UK NATIONAL SCREENING COMMITTEE

Screening for Osteoporosis in Postmenopausal Women

20 March 2013

Aim

1. To agree the UK National Screening Committee’s (UK NSC) formal policy position on screening for osteoporosis in postmenopausal women.

Background

2. A review of screening for osteoporosis in postmenopausal women against the UK NSC criteria was carried out in August 2012 by Lisa Peto, a Specialty Registrar in Public Health and Dr Martin Allaby a Consultant in Public Health Medicine at Solutions for Public Health.

3. The scope of this review 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.

Consultation

4. A public consultation on the screening review took place between 19th September 2012 and 19th December 2012. Four responses were received to the consultation which are available at Annex A. Comments from all four organisations have been reviewed and the screening review has been revised following these comments. Also enclosed at Annex A is Solutions for Public Health’s response to the comments received from the National Osteoporosis Guideline Group and Sectra Medical Systems.

Conclusion

5. 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 quality-associated life-years gained.

• The long term clinical and cost effectiveness of osteoporosis treatment is not known.

6. 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.

Recommendation

7. The UK NSC is asked to agree the policy position on screening for osteoporosis in postmenopausal women as follows:-

A national screening programme to screen postmenopausal women for osteoporosis is not recommended.

8. The UK NSC is asked to agree that the policy should be reviewed in three years time unless there is significant new peer reviewed evidence in the meantime.
Sectra Medical Systems

We read with interest the review over primary osteoporosis screening for the UK National Screening Committee. Although we appreciate the work of the authors in compiling the extensive data available we found that information regarding our screening technology was not presented. Therefore, we would like to present a primary osteoporosis screening solution, which has proven to be efficient and cost-effective in the clinical setting. We intend to show the committee that our screening solution have potential to address many of the difficulties facing a screening programme, like costs, availability and implementation. Furthermore, we attempt to show clinical evidence to support the main function of a screening programme; to reduce the number of osteoporosis related fragility fractures.

In the description of current technologies (section 19) we miss the mention of Digital X-ray Radiogrammetry (DXR) as a method to determine bone mineral density (BMD). DXR is a novel method to determine the increased risk of future fracture by measuring the BMD of the bones in the hand from a single hand x-ray image. Although new, there is substantial evidence that DXR predicts fractures at the hip [1], vertebrae [2] and wrist [3]. In all cases the risk gradient of DXR for fracture prediction (in risk increase per standard deviation change in T-score) was similar to the one reported for hip DXA [4], the current gold standard. In short, BMD retrieved from DXR predicts future fractures. Also, in a recent study it was shown that the area under the curve for the prediction of future hip fracture with DXR was 0.89 for women and 0.84 for men [1]. This is superior to other risk assessment methods like FRAX [5] and comparable to central DXA measurement [6]. However, the omission of DXR in this review is understandable since the technology is not mentioned in the USPSTF, NICE or NOGG reports; likely because the most important clinical evidence [1] was published recently, after the release of these reports. Although novel, DXR is not a fringe technology. It is FDA approved and CE-marked and is in clinical use in the UK at the Holly House Hospital in Essex.

BMD measuring with DXR is easy and quick to do, and do not need any new or specialised hardware making it suitable to use in the screening of large populations. DXR can be measured on any digital X-ray machine, which makes it simple and cheap to implement. One setup for screening involves the combination of DXR with mammography. Since, postmenopausal women are at risk for both breast cancer and osteoporosis the combination of these two programmes makes sense. The mammography X-ray unit is used to collect the hand image for osteoporosis screening. This combination has been shown to work very well in clinical practice, also at high volume screening sites, with only minor effect on the mammography workflow and with high interest from women to take part in the combined screening. Currently this setup exists at several clinics in Norway, Sweden, Germany, Italy and the UK.

The possibility to combine the screening programmes for osteoporosis and breast cancer with minimal effort required from the women attending as well as the staff at
the mammography clinic also makes the screening programmes more cost-effective.

In our externally validated Health Economy model for the UK [7] (described in detail in the next section) primary screening with DXR of women from the age of 65 is cost-saving to society. DXR is intended for primary screening, identifying individuals at risk before they fracture. However, we believe that final diagnosis and treatment initiation should be at the discretion of the physician and based on further evidence like central DXA or known risk factors. With regards to acceptability and risks for the population (section 27) the radiation dose of a DXR measurement is lower than that of a DXA measurement (about 1/10 the dose or less than 1 microsievert [8]) making it suitable to be used in a large population and also enabling subsequent measurements.

We believe that DXR has been overlooked as a screening test for osteoporosis and that there is considerable clinical evidence showing its suitability in identifying persons with an increased fracture risk. Being especially developed as a screening tool, DXR has also other characteristics like easiness of use, low cost and easy implementation into a clinical workflow that makes it suitable as a screening test. We ask the committee to consider this in their evaluation.

Furthermore, we take the opportunity to comment on some of the conclusions and suggestions raised by the review. We believe we found some passages that might need improvement or clarification to properly present the subject for the committee.

Section 2. The scope of the report includes only women whom have not yet suffered a fragility fracture. Theoretically, this is a reasonable limitation to do since the purpose of screening is to introduce primary prevention and individuals that have already suffered a fracture are eligible for secondary prevention. However, it is a known problem that even individuals suffering from a fragility fracture are seldom recognised by the health care and offered treatment and therefore we believe that this limitation is not relevant in practice. The main purpose of the screening programme is not to decide proper treatment but to avoid fractures by identifying those at highest risk of suffering from a fracture. We believe that a reasonable fraction of those taking part in a primary screening programme would have suffered from a previous, unrecognised fracture related to osteoporosis.

Osteoporosis is, as also stated in this report, severely under diagnosed, with studies from the UK suggesting that currently fewer than 30% [9] of patients with a fracture undergo osteoporosis risk assessment and subsequent treatment for secondary prevention of fracture. Additionally, in the review the authors state that it could be as low as 20 % (section 69). The assumption that individuals that have already suffered a fracture have no use of screening is unfortunate and leads to an underestimation of the benefits. In the actual screening setting there will be a large group of people that have suffered a fragility fracture but the healthcare system is not aware of it. These individuals might be identified via screening and thereby get proper access to treatment. This is acknowledged by the authors of the USPSTF report. In their comprehensive review [10] they define the target population to be postmenopausal women and men over age 50 years without known previous osteoporosis related fragility fractures or secondary causes of osteoporosis. However, they recognise that this definition will be compromised because of a presence of a large number of individuals with undiagnosed vertebral fractures. They conclude that many of these
individuals have never been appropriately evaluated for osteoporosis and may be
diagnosed during the course of routine screening [10].

Section 8. USPSTF focus on primary intervention in people that have not yet suffered
a fracture. However, it is difficult to identify clinical trials in which the efficacy of
drugs in this group was investigated. In the large FIT study [11], alendronate has a
poor effect in women without osteoporosis. It is however hard to separate the data
since it includes some women that have already suffered a vertebrae fracture.
USPSTF acknowledge the difficulties this causes in choosing which studies to include
in their review [10].

Section 10. The statement that approximately 16 % of women have suffered a prior
fracture is not supported by the reference [12]. The referred paper is not the original
source, but also in the latter [9] this number is not to be found. However, it states that
an estimate 12-19 % of women over 50 (1.3-2 million out of a 10.6 million
population) in the UK have suffered a prior fragility fracture before suffering a hip
fracture. The claim that 50 % of hip fractures occur in this population may be correct
but is unreferenced in the paper [12]. We found the statement that screening is
irrelevant to those in the population who have already suffered a fracture to be
inappropriate. In osteoporosis it is accepted that many vertebrae fractures go
undetected and even women that suffers wrist fractures are insufficiently picked up by
the healthcare system. This is also stated by the authors in section 11. Screening offers
these individuals a second chance to be identified and receive treatment.

Section 11. The authors specify that 50-70 % of vertebrae fractures receive no clinical
attention. We suggest that primary screening is a good way to identify also these
individuals, and thereby give them access to suitable treatment. Section 32. Only two
studies [11, 13] fit the criteria of primary intervention set by Nelson et al. [14]
However, in the larger FIT study [11] the values for alendronate and hip fractures in
women without a fracture but at T-score < -2.5 was 0.64 (95 % confidence interval,
0.5 – 0.82), which is considerably lower than the values reported when women with a
T-score < -1.6 was included. This is supported by meta-analysis showing
thatalendronate treatment is effective in women with osteoporosis [15, 16]. The report
should state what intervention threshold they found reasonable. For our
DXRtechnology we recommend women with a T-score < -2.5 to be investigated
further and that women with higher T-scores only be investigated if they present
additional risk factors. Hence, bisphosphonate treatment is relevant to those parts of
the population that benefit from screening. Just the same conclusion as the authors
reaches in section 42.

Section 43. The differences between NICE and NOGG guidelines could arise from
the fact that the participants in the NOGG group has been involved in developing the
FRAX tool. Although there appear to be a lack of consensus, the USPSTF report [14]
does not put much emphasis on FRAX and did not find any evidence recommending
it to start treatment on. Also, the validation and benefits of FRAX have recently been
questioned [17].

Section 51. The need of long term randomised controlled studies is a general problem.
Although considered the best form of clinical evidence we believe it is unrealistic to
expect studies spanning decades. This is due to the results loose in relevance as
standard regiments for diagnosis and treatment are changed during the course of the study. This is exemplified in the study by Barr et al.[18], where hormonal replacement therapy was first choice of treatment when the study was initiated and while it is today not recommended in broad use.

The effectiveness and feasibility of screening with DXR is currently being investigated in the Stockholm osteoporosis prevention (STOP) study, which is being finalised during 2013. In this study, women attending mammography is subjected to DXR measurement. A total of 16,000 women are included and followed for 30,000 patient years.

Section 57 and 65. It is unrealistic to expect such a study to be performed as the costs and effort needed are too high. Also, you will run into a similar problem as earlier studies [18, 19], i.e that clinical practice and treatment changes during the course of the study, making the findings difficult to interpret. Furthermore, we believe that health economy models, based on current knowledge and estimates, can give similar information as a trial.

Section 58. We have a health economy model for screening with DXR developed and validated by a third party [7]. We are happy to supply a detailed description of this model if requested. In this UK specific model, women over 65 years of age are screened with DXR when attending mammography (usually every second visit, i.e every 4 years) and those with a T-score < -2.5 are subjected to treatment with bisphosphonates (current standard treatment as recommended by NICE [16, 20] and NOGG [21]) at a yearly cost of €20, the annual intervention cost including physicians meeting is set to €92. The cost for a DXR analysis is set to €35. The risk reduction with bisphosphonates is 38 % for hip fracture, 44 % for vertebral fracture, 34 % for wrist fracture, and 18 % for other fractures. We estimate that at age 65 7.2 women need to be screened to identify 1 with T-score < -2.5.

With the above conditions, screening for osteoporosis in the UK becomes cost-saving for society over a 10 year time period. About 240 hip fractures are avoided per 100,000 women and the cost of the screening program and subsequent treatment is thus paid in full by the reduction in fractures, both at the hip and at other sites. The health economy clearly support screening for osteoporosis in the UK using DXR.

Section 67-68. Since DXR can be implemented on any digital X-ray machine and easily be used together with mammography neither new facilities nor hardware are needed. Individuals at increased risk of fracture identified with DXR can then be further investigated using present DXA facilities. In regions lacking suitable coverage of DXA equipment, DXR in combination with other risk factors can be used to initiate treatment.

Section 72. We do not agree with this conclusion. For the first bullet, please see answer to section 57. For the second, it is unfortunate that the two groups cannot present consensus guidelines. However, it cannot be used as a motivation against screening. In truth, none of the groups recommend screening. For the last bullet we agree that data is missing.
Section 73. We do not question that people having suffered an earlier fracture have a larger risk for a new one. But as we discussed earlier we think it is unwise to exclude this group from the screening population because the rate of treatment and follow up of these patients are poor. In a recent study it was shown that the majority of hip fractures occur in persons with T-scores < -2.5 [1]. Among 5,420 women over 40 years old, 89 suffered a hip fracture, 77 (81 %) occurred in women with a T-score < -2.5 [1]. Strikingly, no woman with a T-score > 0 suffered a fracture. For men it was slightly different, 33 fractures were recorded among 2,837 men over 40 years old. Of these 11 (36 %) occurred in men with a T-score < -2.5 [1]. However, in this study population individuals with prior fractures have not been excluded [1]. Furthermore, data from the FIT trial [11] show that of 390 fractures suffered among 2,218 women in the placebo group, 203 (52 %) occurred in women with a T-score < -2.5. This was a selected population without prior fractures but were all women had a T-score < -1.6. However, the idea that a majority of fractures would occur in women without osteoporosis needs to be further investigated and verified. Current lack of treatment options in women without osteoporosis, doesn't infer lack of screening benefit. The treating physician can decide on treatment options based on the available information. Furthermore, screening is a dynamic process, BMD decreases with age and individuals entering a screening programme might experience a decline in T-score over time. Also, the risk of hip fracture appears to increase drastically around a T-score of -2.5 [1]. Subsequent screening sessions makes it possible to track this deterioration and start treatment in time, before the first fracture occurs.

Section 74. Our main objection to these calculations is that it is too simplistic. The purpose of a screening programme is to identify individuals at risk and treat them before they have a fracture. Hence, people will migrate between the groups and those in group A will end up in group B and C as they age. According to our own model, general screening of the female UK population from 65 years of age would lead to a reduction of 240 hip fractures per 100,000 population over 10 years. We believe the model used in the review underestimates the treatment effect and number of fractures that can be reduced.

Section 76. Looking at the authors own example: If we assume that only 20 % of the women in group C have received follow up and are aware of their disease (as the authors state in section 69) this will increase the pool of possible fractures to reduce to 891 (50 %) hip fractures with 75 % uptake, 75 % compliance and 100 % treated with bisphosphonates. Also, in this group, the risk reduction with bisphosphonates is larger at 0.60 [15]. This would lead to a reduction in the number of hip fractures by 200 over 5 years. With a cohort of 98,108 this would mean that 491 women would need to be screened to avoid 1 hip fracture and 40 women need to be treated to avoid 1 hip fracture. However, we should take into consideration that bisphosphonate treatment reduces all types of fractures [22] so the total effect of the screening is larger than just the effect on hip fractures.

We ask that the above comments are taken into consideration in the evaluation of the report and subsequent decision by the committee. However, should the committee or the authors request additional information and data on DXR please feel free to contact us.
References


8. Sectra documentation on file. Dose calculations DXR, Sundhetsstyrelsen. DOC-JKAN-8UHH8M-1.0


The Royal College of Radiologists (RCR)

Introduction
2) This section should state that 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 (QCT or QUS) or to other anatomical sites (calcaneus).

The condition
9) Although osteoporosis is more common in postmenopausal women, men are also affected as stated in the paragraph (1 in 5 men will sustain an osteoporotic fracture in their lifetime). Why were men not considered in this review?
10) It would seem relevant to optimise identifying and treating those who have already sustained a fracture (secondary prevention) before considering screening for primary prevention.
15) The more accurate term for ‘corticosteroids’ would be ‘glucocorticoids’ as
corticosteroids refer to both glucocorticoids and mineralocorticoids.

The test
19) Quantitive ultrasound (QUS): what is the evidence that supports the statement that QUS is ‘commonly used BMD test’? QUS does not measure BMD directly; measurements obtained using QUS include ultrasound velocity (speed of sound, m/s), broadband ultrasound attenuation (dB/MHz), stiffness index and quantitative ultrasound index. Although the scanners are smaller, less expensive and do not use ionising radiation, there are technical limitations in that measurements are temperature dependent, precision is less than DXA, there are no phantoms for cross calibrating between scanners, the WHO definition of osteoporosis is not applicable to QUS or the calcaneus (heel) and as studies have rarely extended beyond 4-5 years, QUS parameters cannot be added to the FRAX® algorithm\(^1\). Although evidence for the value of QUS of the calcaneus for the prediction of fracture risk is now substantial, how this technique should be used in clinical practice is still not defined. Quantitative computed tomography (QCT) is another method for measuring BMD and is being used by some in the private sector in the UK but involves significant ionising radiation doses and to which the WHO definition of osteoporosis is also not applicable; low BMD is defined as Z score below -2.0.

22) One of the advantages of instruments such as the FRAX tool is that it empowers patients to investigate their own 10 years fracture risk and if a DXA scan would be appropriate for them to have. This might be worth emphasising as encouraging patients to do so can be self-selecting. A limitation of the report is that it emphasises DXA BMD as the screening test, where clinical risk factors such as those used in FRAX may be more cost effective. A recent publication\(^3\) showed that for women aged 55-65 years and men aged 55-75 years without a prior fracture, body weight can be used to select those for whom bone densitometry is cost-effective.

27) As stated, the radiation dose for DXA is extremely low at between 3-9 microSieverts (1-2 days of natural background radiation\(^4\), and carries no significant fatal cancer risk).

The treatment
34) Some differentiation should be made that osteonecrosis of the jaw is a much more frequent complication of IV bisphosphonates administered to patients with cancer than in those being treated with much smaller doses of bisphosphonate, most often administered orally for osteoporosis therapy. Therefore, the risks are not quite the same.

The programme
67) An issue to consider in the UK is that DXA scanners are frequently under-utilised as they are not staffed for full weeks of activity. In a 9am – 5pm working day, at least 10 patients can be scanned (hip and spine) each day and therefore, by using Monday to Friday as working days, at least 50 patients could be scanned to make equipment cost effective. This does not occur on all DXA scanners in the UK.

69) The term ‘plain X-rays’ should read ‘radiographs’.

Implications for policy
72) Overall, I would agree with the conclusions reached here, for the same 3 reasons provided.
Implications for research
78) There was an earlier study⁵ which showed unselective BMD screening for osteoporosis could not be recommended until a specific programme was formulated and justified. There is also a more recent review of evidence for using DXA for screening⁶. There is currently a national research study on osteoporosis screening funded through the Medical Research Council (MRC) which is not referenced⁷.

REFERENCES

National Osteoporosis Society

Thank you for the opportunity to comment on the UK NSC consultation document. We do not disagree with the following arguments used not to recommend screening

1. That there are no good randomised controlled trials looking at clinical and cost effectiveness
2. That there is a lack of consensus concerning guidance (about eligibility) for therapy (i.e. discrepancy between NOGG / NICE)
3. That the long term clinical and cost effectiveness of treatment are not known.

There are however a number of issues that we would like to be considered by the UK NSC.

Introduction

Point 2

The scope of this document is limited to primary prevention in postmenopausal women i.e. the identification and treatment of postmenopausal women who have osteoporosis, but who have not yet had an osteoporotic fragility fracture. We would welcome a considered rationale also for excluding men and premenopausal women.
Point 5

The document aims to interlink both recommendations in NICE TA160, TA161, TA204 and NOGG. Although the reason for the attempted overview is legitimate, the distinction between T-Score thresholds in NICE and 10 year probability of fracture in NOGG is not clear and potentially confusing. This needs to not only be clearer but more considered throughout and where the evidence differs drawn out.

The Test

Point 19 & 20

We would like to make the committee aware of the National Osteoporosis Society Position statement on the use of quantitative ultrasound in the management of osteoporosis (www.nos.org.uk/page.aspx?pid=1074) In particular the following recommendation.

- As quantitative ultrasound (QUS) does not measure bone mineral content or density directly, it cannot be used to diagnose osteoporosis as currently defined by bone mineral content (BMC) or bone mineral density (BMD) assessment.

The Treatment

NICE is currently reviewing the current technology appraisal (TA160, 161, 204) to explore how treatment intervention thresholds can be described in a way that aligns them to the assessment of absolute fracture risk (http://guidance.nice.org.uk/TA160/ReviewUpdateAugust2012). We would like the committee to consider the impact of this review on the implications for policy as the lack of consensus around treatment thresholds is one of the three highlighted issues in the decision not to implement a screening programme.

Point 48,49 & 50

Within clinical management section there are a number of initiatives that have not been considered.

- Department of Health Falls and fractures: effective interventions in health and social care. This document sets out four key interventions that commissioners, working across health and social care, should consider in the context of local services for falls, falls prevention and fractures. It aims to inform local dialogue between commissioners and service providers (http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@pg/documents/digitalasset/dh_109122.pdf)

- There a number of Quality Standards that deal with issues on, or related to, Osteoporosis (http://www.nice.org.uk/guidance/qualitystandards/QualityStandardsLibrary.jsp)
  - Hip Fractures - Published
  - Fractures (excluding head and hip) - referred in March 2012
  - Complex fractures (including compound fractures) - referred in March 2012
Menopause- referred in March 2012

- The Royal College of Physicians National Audit of Falls and Bone Health in Older People (http://www.rcplondon.ac.uk/resources/national-audit-falls-and-bone-health-older-people). Please note that the RCP audit and NHFD have now merged giving one comprehensive national audit programme.

- Scottish Intercollegiate Guidelines Network (SIGN) are currently reviewing Guideline No 71 Management of osteoporosis. (http://www.sign.ac.uk/index.html)

The Programme

Point 51

We would like to make the committee aware of a current RCT which is directly addressing the question of screening.

- SCOOP (Screening of Older Women for Prevention of fracture) is a UK seven centre, pragmatic, randomised controlled trial with 5-year follow-up to assess the effectiveness and cost-effectiveness of a community-based screening programme to reduce fractures. It utilises the FRAX algorithm and DXA to assess the 10-year probability of fracture. An economic analysis will be carried out to assess cost-effectiveness of screening. A qualitative evaluation will be conducted to examine the acceptability of the process to participants (http://link.springer.com/article/10.1007%2Fs00198-011-1876-7).

National Osteoporosis Guideline Group (NOGG)

NOGG is pleased to have the opportunity to comment on this draft report for the UK National Screening Committee. Numbered paragraphs refer to the paragraph numbers of the draft report. Italic font denotes a quotation from the draft report.

2. In the light of the recent report from NICE on fracture risk assessment [NICE 2012] and the wide availability and acceptance and uptake of the guidelines of the National Osteoporosis Guideline Group (NOGG) [Compston 2009], NOGG is disappointed at the scope of this appraisal which is limited to the identification and treatment of postmenopausal women who have osteoporosis, but who have not yet had an osteoporotic fragility fracture. This firmly ties the assessment instrument to the measurement of bone mineral density (BMD).

The limitations of BMD as a screening tool are well recognised and include low sensitivity as well as relatively high cost compared with the cost of intervention [WHO 1994, Kanis 2001a]. The significance of a T-score of -2.5 SD or less is critically dependent on age and the presence or absence of other independent clinical risk factors. This has led to the development of FRAX to assess fracture probability [Kanis 2008a] and other risk engines to assess absolute risk [Nguyen 2008, Hippsley-Cox 2009].

NICE have endorsed the use of FRAX as a risk assessment tool [NICE 2012] and NOGG recommend that treatments are considered appropriate and cost-
effective in postmenopausal women above a given age-specific fracture probability threshold [Kanis 2008c]. The NOGG guidelines reserve the use of BMD to those close to an intervention threshold and therefore make more economic use of assessment than that based on BMD alone. Moreover, FRAX can be used without recourse to BMD to identify patients at high risk. Indeed the performance characteristics of FRAX without BMD are similar, if not better, than that of BMD alone (Table 1) and a substantial body of evidence indicates that the use of FRAX with or without the inclusion of BMD identifies a risk that is responsive to intervention [Kanis 2012a].

Table 1 Gradients of risk (GR) for fracture per SD change in risk score (with 95% confidence intervals; CI) with the use of BMD, clinical risk factors or the combination [Kanis 2007].

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>BMD only</th>
<th>Clinical risk factors alone</th>
<th>Clinical risk factors + BMD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GR</td>
<td>CI</td>
<td>GR</td>
</tr>
<tr>
<td>50</td>
<td>1.19</td>
<td>1.05–1.34</td>
<td>1.41</td>
</tr>
<tr>
<td>60</td>
<td>1.28</td>
<td>1.18–1.39</td>
<td>1.48</td>
</tr>
<tr>
<td>70</td>
<td>1.39</td>
<td>1.30–1.48</td>
<td>1.55</td>
</tr>
<tr>
<td>80</td>
<td>1.54</td>
<td>1.44–1.65</td>
<td>1.63</td>
</tr>
<tr>
<td>90</td>
<td>1.56</td>
<td>1.40–1.75</td>
<td>1.72</td>
</tr>
</tbody>
</table>

Against this background, a more appropriate scope would have been to address the role of algorithms such as FRAX as a screening tool. NOGG is concerned that the present report will throw the baby out with the bathwater and would ask that these limitations of the scope be acknowledged to encourage future appraisals to consider more recent and powerful technologies for screening.

9. The prevalence of osteoporosis depends critically on the precise definition of osteoporosis. Using updated WHO criteria [Kanis 2008b], there are estimated to be more than 2.5 million women in the UK with osteoporosis, rising from a prevalence of 6.3% at the ages of 50-54 years to 36.8% at the age of 80 years and above [Strom 2011].

For the UK, it is estimated that annually there are 447,000 osteoporosis-related symptomatic fractures of which 56,000 are hip fractures, 40,000 are clinical vertebral fractures, 59,000 are forearm fractures and 190,000 are at other vulnerable sites [Strom 2011].

12. The NHS cost of £2.1 billion refers to the annual NHS cost of incident fractures in any one year. This excludes the cost of fractures that occurred before the index year but which still have financial consequences for the index year. Estimates for 2010 suggest that the cost of incident fractures were £3.3 billion
rising to £4.3 billion when the ongoing costs of previous fractures are included (excluding any medical intervention costs) [Strom 2011].

15. The term “glucocorticoids” is more accurate in this context than “corticosteroids”.

19. The scope of the appraisal shuts out the most widely used assessment tool – namely FRAX which has higher sensitivity for any given specificity than BMD. See comments to paragraph 2.

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. These data cannot be compared with data given for QFracture in paragraph 22 above. The area under the curve for osteoporotic fracture is derived from the fracture risk score and not from fracture probability [Kanis 2007]. The importance of this distinction has recently been reviewed [Kanis 2012b].

24. An additional six studies evaluating externally-validated risk assessment instruments. 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. As discussed elsewhere [Cooper 2012], the ‘independent validation’ is drawn from a similar population documented in a similar manner and it is not surprising that high areas under the ROC curve were found. 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). This is misleading. The performance characteristics FRAX in THIN were published in 2007 [Kanis 2007] – now noted on public record [Kanis 2012c].

25. It should be noted that the scope of the present report does not follow the NICE published Clinical Guideline 146 on assessing the risk of fragility fracture in osteoporosis [NICE 2012].

31. Denosumab, strontium and teriparatide have also been shown to reduce the risk of non-vertebral fractures. Non-vertebral fracture reduction has also been shown with alendronate [Wells et al 2008], risedronate [Harris et al 1999] and zoledronate [Black et al 2007]. There have been no trials to date assessing the effect on fracture-related morbidity and mortality (Nelson et al 2010). The report mentions later the study on zoledronate and the more recent meta-analysis in an appropriate perspective (paragraph 54). This could usefully be noted here.

There is no statistically significant evidence that treatment prevents hip fractures (Table 2). This statement is untrue. Statistically significant reductions in hip fracture have been demonstrated for alendronate [Black et al 1996], risedronate [McCling et al 2001], zoledronate [Black et al 2007] and denosumab [Cummings et al 2009]. Strontium ranelate was also shown to
reduce hip fracture in a post hoc analysis of higher risk women [Reginster et al 2008]. All these interventions are approved in Europe for the prevention of hip fracture.

The footnote (6) gives a definition of non-vertebral fracture. This definition is however not consistently applied to phase 3 studies.

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). This is true for all these therapies for vertebral fractures. Not all have shown a reduction in non-vertebral fractures within the first year.

36. Interventions approved for the prevention and treatment of osteoporosis have not been consistently associated with serious adverse events. Strontium ranelate is also associated with increased risk of thromboembolic events. Osteonecrosis of the jaw and atypical fractures should be added for denosumab.

42. The view distilled from the systematic review of Nelson et al [Nelson 2010] is in our view accurate though somewhat misleading in context. There is a strong RCT base to indicate that interventions decrease fracture risk in women without a fragility fracture. The argument is thus whether there is evidence that any treatment prevents fractures in women who have osteopenia. A number of studies have addressed this in addition to the post hoc analysis of alendronate referred to in the draft report [Cummings 1998]. A meta-analysis of studies with alendronate using a less stringent cut-off for BMD indicated efficacy of alendronate on hip fracture risk [Papapoulos 2005]. Other studies, reviewed elsewhere have shown efficacy of intervention in women unselected for BMD [Kanis 2012a]. In other RCTs efficacy has been shown in women more or less independently of BMD including those with risedronate [Kanis 2005, Siris 2008], raloxifene [Kanis 2003, 2010, Johnell 2004], bazedoxifene [Kanis 2009], zoledronate [Eastell 2009], strontium ranelate [Roux 2006, Kanis 2011], denosumab [McCloskey 2012], clodronate [McCloskey 2009] and teriparatide [Marcus 2003, 109]. This large body of data on a wide variety of interventions indicates that treatment effects, with the possible exception of alendronate, are not dependent on baseline BMD. We feel that this is an important issue to articulate if subsequent reviews of screening have a broader scope.

44. The report notes that treatment thresholds provided by NICE are fundamentally different from those recommended by NOGG, because the former are based on identifying groups of women for whom treatment is expected to cost no more than £20,000 per QALY gained. In fairness, it should be noted that the intervention thresholds proposed by NOGG have a cost/QALY gained of £20,000 or less in women with osteoporosis at or above the age of 65 years [Kanis 2008]. It should also be noted that many of the assumptions used by NICE for their health economic assessments have been challenged [Kanis 2010].

45. The report states that ‘NICE’s criteria for initiating treatment with alendronate correspond to substantially higher 10-year fracture risks than those recommended by NOGG’. This is misleading. The 10 year risks calculated from NICE are cumulative 10 year incidence, whereas the thresholds provided by NOGG are based on 10-year probability that additionally take into account of the death hazard [Kanis 2012b]. An analysis using the same metric of risk for
both guidelines indicates that the NICE thresholds, like NOGG are age-dependent. The analysis also indicates that age dependent thresholds derived from NICE’s guidance vary widely whereas those of NOGG do not vary (Table 2). Moreover, NICE thresholds are variously higher or lower than those provided by NOGG.

Table 2. Intervention thresholds expressed as the 10 year probability of a major fracture used by NOGG and thresholds for treatment used by NICE. Data from http://www.shef.ac.uk/FRAX/index.htm

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>NOGG(^a)</th>
<th>NICE(^b)</th>
<th>NICE(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>16.5</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>70</td>
<td>20.3</td>
<td>22</td>
<td>15</td>
</tr>
<tr>
<td>75</td>
<td>23.4</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td>80</td>
<td>27.6</td>
<td>32</td>
<td>18</td>
</tr>
<tr>
<td>85</td>
<td>33.4</td>
<td>36</td>
<td>20</td>
</tr>
</tbody>
</table>

\(^a\) BMI set at 24 kg/m\(^2\); T-score = -2.5 SD  
\(^b\) BMI set at 24 kg/m\(^2\); Parental history of hip fracture; T-score = -2.5 SD  
\(^c\) BMI set at 24 kg/m\(^2\); Current smoking; T-score = -2.5 SD

47. The literature search for this review did not find any studies that compare the numbers of women who would be recommended for BMD measurement or drug treatment if the NICE guidelines or the NOGG recommendations were applied. The report failed to identify a relevant study that compared the old RCP guidelines that are BMD based and the NOGG guidelines that adopt a more parsimonious approach to the use of BMD. NOGG identified similar numbers of women at high risk but with lower numbers of scans required at each age. Compared to the RCP strategy, the FRAX-based NOGG strategy used BMD resources much more efficiently with lower acquisition costs and lower costs per hip fracture averted [Johansson 2012].

The conclusions of Bolland and Grey [2010] are not safe since they misinterpret the NOGG guidance [Compston 2010]. In addition, the data, even if true, are not safe in the context of the report since both guidelines (NOF and NOGG) include resource utilisation in women with a prior fracture.

65. NOGG strongly supports the statement that ‘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.’ One such trial based on the NOGG approach is currently in progress and would be useful to acknowledge [Shepstone 2012].
73 and 74. *If they were screened the first group (women without osteoporosis) would be told that there is no treatment that has been proven to prevent fractures in women like them.* We would challenge this conclusion for the reasons discussed in paragraph 42 above.

**References**


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


