme. Shepstone et al (2012) describe a planned UK seven-centre, unblinded, pragmatic, randomised controlled trial of screening with a 5-year follow-up period. A total of 11,580 women, aged 70 to 85 years and not on prescribed bone protective therapy will be consented to the trial by post via primary care, providing 90% power to detect an 18% decrease in fractures. Decisions about drug treatment will be based on thresholds determined by Kanis et al (2005) rather than current NICE guidance (NICE 2010, 2011a).

There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards

66. Such a plan and standards do not exist yet.

Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme

67. Currently, the UK has insufficient DXA facilities even for a case finding approach for the prevention and treatment of osteoporosis. A review of osteoporosis in the EU estimates that a case finding approach for assessing and monitoring the treatment of osteoporosis requires 10.6 central DXA units per million of the general population and 7,229 scans per year per million population in the UK,¹⁰ rising to an estimated 15,744 scans in 2025 (Strom et al 2011).

⁹ The NHS Drug Tariff price in 2012 is £13.56.

¹⁰ The Royal College of Radiologists commented that DXA scanners in the UK are frequently under-utilised as they are not staffed for full weeks of activity.

An audit of the availability of DXA units in 26 European countries showed that the UK had only 2.5 units per million population in 2007 and this was estimated to rise to 8.2 in 2010. The audit showed that the UK had the fourth lowest density of DXA units in Europe. It also highlighted an inequity of DXA units by geographical location in the UK resulting in long waiting times or long distances to travel (Strom et al 2011).

68. It is not known what the effect of introducing an osteoporosis screening programme would have on the number of DXA scans needed, because this number would depend on which testing approach was used. NICE's approach to eligibility for treatment would increase the number of DXA scans needed because most women would be required to have a scan, whereas assessment using the QFracture tool would reduce the number as it does not require BMD results to calculate risk. NOGG's approach would be somewhere in-between because some women would be at sufficiently high risk to warrant treatment without a DXA scan.

All other options for managing the condition should have been considered (e.g. improving treatment, providing other services), to ensure that no more cost effective intervention could be introduced or current interventions increased within the resources available.

69. As discussed in paragraphs 48-50, there is a need to improve the diagnosis of fracture as it can be easily missed on radiographs (NICE 2011c). In addition, there is a need to improve prescribing of osteoporosis treatment after fracture as currently only two thirds of people are discharged with bone protection treatment after a hip fracture (Currie et al 2011), and historically the figure has been as low as 20% for non-hip fractures (personal communication from Paul Miller). Further improvements needed in managing osteoporosis include improving compliance with treatment and reducing time to surgery after hip fracture. The cost-effectiveness of these suggested improvements, compared to that of a screening programme, is unknown.

Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice.

70. Patient information on osteoporosis is available at patient.co.uk, NHS choices and the National Osteoporosis Society websites.

Public pressure for widening the eligibility criteria, for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.

71. Eligibility criteria for screening and preventative treatment are not yet established.

Implications for Policy

72. It is not appropriate to implement a national screening programme for osteoporosis, because:

- There is no randomised controlled trial assessing the clinical and cost effectiveness of any current approach to screening for osteoporosis. The two trials that were identified (Barr et al 2005, Barr et al 2010) were conducted well over a decade ago and the results are no longer applicable, mainly due to changes in the recommended treatment. Results of the randomised controlled trial of screening described by Shepstone et al (2012) are expected in 2015 (personal communication from Dr Shepstone).
- There is a lack of consensus between two leading UK sources of guidance in this area (the National Osteoporosis Guideline Group (NOGG) and the National Institute for Health and Clinical Excellence (NICE)) regarding which women should be eligible for treatment. NOGG recommends primary preventative treatment for women whose fracture risk is equivalent to the fracture risk of women of the same age who have already had a fracture. NICE's treatment thresholds are fundamentally different, because they are based on identifying groups of women for whom treatment is expected to cost no more than £20,000 per QALY gained.
- The long term clinical and cost effectiveness of osteoporosis treatment is not known.

73. Even if the obstacles to implementation of a national screening programme for osteoporosis were overcome, the impact of screening on the population burden of fractures would be limited. This is because most fractures occur either among women who do not have osteoporosis (each of whom is at low risk of a fracture, but the majority of women are in this group) or among women who have already had a clinically apparent fracture (of whom there are smaller numbers, but each woman in this group is at relatively high risk of a further fracture). If they were screened the first group (women without osteoporosis) would be told that it is not yet clear which of them would be likely to benefit from treatment. Population-based screening is of no relevance to the second group (women who have already had a clinically apparent fracture), because their increased risk of future fractures has already been signaled by their initial fracture, and population screening would provide no new information.

74. The following five charts show, using the same vertical scale, the numbers of fractures that would be expected over a 5 year period in cohorts of 100,000 women in various age bands. The data sources and methods used for these approximate estimates are described in Appendix B. Each cohort of 100,000 women is split into four groups:¹¹

Group A. Women without osteoporosis. (If they accepted an offer of screening, these women would be told that it is not yet clear which of them would be likely to benefit from treatment.)

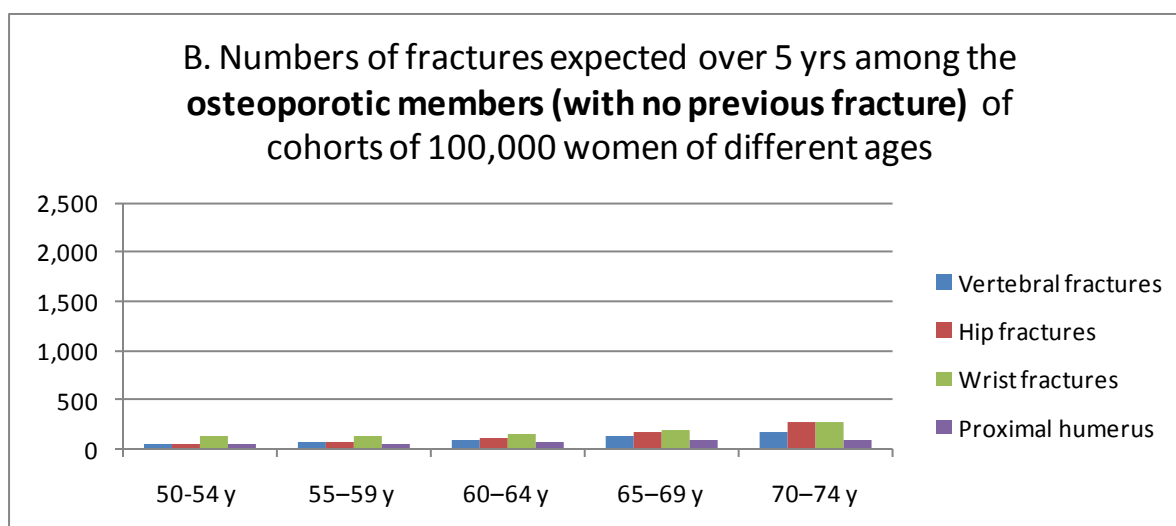
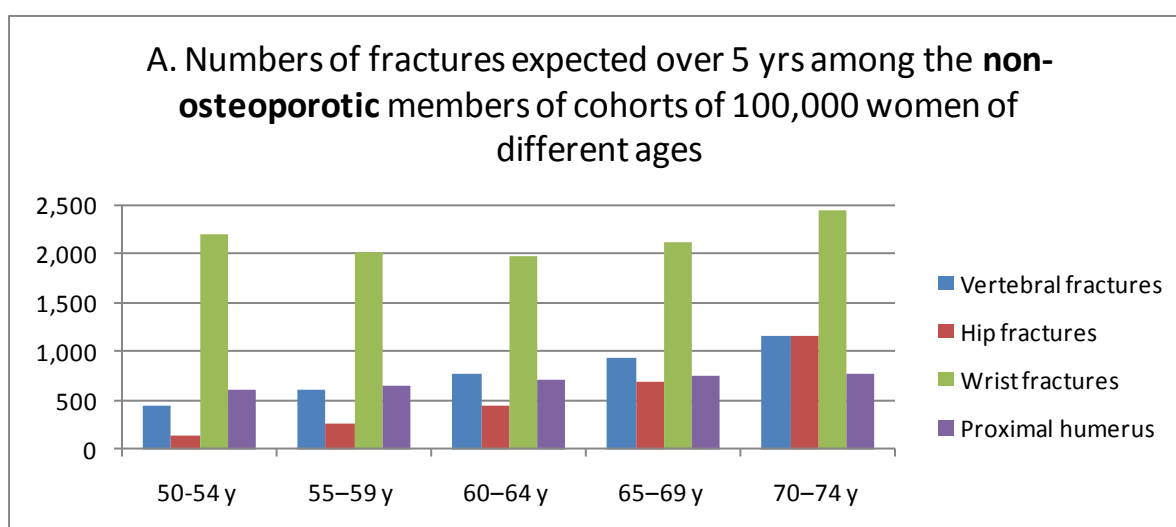
¹¹ The cohort of 100,000 women in each age band is split across groups A-D according to the prevalence of osteoporosis and severe osteoporosis among women of that age. The prevalence of severe osteoporosis rises steeply with age, from 0.49% at 50-54 yrs to 22.4% at 75-79 yrs. The prevalence of non-severe osteoporosis does not increase very much with age: it rises from 2.8% at 50-54 yrs to 6.4% at 75-79 yrs (Stephenson 2009, Table 40).

Group B. Women with osteoporosis but no previous fracture. (These women are one group of potential beneficiaries of screening, because they might be offered treatment on the basis of their screening result.)¹²

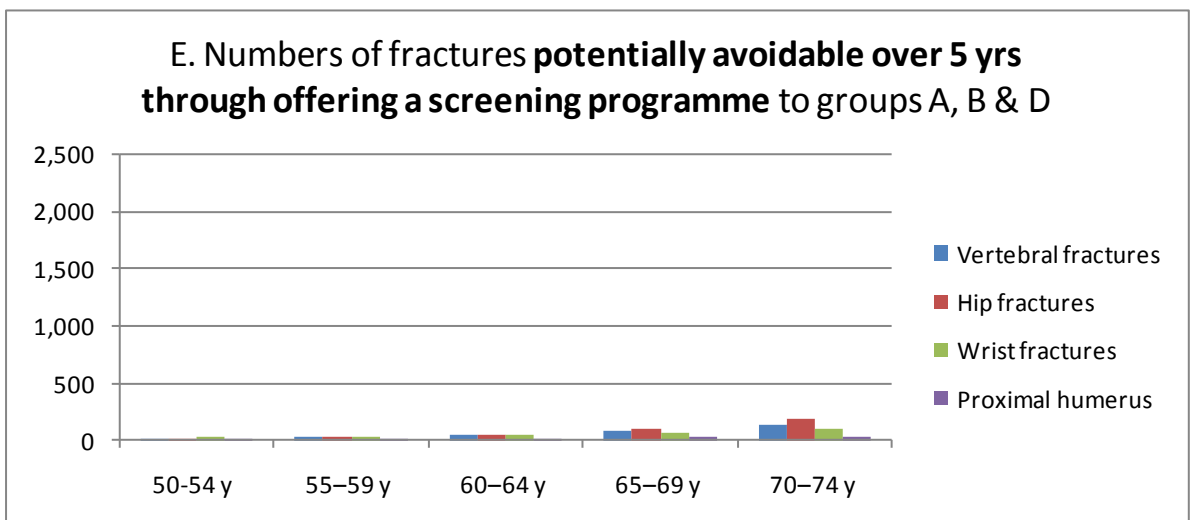
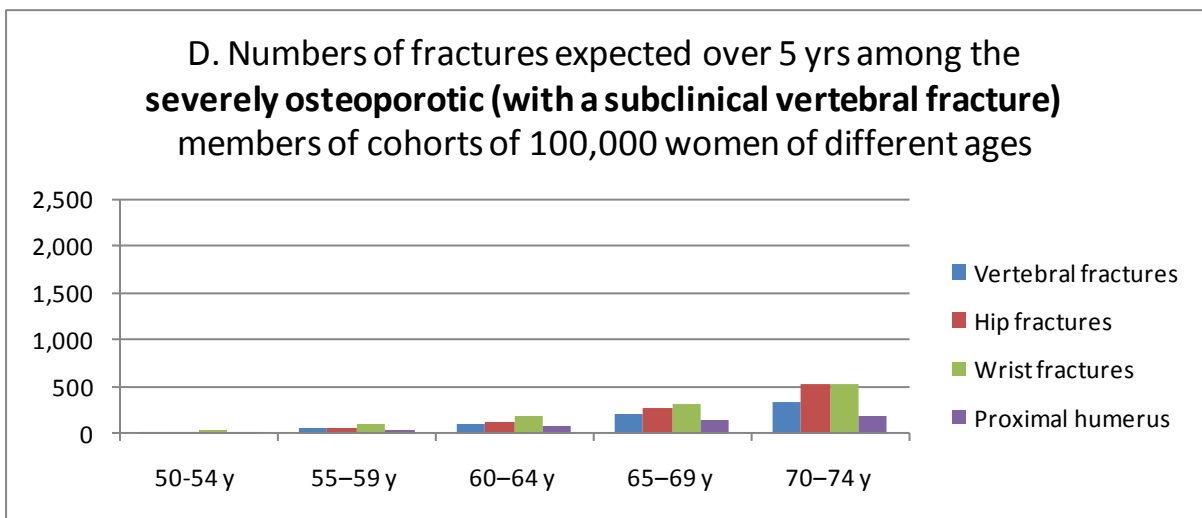
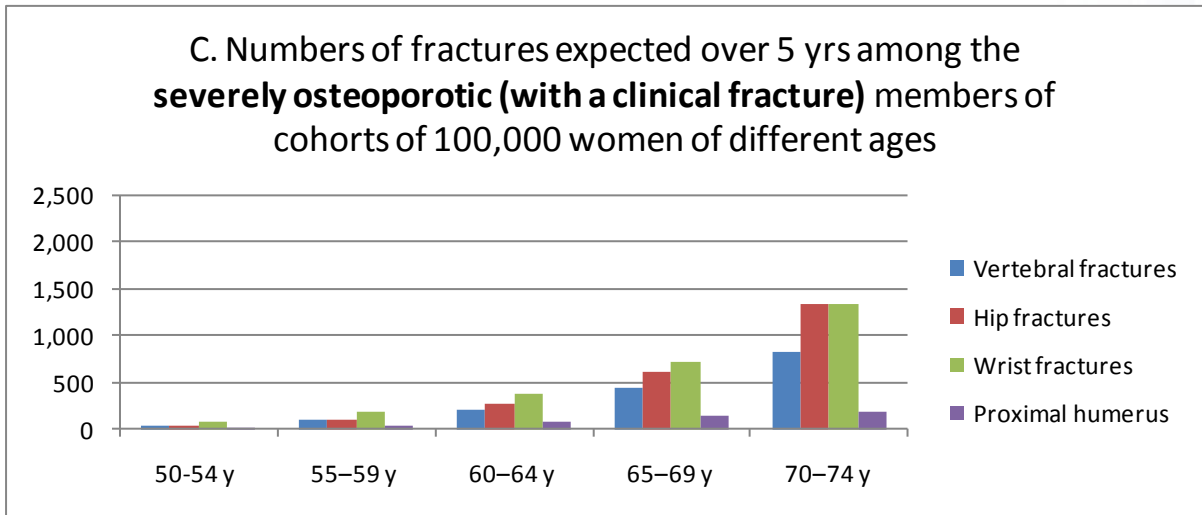
Group C. Women with severe osteoporosis and a previous clinically apparent fracture. (These women will not benefit from screening because their increased risk of future fractures has already been signaled by a previous fracture.)

Group D. Women with severe osteoporosis and a previous subclinical vertebral fracture. (These women are also potential beneficiaries of screening, because neither their osteoporosis nor their previous vertebral fracture is known to their doctors, and they would be candidates for treatment.)

Group E shows the numbers of fractures that might be prevented over 5 years if women of various ages in groups A,B and D were offered screening, and were then offered treatment if they were found to have osteoporosis.



¹² Whether women in Group B are offered treatment or not would depend on whether NOGG's or NICE's guidelines were adopted.



75. When comparing the numbers of fractures in groups E with the numbers in groups A to D, it is evident that only a small proportion of all the fractures in the population could be prevented through a screening programme. For example, a cohort of 100,000 women aged 65-69 years initially could expect the following numbers of hip fractures in each group over the following five years:

	Number of women in group	Number of hip fractures expected over 5 years
Group A. Women without osteoporosis	85,770	694
Group B. Women with osteoporosis but no previous fracture	4,770	174
Group C. Women with severe osteoporosis and a previous clinical fracture	6,386	620
Group D. Women with severe osteoporosis and a previous sub-clinical vertebral fracture	3,075	275
Totals	100,000	1,764

76. Only $(174+275) / 1,764$ (25%) of all the hip fractures expected among women aged 65-69 years would be expected to occur in groups B and D i.e. among women whose osteoporosis would be identified through population-based screening. Assuming 75% uptake of screening (Barr et al 2005), 100% eligibility for treatment among those who are found to have osteoporosis,¹³ 75% compliance with treatment for five years (NICE 2011a), and a relative risk of hip fracture of 0.70 among women taking bisphosphonate treatment (Nelson 2010), 101 hip fractures would be prevented over five years as a result of offering a screening programme to the 93,614 members of this cohort who would be eligible for screening.¹⁴ This example illustrates the general point that a screening programme for osteoporosis will not achieve a substantial reduction in the total number of fractures in the population.
77. Another way of expressing the same result is to say that to prevent one hip fracture about 950 women aged 65-69 years would need to be offered screening, of whom 26 would then need to take bisphosphonate treatment for up to five years.

¹³ This is controversial, as described in paragraphs 43-46. The assumption used here probably exaggerates the likely impact of screening.

¹⁴ The 6,386 women with a previous clinical fracture would not be eligible for population-based screening because they are already known to their doctors, who should therefore try to prevent future fractures as part of good clinical care, not as part of a population-based screening programme.

Implications for Research

78. Given the apparent limitations of a screening programme for osteoporosis in terms of its ability to reduce the total burden of fractures in the population, further research on the efficacy, uptake and compliance of the primary prevention approaches listed in paragraph 16 would be valuable.
79. A screening programme would have greater potential to reduce the population burden of fractures if clearer evidence were available regarding which women with screen-detected osteopenia would benefit from treatment to reduce the risk of fractures. A systematic review of existing RCT data on women with osteopenia, but without clinically apparent fractures, would clarify this.
80. Research questions specific to a screening programme include:
- Is a single screen for osteoporosis sufficient, and if so at what age? If not, what are the optimal ages and intervals for repeated screening, and what is the appropriate age to stop screening?
 - What is the optimal duration of treatment (5-years, 10-years, life-long or until BMD or risk score goes below the treatment threshold)?
 - What are the long term effects of treatment? What is the likely cost-effectiveness of a screening programme for osteoporosis, bearing in mind:
 - the costs of operating a population-based call / recall system
 - the likely levels of uptake of the invitation to be screened and compliance with offers of DXA testing
 - the costs and benefits of the full range of treatments that are likely to be recommended as a result of screening, using up to date prices.

Appendix A

Knowledge update on screening for osteoporosis Paula Coles, Information Scientist June 2011

BACKGROUND: The previous review on osteoporosis screening in postmenopausal women was carried out in February 2002. A knowledge update was then published in February 2005. Both these documents are available on the osteoporosis policy page on the UK National Screening Committee's website: <http://www.screening.nhs.uk/osteoporosis> [accessed 23 June 2011] The literature was therefore searched from January 2005 to June 2011.

The previous review on screening for osteoporosis concluded that there was a major drawback in the use of bone mineral density measurement as a screening tool in that there was "...no value of BMD that discriminates well between patients who get a fracture and those who do not." And "BMD measurements have a high specificity but a low sensitivity. A negative test (suggesting the absence of osteoporosis) indicates low risk of fracture but the low sensitivity of 50% means the prediction that half of all osteoporotic fractures will occur in women who were not detected as having osteoporosis. Hence there is little point in using the BMD as a tool for identifying fracture risk."

The National Institute for Health and Clinical Excellence (NICE) have produced guidance on primary and secondary prevention on osteoporosis:

- TA160 Osteoporosis - primary prevention: guidance (updated January 2011)
- TA161 Osteoporosis - secondary prevention including strontium ranelate: guidance (updated January 2011)
- TA204 Osteoporotic fractures - denosumab: guidance (updated October 2010)

NICE are also in the process of developing the following guidance: Osteoporosis: risk assessment of people with osteoporosis.

All complete and ongoing NICE guidance on osteoporosis is accessible via their website: <http://guidance.nice.org.uk/Topic/Osteoporosis> [accessed 23 June 2011]

The focus of the search strategy was designed to reflect this and therefore as a result was focussed on the test and its ability to predict fracture and the screening programme.

SOURCES SEARCHED: Medline (OvidSP), Embase, Cinahl, Web of Science and the Cochrane Library.

DATES OF SEARCH: January 2005 – June 2011

SEARCH STRATEGY: Medline OvidSP

1. osteoporos\$.tw. (41115)
2. exp Osteoporosis/ (38305)
3. Osteoporotic Fractures/ (235)
4. 1 or 2 or 3 (53192)
5. Mass screening/ (70289)
6. screen\$.tw. (359890)
7. (test or tests or testing).tw. (1226421)

8. detect\$3.tw. (1234177)
9. Risk Assessment/ (130391)
10. Fractures, Bone/pc (2810)
11. Bone density/ (33446)
12. BMD.tw. (15563)
13. bone mineral density.tw. (21057)
14. Risk Factors/ (446963)
15. risk factor\$.tw. (261489)
16. fracture risk\$.tw. (4667)
17. 5 or 6 or 7 or 8 or 9 or 10 (2624000)
18. 11 or 12 or 13 or 14 or 15 or 16 (604373)
19. 4 and 17 and 18 (6448)
20. limit 19 to (clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or controlled clinical trial or evaluation studies or guideline or meta analysis or practice guideline or randomized controlled trial or "review") (2578)
21. limit 20 to yr="2005 -Current" (1318)

Similar searches were also carried out in Embase, Cinahl, Web of Science and the Cochrane Library.

All searches carried out on 6 June 2011

Medline	1318
Embase	806
Cochrane Library	266
Web of Science	1008
Cinahl	647
Total	4045

Inclusions and exclusions

The above search strategies retrieved 4045 references in total. After duplicate references were removed a total of 3069 potentially relevant references were left. The title and abstracts of the remaining citations were scanned for relevance to screening for osteoporosis, focussing on postmenopausal women (aged over 50) and the following NSC criteria (for the reasons outlined in the background note above):

- The test (and its ability to predict fracture risk)
- The screening programme

However, systematic reviews and guidelines on treatment/prevention that were recovered have been included as have references regarding the condition that were recovered.

481 references were deemed to be relevant

Two further studies were published after the search date and are included in the final results, as is the completed NICE guidance referred to above, meaning a total of 486 potentially relevant results that have been classified in to the categories below according to the NSC criteria. There will inevitably be some overlap between categories.

Systematic reviews and meta-analyses <ul style="list-style-type: none"> • Risk factors (6) • Genetics (2) • Management (4) • The test (10) • The treatment (61) - Pharmacological prevention/treatment (48) - Exercise prevention (5) - Supplements (8) • The screening programme (4) • Costs (1) 	88
Guidelines	36
The condition <ul style="list-style-type: none"> • Reviews (19) • Epidemiology (9) • Risk factors (22) • Quality of life (14) • Costs (28) 	92
The test/Fracture risk prediction tools <ul style="list-style-type: none"> • Bone Mineral Density (BMD) reviews (11) • BMD (7) • BMD measured by dual energy x-ray absorptiometry (17) • BMD measured by quantitative ultrasound (10) • BMD measured by computed tomography (3) • BMD measured by magnetic resonance imaging (2) • X-rays (4) • Comparison of BMD measurement tools (8) • BMD and clinical risk factors (49) • FRAX (44) • Clinical risk factors (28) • Previous fracture (11) • Biomarkers (11) • Comparison of risk prediction tools (23) 	228
The screening programme	42
Total	486

Appendix B

Data sources used in paragraphs 74-77 to estimate numbers of fractures in different groups of women

Data	Source
Prevalence of osteoporosis and severe osteoporosis	Stephenson (2009:Table 40)
1-yr risk of fracture in those without osteoporosis	Stephenson (2009:Table 19)
1-yr risk of fracture in those with osteoporosis (excluding those with severe osteoporosis)	Stephenson (2009:Tables 17,18, 20 and 38)
1-yr risk of fracture in those with severe osteoporosis	Assumed to be double the risk of those with osteoporosis (Stephenson 2009:87)

Method used to calculate the age specific risk of each type of fracture in women with osteoporosis (excluding those with severe osteoporosis)

1. Stephenson (2009:Table 20) gives the 1-yr risk of fracture by age for a woman with a T-score of -2.5 and no previous fracture.
2. These figures underestimate the risk for all women with osteoporosis but no previous fracture, because many of them will have a T-score of less than -2.5 .
3. Stephenson (2009:87) describes the following formula for adjusting fracture risks according to changes in T-score:

$$\text{Adjusted fracture risk} = R \times I^Z$$

where

R = unadjusted fracture risk

I = increased risk of fracture per Z-score

Z = fall in Z-score between -2.5 and age-specific average T-score

4. Stephenson (2009:Table 38) gives the average T-score by age for women with a T-score of -2.5 SD or less, which allows Z to be calculated. For example, the average T-score for women aged 65-69 yrs who have a T-score of -2.5 SD or less is -2.84 . For this age group Z is therefore $2.84 - 2.5 = 0.34$.
5. Stephenson (2009:Tables 17 and 18) give values for 'I' (the increased risks of fracture per Z-score).

Caveat

This method may slightly overestimate the risk of fracture in women with osteoporosis, because the T-scores given in Stephenson (2009:Table 38) include some women with severe osteoporosis. The validity of the method was checked by converting the calculated absolute fracture risks to relative risks (relative to the fracture risk among women without osteoporosis), and checking the results against the relative risks given by Stephenson (2009:Table 39). The figures do not match exactly but the discrepancies are small, indicating that the method is sufficiently accurate to support the broad conclusions reached about the impact of screening.

Appendix C

Method for estimating the numbers of women with severe osteoporosis and a previous subclinical vertebral fracture

1. Strom et al (2011, Table 32) gives the numbers of new fractures in women aged 50 years or more in the UK in 2010. The numbers do not include subclinical vertebral fractures, but these can be added to the figures as follows, using the estimate in Strom et al (2011) that only 80% of vertebral fractures in women are subclinical:

Source	Hip	Vertebral (clinical)	Vertebral (all)	Forearm	Other	All sites
Strom (2011) Table 32	56,735	40,369	Not stated	54,309	191,781	343,194
After adding subclinical vertebral fractures	56,735		= 5 x 40,369	54,309	191,781	504,672

2. Among women aged 50 years or more, the proportion of all fractures that are subclinical vertebral fractures is $4 \times 40,369 / 504,672$ (32%).
3. This percentage varies with age. Prevalence data in Strom et al (2011, Table 34) show that the ratio of prevalent hip fractures to prevalent clinical vertebral fractures increases from 0.4 among women aged 50-64 years to 1.4 among women aged 85 years or more. From this it can be estimated that the proportion of all fractures that are subclinical vertebral fractures ranges from approximately 40% in women aged 50-54 years down to approximately 28% in women aged 75-79 years.

Caveats

The results should be treated with caution, because:

1. They combine incidence ratios with prevalence ratios.
2. The adjustment for variation by age in the proportion of all fractures that are subclinical vertebral fractures is only a crude approximation

References

Barr RJ, Stewart A, Torgerson DJ, Seymour DG, Reid DM. Screening elderly women for risk of future fractures--participation rates and impact on incidence of falls and fractures. *Calcified Tissue International* 2005; 76(4):243-248.

Barr RJ, Stewart A, Torgerson DJ, Reid DM. Population screening for osteoporosis risk: A randomised control trial of medication use and fracture risk. *Osteoporosis International* 2010; 21 (4):561-568.

Black D, Bauer D, Schwartz A, Cummings S, Rosen C. Continuing bisphosphonate treatment for osteoporosis — for whom and for how long? *NEJM* 366:2051-53.

Bolland M, Grey A. Disparate Outcomes from Applying UK and US Osteoporosis Treatment Guidelines. *Journal of Clinical Endocrinology & Metabolism* 2010; 95(4):1856-1860.

Bolland M, Grey A, Gamble G, Reid I (2010). Effect of osteoporosis treatment on mortality: A meta-analysis. *J Clin Endocrinol Metab* 95(3):1174–1181.

Cummings SR, Black DM, Thompson DE, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA* 1998;280(24):2077-2082.

Currie C, Partridge M, Plant F, Roberts J, Wakeman R, Williams A. The National Hip Fracture Database National Report 2011. 2011.

Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. *British Medical Journal* 2011; 342:d3651 doi: 10.1136/bmj.d3651.

Damilakis J, Adams J, Guglielmi G, Link T. Radiation exposure in X-ray-based imaging techniques used in osteoporosis. *Eur Radiol.* 2010;20(11):2707-14.

Dennison EM, Cooper C. Where now with NICE? *Osteoporosis International* 2011; 22:1007-1008.

Gauthier A, Kanis J, Jiang Y, Martin M, Compston J, Borgstrom F et al. Epidemiological burden of postmenopausal osteoporosis in the UK from 2010 to 2021: estimations from a disease model. *Archives of Osteoporosis* 2011;1-10.

Hippisley-Cox J, Bayly J, Potter J, Fenty J, Parker C. Evaluation of standards of care for osteoporosis and falls in primary care. QRESEARCH and The Information Centre for health and social care, editor. 2007.

Howe T.E., Shea B., Dawson L.J., Downie F., Murray A., Ross C. et al. Exercise for preventing and treating osteoporosis in postmenopausal women. *Cochrane Database of Systematic Reviews* 2011;(7).

Imaz I, Zegarra P, Gonzalez-Enriquez J, Rubio B, Alcazar R, Amate JM. Poor bisphosphonate adherence for treatment of osteoporosis increases fracture risk: systematic review and meta-analysis. *Osteoporosis International* 2010; 21(11):1943-1951.

Kanis JA, Borgstrom F, Zethraeus N et al (2005) Intervention thresholds for osteoporosis in the UK. *Bone* 36:22–32.

Kanis J.A., Adams J., Borgstrom F., Cooper C., Preedy D., Selby P. et al. The cost-effectiveness of alendronate in the management of osteoporosis. *Bone* 2008; 42:4-15.

Kanis JA, McCloskey E, Jönsson B, Cooper A, Ström O Borgström F (2010) An evaluation of the NICE guidance for the prevention of osteoporotic fragility fractures in postmenopausal women. *Arch Osteoporos* 5: 19-48 DOI 10.1007/s11657-010-0045-5

Kanis JA, Oden A, Johansson H, McCloskey E (2012) Pitfalls in the external validation of FRAX. *Osteoporos Int* 23: 423-31.

Lyles KW, Colón-Emeric CS, Magaziner JS, Adachi JD, Pieper CF, Mautalen C, et al (2007) Zoledronic acid and clinical fractures and mortality after hip fracture.. *N Engl J Med* 357:1799–1809.

Mitchell PJ (2011). Fracture Liaison Services: the UK experience *Osteoporos Int* 22 (Suppl 3):S487–S494.

National Collaborating Centre for Nursing and Supportive Care (2008). Systematic reviews of clinical effectiveness prepared for the guideline 'Osteoporosis: assessment of fracture risk and the prevention of osteoporotic fractures in individuals at high risk'.

National Institute for Health and Clinical Excellence (2010). Denosumab for the prevention of osteoporotic fractures in postmenopausal women. NICE technology appraisal guidance 204.

National Institute for Health and Clinical Excellence (2011a). Alendronate, etidronate, risedronate, raloxifene and strontium ranelate for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE technology appraisal guidance 160 (amended).

National Institute for Health and Clinical Excellence (2011b). Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women (amended). NICE technology appraisal 161 (amended).

National Institute for Health and Clinical Excellence (2011c). The management of hip fracture in adults. NICE Clinical Guideline 124.

National Institute for Health and Clinical Excellence (2012). Osteoporosis: assessing the risk of fragility fracture. NICE Clinical Guideline 146

National Hip Fracture Database (2012). May Newsletter available at <http://www.nhfd.co.uk/003/hipfractureR.nsf/resourceDisplay> accessed 31 August 2012.

National Osteoporosis Guideline Group. Osteoporosis - Clinical guideline for prevention and treatment: Executive Summary, updated July 2010.

Nelson HD, Helfand M, Woolf SH, Allan JD. Screening for Postmenopausal Osteoporosis: A Review of the Evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 2002; 137(6):529-541.

Nelson HD, Haney EM, Dana T, Bougatsos C, Chou R. Screening for Osteoporosis: An Update for the US Preventive Services Task Force. *Annals of Internal Medicine* 2010; 153(2):99-111.

Rabar S, Lau R, O'Flynn N, Li L, Barry P (2012). Risk assessment of fragility fractures: summary of NICE guidance. *BMJ* 345:43-44.

Royal College of Physicians (1999). Osteoporosis: Clinical guidelines for prevention and treatment.

Royal College of Physicians (2000). Osteoporosis: Clinical guidelines for prevention and treatment. Update on pharmacological interventions and an algorithm for management.

Shepstone L, Fordham R, Lenaghan E, Harvey I, Cooper C, Gittoes N et al (2012). A pragmatic randomised controlled trial of the effectiveness and cost-effectiveness of screening older women for the prevention of fractures: rationale, design and methods for the SCOOP study. *Osteoporosis Int.* 10:2507-15.

Slovik DM. Screening for Osteoporosis in Postmenopausal Women. In: Sonya Seigafuse, Ryan Shaw, editors. *Primary Care Medicine: Office Evaluation and Management of the Adult Patient*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2011. 1006-1008

Stevenson M, Jones ML, De Nigris E, Brewer N, Davis S, Oakley J. A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis. *Health Technology Assessment* 2005; 9(22):1-160.

Stevenson M, Davis S, Lloyd-Jones M, Beverley C. The clinical effectiveness and cost-effectiveness of strontium ranelate for the prevention of osteoporotic fragility fractures in postmenopausal women. *Health Technology Assessment* 2007;11(4)

Stevenson M, Lloyd-Jones M and Papaioannou D. Vitamin K to prevent fractures in older women: systematic review and economic evaluation. *Health Technology Assessment* 2009;13(45)

Ström O, Borgström F, Kanis JA, Compston J, Cooper C, McCloskey EV et al. Osteoporosis: burden, health care provision and opportunities in the EU A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Archives of Osteoporosis* 2011; DOI 10.1007/s11657-011-0060-1.

U.S. Preventive Services Task Force. Screening for osteoporosis in postmenopausal women: recommendations and rationale. *Ann Intern Med.* 2002;137:526-8.

U.S. Preventive Services Task Force. Screening for Osteoporosis: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine* 2011; 154:356-364.

van Staa T, Dennison E , Leufkens H, Cooper C (2001). Epidemiology of Fractures in England and Wales. *Bone* Vol. 29(6)17–522.

Whitaker M, Guo J, Kehoe T, Benson G. Bisphosphonates for osteoporosis – where do we go from here? *NEJM* 366:2048-51.