Bone Density Measurement and Treatments of Osteoporosis

David Hailey, Laura Sampietro-Colom, Deborah Marshall, Rosa Rico, Alicia Granados, José Asua, Trevor Sheldon

1996
INAHTA Joint Project

Methods used in the measurement of bone density

Background paper one

David Hailey
Alberta Heritage Foundation for Medical Research

July 1996
Acknowledgements

The initial draft of this paper was prepared under the health technology program of the Australian Institute of Health and Welfare, Canberra. The document was updated and finalized at the Alberta Heritage Foundation for Medical Research, Edmonton.

The author is grateful to Pierre Durieux, Hellen Gelband, David Henry, Olof Johnell and Deborah Marshall for their helpful comments on earlier drafts.
Introduction

Low bone mass is associated with an increased risk of future fracture (1). Bone mass measurement is widely considered to be the most useful objective indicator of fracture risk. However, low bone mass is only one of a number of risk factors some of which have similar predictive ability for fracture. Various approaches have been developed for measurement of bone density which depend on the absorption of radiation by the skeleton, and provide quantitative measures of bone mass. The strength of bone is also dependent on the geometric distribution of bone tissue, and on its material properties. However, while some technologies are capable of providing information on the structure of bone, such measurements have not yet proved useful in routine diagnosis and appraisal of fracture risk.

Definitions

The term bone mass indicates the amount of mineralised tissue (hydroxyapatite) in bone. The term bone density indicates the mass of bone in relation to length (g/cm), area (g/cm²) or volume (g/cm³). The choice of unit depends on the specification of the equipment used. However, g/cm³ is generally preferred and has the advantage of being independent of bone dimensions, removing a source of measurement error.

As hydroxyapatite is insoluble and it is difficult to make homogenous phantoms, various other substances have been used to calibrate bone density measurement (BDM) devices. The most commonly used material is dipotassium hydrogen phosphate, but other materials have also been used, making it difficult to compare measurements. To achieve comparable results, different bone measurement devices should be calibrated against the same substance.

Many of the assessment reports on BDM technologies have discussed their analytical performance in terms of precision and accuracy. Good precision means few random error so that there is small variation between results of measurements on the same sample. Good accuracy means that the systematic error is low, so that the average of a series of measurements on the same sample is close to the true value. Both precision and accuracy for BDM methods are often expressed as coefficients of variation.

Measurements of precision should be performed in vivo so that they are relevant to the setting in which a device will be used. However, a report by the Swedish Council on Technology Assessment in Health Care (SBU) (2) notes that long term measurements of precision (from 6 months to several years) have used phantoms since biological variation is too great to allow reliable in vivo measurements. Most available data refer to measurements that have been undertaken over a short period.

Accuracy of BDM methods has most commonly been measured using ash studies. In an ash study, the density of intact bones is assessed, the bones are then asked to remove carbon and the residue weighed. Accuracy is expressed in terms of the difference between measured and true contents of hydroxyapatite. Further discussion of these analytical performance measures is given below.

Types of BDM methods

A number of techniques have been used to measure bone density. Their adoption in health care systems has depended on several factors, including not only perceptions of analytical performance, cost and ease of use but also payment policies of insurance and other organizations (3).

One of the first techniques to be widely used for BDM was single photon absorptiometry (SPA), based on the measurement of radiation from a radioisotope source after passage through the forearm or wrist. Dual energy photon absorptiometry (DPA), a technique using two radiation energies, is less precise but gives the possibility of undertaking measurement at the femur and spine.

X-ray based methods are now more commonly used. Perhaps the most popular is dual energy X-ray absorptiometry (DXA) which has better accuracy and precision than the earlier methods. Single energy
x-ray absorptometry (SXA), the x-ray equivalent of SPA, has been adopted in a number of health care systems. Quantitative computed tomography (QCT) has also been widely applied. It can provide measurements of bone density at the spine, and information on bone structure, but is less accurate than DXA, more time consuming and delivers a higher radiation dose.

Measurements with ultrasound technology at the heel, and more recently at other sites, are increasingly used, though data on precision, accuracy and predictive value of this approach are still relatively limited.

A summary of BDM methods, indicating principles of operation and status, is given in Table 1.

In addition to the bone density measurement technologies, there is much interest in the measurement of bone formation/resorption using serum and urinary markers of bone metabolism (2). Such approaches are still experimental/developmental. They are a possibly important future approach to fracture risk assessment, alone or in combination with bone absorptiometry.

<table>
<thead>
<tr>
<th>Method</th>
<th>Principle</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single energy photon absorptometry (SPA)</td>
<td>Based on the measurement of gamma radiation from a radioisotope source after passage through the forearm or wrist.</td>
<td>Has been widely used. Limited to measurement of limbs/wrist/heel. Now being replaced by X-ray methods</td>
</tr>
<tr>
<td>Dual energy photon absorptometry (DPA)</td>
<td>Measurement of radiation at two energy levels from one or two radioactive isotopes</td>
<td>Also widely used. Able to measure axial portion of skeleton. More expensive and poorer accuracy than SPA. Use now limited, X-ray methods preferred</td>
</tr>
<tr>
<td>Photodensitometry</td>
<td>Measurement of absorbance of X-ray bone image.</td>
<td>Not in common use, subject to interference from soft tissue, film processing</td>
</tr>
<tr>
<td>Single energy X-ray absorptometry (SXA)</td>
<td>Measurement of radiation from X-ray source in peripheral parts of skeleton. Analogous to SPA</td>
<td>Used in a number of health systems. Comparatively limited popularity, possibly because of greater versatility of DXA</td>
</tr>
<tr>
<td>Dual energy X-ray absorptometry (DXA)</td>
<td>Similar to DPA, but uses an X-ray tube rather than radioisotope as the radiation source</td>
<td>Widespread use; currently the most common method in several countries for routine BDM work</td>
</tr>
<tr>
<td>Quantitative computed tomography (QCT)</td>
<td>Use of CT scanner to measure X-ray absorption. Single and dual energy methods developed</td>
<td>Single energy method quite widely applied in some countries and continues to be used in routine services. Dual energy approach little used in routine work</td>
</tr>
</tbody>
</table>
Characteristics of BDM methods

Some characteristics of the technologies that have had widespread use in BDM are given in Table 2 (2, 4-10). The data for accuracy and precision are derived from a variety of types of study and differing sample sizes. The data tend to reflect efficacy rather than effectiveness.

An important consideration is that the long term and short term precision of a method will be different. This is not always made apparent in the summaries included in various assessments. Data in the SBU report (2) suggest that many studies were undertaken over a period of weeks or a few months. Long term precision will be of most relevance to some routine clinical applications, and the absence of in vivo data is a shortcoming. Only limited numbers of studies on accuracy have been undertaken. These have usually used small sample sizes and are based on in vitro measurements (ash studies). The relationship of such studies to the actual accuracy of BDM methods in a clinical setting is unclear.

This summary illustrates the improved precision of DXA and SXA in comparison with other methods, and the limited accuracy of all modalities. Further work is needed to define the performance of ultrasound methods, the predictive ability of which is uncertain (11).

A retrospective study showed bone density measured by ultrasound at the calcaneus was as good in predicting hip fracture as SPA or DXA measurements at the hip (12). However, Massie et al. (13) in a comparison of DXA of the spine and hip with ultrasound of the heel found that ultrasound was a poor predictor of spinal and hip bone mineral density. A more recent report (14) indicated that correlations
between ultrasound of the calcaneus and DXA of the spine and proximal femur are not high enough to reliably predict bone mineral density at the spine or femur from the ultrasound results. It concluded that the utility of ultrasound has yet to be fully defined, and that its use in screening applications would be premature.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (CV %)</th>
<th>Precision (CV %)</th>
<th>Time of scan (min)</th>
<th>Radiation dose (mrad)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPA</td>
<td>2 - 8</td>
<td>2 - 5</td>
<td>5 - 15</td>
<td>2-5</td>
<td>Simple, relatively inexpensive, small radiation exposure. Decay of source affects performance</td>
</tr>
<tr>
<td>DPA</td>
<td>3-10</td>
<td>2 - 6</td>
<td>20 - 45</td>
<td>5 - 20</td>
<td>Usually used for spine and hip measurements. Somewhat slow. Decay of source affects performance</td>
</tr>
<tr>
<td>SXA</td>
<td>5</td>
<td>1</td>
<td>10 - 20</td>
<td>5 - 10</td>
<td>X-ray equivalent of SPA; currently less popular than DXA</td>
</tr>
<tr>
<td>DXA</td>
<td>2 - 10</td>
<td>1 - 3</td>
<td>3 - 10</td>
<td>3 - 5</td>
<td>Single X-ray source with two energies. Higher photon flux than radionuclide sources, improved detector configuration.</td>
</tr>
<tr>
<td>QCT</td>
<td>5 - 15</td>
<td>2 - 6</td>
<td>10 - 15</td>
<td>200 - 1500</td>
<td>Able to measure bone structure. Need to measure calibration standards simultaneously with the patient.</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>20</td>
<td>2 - 4</td>
<td>5</td>
<td>Nil</td>
<td>Potential to measure bone structure</td>
</tr>
</tbody>
</table>

Source: References 2, 4, 5, 6, 7, 8, 9, 10
Site selection

While cost of the necessary equipment and radiation exposure are matters that may influence the choice of a measurement site, the most important consideration is the ability to indicate fracture risk (4). There are insufficient longitudinal data to determine if any technique or skeletal site is better than others for quantifying bone mass or measuring bone loss rates.

There is still debate as to whether it is necessary to measure bone density of a specific bone to predict its risk of fracture. The Center for Health Care Technology cites several studies which have concluded that risk of fracture of the spine or hip can be predicted on the basis of BDM measurements at appendicular sites (4). The CHCT report notes that the ability to use single-site measurement of bone density to evaluate general fracture risk for all sites has practical importance. Such an approach could eliminate the rationale for multiple-site measurements performed by some clinicians and would put less emphasis on a technique's ability to measure a particular site.

According to a meta-analysis undertaken by SBU (11) measurements at the hip to predict hip fractures, and at the spine to predict spinal fractures, give higher relative risk than measurement at other sites. The practical significance of such findings requires further consideration.

Implications of analytical performance

Some of the assessment reports dealing with BDM mention the need for good precision if a technique is to be used for serial measurements and for good accuracy if a method is to correctly diagnose a particular case, and in screening applications. This point deserves further consideration. In the clinical setting, a decision on management of the patient will typically be taken on the basis of a single result from a BDM examination (ignoring the contribution of clinical diagnosis). In practice, both the accuracy and precision of the method will be relevant to the chance of a single clinical measurement having acceptable reliability.

As neither in vivo precision nor in vivo accuracy data are available for any method in routine use over extended periods of time, there must be considerable doubt as to what their true reliability in a clinical situation might be. The available data give a reasonable basis for making choices between the different methods, but are optimistic estimates of diagnostic performance in the routine situation.

The BDM technologies are far from ideal as measurement devices, particularly if screening is contemplated. An Australian report by the National Health Technology Advisory Panel (5) suggested that SPA would be unlikely to detect those at risk of spine or hip fracture with a sensitivity or specificity greater than 75%. For a population of 80,000 (the approximate number of Australian women aged 50), of whom 10,000 had a bone density low enough to warrant therapy, SPA could fail to identify 2,000 low bone density individuals and incorrectly classify 18,000 as requiring therapy. The technique would have to be 87.5% sensitive before half of those identified as being at risk were in fact in that category. Estimated sensitivity of current methods is considerably lower than this (11).

The limitations of the various techniques in terms of analytical performance also means that it is not feasible to follow up BDM exams over a short period of time. Several health technology assessment agencies have addressed this point.

ANDEM in its 1991 report (8) noted that the best available performance was a reproducibility of 1%, which implied a minimum interval of 1.4 years between examinations to detect a bone loss of 2% (corresponding to the mean annual loss for normal women at the beginning of menopause). The UK review in Effective Health Care (15) noted that 'given the margin of error in measuring bone mass by DXA is of the same order of magnitude as the annual rate of bone loss, reasonably precise estimates of the rate of loss for the whole population would require dealing extending the length of follow up to around four years'. A more recent review with QCT has suggested that adequate follow-up for trabecular bone requires intervals of at least three to five years to discriminate between different loss rates (16).
Table 3 lists a number of estimates published by assessment agencies. These provide an illustration of the limitations of BDM in the clinical setting and, in practice, the minimum intervals will be longer than suggested by these data. As the Center for Health Care Technology notes (4), intervals such as those shown in Table 3 are calculated on the basis that accuracy of the device is invariable. This is unlikely to be the case, particularly when the performance of an analytical instrument is being considered over a lengthy period of time. Accuracy will be affected by changes in the performance of components of the device, in operator performance and in physiological composition. (Also, sequential examinations could well be undertaken in different establishments.)

As indicated in Table 2, the accuracy of BDM methods is not high. The SBU report (2) suggests that currently – used measurement methods have an accuracy of around 10% (CV), whereas accuracy should be better than 2-4% to identify those at risk of fracture.

The total error standard deviation of a BDM method, taking account of both precision and accuracy, might be a more useful measure in practice. Such information does not seem to be readily available. Methods based on following up individuals over a long time will be complicated by changes in bone mineral concentration and fat distribution over that period (6). An option would be to consider measuring the performance of a technique (precision, accuracy and total error standard deviations) across several centres, rather than relying on data from studies involving single BDM machines.

Table 3: Intervals required to reliably detect bone mass loss over time (a)

<table>
<thead>
<tr>
<th>Agency</th>
<th>Precision error (CV,%)</th>
<th>Estimated bone loss (%)</th>
<th>Followup measurement interval (years) (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBU</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>CHCT</td>
<td>1</td>
<td>3</td>
<td>0.9</td>
</tr>
<tr>
<td>ANDEM</td>
<td>1</td>
<td>2</td>
<td>1.4</td>
</tr>
<tr>
<td>CHCT</td>
<td>1</td>
<td>1</td>
<td>2.8</td>
</tr>
<tr>
<td>SBU</td>
<td>5</td>
<td>3</td>
<td>4.7</td>
</tr>
<tr>
<td>CHCT</td>
<td>5</td>
<td>3</td>
<td>4.4</td>
</tr>
<tr>
<td>CHCT</td>
<td>5</td>
<td>1</td>
<td>13.3</td>
</tr>
</tbody>
</table>

(a) Estimates on basis that accuracy is invariable
(b) Sources: References 2, 4, 8
(c) Time frame for a reliable bone mass measurement follow up
Factors noted in the 1986 OHTA report as adversely affecting the analytical performance of DPA included uneven distribution of adipose tissue around the spine, and use of an improper or average baseline (20% or more error in spinal measurements). Other factors which are also relevant to other methods include variation in positioning of patients (mentioned in several other assessments), improper calibration and weak quality assurance programs.

A report by a study group of the World Health Organisation considered that requirements of bone mineral density measurement for screening using a single test, including rapidity, reliability, low radiation dose and a low error of accuracy, are largely met by SPA and DXA, and less so by DPA and QCT (17). From the data available for the present review, such conclusions appear optimistic.

The need for quality control

The National Health Technology Advisory Panel reports (5, 7) noted the need for excellent quality control in bone density measurement services, and suggested some form of accreditation (an option that has so far not been adopted). Quality control should include daily standardisation and calibration procedures, regular maintenance and careful attention to patient positioning. The Center for Health Care Technology (3) also notes the need for strict quality control, including calibration and standardisation procedures and consensus on procedural methods for positioning. Those who publish normal ranges should accurately describe the technique used. Gluer et al. (18) have outlined a concept for quality assurance in the context of clinical trials, which might be helpful for more routine applications of BDM.

Calibration and standardisation of bone densitometers is a complex undertaking that requires more attention since there is little agreement among manufacturers. Even with instruments calibrated according to manufacturer’s instructions, values obtained from imaging of spine phantoms by DPA or DXA have differed by as much as 16% because of differences as great as 8% in both bone mass and bone area (4, 19). There is no industry-wide standard to provide machine-to-machine comparability. Users of bone densitometers need to adopt a uniform procedure for calibrating and standardising instruments, with quality control methods that use an accepted ‘gold standard’. A European Spine Phantom has now been used to calibrate equipment in a number of drug trials and may become more generally useful for routine use. Pearson et al (20) have described its use in cross-calibration of bone densitometers, noting that some machines (a minority) show significant instability over time.

An issue with the increasing use of lower cost portable machines is the availability of adequate technical expertise for quality control in smaller centres (including physician’s office settings).

Cost of examinations

Because of the differences between countries on approaches to reimbursement for BDM exams, cost data from various health care systems are difficult to compare. Cost data cited by the Center for Health Care Technology are listed in Table 4.

SPA and DXA tend to be cheaper than DPA and QCT. Ultrasound may also be cheap, with capital costs possibly of the order of $25,000.

With QCT, the use of the machine for non-BDM work could be a complicating factor.
Table 4: Cost data for BDM examinations in the USA

<table>
<thead>
<tr>
<th>Method</th>
<th>Equipment cost ($US '000)</th>
<th>Scan charges ($ US)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPA</td>
<td>20-30</td>
<td>50-150</td>
</tr>
<tr>
<td>DPA</td>
<td>30-65</td>
<td>150-300</td>
</tr>
<tr>
<td>QCT</td>
<td>5-15&lt;sup&gt;a&lt;/sup&gt;</td>
<td>150-400</td>
</tr>
<tr>
<td>DXA</td>
<td>60-100</td>
<td>150-300</td>
</tr>
</tbody>
</table>

<sup>a</sup> Adaptation of an existing CT scanner; costs refer to software and phantoms

Source: Reference 4

Other approaches

A recent development has been the use of ultrasound to measure the tibia rather than the heel. Fragility is assessed by measurement of the speed of sound along the cortical layer of bone. Early publications have claimed precision of better than 1% and good discrimination between healthy post menopausal women and osteoporotic cases, with about 20% false positives. A preliminary review of this technology by the Health Technology Assessment Unit of Alberta Health (21) suggested that research into the effectiveness in diagnosis and monitoring of osteoporosis over a longer period of time is indicated, using larger number of subjects and randomised trials.

'Peripheral' versions of both DXA and QCT, for use in measurements of the forearm or heel are now available. There are claims from the manufacturers of high reproducibility, but additional data on analytical performance from appropriately designed studies appear to be required. With both techniques, there would be the uncertainty of the usefulness of forearm measurements for assessing the bone density at clinically more important sites.

Some considerations in the clinical use of BDM methods

As noted by the Center for Health Care Evaluation, 'there is a growing tendency to define osteoporosis in terms of a continuum of bone density, with greatest fracture risk in those with lowest absolute density values, rather than in considering osteoporosis as a fracture/non fracture dichotomy' (4). On the other hand, practical use of BDM in health services tends to be directed in terms of a target value for minimum acceptable bone mass.

With probable levels of analytical performance, there is a certainty of substantial numbers of false positives and false negatives in routine practice. The predictive value of BDM measurements will be influenced by the analytical performance of the methods, and the assignment of many women to the wrong category with regard to their risk of fracture.

Such difficulties will be compounded by differences in population mean values for bone mass between ethnic and other groups. For example, Tobias et al. (22) reported that bone mineral density in Asians was lower than that of a reference Caucasian population. They suggest that assessment of risk in Asian women by comparing bone mineral density with such a reference population may have limited validity because of the influence of skeletal size on such measurements.

Further consideration would be desirable on what the consequences of limitations in analytical performance might be for the patient management decision. A point to consider is whether all users of these technologies appreciate the limitations in the performance of these methods, and hence their ability to make a clinically useful discrimination.
More generally, further information on the relationship of BDM data to management decisions would be useful. There appear to be a number of uncertainties, including the weight given to BDM results compared with clinical examination and judgement.

Conclusions

On the basis of the review of assessment reports and recent literature undertaken in preparing this paper, the following points seem to be of particular significance.

- Currently, DXA is a widely-used method for measurement of bone density, and is the most popular technology in a number of health care systems.
- Ultrasound methods are increasingly used, but their performance requires further validation.
- Information on the structure of bone has not yet found application in routine clinical used of BDM.
- Scrupulous quality control is essential for BDM services. Users should be aware that there are no industry-wide standards.
- The analytical performance of all methods is poorly defined in the routine clinical situation. Published data on precision and accuracy tend to be derived from limited studies at centres with expertise in BDM. They probably give an optimistic indication of the diagnostic performance of the BDM methods in routine health care.
- Given the limitations on analytical performance, users of BDM technologies should be aware that there will be large number of false positive and false negative results if women are being assessed against a target value for minimum acceptable bone density. Such difficulties will be compounded if account is not taken of the difference reference values needed for some ethnic groups.
- Further work to establish the analytical performance of BDM methods in routine practice, taking account of both precision and accuracy, would be highly desirable. It would also be desirable for users to define the levels of analytical performance that are required for different applications of BDM measurements.
- Further consideration should be given to the relationship of BDM results to the subsequent patient management decision. The value of BDM measurements in many clinical situations could be quite limited.
References


INAHTA Joint Project

Predictive value of bone densitometry

Background paper two

Deborah Marshall
The Swedish Council on Technology Assessment in Health Care

Trevor Sheldon
NHS Centre for Reviews and Dissemination

July 1996
The initial draft of this paper was prepared at the Swedish Council for Technology Assessment in Health Care, Stockholm, Sweden and subsequently updated at the Alberta Heritage Foundation for Medical Research, Edmonton, Canada.

The authors are grateful to David Hailey, Olof Johnell, Hans Wedel and to the reviewers of the draft INAHTA statement for their helpful comments and suggestions.
Contents

Background ........................................................................................................................................................................ 5

Reports on the applications of bone density measurement .................................................................................................................. 5

Assessment of fracture risk ........................................................................................................................................................................ 6

Present fracture risk .............................................................................................................................................................................. 7

Assessment of future fracture risk .............................................................................................................................................................. 8

Bone density measurement in primary prevention ..................................................................................................................................... 9

Conclusions ....................................................................................................................................................................................... 11

Table 1: Bone density (in hip) in women in Sweden aged 20-89 ........................................................................................................... 13
Table 2: Summary of recent hip fracture case-control studies in women (since 1990) ...................................................................................... 14
Table 3: Characteristics of prospective cohort study populations included in the review .................................................................................. 15
Table 4: Summary of meta-analysis results for women ......................................................................................................................... 17
Table 5: Text characteristics of bone density measurement for hip fracture
Theoretical calculation for 10,000 women measured ............................................................................................................................... 18
Table 6: Potential impact of BDM screening and treatment with HRT in preventing hip fractures in a population of 20,000 menopausal women ................................................................................................................................................................. 19

Figure 1: Bone mineral density distribution at the femoral neck for women with and without fractures ................................................................ 20

Figure 2: Causal pathway of screening for osteoporosis using bone density measurement ........................................................................ 21

References ....................................................................................................................................................................................... 22
Background

Osteoporosis is a histologic diagnosis, meaning “porous bone” where the bone tissue is relatively normal but there is too little of it (1). In practice, osteoporosis is commonly defined as a condition characterised by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk (2). Recently it has been suggested in a report by a World Health Organization (WHO) study group that osteoporosis be defined as a value for bone mineral density (BMD) or bone mineral content (BMC) of 1.5 standard deviations (SD) or more below the young adult mean, and osteopenia as a value for BMD or BMC of more than 1 SD below the young adult mean, but less than 2.5 SD below this value (3).

Whether or not osteoporosis, however defined, actually gives rise to fractures, will depend on a number of other factors, including the amount of bone loss, frequency and type of falls and life expectancy (4). When osteoporosis is manifested clinically as a fracture, it is a considerable health problem. Hip fractures are the most serious and potentially costly consequence of osteoporosis (in terms of morbidity, mortality and economics). The U.S. Office of Technology Assessment estimated that there were about 300,000 hip fracture cases in 1991 in the United States with an associated total cost of about US $5 billion, including nursing home, private and public care services (5). It has been estimated by Cooper et al. That the number of hip fractures occurring in the world each year will increase from 1.66 million in 1990 to 6.26 million by 2050 (6).

Osteoporosis is a problem that is particularly relevant for menopausal women. After the menopause there is a marked decrease in bone density, with a consequent increase in fracture risk. This pattern is thought to be related to the declining levels of estrogen in these women and not entirely age related, since the pattern is different for men during the same ages (7).

It is difficult to be certain how many women can be considered osteoporotic. It is necessary to either consider the number of fractures not due to excessive trauma as a proxy for indicating the prevalence of osteoporosis, or assume a particular cut off value for bone density below which women are identified as having the condition. In Sweden, there are approximately 18,000 hip fractures annually in women over 50 years of age, but it is unlikely that there are all the result of osteoporosis. The use of bone density values to define osteoporosis, as suggested by the WHO report, has profound implications for what proportion of the population is considered to be osteoporotic and thus potentially eligible for intervention. Using Swedish normative data describing bone density values at the hip, the number of women in Sweden between 40 and 89 years who could be considered to have a low bone mineral density varies between about 300,000 (14% of women between 40 and 89) and over 1,000,000 (52% of women between 40 and 89) depending on the cut-off value specified (Table 1) (8).

Reports on the applications of bone density measurement

Over the past ten years, there have been many reports considering the use of bone density measurement (BDM) published by various organisations around the world. Bone density measurement has been discussed for use in several applications: the diagnosis of osteoporosis; monitoring women with a confirmed diagnosis of osteoporosis; monitoring those being treated with drugs that may alter bone metabolism, such as corticosteroids, or hyperparathyroidism; and to identify individuals with low bone density who are at increased risk of fracture.

A survey of organisations from several countries which have produced reports considering the applications of bone density measurement was carried out and has been summarised elsewhere (8, 9). A total of 24 reports – including syntheses produced by health technology assessment agencies, consensus conference reports and summary reviews from expert panels – were identified through the literature and through personal contacts. Each respondent was also asked to characterise the organisation producing the report and the report itself with regard to methodology, data sources and main purpose.
From the survey responses (N=22), it is apparent that there is no clear consensus about the various applications of bone density measurement. When the responses are considered collectively, the application of bone density measurement for the diagnosis of osteoporosis was supported by 11 (65%) reports. Seven reports (44%) supported the application of BDM for the monitoring and follow-up of patients with previously diagnosed disease and 10 (59%) supported its use for monitoring and follow-up of patients receiving treatment that may affect their bone density.

Recommendations about screening applications of BDM varied depending on the specific population group being targeted. None of the reports supported bone density measurement for screening asymptomatic individuals in the general population to identify those with a low bone density who are at a high risk of fracture. However, 5 (25%) reports supported its use for screening asymptomatic menopausal women to identify those with a low bone density. The application of BDM for screening asymptomatic menopausal women with multiple risk factors for osteoporosis (in addition to menopausal status) to identify those with a low bone density who are at a high risk of fracture was supported by 8 (57%) reports. Twelve (71%) reports supported use of BDM for screening women who are uncertain about taking hormone replacement therapy and who may accept treatment if they are identified as having a low bone density.

Overall, the conclusions regarding the applications of bone density measurement are more conservative (i.e. fewer reports support particular applications) in reports produced by official government agencies and public non-profit agencies compared to other types of organisation. Because of the small numbers, this trend only reached conventional levels of significance for assessing women for HRT (p<0.05). Similarly, the conclusions tend to be more conservative in reports that were classified as structured overviews, whether or not the methods were specified in the report, compared to narrative reviews or reports from expert panels or consensus conferences. These differences reached conventional levels of statistical significance for monitoring patients receiving therapy, screening of menopausal women and screening women who are uncertain about taking HRT (p<0.05).

Assessment of fracture risk

It is important to distinguish between bone density measurement as a diagnostic tool and as a prognostic or predictive tool (3, 10). As a diagnostic tool, bone density measurement provides information about the presence or absence of disease (in this case, osteoporosis), however defined. In this situation, the traditional measures of sensitivity (the proportion of individuals who truly have the disease that are identified as diseased by the measurement), specificity (the proportion of individuals who truly do not have the disease and are identified as disease-free by the measurement) and positive predictive value (the proportion of individuals that are identified as diseased by the measurement and truly are diseased) can be used to characterise how well bone density measurement performs (11). This performance is commonly represented in a two by two table, in which the numbers of positive and negative findings (true and false) that will be generated by the test are shown, and from which the predictive value or likelihood ratio can be calculated.

However, BDM, as a prognostic tool, depends on the probability of future adverse events (fractures) being related to current values of bone density. In this situation, the test characteristics of sensitivity and specificity do not refer to the presence or absence of disease at the time of the test, but instead the proportion of individuals who do or do not incur fractures in the future. Sensitivity will be the proportion of individuals who will go on to have a fracture, who are correctly identified by the test; specificity will be the proportion of individuals who will not go on to have a fracture who are correctly identified by the test; and positive predictive value will be the proportion of individuals who are identified as high risk who would go on to have a fracture in the absence of intervention. The likelihood ratio indicates the probability that a low bone density measurement will be found in an individual who goes on to have a fracture in the future compared to that for an individual who does not have a future fracture.
Bone density is one of a number of risk factors for fractures (12). Bone density is a parameter that is relatively easy to measure, and an inverse relationship between bone density and fracture risk has been documented in prospective case-control and cohort studies. Bone density is, however, a continuously distributed variable, and the relationship with risk is logarithmic, with no clear threshold values of bone density below which fractures will occur or a woman can be identified as being at particularly high risk of fracture (10).

Fractures are the key outcome measure for evaluating osteoporosis and the effectiveness of associated interventions in its management. Bone density is used as an indicator of osteoporosis, and this required that a value of bone density is defined below which a woman is identified as osteoporotic. There is no evidence that bone loss in itself necessarily causes any symptoms, and some people with low bone density will never sustain a fracture (3).

Present fracture risk

The risk of sustaining a fracture in the very short term is determined by the current state of skeletal fragility and the likelihood of trauma, predominantly by falling (1). The current state of skeletal fragility can to some extent be assessed directly using bone densitometry and more indirectly from consideration of a number of other risk factors such as increased age, low physical activity, previous fractures or previous hyperthyroidism. It is important to note that BDM is a measure of bone mineral content, which is but one, albeit important, dimension of bone strength. The risk of falling depends on several factors such as reduced visual acuity, reduced hearing and use of medications. Although the risk of falling increases with the number of risk factors, the quantitative relationship is not clear (12, 13, 14).

The relationship between current fracture risk and bone density has been estimated using case-control studies where the bone density of fracture cases is compared to that of controls. A review by Law et al. (15) in 1991 considered case-control studies of women with hip fractures compared to age matched controls, where bone density was measured within 14 days of fracture. Law et al. Reported that the best measurement site for distinguishing between cases and controls was at the neck of the femur, but the weighted average difference between cases and controls was only 0.5 standard deviations. This means that there is a considerable overlap between the distribution of bone density values for individuals with a hip fracture and individuals without a fracture, even though the associated odds ratio is 1.7. Assuming a Gaussian distribution of bone density values, the results of Law et al. translate into a 30% detection rate and a 15% false positive rate using a cut-off value of 1 SD below the mean for controls (15). This work shows that the accuracy of BDM in identifying individuals who have already sustained a fracture is poor.

A similar review of more recent case-control studies (since 1990) was undertaken by the Swedish Council on Technology Assessment in Health Care (8). The results are shown in Table 2 and illustrated in Figure 1. Using a similar methodology to that of Law et al, the weighted average difference between cases and controls measured at the femoral neck was 0.9 SD, suggesting a better separation between the bone mass of individuals with and without a fracture. If a similar analysis to that outlined above is made, a cut-off value of 1 SD below the mean BMD for those without a fracture, would result in a 46% detection rate and a 15% false positive rate using a cut-off value of 1 SD below the mean for controls (15). This work shows that the accuracy of BDM in identifying individuals who have already sustained a fracture is poor.

These estimates of the test characteristics of BDM, derived from case-control studies are of limited value for drawing conclusions about the relationship between bone density and future fractures because the measurements are made after the fracture and so cannot establish a temporal sequence of events.
(16). There is also a potential for bias because of the way subjects are recruited to the study (both cases and controls). There may also be a bias towards reporting lower values for fracture cases since immobilisation (which follows a fracture) tends to decrease bone density.

Assessment of future fracture risk

In order to assess the ability of bone densitometry to predict the risk of future fractures, data are required from prospective cohort studies in which a group of people is followed-up after establishing a baseline BDM.

In a recent meta-analysis of prospective cohort studies which included a single baseline measurement of bone density and subsequent follow up for fractures, 12 study populations were identified (11 of which were populations of women), constituting approximately 90,000 person years of follow up time (8, 17). Only studies published between 1985 and 1995 were a prospective protocol, and a cohort of women, were included in the analysis. Studies published prior to 1985 were excluded since the densitometry technology and methods of analysis have developed considerably since that year. The characteristics of these studies are summarised in Table 3.

The size of the cohort followed and the length of the follow-up time varies substantially amongst these studies. For example, the Creighton University study population included 191 women followed on average for 24 years, while the Study of Osteoporotic Fractures involves over 9,000 women followed for an average of nearly five years. The weighted average follow-up time of these studies is relatively short at 5.8 years.

Several methodological problems were encountered in combining these data from different studies. The populations included in the cohort were recruited in a different manner, and the inclusion and exclusion criteria were not identical. In most cases, women who were unable to walk without assistance, had previous fractures or who had health conditions that might affect bone metabolism were excluded. The methods for measuring bone density in the studies included single photon absorptiometry (SPA), dual X-ray absorptiometry (DXA) and ultrasound. The evolution of the technology over time must also be considered when comparing studies, even if they use the same type of measurement method.

Not all studies reported their results in terms of a 1 SD reduction in bone density, so that in some instances the results had to be recalculated to make them comparable. Sometimes it was difficult to determine if the number of fractures or the number of individuals with fractures were reported in the results. In some cases, the number of individuals with fractures had to be estimated from results reporting total number of fractures. It is an important point to distinguish between these two measurements since, in a clinical context, the critical result is whether or not a specific individual will have a fracture at all. Since very high risk individuals tend to have more than one fracture, use of numbers of fractures rather than numbers of individuals with fractures will over-estimate the predictive accuracy of the test (18).

The results confirm a relationship between reduced bone density and increased fracture risk, with a relative risk of 1.5 for a 1 SD reduction in bone density compared to the mean value for the cohort for all types of fractures measured at any site (Table 4). This means that a woman with a bone density 1 SD lower than the population mean has a risk of fracture one and a half times that of a woman with a bone density corresponding to that mean, assuming other factors are equal.

Measurements at the hip and spine yield higher relative risks for hip and spine fractures (RR = 2.6 and 2.3 respectively). Measurement at the hip and spine for other types of fractures and measurement at other sites besides the hip and spine for any type of fracture yield a similar relative risk of about 1.5.

The sensitivity, specificity and positive predictive value of bone density measurement for future fractures can be calculated from these results (Table 5). Using the relative risk of 2.6 for hip fracture and assuming a Gaussian distribution of bone density values, the test characteristics can be determined for
A theoretical cohort of women at the age of 50 years with a 15% lifetime risk of fracture (19). A cut-off value of 1 SD yields a sensitivity of 38%, a specificity of 88% and a positive predictive value of 36%. If a cut off value of 2 SD is used instead, the specificity (99%) and positive predictive value (56%) are improved, but at the cost of dramatically reduced sensitivity (9%).

Two additional studies in perimenopausal women, published since this meta-analysis, gave similar estimates. For any fractures, RR=1.50; 95% CI, 1.27-1.76 (20) and OR=1.6; 95% CI, 1.16-2.34 (21) for a 1 SD reduction in bone mineral density at the spine; and, RR=1.41; 95%CI, 1.21-1.64 for a 1 SD decrease in bone mineral density at the femoral neck (20). However, follow-up times were short (about two years).

The results of this meta-analysis must be interpreted cautiously. Observational studies are susceptible to bias and combining them may result in biased estimates of the association between bone mass and fracture risk (22).

Although the analysis is based on prospective cohort studies with large numbers of participants, there are few participants with long follow-up time and the mean age of the cohort populations varies considerably. For example, the largest cohort population is that in the Study of Osteoporotic Fractures, but the follow up time is less than five years. The mean age at entry in these cohort studies varies between 57 and 83 years. Therefore, it is not clear how generalisable the results are to a situation where BMD in women about 50 years old is used to predict fractures which occur most commonly between the ages of 70 and 80 years (i.e. 20 to 30 years later). Other risk factors for fracture (e.g. risk of falling) also become more important in older women.

One should also be cautious in applying the results of this meta-analysis to specific subgroups of the population. Most of these studies involve menopausal women with no previous history of fracture, who are mobile and who do not have health conditions that may affect bone metabolism. Therefore, the results may not apply to specific subgroups at increased risk for fracture for various clinical reasons.

**Bone density measurement in primary prevention**

Perhaps the most controversy has been generated by the proposed use of bone density measurement in a primary preventive context in conjunction with interventions such as hormone replacement therapy (HRT). The rationale for this application of bone densitometry is that if women who have low bone density (and so are at an increased risk for future fractures) are treated with an effective intervention which reduces bone loss, numbers of subsequent fractures will be reduced. This sequence or causal pathway of events if illustrated in Figure 2 (10, 23).

Ideally, a bone densitometry screening programme should be based on the scientific evidence derived from randomised controlled trials of women allocated to a screening programme or no screening programme at the age of menopause (around 50 years) and subsequently followed for at least 20-30 years to compare the incidence of fractures (causal link 1). This type of study has not been conducted. A large population screening programme is being piloted in the United Kingdom, but only preliminary results are currently available (24, 25). In the absence of these data, it is necessary to adopt a stepwise approach to the pathway, where evidence for each causal link (marked 2, 3, and 4 in Figure 2) is considered separately and combined to give evidence about the overall causal link 1. However, this approach can result in unrealizable results since estimating each link is subject to some bias and uncertainty, which can be compounded when links are combined to give an overall estimate.

The strongest evidence for causal link 2 comes from the meta-analysis of prospective cohort studies (discussed above). The overall relative risk of any type of fracture was reported as 1.5 for a decrease of 1 SD in bone density (95% CI 1.4-1.7) measured at any site (8, 17).

There is also evidence to support causal link 3 from randomised controlled trials comparing a treatment group that is prescribed estrogens alone or in combination with other drugs and a control group. RCTs
of HRT used as primary or secondary prevent show positive effects through attenuation or reversal of post menopausal bone loss as measured at the spine and forearm, but not all have shown this effect when bone mass was measured at the hip (11). This is discussed in more detail in Background Paper 3 (26).

Evidence to support the last step in the pathway, causal link 4, can be derived from case-control studies where the bone density of fracture cases is compared to that of controls (8, 17). It follows that if bone density can be increased, then the risk of fractures, at least those related to bone density, will be reduced. This could also be inferred from prospective cohort studies and trials of HRT.

There is only one randomised controlled trial that has examined the relationship of treatment with estrogens with a fracture outcome, and so related to causal link 5 in Figure 2. This study reports a 37% reduction in the number of individuals with new vertebral fractures (27). A meta-analysis by Grady et al. (28) reports a pooled estimate of the relative risk for hip fracture of 0.75 (95% CI 0.68-0.84) comparing every use of estrogens with never use. This corresponds to a decrease of about 2% (15.4% to 12.8%) in the lifetime probability of hip fracture for a 50 year old white women treated with long term hormone replacement (28). Since the results from cohort studies are less susceptible to bias than case-control studies, it may be prudent to note that when analysed separately, the pooled RR is 0.85 (95% CI, 0.68-1.07) for cohort studies only and the OR is 0.57 (95% CI, 0.48-0.67) for case-control studies only (22). These results are discussed in more detail in Background Paper 3 (26).

While there is insufficient direct evidence available about the impact of a BDM screening programme, it is possible to get some indication from combining the results for each link. Table 6 presents a number of scenarios for a BDM screening programme linked to treatment with HRT in a hypothetical cohort of 20,000 menopausal women (8, 10, 17, 19, 22, 28, 29, 30).

The overall potential impact of a BDM screening programme would be a reduction of 2% of expected hip fractures (53 actual fractures) – assuming screening uptake of 50%, BDM sensitivity of 38% for a 1 SD cut-off, reduction of fracture risk of 30% which lasts throughout the woman’s lifetime, and a lifetime compliance with HRT of 30%. This means that 393 menopausal women would need to be invited to screening to prevent one hip fracture.

Even if there were strong scientific evidence from randomised controlled trials to support a BDM screening programme, the effectiveness of BDM screening would have to take into account many other factors. These include:

a) **Who to screen and who to treat:** A screening programme should only be established in the context of clearly defined policies on who should be screened and who of those who are identified with or at high risk of the disease should be treated. There are a variety of types of screening, such as population screening where all women are tested; or where women with certain characteristics such as other risk factors are tested. These all have in common the fact that the women tested are asymptomatic and thus strict requirements for evidence of benefit should be applied (31).

There is no clear ‘evidence-based’ threshold value established for defining osteoporosis. For example, screening could be limited to menopausal women with multiple risk factors for osteoporosis and treatment may be reserved for those with a bone density measurement below 1 SD of the mean. The consequences of these different approaches must be evaluated. There are many other risk factors, such as family history of hip fracture and previous fracture after the age of 50 years, that may also play an important role in defining approaches to identifying individuals at increased risk for fracture (12).

b) **Potential risks and benefits:** The potential harm to individuals who are incorrectly labeled as being at high risk of osteoporosis through BDM screening (false positives) must be carefully assessed since they will subsequently receive unnecessary additional testing and treatments, experience anxiety and incur extra costs (32). For example, in the scenario described above, 1,024 women who would not have fractured would be incorrectly measured as having a low bone
density, and would be offered treatment. Thus, nearly two thirds of women advised to take HRT on the basis of their BDM results would be taking medication unnecessarily. Also, 948 women would be incorrectly categorized (using this cut-off) as not having a low bone density (false negatives) and would subsequently have a fracture. Nearly two-thirds of women proceeding to fracture would have been incorrectly reassured.

In addition, there are other potential benefits and costs related to HRT treatment such as the risk reduction for cardiovascular disease, the risk of breast cancer and endometrial cancer, and resumption of menstruation that need to be considered (33, 34).

c) **Practical aspects of screening:** In order for a screening programme to have a public health impact, it is important to achieve a high level of uptake of the programme and compliance with the proposed treatment. Based on experience in the first year, the Aberdeen screening programme reported a compliance of only 50% (21). Estimates of compliance for 10 years suggest that only 30% of women will still be taking therapy (10, 25, 30). Additional issues about HRT treatment are discussed in Background Paper 3 (26).

d) **Alternative approaches:** The relative benefits of a high risk approach to prevention of osteoporosis through a BDM screening programme and subsequent treatment versus a population-based approach to increase the average bone density of the entire population through primary prevention activities should be considered. Alternative strategies could include potentially most cost-effective prophylactic interventions such as exercise, Vitamin D supplementation and prevention of falling in older people (4).

Before the true effects of a screening programme with BDM can be reliably estimated, data are needed on the number of fractures averted, by type, the number of fractures that are simply deferred, and the net costs of the intervention and cost per gain in quality of life. All existing data are based on various models and assumptions and are, therefore, subject to significant uncertainty.

**Conclusions**

On the basis of the information from various assessment reports and primary literature reviewed in this paper, the following important points emerge:

- Fractures are an important problem because of the high incidence and associated morbidity and mortality, particularly hip fractures and particularly in women. Bone density is only one risk factor for fracture – others include the risk of falling.
- Bone density is a continuous variable. There are no obvious threshold values of bone density below which fractures will occur or through which a woman can be identified as being at particularly high risk of fracture.
- A survey of various organisations which have previously produced a report on bone density measurement shows there is no clear agreement about the applications of this technology.
- There is a considerable overlap between the distributions of bone density for individuals with and without fractures. As a result, although it may be the best test available, BDM is poor at determining present fracture risk since it cannot distinguish clearly between patients with fracture (non-traumatic) and those without.
- In evaluating future fracture risk, a meta-analysis of prospective cohort studies shows a relative risk of fracture of 1.5 for a 1 SD reduction in bone density. This is similar for all sites and all fracture types except measurement at the hip for hip fractures (RR = 2.6) and measurement at the spine for spine fractures (RR = 2.3). BDM cannot accurately identify those who will have a fracture in the future, but can identify those at increased risk. The positive predictive value of bone density measurement for future fracture is 36% (based on RR = 2.6, 1 SD cut off, 15% lifetime risk).
- There are no completed randomised controlled trials on the effectiveness of screening programmes using bone density measurement in the prevention of fractures. Also, there is a lack of good data.
on the impact of implementing such programmes.

- It is estimated, under realistic assumptions, that a bone screening programme would lead to the prevention of only 2% of fractures in menopausal women (screening uptake of 50%, sensitivity of 38%, reduced fracture risk of 30% with HRT, lifetime compliance with HRT of 30%).
Table 1: Bone density (in hip) in women in Sweden aged 20-89

<table>
<thead>
<tr>
<th>Age range (years)</th>
<th>BDM in hip neck</th>
<th>Number of women (1992)</th>
<th>WHO definition “Osteopenia” for &lt;2.5 SD and &gt;1 ST below healthy mean*</th>
<th>WHO definition “Osteoporosis” for &gt;2.5 SD below healthy mean*</th>
<th>Lowest quartile of age matched group</th>
<th>For 1 SD below the age matched mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean BDM</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>1.01</td>
<td>0.13</td>
<td>599,597</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-39</td>
<td>0.95</td>
<td>0.11</td>
<td>567,580</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>0.99</td>
<td>0.15</td>
<td>629,054</td>
<td>120,904 (19%)</td>
<td>12,078 (2%)</td>
<td>157,264 (25%)</td>
</tr>
<tr>
<td>50-59</td>
<td>0.89</td>
<td>0.14</td>
<td>458,964</td>
<td>191,296 (42%)</td>
<td>30,659 (7%)</td>
<td>114,741 (25%)</td>
</tr>
<tr>
<td>60-69</td>
<td>0.79</td>
<td>0.14</td>
<td>436,904</td>
<td>302,119 (69%)</td>
<td>93,847 (22%)</td>
<td>106,726 (25%)</td>
</tr>
<tr>
<td>70-79</td>
<td>0.74</td>
<td>0.12</td>
<td>407,160</td>
<td>342,544 (84%)</td>
<td>125,609 (31%)</td>
<td>101,790 (25%)</td>
</tr>
<tr>
<td>80-89</td>
<td>0.75</td>
<td>0.19</td>
<td>220,534</td>
<td>158,564 (72%)</td>
<td>78,444 (36%)</td>
<td>55,134 (25%)</td>
</tr>
<tr>
<td>Sum 40-89</td>
<td>2,152,616</td>
<td>1,115,427 (52%)</td>
<td>304,637 (14%)</td>
<td>535,655 (25%)</td>
<td>344,419 (16%)</td>
<td></td>
</tr>
</tbody>
</table>

* The bone mineral density for the healthy mean in this population is taken as the mean for the two age groups 20-29 and 30-39 years.

Source: Reference 8.
Table 2: Summary of recent hip fracture case-control studies in women (since 1990)

<table>
<thead>
<tr>
<th>Measurement site</th>
<th>Weighted average difference between cases and controls (standardized)*</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femoral neck</td>
<td>-0.90</td>
<td>2.68</td>
</tr>
<tr>
<td>Trochanter</td>
<td>-0.97</td>
<td>2.79</td>
</tr>
<tr>
<td>Ward’s Triangle</td>
<td>-0.68</td>
<td>2.10</td>
</tr>
<tr>
<td>Lumbar spine</td>
<td>-0.57</td>
<td>1.81</td>
</tr>
</tbody>
</table>

* Expressed in terms of SDs of the mean value for controls

Source: Reference 8.
<table>
<thead>
<tr>
<th>Study population</th>
<th>Fracture type</th>
<th>Measurement site</th>
<th>Average years of follow-up</th>
<th>Mean age at entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana Medical Centre</td>
<td>Forearm</td>
<td>Proximal radius</td>
<td>6.7</td>
<td>57</td>
</tr>
<tr>
<td>Malmo General Hospital</td>
<td>Hip</td>
<td>Proximal radius</td>
<td>15.0</td>
<td>63</td>
</tr>
<tr>
<td>Lund, Sweden</td>
<td>Hip</td>
<td>Distal radius</td>
<td>11.0</td>
<td>63</td>
</tr>
<tr>
<td>Malmo General Hospital</td>
<td>Frailty</td>
<td>Proximal radius</td>
<td>4.7</td>
<td>57</td>
</tr>
<tr>
<td>Kuokini Medical Centre</td>
<td>Vertebral</td>
<td>Proximal radius</td>
<td>4.7</td>
<td>74</td>
</tr>
<tr>
<td>Honolulu, USA</td>
<td>Hip</td>
<td>Distal radius</td>
<td>6.7</td>
<td>304</td>
</tr>
<tr>
<td>San Francisco, USA</td>
<td>Vertebral</td>
<td>Proximal femur</td>
<td>8.3</td>
<td>304</td>
</tr>
<tr>
<td>Mayo Clinic &amp; Mayo Foundation, Rochester, USA</td>
<td>All non-spine</td>
<td>Femoral neck</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Characteristics of prospective cohort study populations included in the review.
<table>
<thead>
<tr>
<th>Study population</th>
<th>Size of cohort followed</th>
<th>Average years of follow-up</th>
<th>Measurement site</th>
<th>Fracture type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Royal Adelaide Hospital</td>
<td>492</td>
<td>5.0</td>
<td>Forearm</td>
<td>All</td>
</tr>
<tr>
<td>Adelaide, Australia</td>
<td>383</td>
<td>4.0</td>
<td>Forearm</td>
<td>Wrist</td>
</tr>
<tr>
<td>University of North Carolina</td>
<td>1,414</td>
<td>2.0</td>
<td>Forearm</td>
<td>Hip</td>
</tr>
<tr>
<td>Chapel Hill, USA</td>
<td>1,080</td>
<td>3.2</td>
<td>Calcaneous neck</td>
<td>Hip</td>
</tr>
<tr>
<td>Aberdeen, UK</td>
<td>69</td>
<td>2.8</td>
<td>Femoral neck</td>
<td>All</td>
</tr>
<tr>
<td>St. Vincent's Hospital</td>
<td>76</td>
<td>24.0</td>
<td>Calcaneous</td>
<td>All</td>
</tr>
<tr>
<td>Sydney, Australia</td>
<td>191</td>
<td></td>
<td>Proximal radius</td>
<td></td>
</tr>
</tbody>
</table>

Source: Reference 8
| Measurement site | Fracture type | Vertebral | Hip | Proximal radius | Distal radius | Lumbar spine | Calcan 
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Forearm</td>
<td>All</td>
<td>1.5 (1.3-1.6)</td>
<td>1.5 (1.3-1.6)</td>
<td>1.5 (1.3-1.6)</td>
<td>1.5 (1.3-1.6)</td>
<td>1.5 (1.3-1.6)</td>
<td>1.5 (1.3-1.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The pooled (weighted) estimates of relative risk were calculated using a fixed effects model. Source: Reference 8 and 17.
Table 5: Text characteristics of bone density measurement for hip fracture
Theoretical calculation for 10,000 women measured

These calculations are based on the following assumptions:

Lifetime risk of about 15% for hip fracture for women aged 50 years (Reference 19).
Gaussian distribution of bone density values in the population.
Logistic distribution of risk.

Using a cut-off value of 1 SD below the mean

Sensitivity: 38%
Specificity: 88%
Positive Predictive Value: 36%
Relative Risk is 2.6 for 1 SD reduction in BDM (therefore the relative risk in the table is 3.2 and the odds ratio is 4.4)

<table>
<thead>
<tr>
<th></th>
<th>Truth +</th>
<th>Truth -</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1 SD BDM</td>
<td>576</td>
<td>1,024</td>
<td>1,600</td>
</tr>
<tr>
<td>&gt; 1 SD BDM</td>
<td>948</td>
<td>7,453</td>
<td>8,400</td>
</tr>
<tr>
<td>Total</td>
<td>1,524</td>
<td>8,476</td>
<td>10,000</td>
</tr>
</tbody>
</table>

With a cut-off value of 2SD below the mean

Sensitivity: 9%
Specificity: 99%
Positive Predictive Value: 56%
Relative Risk is 2.6 for 1 SD reduction in BDM (therefore the relative risk in the table is 3.9 and the odds ratio is 7.5)

<table>
<thead>
<tr>
<th></th>
<th>Truth +</th>
<th>Truth -</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2 SD BDM</td>
<td>135</td>
<td>108</td>
<td>243</td>
</tr>
<tr>
<td>&gt; 2 SD BDM</td>
<td>1,387</td>
<td>8,370</td>
<td>9,757</td>
</tr>
<tr>
<td>Total</td>
<td>1,522</td>
<td>8,478</td>
<td>10,000</td>
</tr>
</tbody>
</table>

Source: Reference 8.
Table 6: Potential impact of BDM screening and treatment with HRT in preventing hip fractures in a population of 20,000 menopausal women
(Under realistic assumptions)\textsuperscript{b c d e}

<table>
<thead>
<tr>
<th>Fracture risk reduction from HRT</th>
<th>Compliance = 30% s</th>
<th>Compliance = 50% s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N hip fractures avoided</td>
<td>% hip fractures avoided</td>
</tr>
<tr>
<td></td>
<td>N needed to invite to screen</td>
<td>per hip fracture avoided</td>
</tr>
<tr>
<td>Screening Uptake = 50% h</td>
<td>15% (RR=0.85)</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>803</td>
</tr>
<tr>
<td></td>
<td>30% (RR=0.70)</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>393</td>
</tr>
<tr>
<td></td>
<td>50% (RR=0.50)</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>234</td>
</tr>
<tr>
<td>Screening Uptake = 70% h</td>
<td>15% (RR=0.85)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>560</td>
</tr>
<tr>
<td></td>
<td>30% (RR=0.70)</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>278</td>
</tr>
<tr>
<td></td>
<td>50% (RR=0.50)</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>166</td>
</tr>
</tbody>
</table>

These scenarios were calculated using the following assumptions:
- Assume a cohort of 20,000 menopausal women are invited to a BDM screening programme and those identified with a bone density < 1 SD below the health adult mean are treated with HRT.
- Bone density values follow a Gaussian distribution in the population
- The lifetime risk of hip fracture is 15.25\% for women over 50 years

a) Hip fracture RR=2.6 for 1 SD decrease in BMD below age adjusted mean
b) For those who comply, compliance continues for their remaining lifetimes (about 30 years)
c) There is no decrease in the protective effect of HRT over time
d) No side effects are taken into account
e) The reduced beneficial effect on those who would be treated because some women would already be taking HRT for other reasons is not taken into account
f) The range considered for fracture risk reduction is 15-50\%
g) The range considered for compliance is 30-50\%
h) The range for attendance to a screening programme is 50-70\%

Source: References 8, 10, 17, 19, 21, 22, 24, 26, 28, 29, 30, 34
Figure 1: Bone mineral density distribution at the femoral neck for women with and without fractures

1 SD cutoff
46% detection rate
16% false positive rate

2 SD cutoff
14% detection rate
3% false positive rate

Patients:
Hip Fracture
Control

Source: Data from recent case-control studies; adapted from references 8 and 15.
Figure 2: Causal pathway of screening for osteoporosis using bone density measurement

1. Asymptomatic women menopausal
2. Measure bone density
3. Identify risk groups based on bone density levels
4. Reduce bone loss/increase bone density
5. Reduce fractures

Source: Adapted from Reference 23.
References


A review of the evidence for reduction in bone loss and fractures through use of hormone replacement therapy and calcitonin

Background paper three

Laura Sampietro- Colom, Alicia Granados
Catalan Agency for Health Technology Assessment

Rosa Rico, José Asua
Basque Office for Health Technology Assessment

September 1996
Acknowledgements:

To Dr. David Hailey for all the documentation in the HRT section.

To Dr. J. Blanch, Dr. A. D. Rez, Dr. J. Foz, Dr. J.M. Segur of the Catalan Calcitonin Working Group for their collaboration in the CAHTA document of calcitonin.

To all members of CAHTA for their support in the analysis process and in the elaboration of the document, especially to Ms. Isabel Parada for her administrative support.
Further details of the work undertaken in preparation of this background paper are provided in Appendices, which are available on request from SBU - Box 16158, S-10324, Stockholm, Sweden or from AHFMR - 3125 ManuLife Place, 10180 - 101 Street, Edmonton, Alberta T5J 3S4 Canada.
Introduction

Osteoporosis has been defined as a metabolic disease characterized by a reduction in bone mass that compromises the skeleton’s biomechanical integrity and leads to an increased risk of fractures. However, bone mass loss is a physiological condition due to the ageing process and its presence is only another of the 16 independent risk factors currently identified that can contribute to the development of fractures. As part of a physiological condition, at any age bone density is in average lower in women than in men. Most bone loss occurs within the 3 to 6 years after menopause (2% each year), and accelerated bone loss diminishes exponentially over time (1% each year). During their lifetimes, women and men lose mainly cancellous bone (50% for women and 30% for men), which is mostly located in the spinal column and at the end of long bones.

Several studies have shown an inverse relationship between bone mass density and the risk of fracture. Fractures of the proximal femur, vertebrae and forearm are often linked to a decrease in bone mass. Thus, several strategies focused on bone mass preservation have been described for fracture prevention. These strategies may include reach the highest peak bone mass and the prevention of bone mass loss around menopause and with ageing. Hence, most of the pharmacological therapeutic options proposed for the prevention of progressive osteopenia have been focused either in decreasing bone resorption or in increasing bone formation. Estrogen, calcitonin, calcium, and biphosphonates and calcitriol have been proposed among the drugs that decrease bone resorption. Sodium fluoride, parathyroid hormone and growth factors are drugs that increase bone formation. All these pharmacological agents are in different stages of research. Agents that increase bone formation are, nowadays, considered under clinical investigational status. New types of biphosphonates (such as pamidronate, alendronate) and calcitriol are currently being tested under experimental protocols.

Hormone replacement therapy (HRT), by means of estrogen replacement therapy, so called unopposed therapy, or estrogen plus progestagen, so called opposed therapy, have been used in the last decades for the prevention of bone mass loss mainly in postmenopausal women. Also, in the last years, calcitonin (CT) has been used for the same purpose.

This pharmacological approach to the treatment of osteoporosis is based on the hypothesis that if bone mass loss is correlated to fracture risk, the stabilization or increase of bone mass would be expected to reduce the incidence of future fractures. However, although low bone mass is one of the risk factors for fracture, one cannot assume “a priori” that a regimen that maintains or increase bone mass will reduce the incidence of fractures. Moreover, as it was already mentioned above, there is scientific evidence indicating that fractures have a multifactorial etiology. This point is important taking into account that the ultimate goal of therapy has to be the prevention of fractures, especially hip fractures. Although vertebral fractures are frequent among postmenopausal women, hip fractures in older women and men is a major and growing health care problem in ageing societies and is associated with an important socio-economic impact on western societies and, hence, is the outcome with a greatest impact from the public health viewpoint.

The following report is addressed to review the current state of scientific knowledge on the effects (efficacy and/or effectiveness) of HRT and intranasal salmon calcitonin (SCT(N)) in preserving bone mass and lowering the incidence of future fractures. These two types of therapies have been chosen for several reasons. As regards HRT, there are considerable amount of scientific evidence to look at through. Additionally, currently there is a controversy in the scientific community about its benefits regarding the prevention of fracture events later in life. Finally, it is important to carry out an assessment upon one of the potential benefits attributed to HRT considering that this therapy could have some benefits and risks upon other pathological conditions (f.ex. breast
cancer, cardiovascular diseases...). Regarding SCT(N), it’s important to point out that it is been particularly prescribed for osteoporosis in the Mediterranean countries during the last five years leading to an important increased health care expenditures.12-14

Definitions and end points

The measures addressed in the following review have been the efficacy and the effectiveness of HRT and nasal salmon calcitonin (SCT(N)), in primary and secondary prevention for both outcomes, preserving bone mass and preventing fractures. Primary prevention is defined by women with natural or surgical menopause with normal skeletal status without a history of fracture. Secondary prevention is defined by women with natural or surgical menopause with one or more non-traumatic fracture or bone mass below more than one or two standard deviations (SD), from both young adults or from age-matched controls depending on the studies. Therapy started early or late after menopause refers, respectively, to the initiation of therapy within or after the first 10 years following last menopausal period.

Efficacy is defined as obtain the clinical benefit in a defined population from a medical technology applied for a given medical problem under ideal conditions of use (generally assessed in controlled trials). Effectiveness differs from efficacy in that the benefits are obtained under an average or actual conditions of use (generally assessed by designs other than controlled trials, however lastly multicenter RCT are also considered a measure of effectiveness).15

As regards the effect size (a summary measure of the effect of treatment on bone mass), it is represented by the number of standard deviation (SD) units which the average annual decline in bone mass in the control group exceeds that in the treatment group. A positive number indicates that the decline is greater in the control group than in the treated group. SD is statistically significant at alfa of 0.05 if 95% confidence interval does not include 0.

Material and methods

a) Hormone Replacement Therapy: literature search and data extraction

Different reports from organizations who are members of INAHTA which address the subject of hormone replacement therapy (HRT) in postmenopausal osteoporosis were reviewed.16-19 Data extraction process from the descriptive studies matrix included in the technology assessment reports mentioned was performed. An updating of the scientific literature by computer search identified through MEDLINE from 1993-1994 was carried out. Articles of relevant interest appeared up to May 1995 were also selected for this report.

b) Intranasal calcitonin: literature data extraction

There is only one technology assessment report addressing the use of intranasal calcitonin for osteoporosis.20 Data identification in that report was carried out through a literature search in English, French, Italian and Spanish languages using MEDLINE and Embase (1990-1993). For the current report, an update of the scientific literature by computer search using MEDLINE and Embase 1993 to 1994 was performed. Additional studies were obtained through a carefully checking of book chapters and references from review articles.

c) Inclusion criteria

Only original studies carried out in human subjects were selected. All the articles identified trough the search have been included. Additionally there have been tacked into account the results derived
from the technology assessment reports from different Agencies for Health Technology Assessment (considered as a *grey* literature).

d) Exclusion criteria

Studies addressing exclusively biochemical parameters related to bone metabolism were not included. An exhaustive analysis of the studies that addressed major adverse health events affected by HRT, such as coronary heart disease, breast cancer, endometrial cancer, gallbladder disease or colorectal cancer was not performed.

Regarding nasal calcitonin (CT), two studies were not considered for this report because of the type of CT used, human calcitonin (hCT) and elcatonin.

e) Analysis and data synthesis

All HRT and SCT (N) studies included in the following analysis are of different methodological quality, with ranging from randomized controlled trials to observational studies. However, when conclusions were made, the specific value given to each of the study's results was in accordance with its scientific level of quality. Tables 1 and 2 summarize the results of the literature search as regards to hormone replacement therapy (HRT) and nasal calcitonin (CT) with studies clarified according to the type of design. As regards HRT, all these studies identified through the update and not included in the list of articles subject to quality assessment by Henry et al in their meta-analysis have been object of quality assessment. This quality assessment have been based on the quality criterion used by Henry et al. and in the criterion proposed by Chalmers/Jadad score. Regarding SCT (N) the latter approach have been used.

Role of hormone replacement therapy and calcitonin in the prevention of bone mass loss

a) Effects of estrogens on bone mass

As an antiresorptive agent, estrogens may act by decreasing the imbalance between bone resorption and formation, by decreasing the overall rate of bone turnover, mainly when turnover is increased, or by both mechanisms. The effect of estrogens is greater on trabecular bone than in cortical bone. This is important because trabecular bone has a much higher rate of bone turnover, due to estrogen-deficiency and, thus, it is more responsive to hormonal changes.

Bone loss at menopause is associated with an increased bone turnover and a negative remodelling imbalance, i.e. bone resorption is higher than bone formation. The bone loss is more accelerated during the first 3 to 6 years after menopause, diminishing exponentially over time. The major effect of estrogen replacement therapy (ERT) in postmenopausal women (POM) is to reduce bone turnover, and decreasing bone resorption at a rate higher than bone formation. Thus, an antiresorptive agent will increase bone mass. However that increase is sustained only until a new bone mass state is reached. At the steady state, however, the decrease in bone resorption is matched by a decrease in bone formation; thereafter bone mass is maintained, not increased. The plateau is reached within the first two years of treatment and, then, the bone mass remains stable for the rest of the treatment period. Thus, the final effect of estrogen therapy in bone mass will differ at different stages of postmenopausal women.

b) Effect of calcitonin in bone mass

Calcitonin inhibits osteoclastic activity (parathyroid hormone-induced bone resorption), and decreases the plasma concentrations of calcium if bone turnover is high. This property makes CT potentially useful in disorders of bone metabolism characterized by excessive bone resorption such as in postmenopausal osteoporosis. Four types of calcitonin have been used until now in
humans: porcine (PCT), human (HCT), salmon (SCT) and elcatonin (ECT). Until recently, only injectable forms (subcutaneous or intramuscular) of CT have been available for clinical use; however, during the past three to four years nasally administered CT has been used in clinical practice. Recently rectal administration are under experimental investigation. The administration of calcitonin can be either continuous or intermittent (day/months on and day/months off).

The first therapeutic indication of CT was the treatment of Paget’s disease. Since PCT was associated to an antibody formation and hCT had a worse side effects profile, a higher cost and a more difficult quality control system, SCT became the treatment of choice for Paget’s disease in North America, being approved by the Food and Drug Administration (USA) for marketing in 1978. SCT was licensed for the treatment of osteoporosis in 1984.

SCT is the most powerful of all types of CT. It has a very low metabolic clearance rate and a high receptor-ligand affinity. Its prolonged use may lead to an increased antibody titre that may decrease responsiveness of bone mineral loss.

Results

A. Hormone replacement therapy (HRT)

a.1. Evidence of HRT on bone mass
   a.1.1. Bone mass and duration of treatment
   a.1.2. HRT started early after menopause
   a.1.3. HRT started late after menopause
   a.1.4. Effect of HRT at different measurement sites
   a.1.5. Route or pattern of administration of HRT
   a.1.6. Effect of HRT after cessation of therapy

a.2. Evidence of HRT and fracture risk
   a.2.1. Effect on overall fracture risk
   a.2.2. Effect on hip fracture risk
   a.2.3. Effect on vertebral risk
   a.2.4. Effect on forearm and wrist fracture risk
   a.2.5. Effect depending on age
   a.2.6. Effect according duration and past or recent use of treatment
   a.2.7. Effect according time since cessation of therapy

B. Intranasal salmon calcitonin (SCT(N))

b.1. Evidence of intranasal SCT on bone mass
   b.1.1. Duration of treatment
   b.1.2. Effect of SCT(N) when started early after menopause
   b.1.3. Effect of SCT(N) when started late after menopause
   b.1.4. Effect of SCT (N) at different measurement sites
   b.1.5. Effect of SCT (N) at different doses
b.2 Evidence off SCT(N) in fracture risk

A. Hormone replacement therapy

a.1. Evidence of HRT on bone mass

A large number of randomized controlled trials (n=55) have demonstrated the effect of HRT, used alone or in combination with progestogens and/or calcium, in reducing the rate of bone mass loss and/or preserving bone mass in postmenopausal women (POM), both for primary (n=43/35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,184,190,196,197,198,200) and secondary prevention (n=12/72,73,74,75,76,77,78,79,80,189,191,194).1

a.1.1. Bone mass and duration of treatment

Most of the studies addressing the effect of HRT in bone mass have a follow-up period of less than 24 months. Most of these short-term studies2 showed an increase in forearm and spine bone mass when women are treated with different doses and presentations of HRT. Only one RCT have lasted for 4 years and it shows a positive effect both in bone mass vertebrae and femur184. Two RCT,36,54 and one case-series study81 on primary prevention have lasted for 36 months. These studies have shown an increase in forearm and vertebral bone mass density (BMD) in the treatment group as compared to the control group, being statistically significant only for the two RCT.

One prospective cohort study, also in primary prevention, has had a duration of 25 years82 showing a positive effect on spine and forearm being highest in cortical bone (BMD forearm-cortical= increase 12.0% (p<0.02); BMD spine-cortical=816.5% (p<0.03); BMD spine-trabecular=810.4-17.4% (NS)). One case-control study, nested in the Framingham cohort, studied women (n=670, mean age:68 years) who had been taking ERT for a period from one to ten years. Women who had been taking estrogens at least seven years or more had higher values of bone mineral density for femoral neck, spine and forearm bone density than the untreated women (<75 years=11.2% (p<0.001); >75 years=3.2% (only statistically significant on the radius shaft, p<0.01)). Bone mass density increased with an increasing duration of estrogen therapy. However, ten years were needed for HRT to have an effect in femoral neck and radius (shaft and ultradistal) (p<0.05). Positive effect of HRT in bone mineral density of the trochanter, ward’s triangle, and spine were only seen after 7 to 9 years of treatment compared with women non treated.83

a.1.2. HRT started early after menopause

Several RCTs reported that HRT started early after menopause (within the 10 years following last menstruation period) maintained or increased bone mass within the first five years after menopause35,37,38,39,40,41,42,43,45,46,47,48,49,50,51,52,53,55,58,59,61,62,67,70,79,81,84,85,190,191,194. These increases in bone mass range, approximately, from 1.2% to 5% in forearm, and from 3% to 6% in vertebrae depending the studies4. Two prospective cohort studies, including

1 See background document. Appendices.

2 The Food and Drug Administration and the European Foundation for Osteoporosis and Bone disease have established a minimum follow-up of 2-3 years in order to assess adequately the effect of any drug in bone mass.

3 See background paper. Appendices.
women requesting for HRT, showed an statistically significant (p<0.05) increase of 3%\(^{86}\) to 4%\(^{87}\) in hip bone mass, in those treated women as compared with control and baseline and control values, respectively. Most of these studies analyzed HRT for primary prevention and followed women for a period of time of less than three years after the initiation of the therapy.

The meta-analysis of RCTs carried out by Henry et al.\(^{19}\) showed that HRT, opposed or unopposed, with or without calcium, prescribed for primary prevention early after menopause had a positive impact in reducing bone mass loss at spine and forearm sites. Table 3 describes the effect sizes of different types of hormone regimens in primary prevention. As regard hip bone mass, one RCT for secondary prevention was identified.\(^{33}\) Its results show an statistically significant positive effect in treated women compared with the group of placebo and calcium.

The review carried out by the Office of Technology Assessment (OTA)\(^{18}\) has identified three studies including women who initiated therapy early after menopause followed for more than ten years.\(^{88-90}\) These studies showed that HRT maintained bone mass or reduced the rate of bone loss in POM better than placebo. One of them, the only published long-term clinical trial of HRT versus placebo, showed that after 10 years women who begun HRT within three years of menopause had a small but statistically significant increase in metacarpal bone mass as compared with the placebo group\(^{4}\). In those who started therapy after three years, a statistically significant decrease was observed\(^{5}\). Women assigned to the placebo group had a sustained decrease in bone mass\(^{6,90}\). The retrospective cohort study with an average follow-up period of six years showed no statistically significant change in bone mass and cortical thickness in natural menopause treated women when compared with baseline values (81.12% NS), whereas untreated women had a statistically significant decrease in bone mass and cortical thickness compared with baseline values (96.30%, p<0.05).\(^{88}\) Finally, the prevalence study of women who had begun therapy within 5 years of menopause and who continued it for at least 10 years has shown a significant difference in spinal bone mean mineral density between long-term estrogen users (1,219 g/cm\(^2\)) and non users (1,092 g/cm\(^2\)), after controlling for age and type of menopause.\(^{89}\)

Recently, a case-control study on the effect of postmenopausal estrogen therapy on bone density in 670 women of white race from the Framingham study cohort have been published.\(^{83}\) This study showed that women under 75 years of age having taken estrogens for more than seven years the past had a bone density 11.2% greater than the women who had never received estrogen, averaging all sites. The positive effect was observed in radius (shaft, ultradistal), ward’s triangle and trochanter (p<0.001), as well as at femoral neck (7.6%; p<0.004) and lumar spine (8%; p<0.018). Bone density was 3.2% higher among 75 year old and older women with the same therapy and duration than in women who had never taken estrogen. However, positive effects were only statistically significant at the radius shaft (88.5% p<0,02).

The long-term effect of HRT started around menopause or soon after has not been studied through a RCT.

\textbf{a.1.3. HRT started later after menopause}

Several RCT,\(^{56,72,74,76,48,60}\), three non-randomized controlled trials,\(^{193,195,199}\) and one retrospective cohort study,\(^{91,75}\) one case-control\(^{83}\) and one cross-sectional study\(^{92}\) have shown that HRT is able to halt or possibly reverse bone loss, even if its start is delayed late after
menopause, generally for secondary prevention\textsuperscript{72,92,83,74,76,60,93,193}. Table 4 summarize the results of some of these studies.

In the meta-analysis performed by Henry et al.\textsuperscript{19} HRT for secondary prevention initiated late after menopause showed a low protective effect on BMD spine when oral unopposed estrogens were used. However, when calcium was added to estrogens there was a large positive effect on spine BMD (2.27 SD; 95\%CI:1.17-3.37). However, when the effects of age on skeletal response was analyzed the authors conclude that with the available evidence there is no data to support that age attenuates the response to estrogens.\textsuperscript{94,95} Table 4 shows the effect of age and HRT bone mass response. Table 5 describes the effect sizes of different hormonal regimens in secondary prevention.

\textbf{a.1.4. Effect of HRT at different measurement sites}

When individual studies were analyzed the increases in bone mass were higher in lumbar spine than in forearm or hip. However, when meta-analytical techniques were used the effect appears to be stronger at the forearm than the spine\textsuperscript{7,19}. Nevertheless, the authors suggest that although the results of the meta-analysis shows a marginal stronger effect at forearm than at the spine, the 95\% confidence intervals are wide and overlap (i.e. the difference observed is not statistically significant), suggesting that there may not be a real difference between sites.\textsuperscript{19}

When the effect on bone mass in the hip was analyzed, the comparison of the results of four RCTs\textsuperscript{74,80,73,78} in secondary prevention showed a moderate size positive effect (control difference 0.92 SD; 95\%CI:0.34-1.50).\textsuperscript{19} Additionally, the effect sizes were consistently two-fold higher across the individual studies for the spine bone mass than for hip. Table 6 shows the effect of HRT at different measurement sites.

One cross-sectional study\textsuperscript{92} in primary prevention, a case-control study nested in a cohort\textsuperscript{83} and a retrospective cohort study\textsuperscript{93} in secondary prevention showed a higher increases in trochanteric bone mass than in femoral bone mass, being these results statistically significant.

\textbf{a.1.5. Hormone regimen and route of administration}

Considering either the route (oral, transdermal) or the pattern of administration (sequential, continuous) of estrogen and/or progestagen, HRT has been found to reduce postmenopausal bone loss whatever hormonal regimen or route of administration was chosen.\textsuperscript{19,18}

\textbf{a.1.6. Effect of HRT after cessation of therapy}

One RCT\textsuperscript{54}, two case-series studies\textsuperscript{96,97} and one retrospective cohort study\textsuperscript{192} have shown that after cessation of therapy bone loss accelerated to a rate equivalent to that of untreated women at menopause. The RCT studied the effect on forearm of initiation and withdrawal of HRT in healthy women, six month to three years after menopause. During three years

\textsuperscript{4} Time since last menopause use Tm < 3 years = \#8.67\%; p<0.001.
\textsuperscript{5} Tm > 3 years = \$0.5\%; p<0.001.
\textsuperscript{6} Tm < 3 years = \$9\% / Tm > 3 years = \$11.29\%; p values versus baseline not reported.
\textsuperscript{7} Primary prevention: number of studies = 8 / Effect size: a) spine = 1.17 (0.63 - 1.70); b) forearm = 1.38 (0.93 - 1.84).
of HRT bone mass increase (3.7%, p<0.001), however the annual rate of bone loss after discontinuation of HRT was identical with the bone loss in the placebo group.\textsuperscript{54} One of the case-series studies showed that eight years after the initial women attendance for treatment, there was no difference between the women treated for four years and patients who had received placebo for the full 8 years of the study.\textsuperscript{97} The other case-series reported a percentage of bone loss rate in those women who previously used estrogen and discontinued it after age 65 years similar to those women never users aged 56 to 70 years (more than 2.5% per year; p<0.05).\textsuperscript{96} The retrospective cohort study shows that bone loss was accelerated among women who stopped their estrogen use, being annual losses increased 0.35 to 0.60%\textsuperscript{192}. Table 7 summarizes the results from these studies.

**a.2. Evidence of HRT on fracture risk**

**a.2.1. Effect of HRT on overall types of fracture risk**

The meta-analysis carried out by Henry et al.,\textsuperscript{19} which included the observational studies used by Grady et al in their meta-analysis,\textsuperscript{98} included seven cohort studies,\textsuperscript{99-105} and 11 case-control studies\textsuperscript{106-116} performed from 1979 to 1993. All the studies but one\textsuperscript{103} reported a reduction in the relative risk for all type of osteoporotic fractures.

A case-control study which assess risk factors associated with fractures among a group of 3140 perimenopausal women, concluded that HRT is protective in this respect\textsuperscript{185}. For all types of fractures, perimenopausal women with past or present history of HRT had 30% less risk of having a fracture (OR=0.70; 95%CI:0.50-0.96). Of the fractures cases, 44 (28%) of them were undergoing HRT at the time of fracture.

**a.2.2. Effect of HRT on hip fracture risk**

The studies identified were of observational design. All non-experimental studies, but one prospective cohort,\textsuperscript{101} indicated a decreasing trend in the relative risk of suffering hip fractures. However, four cohort studies\textsuperscript{99,102,103,117} and six case-control studies\textsuperscript{107,108,111,112,114,115} did not show statistical significant differences in the risk of hip fractures between ever and never users of HRT.

The largest and most complete cohort study of HRT carried out\textsuperscript{100} in a cohort of 23,000 Swedish users (average age=53.7) followed-up for 5.7 years, reported a greater protective effect of HRT for trochanteric hip fractures (RR=0.60; 95%CI:0.35-0.96) than for cervical fractures (RR=0.73; 95%CI:0.55-0.95).

A cohort study with the largest follow-up (approx.30 y) shows a decreasing trend in the relative risk of suffering peripheral fractures (RR=0.23; 95%CI:0.06-0.97). There were 12 patients with peripheral fractures per 1000 patients year in the control group (n=11), three of them hip fractures, an no fracture among the patients in the HRT group. Although there is a decrease in the relative risk, the superior limit of the confidence interval almost reach 1, that may lead to a non statistically significant result.

Two meta-analyses have studied the relationship between HRT and the decrease in fracture risk.\textsuperscript{19,98} Grady et al included in its meta-analysis 11 epidemiological studies\textsuperscript{8} performed from 1979 to 1991.\textsuperscript{98} All these studies, but one,\textsuperscript{101} reported a reduction in the risk of hip fracture.

\textsuperscript{8} Case-control = 6\textsuperscript{106-110,118}; cohorts = 5\textsuperscript{99-101,105,117}.
fracture. The estimated pooled relative risk of hip fracture comparing ever-users with non-users of estrogens indicated a risk reduction of about 25% (RR=0.75; 95%CI:0.68-0.84). That resulted in a 2% decrease in the life-time probability of a hip fracture for a 50 year old white women.

A more recent meta-analysis\(^9\) reported a reduction of the risk of hip fracture of about 15% in ever-user compared with non-users of hormone replacement therapy (estimated pooled RR=0.85; 95%CI:0.68-1.07) for cohort studies, and about 43% (estimated pooled OR=0.57; 95%CI:0.48-0.67), for case-control studies. However, when the case-control studies were analyzed by level of quality, those with highest level were associated to a relative risk closed to the ones presented by the cohort studies.\(^9\)

### a.2.3. Effect of HRT on vertebral fractures risk

There is only one randomized controlled trial identified.\(^80\) This trial assessed the effect of transdermal estrogen (0,1mg of 17-b-estradiol plus medroxyprogesterone acetate) in 75 post-menopausal women (age:45-75 years) with one or more vertebral fractures due to osteoporosis. Treatment was associated with a decrease rate of 61% in the incidence of new vertebral fractures (RR=0.39; 95%CI: 0.16-0.95). Eight new fractures occurred in seven women in the estrogen group, whereas 20 occurred in 12 women in the control group.

Two prospective cohort studies analyzed vertebral fractures. One of them showed a reduction in wedge deformities,\(^119\) and the other showed a lesser rate of vertebral fractures in users versus controls (181/1000 person-years vs 834/100-person-years, respectively).\(^120\)

### a.2.4. Effect of HRT on forearm and wrist fracture risk

For forearm, the studies showed a decreased in risk of fracture, but the effect was not always statistically significant (cohort\(^99,103,104\); case-control\(^106,115\)).

A meta-analysis of observational studies\(^9\) also analyzed the impact of HRT in the risk of wrist and forearm fractures from cohort studies,\(^99,103,104\) indicating a reduction in fracture risk of 30% (estimated pooled RR: 0.70; 95%CI:0.52-0.93) in ‘ever users’. When case-control studies were considered,\(^106,109,115\) there was also a significant decrease in fracture risk of 40% for ever users of estrogen (estimated pooled OR=0.58; 95% CI:0.42-0.79).

### a.2.5. Effect of HRT on hip fractures according to age

Some of the studies included in the mentioned above meta-analyses\(^9,98\) reported a decrease in the potential protective effect for hip fracture with age and at elderly ages, when most hip fractures occur, there was no statistically significant difference in fracture risk between ever and non-ever users of HRT (cohort\(^100,101,117\); case-control\(^113\)).

A recent published prospective cohort study (n= 9704 ambulatory white women) showed that women over 75 years who were current users of estrogen therapy had a protective effect in the risk of hip fracture (RR=0.18; 95%CI:0.04-0.77). However, 75 year old or younger women, who were also current users reported no effect (RR=0.94; 95%CI:0.52-1.69), both compared with women who had never used estrogens.\(^121\) Table 8 summarizes the relative risk of hip fracture associated to different ages.

### a.2.6. Effect according past or recent use and duration of treatment

One prospective cohort study\(^101\) indicated the lack of relationship between the duration of estrogen use and hip fracture (RR for ever use = 1.02; 95%CI:0.81-1.27).

Regarding duration, the different estimated relative risk of this study were 1.19 (95%CI:0.89-1.60) for no more than three years of use, 0.89 (95%CI:0.63-1.23) for 4-14 years of use
and 0.88 (95%CI:0.63-1.24) for at least 15 years of use compared with never users. A recent published prospective cohort study\textsuperscript{121} analyzed the relationship between short-term (<10 years) and long-term (>10 years) duration of use in 65 years of age women who were current and previously users of ERT. In current users, short-term duration of treatment was associated with a decrease of 30% (RR=0.67; 95%CI:0.49-0.92) in the risk of all non-spinal fractures, whereas for long-term users this reduction was 40% (RR=0.60; 95%CI:0.45-0.83). As regards hip fractures, the decrease in risk in current users was 19% (RR=0.81; 95%CI:0.40-1.65) and 73% (RR=0.27; 95%CI:0.08-0.85) for short and long term users, respectively.

Nevertheless, long term use among pass users was not associated with a reduction in the risk for hip (RR=1.67; 95%CI:0.92-3.10), wrist (RR=0.90; 95%CI:0.50-1.64) or all non-spinal (RR=1.00; 95%CI:0.75-1.35) fractures. That fact also was observed in women who were previous users for less than 10 years. Table 9 summarizes the data from studies that analyzed the effect of previous and recent use of HRT and years of duration in risk fracture.

In the meta-analysis carried out by Henry et al, the four case-control studies\textsuperscript{107,113,115,116} that examined the relationship between extended use of estrogen (use for five or more years) and hip fracture risk compared to never use estimated a trend to 66% reduction in the relative risk of hip fracture (OR=0.34; 95%CI:0.20-0.55). When compared to a shorter term use (0-60 months) the estimated pooled relative risk reduction was 61% (OR=0.39; 95%CI:0.25-0.62).

When the results from different cohort-studies\textsuperscript{99,103,104} with a duration of use higher than 5 years were pooled, the risk reduction in wrist and forearm fracture reported was about 15% (estimated pooled RR=0.85; 95%CI:0.73-0.99). This figure was not significantly different from that observed in never users.\textsuperscript{19}

a.2.7. Effect according the time since cessation of therapy

A prospective cohort study\textsuperscript{101} shows no difference in hip fracture risk between ever and never users of HRT whatever was the duration of therapy (from less than three years to more than 15 years) and the time since the last use (from less than two years through more than 15 years).

Pooling the results from two cohort studies\textsuperscript{101,102} reported an estimated RR effect on hip fracture, between former versus never users, of 0.88 (0.67-1.15) and 1.07 (0.85-1.34) when 2-14 years or more than 15 years exist since last estrogen use, respectively. These results showed also no statistically significant differences in hip fracture risk between ever and never users of HRT (RR=0.73; 95%CI:0.5-1.07).\textsuperscript{19}

B. Intranasal salmon calcitonin SCT(N)

b.1. Evidence of SCT(N) on bone mass

Several studies have shown that intranasal salmon calcitonin [SCT(N)] has a positive effect on decreasing bone loss or and preserving bone mass in both primary (RCT\textsuperscript{122,125,124,125,126,127,128,129,130,131,187}) and secondary (RCT\textsuperscript{26,133,134,135,136,137,138,188}; RCT+case-series\textsuperscript{139,140}; observational+RCT\textsuperscript{142}) prevention in postmenopausal women, either for natural or surgical (RCT\textsuperscript{143,144}) menopause. Only three studies in primary prevention,\textsuperscript{128,129,187} four studies in secondary prevention\textsuperscript{133,134,136,137} and one on surgical menopause\textsuperscript{143} were of good or good enough quality.
b.1.1. Duration of treatment

Most of identified studies analyzed the effect of SCT (N) in a follow-up period lesser or equal to two years, starting treatment either early after menopause or at elderly ages. Two RCT in primary prevention have had a duration of 36 months, both of them randomly assigned patients who started menopause (Tm=3-36 months) to 50 mg of SCT(N) plus 500 mg Ca or 500 mg Ca, showing a statistically significant increase in bone mass at vertebral in the SCT (N) treated group (p<0.01; p<0.05). One RCT in primary prevention had last for 5 year and also showed an increase in vertebral bone mass after 42 months of treatment. However, the increase was not statistically significant at the end of the five years.

b.1.2. Effect of SCT(N) when begun soon after menopause

An increase or preservation in bone mass was observed when treatment with SCT(N) start early after menopause. The RCT with the highest follow-up period of time (five years) in this group of patients assigned randomly to 50 IU SCT (N) intermittent (five days per week) plus 500 mgr of calcium to one group of patients (nfinal=42) and 500 mgr of calcium to the other group (nfinal=45). This trial reported a bone sparing effect of 7.8% in the lumbar spine in the SCT group compared to women who received calcium alone, concluding that intermittent nasal administration of low dose of SCT(N) prevents postmenopausal bone loss in the spine if used during the first five years after menopause. However, the authors also pointed out that, during the last two years, the group receiving calcium alone did not experience further bone loss suggesting a possible protective effect of calcium alone against trabecular bone loss after the years immediately following the menopause. There are no data available about the effect of SCT(N) on bone mass in women later in menopause (when bone turnover is lower) if therapy started early after menopause.

b.1.3. Effect of SCT(N) when started late after menopause

Several studies that assessed the impact of SCT(N) in osteoporotic postmenopausal women who started treatment later after menopause, reported a statistically significant positive effect on bone mass, mainly at spine (RCT134,136; RCT+case-serie140,139). All these studies had a follow-up period equal or lesser than two years. Thus, there is no evidence of the effect of the therapy in bone mass in women aged 75 years or over when the risk of hip fracture is the highest.

b.1.4. Effect of SCT(N) at different measurement sites

Several studies that analyzed the effect of SCT(N) in bone mass at vertebrae and forearm, both for primary (RCT125,131) and for secondary (RCT134,136,137; RCT+case-serie140) prevention, reported the largest increment in bone mineral density at the level of the spine. However, some other studies did not show differences in bone mass increments between the treated and the control groups, either for primary (RCT129) or secondary (RCT133,188) prevention.

Three studies in secondary prevention assessed the effect of SCT(N) on the hip. Two of them, a multicentre RCT (n=9 centres), analyzed the effect of 100 IU SCT(N) (n=75) on the femoral neck and trochanter as compared with 10 (n=68) or 20 mg (n=72) of alendronate or placebo (n=71) after one26 and two 188 years of follow-up. The former study reported no statistical significant differences in the percent of increase in bone mass at hip (femoral neck and trochanter) from baseline levels and placebo when SCT (N) is used for one year. The same results were observed in the later188
The other RCT, performed in women with normal and high bone mass turnover,\textsuperscript{142} reported that 200 IU SCT(N) in high bone turnover osteoporotic women was more effective in increasing bone mass at spine than 100 IU SCT(N) (p<0.005) after two years of treatment. That differential increment was less evident in femoral shaft BMD, although it remains statistically significant (p<0.005).

Another study (RCT+case-serie)\textsuperscript{139} analyzing a discontinuous calcitonin treatment on osteoporotic women reported that patients with initially high bone turn-over (serum alkaline phosphatase and FuHpr/Cr) and low bone mass at the forearm had a statistically significant higher bone response than those with initially low turn-over and high bone mass (p<0.02) (data not showed in the study).

b.1.5. Effect of SCT(N) at different doses and patterns of administration

Those studies focused on the impact of different doses of SCT(N) in bone mass of the spine showed an statistically significant positive effect at high (200 IU) (RCT\textsuperscript{134,137}) and low (50 IU) (RCT\textsuperscript{123,130,131}) doses, when given for less than 3.6 years. Higher units were given in secondary prevention whereas low doses were prescribed in primary prevention.

Several RCT compared different doses of SCT (N) both in elderly osteoporotic (RCT\textsuperscript{134,136}; observational+RCT\textsuperscript{142}) and young healthy (RCT\textsuperscript{125}) postmenopausal women. In the former group of studies, a non statistically significant increase in spinal bone mass was observed after one year of treatment with 100UI SCT(N)\textsuperscript{134} or after two years of treatment with 200 IU SCT(N)\textsuperscript{136,142}. One of them showed a dose related response to salcatonin, manifested by an increase of 1% per 100 IU (0.2-1.7%; p=0.008).\textsuperscript{136} The other study also showed a positive effect on the femoral shaft after two years of treatment (p<0.05 100 vs 200).\textsuperscript{142} However, one of the RCT\textsuperscript{134} reported a decrease of vertebral bone mass during the last three months of treatment when a dose of 200 IU of SCT(N) was used.

In the group of healthy POM a statistically significant increases in spinal bone mass were also observed in a 200 IU dose after two years of treatment (increase 3%, p=0.003).\textsuperscript{125} A recent RCT of low-dose of SCT (N) (50 IU/5d/w) plus calcium compared with calcium alone (500 mg) showed a positive effect in spine bone mass after 42 months of treatment without develop resistances (82.5%; p<0.01 vs base lines levels); nevertheless there was no increase at the end of the five years (81.1(1.1)%; NS vs base lines levels). The difference in the evolution of spine BMD between the two groups of patients was statistically significant (p<0.001).\textsuperscript{131}

As regards as intermittent pattern of administration, two studies in primary prevention (RCT\textsuperscript{131,124}) and four in secondary prevention (RCT\textsuperscript{143,138}; RCT+case-serie\textsuperscript{139}) have been identified. Only one RCT included in the latter group have compared continuous versus intermittent pattern of administration. This study shows not statistically significant differences at two months of treatment versus baseline levels in neither of the two groups.\textsuperscript{143}

For primary prevention, only one of the studies have shown an statistically significant increase in spine bone mass after 42 months of treatment when 50 IU of SCT(N) was given (p<0.01 versus baseline levels).\textsuperscript{131} For secondary prevention three studies showed a statistically significant effect in spine bone mass when 200IU of SCT(N) was given discontinuously (RCT+case-serie\textsuperscript{139,140}). On of them, concluded that the optimum response in forearm was achieved using discontinuous therapy.

b.2 Evidence of SCT (N) in fracture risk

Three randomized controlled trials on secondary prevention\textsuperscript{133,135,146} and two case-control
studies analyzing the effect of calcitonin in fracture rates have been identified. Two of the RCT referred to SCT(N) while the other assessed SCT (IM). All of them measuring the effect on vertebral fractures.

One of the RCTs that analyzes the effect of SCT(N) in osteoporotic women late after menopause, showed a decrease in the relative risk of patients with new vertebral fractures (RR=0.23;95%CI:0.07-0.77) and rate of new fractures (RR=0.37;95%CI:0.14-0.95) in salcatonin users compared with non-users in fracture vertebral rate when salcatonin is used. Data from the other show no differences after three years of treatment.

As shown by Rico et al., there is a statistically significant decrease in the rate of new vertebral fractures in the CT group (960%) as compared with the group of patients treated with 500 mg of calcium alone (835%; p<0.025 between groups) after two years of treatment in secondary prevention.

Two case-control studies reported the risk for hip fracture between ever and never users of CT. The results indicated that having taken calcitonin in the past significantly decreases the risk of hip fracture. Data from the Mediterranean Osteoporosis Study suggested a reduction in the risk of hip fracture from 29% (RR=0.71; 95%CI:0.52-0.9, when adjustment for previously estrogens intake was made) to 31% (RR=0.69; 95%CI:0.51-0.92; when adjustment for centre, age, BMI, and previous fragility fracture were made). Other study found a 53% decrease in the risk of hip fracture (RR=0.47; 95%CI:0.30-0.74) when previous use of calcitonin plus calcium was made.

Discussion

Antiresorptive agents, such as HRT and CT, are addressed to preserve or reduce bone mass loss. The final potential aim of these treatments is to reduce the incidence of future fractures, specially hip fracture, at the period of age when most of them tend to appear (75 y).

Available scientific evidence indicated that hormone replacement therapy has an effect in the short-term to preserve or increase bone mass both in primary and secondary osteoporosis. For both types of prevention, the impact was higher in lumbar spine than in forearm and hip. At the hip, the bone mass increment observed was also higher in the trochanter than in the femoral neck. These results confirmed the fact that spine and trochanter were more sensitive to estrogen therapy due to the fact that in that sites trabecular bone is present in higher rates than cortical bone.

When started early after menopause, within three years from the last menstrual period, an small but statistical significant increase in bone mass was reported after 10 years of treatment (p<0.001). However, the results from Felson et al. suggested that ever use of estrogen replacement therapy (ERT) initiated soon after menopause, during at least seven years of duration, have a low protective effect in preserving bone mass among women older than 75 or 80 years (11.2 % vs 3.2%). If this figures will be confirmed by other studies it would be necessary to ask for the usefulness of HRT given early after menopause to preserve bone mass at elderly ages.

When started later after menopause, several studies reported that HRT was able to halt or possible reverse bone loss. However, the meta-analysis performed by Henry et al. showed a trend to the largest positive effect on decreasing bone mass loss of oral unopposed estrogens when calcium was added. This results indicated that calcium could play an important role in preserving bone mass in secondary prevention.

Several studies showed that the effect of HRT in bone mass disappeared progressively after cessation of therapy, reaching a rate of bone loss equal to the rate of untreated women within few years after withdrawal of treatment. Thus, these studies indicated that the effect of HRT was only
maintained while it is used. This is of most importance due to the fact that most HRT are prescribed around or early after the beginning of menopause, and generally for not more than ten years. Therefore, leading a gap of 15-20 years since the period of life when more hip fractures occurs (>75y).\textsuperscript{145}

Taking into account all that results, two ways of action could be formulated in order to preserve bone mass during all long women life, to give HRT over a life time when menopause started, or start the therapy later in life. Both of them have its limitations due to the increase in the risk of adverse events after few years of treatment (f.ex breast cancer), and the side effects (f.e. bleeding) associated with HRT which probably limits its use in elderly ages. Another limitation is the current controversial issue that HRT started late in menopause despite increasing bone mass can not restore the potential impairment of bone micro-architecture.

Calcitonin is another antiresorptive agent that have been proposed during the last decade for the treatment of bone mass loss. The drug has a short-term safety profile and no severe adverse effects have been already described.\textsuperscript{26,13,12,134,149} The comfortable and easy administration of intranasal SCT and its less rate of side-effects as compared to parenteral administration,\textsuperscript{149} even in patient known to be intolerant to parenteral administration,\textsuperscript{150} have lead to an increase in its consumption in the last five years, mainly in the Mediterranean countries.\textsuperscript{12,13,14} Several RCT have proved the short-term efficacy of calcitonin in preserving bone mass at young or elderly ages, mainly at the spine.\textsuperscript{125,129,133} At hip, 200 IU SCT(N) showed more responsiveness among women with high turnover after two years of treatment.\textsuperscript{142} These results confirmed previous observations that patients with high turnover were more likely to respond better to therapy.\textsuperscript{131} That characteristic is common to all antiresorptive agents which have a higher impact when bone turn-over is high, as within the first 6 years of menopause. No data are available on the long term efficacy of SCT (N) in preserving bone mass in a large population of patients. Because of the short follow-up of the studies, either for those that started therapy early or late after menopause, there are still uncertainties about the long-term efficacy of SCT (N) in preserving bone mass in women over 75 years of age.

The lack of evidence about the long-term efficacy of SCT(N) is important if we take into account that the prolonged use of CT have been associated to a decrease in bone mineral responsiveness.\textsuperscript{28,34} That clinical resistance have been attributed to a possible passive immunization, i.e. antibodies to heterologous synthetic CT.\textsuperscript{28,34} Other mechanisms proposed to account for the resistance has been a receptor down regulation or a compensatory increase in the secretion of parathyroid hormone or in the formation of 1,25-dihydroxycholecalciferol.\textsuperscript{139,151} Therefore, two hypotheses could be suggested to overcome clinical resistance: 1) establishing a SCT (N) intermittent regimen; 2) decrease the doses to be administered. The largest follow-up RCT\textsuperscript{9} of intermittent therapy at low doses (50 IU) have proved its effect in preserving bone mass when administered for 42 months. However, at the end of the period, the increases in bone mass observed were not statistically significant different from the ones observed in the control group (500mg Calcium). Further studies are needed in order to define the real long-term effects of SCT(N) and the therapeutic approach that might achieve the highest effect, such as length and size cycles.

There is no publications dealing with prospective controlled studies comparing the effect of HRT and SCT(N) in bone mass.

Since most of the studies published were performed under ideal conditions of use and for short-term periods of time, no data is available from studies about the long-term efficacy of HRT or

\textsuperscript{9} 5 years.
SCT (N) in preserving bone mass under actual conditions of use in a large population of patients, i.e., its effectiveness.

The scientific literature has indicated an inverse relationship between bone mass and fracture risk, i.e. the lower the bone mass, the higher the risk of fracture, and vice versa. Thus, low bone mass have been considered as a risk marker for fracture. Hypothetically, if antiresorptive agents preserve bone mass or decrease the rate of bone loss, then, the fracture risk will be diminished. Two meta-analysis from case-control studies and cohort studies showed a trend to a decrease in risk of fracture among ever users of HRT. The figures reported in meta-analyses for the hip fracture risk reduction, when ever and never users of HRT were compared, were 15% for cohort studies and 43% per case-control studies in one metaanalyses and 25% in the other. In the former met-analysis, those case-control studies with better quality assessment punctuation had results that were more consistent with those of the cohort studies, hence decreasing the risk reduction. Also, a 30% reduction in forearm and wrist fractures have been observed.

Although these metaanalyses showed a trend to a reduction in the risk of hip fracture, some of the studies included in these meta-analysis reported an increase in risk of hip fracture with advancing age and the risk reduction achieved at that later ages (>60 years) are not statistically significant. Some facts can explain this non significant results. First of all, sample sizes at this age group are small which it raise the problem of statistical power. Also, the effect size of treatment depends upon the length of follow-up and the rate of fractures in the control group. Finally, with an advancing age and progressive bone loss, the relative importance of other skeletal (e.g. discontinuity of trabecular architecture) and extraeskeletal (e.g. propensity to fall) risk factors may increase. Moreover, a recent published study in women over 65 years reported no substantial effect on the risk of fractures when women report a previous use of more than 10 years of ERT started early after menopause (RR=1.00; 95%CI:0.75-1.35). This study also reported that only current use for more than 10 years confers a decrease in the risk of hip fracture. Moreover, the Leisure World Study reported no relationship between HRT and protection in the risk of fracture in women over 73 years who have stopped therapy for 2 years or more.

As regards SCT(N), three RCT have studied its impact on decreasing vertebral fractures in osteoporotic women. Although the results from all these studies could suggest a positive impact of calcitonin in decreasing the incidence risk of vertebral fracture, the data available do not allow us to conclude definitively that SCT (N) has en effect in decreasing the incidence of vertebral fracture in primary prevention as well as quantify its real effect for several reasons. First at all, all the studies have been performed in osteoporotic women with one or more forearm or vertebral fractures. Ross et al. have observed, in a prospective cohort (n=1098) study of 4,7 follow-up, that the presence of a single vertebral fracture at the baseline examination increased fivefold the risk of subsequent fractures (RR=5.3; 95%CI:1.9-15.2). Two or more vertebral fractures at baseline increase the risk by 12-fold (RR=11.8; 95%CI:5.1-26.8). Thus, the initial number of fractures should be taken into account when the effect of treatment is studied, there could be a possible differential efficacy of the treatment among patients with small or large numbers of vertebral fractures at baseline. None of the identified studies have randomized patients according to baseline number of vertebral fractures. Thus, the net effect of CT in the reduction of the fractures rate derived from these studies could be overestimated. Also, another confounding factor could be the fact that some of the study subjects with vertebral fractures were already enrolled in physiotherapy and rehabilitation programmes which might contribute to lowering the fracture rate.

Two case-controls that analyze the impact of CT in decreasing hip fracture, showed a statistically significant positive impact when previous use of CT was reported. Nevertheless, this type of
studies do not allow to establish a causal relationship between the effect of calcitonin and the decrease in the incidence of fracture due to the nature of the design.

Low bone mass can be considered as a risk marker of fracture. A risk marker of fracture risk have been defined as an attribute or exposure that is associated with an increased probability of occurrence of fracture, but not as an absolute predictor of the untoward event. At that point, two questions arise. The first question refers to the appropriateness of low bone mass as an adequate surrogate of fracture, and hence, the appropriateness to prescribe HRT and or SCT (N) top prevent hip fractures when its incidence is higher. The second refers to the importance of low bone mass as a risk factor when compared with other known risk factors for fracture.

Regarding the first question, the results of five case-control studies that showed only a 0.5 standard deviation reduction in the bone mass of the neck of the femur, among patients with hip fractures and controls. A similar review using case-control studies since 1990, give a weighted average difference between cases and controls of 0.09 SD. Hence, these studies suggest that fractures might appear or not in patients with the same specific bone-mass level, questioning the usefulness of low bone mass as a proxy for fractures.

Given the hypothesis that low bone mass would be considered as an adequate proxy of fracture incidence, therapy should preserve or increase bone mass since its initiation and, therefore, prevent fractures. The type of study that might provide a definitively conclusion on the relationship between the impact of HRT and SCT(N) on bone mass, started early or later after menopause, and subsequent prevention of fractures at elderly ages (>75) is the RCT. Currently, there is not any RCT identified for that purpose. The main limitations to carry out a trial of that characteristics are: the need of a big sample size to have enough cases of hip fractures and the long-term period of follow-up between menopause and the age when most hip fractures occur; the great amount of patients and the economic resources needed.

Taking into account that HRT usually is recommended for no more than 10 years, mainly due to the adverse effects associated to HRT, there will be a gap of 15 years between the moment in which therapy is stopped and the period of time when most hip fractures occur (>75 y). As regards as SCT(N), hen compared to the potential hazards associated with the use of pharmacological doses of estrogen, it appears relatively innocuous. Thus, women might be willing to follow long-term therapy. However, a decrease in responsiveness to long-term administration of CT have been observed in several studies. Thus, whether the risk protection remains at elderly ages when SCT(N) is used is not known, questioning the usefulness of this therapy for bone mass preservation. Additionally, considering that most of the randomized controlled trials of the effect of HRT and SCT(N) lasted for not more than two years, and that the protective effect in the case of HRT wears off early after cessation of treatment, ever users (past users) women of HRT at advance ages will probably have the same risk of fracture that never treated women. The same conclusion might apply for SCT(N).

The impact of estrogen treatment and bone densitometry measurements on occurrence of hip fractures among Australian women have been calculated through modelling techniques using 500 person years of treatment as the maximum number of what ‘society’ regards as an acceptable for preventing one hip fracture. The results of the study suggest that the person years of treatment required to prevent one hip fracture varies between more than 500 for the highest risk group (BMD< 2SD) and 6850 for those at lowest risk. In women older than 80, the number needed to be treated are below the cut-off level established. The authors concluded that these number are to large to justify treatment with estrogen therapy in women under 80 years and in women over 80 estrogen replacement therapy might be recommended; however this conclusion have to be tested in a RCT.
It is necessary to take into account that one factor that influences the potential effectiveness of any therapy is the level of compliance with it. As shown in different surveys, long-term compliance with HRT is low, and it has been reported to be around 30%. Short-term compliance (one year after screening) have also been low, around 50% (48.9%) in risk groups, defined as belonging to the lowest premenopausal quarter of BMD at either spine or neck of the femur. However, other study shows a higher level of short-term compliance being of 66-93% and 49-73% at the first and at the second year depending on the HRT pattern, sequential or continuous respectively. The low compliance with HRT has been associated to the presence of side effects (eg. breast tenderness, bleeding, depression...), the fear of cancer, dislike of taking tablets, and that women spontaneously withdraw the treatment when climacteric symptoms disappeared.

The second question refers to the relative importance of low bone mass, as a risk marker, when compared with other risk factors associated with hip fracture. Several studies reported other important risk factors that either contribute to low bone mass or determine the presence of fractures.

In a recent prospective cohort study of 9,516 white women aged 65 years or older, there have been identified 16 independent risk factors for hip fractures besides bone density. The highest risk was associated to a history of maternal hip fracture (RR=1.8; 95%CI:1.2-2.7), followed by other risk factors such as previous hyperthyroidism (RR=1.7;95%CI:1.2-2.5), on feet < 4h/D (RR=1.7; 95%CI:1.2-2.4), inability to raise from chair (RR=1.7; 95%CI:1.1-2.7), among others. Age was also an independent risk factor, that is, for each five years of increase in age the relative risk increases 1.4 (95%CI:1.2-1.6). Decrease of bone mass in one standard deviation was associated to an increase in risk similar to the one predisposed by other risk factors (RR=1.6; 95%CI:1.3-1.9).

The number of risk factors also influences in the subsequent risk of hip fracture. The incidence of hip fracture ranged from 1,1 (95%CI:0.5-1.6) per 1000 woman-years among women with no more than two risk factors and normal calcaneal bone density for their age to 27, (95%CI:20-34) per 1000 woman-year among those with five or more risk factors and bone density in the lowest third for their age. Additionally, the results from the population model mentioned above suggest that the number needed to be treated in women with five or more risk factors (including age over 80 years) would be in a societal point of view reasonable.

These results reinforce the fact that with advanced age and progressive bone loss, the relative importance of other skeletal and extraskeletal risk factors increase, as mentioned above. Thus, it could be postulated that modifying several risk factors (besides low bone mass), might reduce the incidence of hip fractures in elderly. Cummings et al reported that among the steps that may contribute to decrease the risk of hip fractures are: maintaining body weight, walking for exercise, avoiding long-acting benzodiazepines, minimizing caffeine intake, and treating impaired visual function. A study that analyzed the impact of a multiple-risk-factor intervention strategy resulted in a significant reduction of the risk of falling among the elderly people in the community.

After this discussion it can be stated that antiresorptive agents influences in only one of the several risk factors predisposing to future fractures. However, when considering the value of a specific therapy it is also necessary to analyze the relation between its risks and benefits. HRT is effective for the relief of menopausal symptoms and the reversal of atrophic urogenital changes. Case-control and cohort studies generally support the concept of a cardioprotective effect of estrogen therapy. A quantitative assessment of the epidemiologic evidence indicated that estrogen therapy is associated with a 40-50% reduction in the risk of coronary heart disease among postmenopausal women. It has been pointed out that women at lower risk have more tendency to use estrogen supplements, thus the cardiovascular protective effect expressed in these studies could have been overestimated. The hypothesis that the addition of progestagens could reverse the favourable
Effects of estrogens in cardiovascular diseases have been rejected recently. Unopposed estrogens lead to an increase in endometrial cancer; however the addition of cyclic progestogen therapy diminishes this risk. When progesterin is added to estrogens, some unpleasant side effects have been described, though, such as depression and bleeding.

Currently, there is a controversy in the scientific community about the relationship between estrogen and breast cancer. Recent published metanalyses support the association between HRT use and breast cancer risk. The trend of an increased risk associated with an increased duration of use appears repeatedly and its rate ranges between 63% (RR=1.35; 95%CI:0.89-2.04) when used more than 12 years to a 30% increase after 15 years of estrogen use (RR=1.3; 95%CI:1.2-1.6). However, another meta-analysis concluded that the treatment of postmenopausal women with standard doses of estrogens did not increase the risk of breast cancer. A recent published prospective cohort study (725,550 person-years of follow-up) has identified a high risk of invasive cancer among POM who were currently taking HRT (RR=1.32; 95%CI:1.14-1.54 for ERT; RR=1.41; 95%CI:1.15-1.74 ERT+progestagens) as compared with POM who never used hormones. This risk increases with five or more years of POM hormone therapy mainly in older women (RR=1.71;95%CI:1.34-2.18). The relative risk of death due to breast cancer in that cohort was 1.45 (95%CI:1.01-2.09). The parallelism between the risk incidence of breast cancers and death suggest that these breast cancer was clinically important. Other side effects associated with HRT described have been reactivation of endometriosis and gallbladder disease.

It is not clear that benefits outweigh risk in POM women, particularly women with few risk factors for cardiovascular diseases. Furthermore, as seen above ERT for up to seven years in the decade after menopause and what’s more after two years cessation of therapy can not be expected to protect against osteoporotic fractures years later in life.

Calcitonin has a good short-term safety profile. Calcitonin have showed its efficacy in decreasing the pain associated with bone disorders. This effect is higher with nasal SCT than with the parenteral administration. These finding could suggest a direct central effect of CT nasal spray, as nasal mucosa is close to hypothalamic pain-regulation area. Although the evidence available suggest a significant analgesic effect with CT, studies supporting its use in osteoporosis only for pain relief have not been identified. Also, there are no scientific evidence comparing the efficacy of CT with other major and minor current analgesics.

There are some limitations associated to both the overall report analysis and the design of the studies reviewed for the report. First of all, this document deals mainly with intranasal salmon calcitonin [SCT (N)] due to the increased use of that type of calcitonin in the Mediterranean countries in the last five years. It is necessary to point out that the studies found in the period of time reviewed were mainly analyzing this type of calcitonin. A possible reason that might explain this situation could be the recent market introduction of the intranasal presentation due to the fact that intramuscular salmon calcitonin was associated to side effects and resistance, stimulating the interest of the pharmaceutical industry to look for alternative delivery routes. For hormone replacement therapy, an exhaustive review of the adverse effects associated to this type of therapy was not performed because the aim of the overall INAHTA project was to assess the effect of HRT and CT in bone mass density and in the prevention of future fractures.

Although there are RCTs analyzing the efficacy of HRT and SCT(N) on bone mass, some limitations affect the comparability among studies. As regards HRT studies, most of them and the rest of studies reviewed, were heterogeneous in several factors such as sample size (most of them included a low number of participants), type, doses, and route of administration, period of follow-up, time since menopause at the beginning of treatment, and baseline status of the participants (healthy or...
osteooporotic). Regarding SCT(N), the studies have small sample sizes, there is a lack of explicit reference to the baseline prognostic factors of the patients enrolled in the RCTs, and there is an heterogeneity of the doses an therapeutic approaches (continuous, cyclic) used in the different studies analyzed. Moreover, for both types of drugs most of the studies have a follow-up less than two years. The Food and Drug Administration\textsuperscript{178} and the European Foundation for Osteoporosis and Bone Disease\textsuperscript{179} recommended a minimum of 2-3 years of follow-up in order to be able to measure the real effect of any therapy on bone mass. This requirement is based in the fact that the therapeutic intervention induce transient changes in skeletal mass until a new steady state of rate changes is achieved (after 2 to 3 years of therapy). Therefore, the conclusions extracted from the results of all these studies should be taken into account cautiously.

The scientific evidence addressing the relationship between HRT and hip fracture events came from prospective cohort and case-control studies. It is difficult to distinguish between association and causation with these type of studies due to the confounding factors and source of biases associated to these type of design. Moreover, as stated by Eisman et al.\textsuperscript{180}, most studies about the efficacy of the treatment on osteoporotic fractures have been performed in selected groups of patients within the community (f.ex volunteers asking for HRT), and the extrapolation to wider clinical subsets of the population is not straightforward. Also, it is not clear whether or not studies on carefully selected healthy elderly women can be simply extrapolated to the less healthy cohorts. Besides, for vertebral fractures, most of the studies have been carried out in osteoporotic women (presence of fractures), hence, the results from these studies should be extrapolated to healthy cohorts with caution. However, there are ongoing studies which are expected to provide an answer upon high quality evidence for the impact of HRT upon fractures, such as the Women's Health Initiative and the ongoing screening studies in UK and Sweden.

Finally, it is necessary to take into account that menopause is a physiological condition associated to the ageing process in women that, if necessary, requires a multiple health and social intervention approach. Bone loss is also a physiological consequence of the ageing process. Ethical issues might appear if medicalization of all postmenopausal women with therapies of doubtful effectiveness are carried out in order to prevent the adverse consequences in a specific system, i.e. bone. Assuring the appropriate use of a therapy could overcome at some extent the ethical dilemmas associated with its use. Hence, the prescription of any type of drug therapy should be based on weighting the overall benefits and risk for each woman and comparing them to the risk-benefit of alternative available therapies. It is necessary to point out that some of the health approaches to prevent the consequences of estrogen deficit on bone mass have virtually no risk and have been proven effective, such as physical weight-bearing exercise.\textsuperscript{181} Thus, a multiple-risk intervention strategy probably will give the best benefits at a reasonable health and social costs.

Conclusion

Based on the available evidence the following conclusions can be suggested according the classification of the level of quality of the evidence used at CAHTA.\textsuperscript{23} Tables 10 and 11 summarise the results of the quality of the evidence ratings.

a) Hormone Replacement Therapy and preservation of bone mass

Primary prevention: There is good evidence to support that HRT preserves bone mass (Meta analyses of RCT=1; RCT=25)\textsuperscript{19,38,41,42,44,46,48,49,50,51,52,53,55,87,58,82,41,67,68,70,69,184,196,197,198}

Secondary prevention: there is good evidence to support that HRT preserves bone mass (Meta-
analysis of RCT=1; RCT=5)\textsuperscript{19,72,148,74,191}

Starting early after menopause (time since last menopausal period < 10 years): there is good evidence to support that HRT preserves bone mass soon after menopause (RCT=29)\textsuperscript{38,41,42,44,46,48,49,50,53,55,52,58,82,41,67,68,85,51,52,67,69,106,73,196,197,198,184}

Starting within three years after menopause plus long-term follow-up (10 years): there is fair evidence regarding the positive effect of HRT in bone mass (RCT=1)\textsuperscript{90}

Starting over three years after menopause plus long-term follow-up (10 years): there is good evidence regarding the lack in preserving bone mass (RCT=1)\textsuperscript{90}

Starting late after menopause (> 10 years): There is good evidence to support that HRT has a small effect in bone mass preservation even if its start is delayed late after menopause, generally for secondary prevention. When calcium is added the effect is stronger (Meta-analysis of RCT=1, RCT=1, CT=3).\textsuperscript{19,192,193,195,199}

Ever use in women 75 years old or younger: There is fair evidence to support a protective effect of HRT in bone mass loss (case-control study nested in a cohort=1)\textsuperscript{83}

Ever use in women over 75 years of age: There is fair evidence suggesting a small effect of HRT in radius shaft bone mass bone mass preservation after 75 years (Case-control in a cohort=1)\textsuperscript{83}

Cessation of therapy: There is fair evidence to support that withdrawal of therapy induces rates of bone loss similar to those in untreated or placebo treated women. (RCT=1, case-series=2, cohort=1)\textsuperscript{54,96,97,192}

b) Nasal salmon calcitonin and preservation of bone mass

Primary prevention: There is good evidence to support that SCT (N) preserves bone mass (RCT=1),\textsuperscript{187} however, there is fair evidence for the recommendation regarding the exclusion of this technology (RCT=8).\textsuperscript{122,123,124,125,126,127,128} To ascertain the overall pull-effect it would be necessary to carry out a meta-analysis.

Secondary prevention: there is good evidence to support that SCT(N) preserves bone mass seen mainly at spine (RCT=8)\textsuperscript{133,134,136,137,143,128,26,188}

Started early after menopause (time since last menopausal period < 10 years): there is some good evidence to support that SCT(N) preserves bone mass soon after menopause (RCT=4)\textsuperscript{128,137,129,187}, however, there is more fair evidence for the recommendation regarding the exclusion of this technology (RCT= 6)\textsuperscript{123,125,127,131,130,142} To ascertain the overall pulled effect it would be necessary to carry out a meta-analysis.

Long-term bone mass preservation when started early after menopause: There are no data available about the effect of SCT(N) at elderly ages when started early after menopause.

Started long after menopause (> 10 years): There is good evidence to support that SCT(N) preserve bone mass in osteoporotic women (RCT=1).\textsuperscript{136}

c) Hormone replacement therapy and risk for fractures

Risk for overall types of fractures: there is fair evidence to support a trend to decrease in risk for overall types of fractures when HRT is used (meta-analysis of cohort and case-control studies=1, case-control=1)\textsuperscript{19,185}

Risk for hip fracture: there is fair evidence to support a trend to decrease in hip fracture risk if ever used of HRT (meta-analysis of cohort and case-control studies=2)\textsuperscript{19,98}

Risk for wrist and forearm fractures: there is fair evidence to support a trend to decrease in wrist
and forearm fracture risk (meta-analysis of cohort and case-control studies=1)\textsuperscript{19}

Risk for vertebral fractures: There is fair evidence to support the decrease in vertebral fracture risk in primary and secondary prevention (RCT=1, prospective cohort=2)\textsuperscript{80,119,120}

Risk of hip fracture at elderly ages if past used of HRT exist: there is fair evidence supporting the lack of decreasing risk for hip fractures at elderly ages (>60 years) when ever use of HRT (prospective cohort=4; case-control=1)\textsuperscript{101,121,100,117,113}

Risk of hip fractures at elderly ages (>60) when current use exist: there is fair evidence suggesting the decrease in hip fracture risk for current use for more than 10 years (Prospective cohort=1)\textsuperscript{121}

Risk fracture according cessation of therapy: There is fair evidence supporting that the decrease in hip fracture risk disappears after cessation of therapy (prospective cohort=2; case-control=1)\textsuperscript{101,121,118}

d) SCT (N) and fracture risk

Vertebral fracture risk: There is fair evidence to support that SCT(N) decrease the risk for new vertebral fractures in secondary prevention (RCT=1)\textsuperscript{136}

Hip fracture risk: there is no evidence to support that SCT(N) decrease the risk for hip fracture. There is fair evidence that calcitonin decrease the risk for hip fracture (case-control=2).\textsuperscript{113,147}

Addendum:

During the phase of external peer-review of this background document, a meta-analysis of calcitonin in the treatment of post-menopausal osteoporosis have been performed.\textsuperscript{24} The interventions included were intranasal or intramuscular calcitonin at variable doses (50-400 IU) plus calcium versus placebo. The conclusion reached do not differ from the ones expressed in the current background document, i.e. calcitonin is efficacious in improving BMD, specially vertebral and femoral neck, and antifracture efficacy still remains to be established.
<table>
<thead>
<tr>
<th>Design</th>
<th>Hormone replacement therapy</th>
<th>Nasal calcitonin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meta-analysis</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>RCT</td>
<td>60</td>
<td>24’</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>Observation+RCT</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>RCT+case-series</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Cohort</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>case-control</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>cross-sectional</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>case-series</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>79</td>
<td>28</td>
</tr>
</tbody>
</table>

* 2/24 studies do not use SCT (1 hCT (N)\textsuperscript{21} and 1 CTc (N)\textsuperscript{22})
### TABLE 2. Evidence of HRT/CT (N) effect on fracture

<table>
<thead>
<tr>
<th>Design</th>
<th>Hormone replacement therapy</th>
<th>Nasal calcitonin</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCT</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>RCT + case-series</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Cohort</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>case-control</td>
<td>13</td>
<td>2*</td>
</tr>
</tbody>
</table>

* The studies do not specify the type of calcitonin used.

### TABLE 3. Effect size of different hormone regimens according site of measurement in the primary prevention of bone loss

<table>
<thead>
<tr>
<th></th>
<th>Forearm</th>
<th>Spine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral unopposed</td>
<td>1.43 SD (1.12-1.78)</td>
<td></td>
</tr>
<tr>
<td>Oral unopposed + Ca</td>
<td>0.92 SD (0.33-1.51)</td>
<td>0.77 SD (0.33-1.51)*</td>
</tr>
<tr>
<td>Oral opposed</td>
<td>1.33 SD (0.74-1.92)**</td>
<td>1.23 SD (0.75-1.72)***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.18 SD (0.65-1.70)****</td>
</tr>
<tr>
<td>Oral opposed + Ca</td>
<td>1.52 SD (1.22-1.81)</td>
<td>0.94 SD (0.18-1.69)**</td>
</tr>
<tr>
<td></td>
<td>1.46 SD (0.95-1.97)+</td>
<td></td>
</tr>
<tr>
<td>Transdermal opposed</td>
<td>1.10 SD (-0.09 to 2.29)*</td>
<td>1.02 SD (0.67-1.36)</td>
</tr>
<tr>
<td>Transdermal opposed + Ca</td>
<td>1.70 SD (1.09-2.31)</td>
<td>1.37 SD (0.45-2.30)</td>
</tr>
</tbody>
</table>

Source: Henry D et al.19

* Standard dose (conjugated estrogens=0.6-0.625 mg; estrone=1.25 mg; estradiol=2 mg)
** Low dose (conjugated estrogen=0.30mg; estrone=0,625mg;estradiol=1)
+ Heterogeneity: p=0.05
** Heterogeneity: p= 0.00003
*** Heterogeneity: p= 0.0000005
### TABLE 4. Effect of age and HRT in bone mass response (range of effect sizes identified from different articles)\textsuperscript{84,85.}

<table>
<thead>
<tr>
<th></th>
<th>Spine (SD,range)</th>
<th>Forearm (SD,range)</th>
<th>Hip (SD,range)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary prevention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60 years old</td>
<td>0.48-2.51</td>
<td>0.50-2.38</td>
<td></td>
</tr>
<tr>
<td>=70 years old (1 study)</td>
<td></td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td><strong>Secondary prevention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60 years old</td>
<td>0.73-3.84</td>
<td>0.97</td>
<td>0.32-1.68</td>
</tr>
<tr>
<td>&gt;60 years old (4 studies)</td>
<td>0.095-3.84</td>
<td>1.00-4.94</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Source: Henry D et al\textsuperscript{94}

### TABLE 5. Effect size of treatment in secondary prevention.

<table>
<thead>
<tr>
<th></th>
<th>Forearm</th>
<th>Spine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unopposed oral</td>
<td></td>
<td>0.95 (0.21-1.70)</td>
</tr>
<tr>
<td>Unopposed oral+ Ca</td>
<td></td>
<td>2.27 (1.17-3.37)</td>
</tr>
<tr>
<td>Opposed oral+Ca</td>
<td>2.94 (-0.92-6.80)</td>
<td>3.16 (1.83-4.49)</td>
</tr>
<tr>
<td>Opposed transdermal+Ca</td>
<td>0.97 (0.46-1.48)</td>
<td>0.73 (0.24-1.21)</td>
</tr>
</tbody>
</table>

Source: Henry et al\textsuperscript{19}

### TABLE 6. Effect of HRT at different measurement sites

<table>
<thead>
<tr>
<th></th>
<th>No. Studies</th>
<th>Effect Size (95%CI)</th>
<th>Heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Prevention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine</td>
<td>8</td>
<td>1.17 (0.63-1.70)</td>
<td>0.0000005</td>
</tr>
<tr>
<td>Forearm</td>
<td></td>
<td>1.38 (0.93-1.84)</td>
<td>0.0008</td>
</tr>
<tr>
<td><strong>Secondary Prevention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine</td>
<td>2</td>
<td>2.23 (-0.81-5.28)</td>
<td>0.000002</td>
</tr>
<tr>
<td>Forearm</td>
<td></td>
<td>2.90 (-0.99-6.79)</td>
<td>0.000002</td>
</tr>
<tr>
<td>Spine</td>
<td>4</td>
<td>2.12 (0.89-3.34)</td>
<td>0.00002</td>
</tr>
<tr>
<td>Hip</td>
<td></td>
<td>0.92 (0.34-1.50)</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Source: Henry et al\textsuperscript{19}
<table>
<thead>
<tr>
<th>Design</th>
<th>On therapy</th>
<th>Cessation of therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lindsay et al.</strong>&lt;sup&gt;97&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case-series</td>
<td>PL&lt;sub&gt;4 year&lt;/sub&gt;= 9.26%</td>
<td>ERT&lt;sub&gt;4 year&lt;/sub&gt;= 9.25% (p&lt;0.01)</td>
</tr>
<tr>
<td>N=43</td>
<td>ERT&lt;sub&gt;4 and 8 year&lt;/sub&gt;= no changes (p&lt;0.001 vs placebo)</td>
<td>ERT&lt;sub&gt;8 year&lt;/sub&gt;= no evidence of bone loss (current treatment when analyzed)</td>
</tr>
<tr>
<td>POM oophorectomized (&lt;47 years)</td>
<td>PL=second 4 year period= 9.75% (p&lt;0.025 vs ERT&lt;sub&gt;4 year&lt;/sub&gt;)</td>
<td></td>
</tr>
<tr>
<td>PL&lt;sub&gt;4 year&lt;/sub&gt;= 14</td>
<td>ERT&lt;sub&gt;4 year&lt;/sub&gt;= 14</td>
<td>ERT&lt;sub&gt;8 year&lt;/sub&gt;= 15</td>
</tr>
<tr>
<td>ERT start within 3 years of menopause</td>
<td>Metacarpal BMC</td>
<td></td>
</tr>
<tr>
<td><strong>Christiansen C et al.</strong>&lt;sup&gt;54&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td>HRT = 83.7% after three years of treatment (p&lt;0.001)</td>
<td>HRT = 9.23% annual (p&lt;0.01)</td>
</tr>
<tr>
<td>N=94</td>
<td>PL = 91.9% per year (p&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>HRT=43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo=51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POM healthy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age:44-54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tm=6m-3y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forearm BMC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Quigley ME et al.</strong>&lt;sup&gt;96&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case-series</td>
<td>Percentage decrease bone loss/year by age group</td>
<td>previous users who discontinued therapy after 65 y= 9.26% /year (p&lt;0.05)</td>
</tr>
<tr>
<td>N=397</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age:51-80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous or non previous estrogens users who discontinued or not ERT after 65 years of age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Users:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-60y=90.5% (p&lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>61-70y=90.7% (p&lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71-80=90.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never users:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-60y=92%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>61-70=92.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71-80y=91%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 65 years:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- started therapy at 65 year= 90.4% (p&lt;0.01 vs never users)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Never users: 91.4% (p&lt;0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Always users: 90.6% (p&lt;0.05)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 7. Effect of HRT on bone mass after cessation of therapy

<table>
<thead>
<tr>
<th>Design</th>
<th>Treatment</th>
<th>Cessation of therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis JW et al. (Bone) (1995) (ref.192)</td>
<td>Retrospective cohort Estrogens + calcium (75% of women took 0.65 mg conjugated estrogens) Mean yr use/subject: Estrogen use singly: 2.7yr Estrogen+Ca:3.3</td>
<td>No treatment= decrease 1% Estrogen= 0.75-0.85% per year decrease bone loss Cessation estrogen use= 0.35-0.60 increase in bone loss (greater than average loss rates)</td>
</tr>
</tbody>
</table>

TABLE 8. Relative risk of hip fracture according to age

<table>
<thead>
<tr>
<th>Study</th>
<th>Relative risk (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paganini-Hill (1991)</td>
<td>Average age 73 y: RR= 1.02 (0.81-1.27)</td>
</tr>
<tr>
<td>Kiel et al. (1987)</td>
<td>Age &lt; 75 y (65-74): RR= 0.37 (0.05-2.46)† Age &gt; 75 y: RR= 0.82 (0.21-3.24)‡</td>
</tr>
<tr>
<td>Kanis et al. (1992)</td>
<td>Age &lt; 80 y: RR= 0.51 (0.31-0.84)‡ Age &gt; 80 y: RR= 0.70 (0.29-1.66)‡</td>
</tr>
<tr>
<td>Cauley JA et al. (1995)</td>
<td>Age &lt; 75 y (current users): RR= 0.94 (0.52-1.69)³ Age &gt; 75 y (current users): RR= 0.18 (0.04-0.77)³</td>
</tr>
<tr>
<td>NaessJn (1990)</td>
<td>Age &lt;60 y: RR_{rocheul}=0.37 (0.13-0.79) RR_{cervical hip}=0.58 (0.41-0.80) Age &gt; 60 y: RR_{rocheul}=1.03 (0.74-1.40) RR_{cervical hip}=0.95 (0.75-1.18)</td>
</tr>
</tbody>
</table>

Observation: the non statistically significative results could be due to a small ample size or to a cohort-effect.
† recent estrogen use (# 2 years)
‡ adjusted by age, previous fractures, BMI
³ multivariate-adjusted relative risk

TABLE 9. Risk of hip fracture according to past or recent use (95%CI)

<table>
<thead>
<tr>
<th>Study</th>
<th>past use</th>
<th>recent/current use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kiel et al. (1987)</td>
<td>RR_{(&gt;2 y)}= 0.74 (0.49-1.14)</td>
<td>RR_{(&lt;2 y)}= 0.34 (0.12-0.98)</td>
</tr>
<tr>
<td>(Age:75+/-.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paganini-Hill (1991)</td>
<td>RR_{(&gt;14 y)}= 0.88 (0.63-1.23) RR_{(&gt;15 y)}= 1.15 (0.88-1.50)</td>
<td>RR_{(&gt;2 y)}= 0.80 (0.53-1.21)</td>
</tr>
<tr>
<td>(Average age:73 y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiel et al. (1992)</td>
<td>RR_{(&lt; equal 2 y)}= 0.87 (0.56-1.34)</td>
<td>RR_{(&lt;2 y)}= 0.39 (0.13-1.13)</td>
</tr>
<tr>
<td>(Age:55-76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cauley et al. (1995)</td>
<td>RR_{(&gt;10 y)}= 1.67 (0.92-3.01) RR_{(&gt;10 y)}= 0.97 (0.65-1.46)</td>
<td>RR_{(&gt;10 y)}= 0.27 (0.08-0.85) RR_{(&lt;10 y)}= 0.81 (0.40-1.65)</td>
</tr>
<tr>
<td>(Age &gt; 65 y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>Type of design</td>
<td>Conditions of scientific rigour</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>I</td>
<td>Meta-analysis of randomized controlled trials</td>
<td>Analysis of patient individual data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Meta-regresion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Different techniques of analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absence of heterogeneity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the studies</td>
</tr>
<tr>
<td>II</td>
<td>Large sample randomized controlled trials</td>
<td>Assessment of statistical power</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multicenter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>III</td>
<td>Small sample randomized controlled trials</td>
<td>Assessment of statistical power</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>IV</td>
<td>Non-randomized controlled prospective trials</td>
<td>Concurrent controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multicenter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>V</td>
<td>Non-randomized controlled retrospective trials</td>
<td>Historical controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>VI</td>
<td>Cohort studies</td>
<td>Multicenter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Matching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>VII</td>
<td>Case-control studies</td>
<td>Multicenter studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>VIII</td>
<td>Non-controlled clinical series Descriptive studies: surveillance of disease, surveys, registers, data bases, prevalence studies</td>
<td>Multicenter</td>
</tr>
<tr>
<td></td>
<td>Experts' committees, conferences of opinion</td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td>Anecdotes or cases reports</td>
<td></td>
</tr>
</tbody>
</table>

* Complemented by the characteristics described in the evidence table. From higher (I) to lower (X) scientific rigour. Quality of the study assessed by specific protocols. From: Jovell AJ, Navarro-Rubio MD. La evaluaci\n de la evidencia cientific \na (The evaluation of scientific evidence) Med Clin (Barcelona) (In press).23
### TABLE 11. Appropriateness of the recommendations according to the quality of scientific evidence

<table>
<thead>
<tr>
<th>Strength of the recommendations</th>
<th>Quality levels</th>
</tr>
</thead>
</table>
| **A:** There is GOOD evidence to support the recommendation of the adoption of the technology | I  III<sup>a</sup>  
II  IV<sup>a</sup>  |
| **B:** There is FAIR evidence to support the recommendation of the adoption of the technology | III<sup>b</sup> V  
IV<sup>b</sup> VI  |
| **C:** There is POOR evidence, therefore the decision of adopting the technology should be based in other criteria | VIII  IX  |
| **D:** There is FAIR evidence to support the recommendation that the condition be excluded | III<sup>c</sup> V  
IV<sup>c</sup> VI  |
| **E:** There is GOOD evidence to support the recommendation that the condition be excluded | II  IV<sup>d</sup>  
I  III<sup>d</sup>  |

* Depending of the conditions of scientific rigor

- Good quality + statistical significance
- Good quality + non statistical significance
- Poor quality + statistical significance
- Poor quality + non statistical significance
References


12. *Divisi\ d’Atanci\ FarmacPutica i Prestacions ComplementBries*. Catalan Health Service (personal communication)

13. La calcitonina, un producto de dudosa efectividad. Osteba Bulletin 1994; n° 1


56. Stevenson JC, Crook D, Godsland IA, Lees B, Whitehead MI. Oral versus transdermal
hormone replacement therapy. *Int J Fertil* 1993; 38 Suppl 1:S30-S35.


125. Meunier PJ, Gozzo I, Chaumet-Riffaud Ph.D, Delmas PD, Guignard M, Chapuy MC,


INAHTA Joint Project

Bone Density Measurement and Treatments of Osteoporosis

Statement of findings

David Hailey, Laura Sampietro-Colom, Deborah Marshall, Rosa Rico, Alicia Granados, José Asua, Trevor Sheldon

September 1996
Acknowledgements

The authors and agencies involved in the preparation of this manuscript gratefully acknowledge and thank the following individuals, who reviewed this document in draft form and made many helpful comments and suggestions:

Ann Cranney, Beverly Shea, Peter Tugwell, University of Ottawa, Canada; Pierre Durieux, Assistance Publique Hôpitaux de Paris, France; David Feeny, McMaster University, Canada; Frederic Fleurette, Agence Nationale pour le Développement de l'Évaluation Médicale, France; Hellen Gelband, Robert McDonough, Office of Technology Assessment, USA; Line Gariépy, Guy Régnier, Conseil d'évaluation des technologies de la santé du Québec, Canada; Tom Holohan, Office of Health Technology Assessment, USA; Olof Johnell, Lund University, Sweden; Lawrence Joseph, McGill University, Canada; David Henry, The University of Newcastle, Australia; Albert Jovell, Catalan Agency for Health Technology Assessment, Spain; Sverker Ljunghall, Uppsala University, Sweden.

The authors are especially grateful for the assistance of Dr. David Henry and his colleagues at the University of Newcastle, Australia, who shared their unpublished work on HRT. Dr. Henry is also coordinating editor of the Cochrane Review Group that maintains systematic reviews on the prevention and treatment of osteoporosis.

They also acknowledge the valuable assistance of the following persons who participated in meetings to develop the draft statement and helped to develop the background documents:

David Banta, TNO Prevention and Health, The Netherlands; Renaldo Battista and Jean-Marie Lance, Conseil d’évaluation des technologies de la santé du Québec, Canada; Frederic Fleurette, Agence Nationale pour le Développement de l'Évaluation Médicale, France; Albert Jovell, Catalan Agency for Health Technology Assessment, Spain; Pedro Koch, Ufficio federale delle assicurazioni sociali, Switzerland; Robert McDonough, Office of Technology Assessment, USA; Devidas Menon, Canadian Coordinating Office for Health Technology Assessment; Penny Rogers, Australian Health Technology Advisory Committee; Robert Segaar, Health Council of the Netherlands; Lars Werkö, The Swedish Council on Technology Assessment in Health Care.
This document is endorsed by the following organisations which are members of the International Network of Agencies for Health Technology Assessment (INAHTA):

Agencia de Evaluación de Tecnologías Sanitarias, Madrid (AETS)
Alberta Heritage Foundation for Medical Research, Edmonton (AHFMR)
Basque Office for Health Technology Assessment, Vitoria – Gasteiz (Osteba)
Catalan Agency for Health Technology Assessment, Barcelona (CAHTA)
Center for Health Care Technology, Agency for Health Care Policy and Research, Rockville (CHCT)
Danish Institute for Health Services Research and Development, Copenhagen (DSI)
NHS Centre for Reviews and Dissemination, York (UK NHS CRD)
Finnish Office for Health Care Technology Assessment, Helsinki (FinOHTA)
Swedish Council on Technology Assessment in Health Care, Stockholm (SBU)
TNO Prevention and Health, Leiden (TNO)
Contents

Summary ............................................................................................................................. 5
Introduction ......................................................................................................................... 6
Scope ................................................................................................................................. 6
Definitions and endpoints ............................................................................................... 6
Methods ............................................................................................................................. 6
Results ................................................................................................................................. 6
  1 Methods for bone density measurement................................................................. 6
  2 The use of bone density measurement to predict fractures in individuals.............. 7
  3 Effect of HRT in preventing fractures and preserving bone mass............................ 8
  4 Effect of intranasal salmon calcitonin in preventing fractures and preserving bone mass............................................................... 6
Discussion .......................................................................................................................... 5
  Table 1: Level of scientific evidence.......................................................................... 6
  Table 2: Characteristics of common methods for measuring bone density.................. 7
  Table 3: Relative risk of hip fracture according to age............................................... 8
  Table 4: Results of a meta-analysis of the effect of HRT on bone mass....................... 6
  Table 5: Potential impact of BDM screening and treatment with HRT in preventing hip fractures in a population of 20,000 menopausal women............................................. 7
References ........................................................................................................................ 6
Summary

Objective: To provide a summary of the available scientific evidence regarding the performance of bone density measurement (BDM) techniques and the effectiveness of BDM screening and related interventions (hormone replacement therapy (HRT) and intranasal salmon calcitonin SCT(N)) in menopausal women to prevent fractures in later life.

Methods: Synthesis of systematic reviews of evidence on BDM, HRT and SCT(N) undertaken previously by health technology assessment agencies, and other relevant systematic reviews. These sources were updated by adding primary studies identified through additional literature searches. The evidence was evaluated using an internationally accepted classification system incorporating study design and quality.

Main Findings: The analytical performance of BDM technologies in the routine clinical situation has not been adequately assessed. Fair evidence from prospective cohort studies suggests that BDM can predict the risk of fractures, but not with high accuracy. Although good evidence exists to support the efficacy of HRT and SCT(N) in preserving bone mass during treatment, there is also fair evidence that the effect wears off after cessation of therapy. Fair evidence, from low quality RCTs and observational studies, suggests that these therapies are efficacious in preventing fractures. However, when this evidence is used to evaluate the potential effectiveness of BDM screening of menopausal women in combination with these therapies it is estimated, using optimistic assumptions, that only 1-7% of hip fractures might be prevented.

Conclusion: The currently available evidence does not support the use of BDM screening in combination with HRT or SCT(N) treatment.
Introduction

Bone fractures represent a serious health problem for older women. Osteoporosis - the natural loss of bone that occurs with age and especially within 3-6 years after the menopause - predisposes women to bone fractures that occur most commonly at the hip, wrist and spine. Hip fractures are of particular concern because of their high costs in terms of morbidity, mortality and their economic and social burden. They are estimated to increase worldwide from 1.66 million annually in 1990 to over 6 million by the year 2050.

Consequently, there is growing international interest in approaches to identify individuals at high risk for fractures and in interventions that might help to prevent these events. Various techniques to measure bone density have been developed for detecting those at high risk of having a fracture. These individuals are commonly prescribed treatments in the form of hormone replacement therapy (HRT), and in some Mediterranean countries, calcitonin (CT).

Scope

The objective of this review is to assess the available scientific evidence regarding the performance of bone density measurement (BDM) and its effectiveness in preventing fractures when used in conjunction with prophylactic treatments (HRT and intranasal salmon calcitonin) in menopausal women. The review has been produced through a collaboration between publicly-funded national and regional agencies from several countries, which undertake health technology assessment, to consolidate and critically evaluate the evidence on these topics. It is intended as a resource to those responsible for funding and using these technologies, which have been selected for review because:

- they are of interest to a number of national and regional health care systems and have been the subject of independent reports published by health technology assessment agencies in several countries;
- programmes to measure bone density in menopausal women to identify those at risk of fracture and who may benefit from treatment would have a considerable impact on health care systems if this were adopted as a widespread practice;
- HRT is prescribed widely in many countries and is the subject of current controversy regarding potential benefits in preventing fractures;
- calcitonin, especially in the form of intranasal salmon calcitonin (SCT(N)), is of particular interest in Mediterranean countries, where its use has increased substantially and the associated costs and effectiveness continue to be debated.

For the purposes of this analysis, it is assumed that the effectiveness of different types of HRT treatment is the same. This paper does not assess other approaches to identify individuals at high risk for fracture nor any alternative interventions such as exercise, hip protector pads and vitamin D and bisphosphonates. These topics will require separate review. The impact of BDM and associated treatments is considered only in terms of their effects on the risk of fractures in women. Effects of HRT on other clinical conditions such as cardiovascular disease and breast cancer are not addressed, nor are resource implications discussed. Such important factors require consideration when addressing the clinical application of these technologies or formulating local policies on their use.

Definitions and endpoints

The diagnosis of osteoporosis is histologic. In practice, osteoporosis is commonly defined as a condition characterised by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk. More recently, it has been suggested in a report by a World Health Organization expert group that osteoporosis in an individual be defined solely on the basis of the level of bone density (as < 2.5 standard deviations (SD) below the young adult mean). The choice of this definition has profound effects on the proportion of women
in the population who are considered to be osteoporotic and thus potentially eligible for intervention.

Primary prevention with HRT or calcitonin is defined for the purposes of this paper as intervention in women with natural or surgical menopause with normal skeletal status and without a history of fracture. Secondary prevention is defined as intervention in women with natural or surgical menopause with one or more non-traumatic fractures or whose bone density is more than one or two standard deviations from young adult values or from age-matched controls, depending on the study.

Efficacy refers to the performance of a health technology under ideal clinical conditions of use in a defined population and effectiveness to its performance under average or routine conditions of use. The main outcome measures used in this paper are bone density and fractures. Evaluations is in terms of the relative risk of fracture for a 1 SD decrease in bone mineral density below the age adjusted mean; the percentage of fractures potentially prevented by BDM screening linked to treatments; and the number of individuals who would need to be invited for screening in order to prevent one hip fracture.

Treatment effects of HRT and SCT(N) for preserving bone density are expressed as the number of SD units by which the average annual decline in bone mass in the control group exceeds that in the treatment group and as relative risks or odds ratios for fractures. An a level of 0.05 was used as the cut-off for statistical significance. The effects of SCT(N) for preserving bone density are expressed as the difference in the percent change between the treatment and control groups.

Methods

This paper is based primarily on systematic reviews of evidence undertaken by health technology assessment agencies and other systematic reviews that were relevant. These sources were updated by adding primary studies identified through additional literature searches.

For studies on bone densitometry, Medline, SweMed and EMBASE from September 1994 to May 1996 were searched. The main subject headings used in the searches were: ‘bone and bones’, ‘bone density’, ‘densitometry’, ‘osteoporosis’ and the names of the BDM technologies. Reference lists of review articles were also checked. Trials of screening programmes, prospective cohort studies examining the predictive value of BDM, case control studies of hip fractures (to September 1994) and studies evaluating or comparing methods of BDM in human subjects in English, French, German or Swedish were included for further review.

For studies on HRT, Medline was searched from 1993 to May 1996 using the main keywords ‘osteoporosis’, ‘postmenopausal’, ‘postmenopause’, ‘bone and bones’, ‘bone density’, ‘fractures’, ‘estrogen replacement therapy’. In the case of calcitonin, both Medline and EMBASE were searched between 1990 up to 1994, and Medline only to May 1996 using the same keywords as for HRT, but with calcitonin as the therapy. Publications reporting randomised controlled trials, cohort studies and case control studies on SCT(N) in English, French, Italian and Spanish were selected for detailed review. Additional studies were obtained by checking reference lists from book chapters and review articles. Only original studies in human subjects were selected.

The conclusions in this paper are based on a classification system that considers the type of study design and conditions of scientific rigour, similar to that used by the Canadian Task Force on the Periodic Health Exam, the U.S. Preventive Services Task Force, and by Sackett (Table 1). The search for relevant research was supplemented by two surveys of organisations that had produced reports addressing the issues of BDM and HRT. One survey included 24 organisations that had published reports on bone density measurement; the other included six health technology assessment agencies that had produced reports on the issue of HRT in primary and secondary prevention of osteoporosis. Both surveys asked questions about the process followed in preparing the report, its main purpose, and the conclusions made about the clinical applications of BDM and HRT.
Further details about the methodology are available in the background documents.

Results

1 Methods for Bone Density Measurement
1.1 Various methods are currently used for measuring bone density, but x-ray based methods (dual x-ray absorptiometry (DXA) in particular) dominate the market. Ultrasound methods are gaining in popularity, but their analytical performance still requires validation.

1.2 The precision and accuracy of BDM technologies are measures of their performance as analytical methods. These measures are generally poorly defined in the routine clinical situation for all available methods.

Table 2 describes the performance characteristics of BDM methods in common use. Good precision means low random errors, so that there is small variation between results of measurements on the same sample. Good accuracy means that the systematic error is low, so that the average of a series of measurements on the same sample is close to the true value. Both precision and accuracy can be expressed as coefficients of variation (CV), and the total error associated with a BDM will be the sum of these errors. Variations in patient positioning, operator performance and machine characteristics will contribute to these errors.

Most available data are from measurements taken in a single centre over a short period by expert operators. Consequently, they tend to reflect the efficacy of the individual methods, and are likely to underestimate the error which will occur in routine clinical practice. Standards to define the levels of analytical performance required for different applications of BDM are essential.

1.3 The performance characteristics of BDM methods impose constraints on specific applications of these techniques.

Even with a precision error as low as 1% SD, serial measurements using BDM would require a minimum follow up interval of 1 to 1.5 years to detect a bone loss of 2-3% (the average loss per year for a normal woman at menopause) or nearly 5 years with a precision error of 5% SD. The interval needed would be much greater if BDM were used to monitor therapy or in other situations where the rate of bone loss is likely to be lower.

The usefulness of a single BDM, from a screening test for example, is affected by both accuracy and precision of the method. Given the performance levels of currently available methods, many individuals will be wrongly classified with regard to their risk of fracture, either as false positives or false negatives. This problem is compounded by systematic differences in ‘normal’ values in population (e.g. ethnic groups) from those used to set reference values for BDM equipment.

1.4 If bone density measurements are to be undertaken, it is essential that scrupulous quality control is followed.

Calibration and standardisation of bone densitometers is a complex undertaking that requires close attention since there is little agreement among manufacturers. Even with instruments calibrated according to manufacturer’s instructions, values obtained from imaging of spine phantoms by DPA or DXA have differed by as much as 16% because of differences as great as 8% in estimated values for both bone mass and bone area. Instrument performance also may vary over time.

Good quality control is essential and should include daily standardisation and calibration procedures, regular maintenance, careful attention to patient positioning and possible accreditation of units. There are currently no industry-wide, nor clinical standards for BDM technologies, although in Europe a spine phantom is now being used to calibrate equipment in various drug trials.
2.0 The use of bone density measurement to predict fractures in individuals

2.1 There are no randomised controlled trials which have evaluated the efficacy of using BDM to screen menopausal women and prevent fractures.

Ideally, there should be data available from controlled trials where menopausal women are randomised to a screening programme or no screening programme and subsequently followed for 20-30 years (at the age when most fractures occur) to determine the effect on the number of individuals in whom fractures are prevented. No data from studies with designs corresponding to Levels I-V (Table 1) are currently available, though a BDM screening programme is being piloted in the United Kingdom. 37

2.2 There is FAIR evidence that BDM can predict the risk of fracture in menopausal women. However, because of the considerable overlap between the distribution of bone mineral density for individuals who have and do not have fractures, BDM cannot reliably distinguish those who will have a fracture from those who will not.

BDM techniques are able to estimate, with various degrees of error, low bone mineral density which is a risk factor for future fracture. However, the main outcome measure of interest is fracture. A recent meta analysis of prospective cohort studies (Level VI), shows that there is an inverse association between bone density and the risk of future fractures.10 38 Eleven study populations, constituting about 90,000 person years of follow up time, were identified. The relative risk for all types of fractures at all sites of a decrease in bone density of 1 SD below the age adjusted mean was 1.5 (95% CI 1.4 to 1.6).

Some sites had a better predictive ability - measurement at the spine for predicting vertebral fractures (RR=2.3, 95% CI 1.9 to 2.8) and measurement at the hip for predicting hip fractures (RR=2.6, 95% CI 2.0 to 3.5). Using the relative risk value of 2.6 for hip fracture, derived from a meta-analysis 38, and assuming a Gaussian distribution of bone density values, the test characteristics can be determined for a theoretical cohort of women at the age of 50 years with a 15% lifetime risk of fracture 39. A cut-off value of 1 SD yields a sensitivity of 38%, a specificity of 88% and a positive predictive value of 36%.

However, most of these studies have a relatively short follow up time (weighted average 5.8 years), and the extent to which the results can be extrapolated to the prediction of fractures in individuals which will occur in 20 to 30 years in the future is not known. The ability of BDM to predict fracture risk will not be the same for all age groups because of the increased importance of other skeletal and extraskeletal risk factors with increasing age. 40 41

Two additional studies in menopausal women, published since this meta analysis, come to similar conclusions. For any fracture, RR=1.50; 95% CI, 1.27 - 1.76 42; and OR =1.6; 95% CI, 1.16 - 2.34 43 for 1 SD reduction in bone mineral density at the spine; and, RR=1.41; 95% CI, 1.21 - 1.64 for 1 SD decrease in bone mineral density at the femoral neck. 42 However, these studies also had short follow up periods.

There are no threshold values of bone density below which fractures will necessarily occur, the relationship between risk and bone density being continuous. On the basis of data from case control studies (Level VII) of hip fracture, BDM does not accurately distinguish between patients with recent (non-traumatic) fractures and those without fractures.

A recent review of case control studies of hip fracture, using the same approach as that of Law et al. 44, determined the weighted average difference in bone mineral density between cases and controls to be 0.9 SD. 10 Error! Bookmark not defined. Using these figures, a 1 SD cut-off below the mean bone mineral density for those without a fracture would result in a 46% detection rate with a 16% false positive rate.

2.3 Low bone density is only one of a number of risk factors for fracture in menopausal women, some of which have similar estimates for risk association with fractures.

Scientific evidence of similar strength (Level VI) exists regarding many other risk factors in menopausal women that have similar independent predictive ability for fracture to that of bone mineral density.
For example, data from the Study of Osteoporotic Fractures identify a history of maternal hip fracture (RR=2.0; 95% CI, 1.4 - 2.9), previous fractures of any type after the age of 50 (RR=1.5; 95% CI, 1.1 - 2.0), self-rated health as fair to poor (RR=1.7; 95% CI, 1.3 - 2.2), and previous hyperthyroidism (RR=1.8; 95% CI, 1.2 - 2.6) as independent risk factors for hip fracture.

Other characteristics that are observable in a physical examination that were identified as risk factors included the inability to rise from a chair without using one’s arms (RR=2.1; 95% CI, 1.3 - 3.2), a faster resting pulse rate (RR=1.8; 95% CI, 1.3 - 2.5), and poorer depth perception (RR=1.5; 95% CI, 1.1 -2.0). The presence of five or more of the risk factors dramatically increased the incidence of fractures by about 18 times that for women with two or fewer risk factors.

In a study which involved over 25,000 younger women, body height (age-adjusted RR=3.2; 95% CI, 1.46 - 8.97) and history of diabetes mellitus (age-adjusted RR=5.81; 95% CI, 2.15 - 15.71) were also identified as risk factors for hip fracture.

In comparison with BDM, many of these other risk factors are easily and cheaply measured. A question here is the extent of any added value offered by BDM.

2.4 There appears to be no consensus about the appropriate applications of BDM between organisations that had published a report on this topic.

A survey (Level VIII) of reports on the applications of BDM technologies suggests that there is wide variation in the views taken on the appropriate use of BDM for clinical and screening purposes. Studies based on more systematic review methods tended to be more conservative in their conclusions about the potential uses of BDM.

3.0 Effect of HRT in preventing fractures and preserving bone mass

3.1 There is FAIR evidence that ever-use of HRT is associated with a decrease in fractures of all types.

Data from only one small RCT (Level III) are available, which show a reduction in new vertebral fractures with HRT used for secondary prevention. There were 8 new fractures in 7 women in the estrogen group compared to 20 fractures in 12 women in the placebo group, yielding a lower rate of vertebral fracture (61%) in the estrogen group (RR=0.39; 95% CI, 0.16 -0.95)). However, the number of individuals who experienced a new vertebral fracture was reduced by a smaller amount (37%) and was not statistically significant. Data measuring treatment effects in terms of fractures must be interpreted cautiously - some studies use the number of fractures instead of the number of individuals with fractures as the endpoint which will overestimate the effectiveness of treatments. In the clinical context, the important result is whether or not an individual will have a fracture at all.

For all other types of fracture, data are available only from cohort and case control studies (Levels VI and VII). Pooled estimates from these observational studies show a tendency to a modest reduction in relative risk for hip fracture with any use of HRT. These estimates give RR=0.75; 95% CI, 0.68-0.84 for all observational studies; RR=0.85; 95% CI, 0.68-1.07 for cohort studies only (Level VI); and OR=0.57; 95% CI, 0.48-0.67 for case control studies only (Level VII). The results from the cohort studies are more reliable since they are less susceptible to bias than case control studies. The higher quality case control studies showed a trend to risk reduction similar to that for the cohort studies.

A similar protective effect was observed for forearm and wrist fractures (RR=0.70; 95% CI, 0.52-0.93 for cohort studies only and OR=0.60; 95% CI, 0.41-0.88 for case control studies only).

3.2 There is FAIR evidence that current long-term use of HRT has a protective effect for fractures.

A prospective cohort study (Level VI) analysed the risk of fracture for short-term (<10 years) and
long-term (>10 years) duration of use in women who were 65 at the time of entering the study and who were current and previous users of HRT. In current users, short-term duration of treatment was associated with a decrease of 30% (RR=0.67, 95% CI, 0.49-0.92) in the risk of all non-spinal fractures, whereas for long-term users this reduction was 40% (RR=0.60, 95% CI, 0.45-0.83). The decrease in risk of hip fractures in current users was 19% (RR=0.81, 95% CI, 0.40-1.65) and 73% (RR=0.27; 95% CI, 0.08-0.85) for short and long-term users respectively.

Woman over 75 years who were current users of estrogen therapy had a reduced risk of hip fracture (RR=0.18; 95% CI, 0.04-0.77). However, for women 75 years old or younger, who were also current users, there was no effect (RR=0.94; 95% CI, 0.51-1.69), both compared with women who had never used estrogens 50.

Pooling the results of four case control studies (Level VII) that examined the relationship between extended use of estrogen (> 5 years) and never use, a trend to a 66% reduction in the relative risk of hip fracture (OR=0.34; 95% CI; 0.20-0.55) was found 18. When compared to a shorter term use (0-60 months) the estimated pooled relative risk reduction was 61% (OR=0.39; 95% CI, 0.25-0.62).

When the results from different cohort studies 51 52 53 (Level VI) with a duration of use longer than 5 years were pooled 18, the risk reduction reported in wrist and forearm fracture was about 15% (estimated pooled RR=0.85; 95% CI, 0.73-0.99). This figure was not significantly different from that observed in never users.

A prospective cohort study 54 (Level VI) found there was not a statistically significant difference in hip fracture risk between ever and never users of HRT regardless of the duration of therapy (less than 3 years, RR=1.19; 95% CI, 0.89-1.60; 4-14 years, RR=0.89; 95% CI, 0.63-1.23; > 15 years, RR=0.88; 95% CI, 0.63-1.23).

3.3 There is FAIR evidence that there is no decrease in risk for hip fracture at older ages with ever use of HRT.

Three cohort studies 54, 55 56 (Level VI) and one case-control study 57 (Level VII) reported a decrease in the potential protective effect for hip fracture with age and at older ages, when most hip fractures occur, there was no statistically significant difference in fracture risk between ever and never users of HRT (Table 3).

A prospective cohort study 59 (Level VI) found not statistically significant decrease in risk of hip fracture for women over 65 years who were previous users of estrogens (RR = 1.03; 95% CI, 0.69-1.55). In comparison, the RR for all current users was 0.60 (95% CI, 0.36-1.02). For current users with a history of osteoporosis the RR was 0.86 (95% CI, 0.42-1.75), and for women with no history of the condition the RR was 0.45 (95% CI, 0.20-0.99).

3.4 There is FAIR evidence that the longer the period since cessation of therapy, the smaller the reduction in risk of fracture.

Pooling the results from two cohort studies not defined. 54,56 (Level VI) gives an estimated protective effect on hip fracture between former and never users of RR=0.88 (95% CI, 0.67 - 1.15) when 2-14 years have elapsed since last estrogen use. When more than 15 years have lapsed, there is no evidence of any benefit (RR=1.07 (95% CI, 0.85-1.34)). 18

3.5 There is GOOD evidence that HRT, used alone or in combination with progestogens and/or calcium, for primary and secondary prevention, has a protective effect against bone mass loss, as measured by various BDM techniques at the forearm, spine and hip.

A meta analysis of RCTs (Level I) showed a trend towards a positive effect of HRT on bone mass both in primary and secondary prevention (Table 4). 18 Effect sizes in the forearm and spine in studies of secondary prevention were larger than in studies of primary prevention and had wider confidence intervals.
3.6 The effect of HRT in reducing bone loss has mainly been studied in women shortly after menopause. However, there is GOOD evidence that age does not attenuate the short term response to treatment.

When the results of 43 RCTs examining skeletal response in women of various ages are examined, the protective effect of HRT appears to be the same for women who are under 60 years and those who are over 60 years old. Although results from different studies show that this effect may be greater in the lumbar spine than in the forearm or hip, the pooled effect size from these studies showed a tendency to be marginally greater for forearm than for the spine.

3.7 There is FAIR evidence that the protective effect of HRT on loss of bone mass may decline over time when therapy is started soon after the menopause, and that the protective effect wears off after cessation of treatment.

In a nested case control study, a statistically significant difference of 11.2% in bone mass averaged over all sites was found in women under 75 years with greater than 7 years of HRT therapy compared to those who had not received therapy. In women over 75 years with greater than 7 years of therapy compared to those with no therapy, a difference in bone mass of only 3.2% was found, which was only statistically significant at the radius shaft (8.5%, p<0.02).

Other studies indicate that the protective effect of HRT on bone loss is only maintained when currently used. The protective effect appears to disappear progressively after cessation of therapy, reaching a rate of bone loss equal to that in untreated or placebo treated women (2-3% per year) within a few years after withdrawal of treatment. This point is critically important since therapy with HRT is generally prescribed around or soon after the menopause and generally not for more than ten years, consequently leaving a period of 15-20 years between the cessation of treatment and the time when most fractures occur (> 75 years of age).

3.8 Long term compliance with HRT is likely to be less than 50% for menopausal women.

Data from different surveys suggest that long term compliance with HRT is low (approximately 30%). This is mainly because of the presence of various side effects (e.g. breast tenderness, bleeding, depression), fear of cancer, dislike of taking tablets and failure to continue treatment when climacteric symptoms disappear. Recent survey data of women in the United Kingdom who had a BMD in a population screening programme 1 year before, are somewhat more optimistic, suggesting short-term compliance rates of 48% for postmenopausal women and 59% for women with a simple hysterectomy. This is an important factor to take into account when considering the likely effectiveness of HRT, given the findings of reduced effect of HRT after cessation of therapy.

4.0 Effect of intranasal salmon calcitonin in preventing fractures and preserving bone mass

4.1 There is FAIR evidence to support the efficacy of intranasal salmon calcitonin in decreasing the risk of fractures.

Three RCTs (Levels II and III) on secondary prevention analysed the efficacy of SCT in decreasing the risk of fractures. Two of these concerned the effect of SCT (N) on vertebral fractures in women late after menopause with osteoporosis. One showed a decrease in the risk of patients with first vertebral fractures (RR=0.23; 95% CI, 0.07-0.77) as well as in the rate of new fractures (RR=0.37; 95% CI, 0.14-0.95) in current users compared with non-users. The other RCT showed no statistically significant differences after three years of treatment.

Two case control studies (Level VII) and one meta analysis of RCTs (Level I) analysing the effect of all types of calcitonins on fracture rates were identified. Both case-control studies show evidence of reduced hip fracture risk comparing ever and never users of calcitonin. Data from the Mediterranean Osteoporosis Study suggest a protective effect of about 30% (RR=0.71;
95% CI, 0.52-0.90, adjusted for previous estrogen intake). The other study found a 53% decrease in the risk of hip fracture (RR=0.47; 95% CI; 0.30-0.74) with previous use of calcitonin plus calcium. The meta analysis identified only two prospective trials and concluded that the protective effect of calcitonin still remains to be established.

4.2 There is GOOD evidence demonstrating the short term efficacy of intranasal salmon calcitonin (SCT(N)) in preserving bone mass in both primary and secondary prevention in postmenopausal women.

Several RCTs have shown that SCT(N) decreases bone loss and/or preserves bone mass in postmenopausal women. This is the case in both primary and secondary prevention, after both natural and surgical menopause. Only three of the RCT studies in primary prevention and four in secondary prevention in natural postmenopausal women were of good quality design. For primary prevention, one study showed an increase of 2.6% in vertebral bone mass (p<0.01 vs. baseline levels). Another study showed an increase of 2.9% at the forearm (first 6 months, 1.5%, p<0.005; last 6 months, 1.5%, p<0.002). One of the studies showed an increase in vertebral bone mass that was not statistically significant. For secondary prevention, statistically significant increases in bone mass were observed at vertebrae ranging from 3% to 8.6% depending on the dosage.

4.3 There is FAIR evidence demonstrating the long term preservation (5 year) of bone mass using SCT (N), but no data are available about the long term effect when treatment is started early after menopause.

One RCT (Level III) in primary prevention with a duration of 5 years shows a statistically significant increase in vertebral bone mass after 42 months of treatment compared with those women not treated (2.5%, p<0.001), but not at the end of the five years.

4.4 There are no prospective studies comparing the efficacy or the effectiveness of HRT and intranasal calcitonin.

Discussion

This paper has reviewed the current scientific evidence available regarding the ability of BDM to predict fractures, and the efficacy and effectiveness of some common associated treatments for low bone density. In summary, using the classification system for evidence outlined in Table 1, there is:

FAIR evidence from prospective cohort studies suggesting that BDM can assess the risk of future fracture occurrence in populations over the short term, but not with a high degree of accuracy;

FAIR evidence, from low quality RCTs and observational studies, showing the efficacy of HRT and SCT(N) in preventing fractures while therapy is continued;

GOOD evidence to support the efficacy of HRT and SCT (N) in preserving bone mass during therapy;

FAIR evidence that the protective effect of HRT diminishes and may eventually wear off after cessation of therapy.

Common limitations of the studies that have been undertaken are that sample sizes were often small and the follow up periods too short. These studies are subject to various errors and biases. Cohort studies are particularly subject to the effects of confounding and case-control studies to recall and observer bias, (and, when bone mass is measured after the fracture, to problems of interpreting the direction of causality). Because many of these summaries of the evidence are based only on observational studies, the degree to which they reflect a causal relationship is not certain. More research is needed to corroborate the results of these studies, given the methodological limitations.
Due to the relatively short follow up of these cohort studies, the accuracy with which BDM predicts fractures which occur many years after measurement is not known. Similarly, the long term positive and negative effects of HRT and SCT(N) are not accurately known. The follow-up period in most trials of these interventions is shorter than that recommended by the U.S. Food and Drug Administration or the European Foundation for Osteoporosis and Bone Diseases (2-3 years). 105 106

These points are critical when considering BDM screening of menopausal women linked to subsequent treatment with HRT or SCT(N), since these interventions have not been shown to be beneficial at the ages when most fractures occur (>75 years).

While BDM has been proposed for a number of applications,24 28 there is particular interest in its potential use for population screening programmes and in opportunistic screening for women around the menopause who, for example, seek advice from medical practitioners on whether to take HRT. How should the individual components of evidence summarised above be interpreted in a broader context, in terms of health policy and for routine clinical practice?

In general, screening programmes must use a reliable diagnostic test and be offered in conjunction with treatment that has demonstrated effectiveness. There must be “conclusive evidence that screening can alter the natural history of the disease in a significant proportion of those screened.” 107 When interventions are offered to patients who are not ill or who have not specifically sought assistance, the onus is on those offering the intervention to be certain that the patient will benefit. Current evidence is insufficient to reach conclusions about the value of BDM screening or HRT or SCT(N) therapy.

However, it is possible to get some indication of the potential population impact of BDM screening in preventing hip fractures. Table 5 presents a number of scenarios, using realistic assumptions, for a BDM screening program linked to treatment with HRT in a hypothetical cohort of 20,000 menopausal women. 2 10 18 38 49 74 75 77 108 For the scenario with 50% screening uptake, 30% long term compliance with treatment and 30% reduction in lifetime fracture risk from HRT, one hip fracture would be avoided for every nine women identified through screening as being at risk, and who comply with therapy for 30 years. This apparently promising estimate represents the optimum benefit achievable, and has to be put in the context of the overall screening process.

Of the 20,000 women invited for screening, 10,000 are likely to attend for BDM. Of these, 1,600 will be identified as having a bone density of less than one SD below the population mean (assuming a Gaussian distribution) and will be offered HRT. Of those offered HRT, 576 will have been correctly identified (true positives) and 173 (30%) of them will comply with therapy, so preventing 52 fractures. There will, however, also be 1,024 false positives, so that almost two thirds of women advised to take HRT would be unnecessarily using this treatment since they would not have had a fracture, under the assumptions of the model used here.

A further summary measure of some interest is the number of false negative cases. Under this scenario, of 10,000 women who present for screening, 948 who will go on to have a hip fracture will be advised that they are not at high risk. Nearly two thirds of those who will sustain a fracture and who have a BDM will, therefore have been falsely reassured.

When this broader perspective is considered, 393 women would need to be invited for screening and 197 actually attend in order to avoid one fracture. Thus, the overall impact of the program would be to reduce the number of fractures over the remaining lifetime of the cohort of 20,000 women from 3,050 to 2,998, a reduction of 1.7%.

When all the scenarios presented in Table 5 are considered, a BDM screening program aimed at menopausal women might prevent between 1 and 7% of fractures. Taken together, these estimates of the effectiveness of such a program are not particularly encouraging from a public health perspective and are unlikely to represent good value for money. Similar indications might be expected to follow from scenarios in which SCT(N) was the available treatment.
For those formulating policy and considering offering such interventions to their patients, there are other potential effects to be considered. Current data suggest that HRT is associated with a 40-50% reduction in the risk of coronary heart disease among postmenopausal women and an increased risk of 30-70% for breast cancer, independent of levels of bone mass. Evaluation of incremental cost per health gain, taking account of these factors, is needed to further inform health policy makers of the worth of screening programs.

A number of important social and ethical concerns must also be addressed. At what point does a decrease in bone density with age become a medical problem requiring treatment? What are the implications of using a definition of disease (a BMD <2.5SD below the young adult mean proposed by the WHO) where, consequently, over 30% of women between 70 and 79 years of age could be considered in need of therapy because of a low bone mineral density.

Alternative approaches to preventing and treating osteoporosis also require further consideration. Other risk factors, observable in a clinical exam, might be used to identify high risk groups without using BDM. BDM and appropriate treatment might be better targeted to older women at ages when most fractures occur. Attention could also be given to alternative strategies to prevent fractures. These could include potentially more cost-effective prophylactic interventions such as annual Vitamin D injections or the prevention of accidents in older people.

Ongoing studies, such as the RCT of population screening being undertaken in the United Kingdom, and studies of alternative approaches to preventing osteoporosis, may provide better information upon which such decisions can be made in the future. The currently available evidence does not support BMD screening of menopausal women in combination with HRT or SCT(N) in the context of population or opportunistic screening for the prevention of fractures, and estimates based on what data are available are not encouraging about its potential effectiveness.
<table>
<thead>
<tr>
<th>Level</th>
<th>Strength of evidence</th>
<th>Type of study design</th>
<th>Conditions of scientific rigour*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest (I) to Lowest (IX)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Good</td>
<td>Meta-analysis of randomised controlled trials</td>
<td>Analysis of patient individual data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Meta-regression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Different techniques of analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absence of heterogeneity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quality of the studies</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>Large sample randomised controlled trials</td>
<td>Assessment of statistical power</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multicentre</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>III</td>
<td>Good</td>
<td>Small sample randomised controlled trials</td>
<td>Assessment of statistical power</td>
</tr>
<tr>
<td>IV</td>
<td>to</td>
<td>Non-randomised prospective trials</td>
<td>Quality of the study</td>
</tr>
<tr>
<td></td>
<td>Fair</td>
<td></td>
<td>Concurrent controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multicentre</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>V</td>
<td></td>
<td>Non-randomised retrospective trials</td>
<td>Historical controls</td>
</tr>
<tr>
<td>VI</td>
<td>Fair</td>
<td>Cohort studies</td>
<td>Quality of the study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Concurrent controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multicentre</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>VII</td>
<td></td>
<td>Case-control studies</td>
<td>Multicentre studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quality of the study</td>
</tr>
<tr>
<td>VIII</td>
<td>Poor</td>
<td>Non-controlled clinical series</td>
<td>Multicentre</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Descriptive studies: surveillance of disease,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>surveys, registers, data bases, prevalence</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>studies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expert committees, consensus conferences</td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td></td>
<td>Anecdotes or case reports</td>
<td></td>
</tr>
</tbody>
</table>

* Quality of the study assessed by specific protocols and conditions of scientific rigour.
Source: Adapted from reference 20.
<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy CV (%)</th>
<th>Precision CV (%)</th>
<th>Time of scan (minutes)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPA Single Photon Absorptiometry</td>
<td>2 - 8</td>
<td>2 - 5</td>
<td>5 - 15</td>
<td>Simple, relatively inexpensive, small radiation exposure. Decay of source affects performance</td>
</tr>
<tr>
<td>DPA Dual Photon Absorptiometry</td>
<td>3-10</td>
<td>2 - 6</td>
<td>20 - 45</td>
<td>Usually used for spine and hip measurements. Decay of source affects performance</td>
</tr>
<tr>
<td>SXA Single X-ray Absorptiometry</td>
<td>5</td>
<td>1</td>
<td>10 -20</td>
<td>X-ray equivalent of SPA</td>
</tr>
<tr>
<td>DXA Dual X-ray Absorptiometry</td>
<td>3-6</td>
<td>1 - 3</td>
<td>3 - 10</td>
<td>Single X-ray source with two energies. Higher photon flux than radionuclide sources, improved detector configuration.</td>
</tr>
<tr>
<td>QCT Quantitative Computed Tomography</td>
<td>5 - 15</td>
<td>2 - 5</td>
<td>10 - 15</td>
<td>Able to measure bone structure. Need to measure calibration standards simultaneously with the patient.</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>20</td>
<td>2-4</td>
<td>5</td>
<td>Potential to measure bone structure</td>
</tr>
</tbody>
</table>

CV = Coefficient of variation

Source: References 9, 10, 12, 13, 16, 28, 31
<table>
<thead>
<tr>
<th>Study</th>
<th>Age group</th>
<th>Relative risk (95% CI)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
<td>Younger</td>
<td>Older</td>
</tr>
<tr>
<td>Pagnini-Hill,</td>
<td>avg age=73</td>
<td>1.02</td>
<td>(0.81-1.27)²</td>
<td></td>
</tr>
<tr>
<td>1991¹⁴</td>
<td>&gt;60 (trochanter)</td>
<td>1.03</td>
<td>(0.74-1.40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;60 (cervical hip)</td>
<td>0.95</td>
<td>(0.75-1.18)</td>
<td></td>
</tr>
<tr>
<td>Naessén, 1990</td>
<td>&lt;60 (trochanter)</td>
<td>0.37</td>
<td>(0.13-0.79)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;60 (cervical hip)</td>
<td>0.58</td>
<td>(0.41-0.80)</td>
<td></td>
</tr>
<tr>
<td>Kiel, 1987²⁶</td>
<td>&gt;75</td>
<td>0.37</td>
<td>(0.05-2.46)¹</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.82</td>
<td>(0.21-3.24)³</td>
<td></td>
</tr>
<tr>
<td>Kanis, 1992²⁷</td>
<td>&lt;80</td>
<td>0.51</td>
<td>(0.31-0.84)²</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;80</td>
<td>0.70</td>
<td>(0.29-1.66)³</td>
<td></td>
</tr>
<tr>
<td>Cauley, 1995²⁰</td>
<td>≤75</td>
<td>0.94</td>
<td>(0.52-1.69)⁴</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;75</td>
<td>0.18</td>
<td>(0.04-0.77)⁴</td>
<td></td>
</tr>
<tr>
<td>Cauley, 1995²⁰</td>
<td>&gt;65, current users</td>
<td>0.60 *</td>
<td>(0.36 - 1.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;65, previous users</td>
<td>1.03 *</td>
<td>(0.69 - 1.55)</td>
<td></td>
</tr>
</tbody>
</table>

¹ = age - adjusted RR
² = recent estrogen use (<= 2 years)
³ = adjusted by age, previous fractures, body mass index
⁴ = current users, multivariate-adjusted relative risk
⁵ = multivariate - adjusted, includes adjustment for history of osteoporosis (yes or no)
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Forearm (SD, 95% CI)*</th>
<th>Spine (SD, 95% CI)*</th>
<th>Hip (SD, 95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oral unopposed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary prevention</td>
<td>1.45 (1.12-1.78)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>0.95 (0.21-1.70)</td>
<td>0.77 (0.33-1.51)*</td>
<td>-</td>
</tr>
<tr>
<td><strong>Oral unopposed + Calcium</strong></td>
<td>0.92 (0.33-1.51)</td>
<td>2.27 (1.17-3.37)</td>
<td>-</td>
</tr>
<tr>
<td>Primary prevention</td>
<td>-</td>
<td>0.77 (0.33-1.51)*</td>
<td>-</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>-</td>
<td>2.27 (1.17-3.37)</td>
<td>-</td>
</tr>
<tr>
<td><strong>Oral opposed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary prevention</td>
<td>1.33 (0.74-1.92)b</td>
<td>1.23 (0.75-1.72)b</td>
<td>-</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>1.18 (0.65-1.70)b,c</td>
<td>1.18 (0.65-1.70)b,c</td>
<td>-</td>
</tr>
<tr>
<td><strong>Oral opposed + Calcium</strong></td>
<td>1.52 (1.16-1.84)</td>
<td>0.94 (0.18-1.69)d</td>
<td>-</td>
</tr>
<tr>
<td>Primary prevention</td>
<td>1.46 (0.95-1.97)c</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>2.94 (-0.92-6.80)b</td>
<td>3.16 (1.83-4.49)b</td>
<td>-</td>
</tr>
<tr>
<td><strong>Transdermal opposed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary prevention</td>
<td>1.10 (-0.09 to 2.29)a</td>
<td>1.02 (0.67-1.36)</td>
<td>-</td>
</tr>
<tr>
<td><strong>Transdermal opposed + Calcium</strong></td>
<td>1.70 (1.09-2.31)</td>
<td>1.37 (0.45-2.30)</td>
<td>-</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>0.97 (0.46-1.48)</td>
<td>0.73 (0.24-1.21)</td>
<td>-</td>
</tr>
<tr>
<td>Overall Primary Prevention</td>
<td>1.38 (0.93-1.84)b</td>
<td>1.17 (0.63-1.70)b</td>
<td>-</td>
</tr>
<tr>
<td>(8 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Secondary Prevention</td>
<td>2.90 (-0.99-6.79)b</td>
<td>2.23 (-0.81-5.28)b</td>
<td>0.92 (0.34-1.50)a</td>
</tr>
<tr>
<td>(2 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Secondary Prevention</td>
<td>2.90 (-0.99-6.79)b</td>
<td>2.23 (-0.81-5.28)b</td>
<td>0.92 (0.34-1.50)a</td>
</tr>
<tr>
<td>(4 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*SD = treatment effect size in terms of standard deviation units by which the average annual decline in bone mass in the control group exceeds that in the treatment group.

* Heterogeneity: p = 0.05;  b Heterogeneity: p < 0.05

* Standard dose (conjugated estrogens = 0.60-0.625 mg; estrone = 1.25 mg; estradiol = 2.0 mg)

* Low dose (conjugated estrogen = 0.30 mg; estrone = 0.625 mg; estradiol = 1.0 mg)

Source: References 18, 19.
## Table 5: Potential impact of BDM screening and treatment with HRT in preventing hip fractures in a population of 20,000 menopausal women (under realistic assumptions)\(^{a,b,c,d}\)

<table>
<thead>
<tr>
<th>Fracture risk reduction from HRT (^{i})</th>
<th>Compliance = 30% (^{e})</th>
<th>Compliance = 50% (^{f})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N hip fractures avoided</td>
<td>N hip fractures avoided</td>
</tr>
<tr>
<td></td>
<td>% hip fractures avoided</td>
<td>% hip fractures avoided</td>
</tr>
<tr>
<td></td>
<td>N needed to invite to screen per hip fracture avoided</td>
<td>N needed to invite to screen per hip fracture avoided</td>
</tr>
</tbody>
</table>

Screening uptake 50% \(^{h}\): N false negatives = 948, N false positives = 1,024, N offered HRT = 1,600

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15% (RR=0.85)</td>
<td>27</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>0.9%</td>
<td>1.5%</td>
</tr>
<tr>
<td></td>
<td>803</td>
<td>474</td>
</tr>
<tr>
<td>30% (RR=0.70)</td>
<td>53</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>1.7%</td>
<td>2.9%</td>
</tr>
<tr>
<td></td>
<td>393</td>
<td>234</td>
</tr>
<tr>
<td>50% (RR=0.50)</td>
<td>87</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>2.9%</td>
<td>4.8%</td>
</tr>
<tr>
<td></td>
<td>234</td>
<td>140</td>
</tr>
</tbody>
</table>

Screening uptake = 70% \(^{h}\): N false negatives = 1,327, N false positives = 1,434, N offered HRT = 2,240

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15% (RR=0.85)</td>
<td>38</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>1.2%</td>
<td>2.0%</td>
</tr>
<tr>
<td></td>
<td>560</td>
<td>334</td>
</tr>
<tr>
<td>30% (RR=0.70)</td>
<td>74</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>2.4%</td>
<td>4.0%</td>
</tr>
<tr>
<td></td>
<td>278</td>
<td>166</td>
</tr>
<tr>
<td>50% (RR=0.50)</td>
<td>122</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>4.0%</td>
<td>6.7%</td>
</tr>
<tr>
<td></td>
<td>166</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 5 – Notes

These scenarios were calculated using the following assumptions:

A cohort of 20,000 menopausal women are invited to a BDM screening programme and those identified with a bone density < 1 SD below the healthy adult mean are treated with HRT.

Bone density values follow a Gaussian distribution in the population

The lifetime risk of hip fracture is 15.25% for women over 50 years 39

a Hip fracture RR=2.6 for 1 SD decrease in BMD below age adjusted mean 10 38

b For those who comply, compliance continues for their remaining lifetimes (about 30 years)

c There is no decrease in the protective effect of HRT over time

d No side effects are taken into account 109 110 111 112

e The reduced beneficial effect on those who would be treated because some women would already be taking HRT for other reasons is not taken into account.

f The range considered for fracture risk reduction is 15%-50% 17 18 49

The range considered for compliance is 30-50% 2 77 108

h The range for attendance to a screening programme is 50-70% 2 108
References


19. Henry D, Robertson J, O’Connell D, Gillespie W. *The skeletal effects of estrogen therapy in post-


81. Mazzuoli GF, Gennari C, Passeri M, Acca M, Camporeale A, Pioli G. Hip fracture in Italy:


100. Christiansen C. Use of nasally administered salmon calcitonin in preventing bone loss. Calcif


