

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

AHRQ is the lead Federal agency charged with supporting research designed to improve the quality of health care, reduce its cost, address patient safety and medical errors, and broaden access to essential services. AHRQ sponsors and conducts research that provides evidence-based information on health care outcomes; quality; and cost, use, and access. The information helps health care decisionmakers—patients and clinicians, health system leaders, and policymakers—make more informed decisions and improve the quality of health care services.

Screening for Postmenopausal Osteoporosis

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
2101 East Jefferson Street
Rockville, MD 20852
<http://www.ahrq.gov>

Contract No. 290-97-0018

Task Order No. 2

Technical Support of the U.S. Preventive Services Task Force

Prepared by:

Oregon Health Sciences University
Evidence-based Practice Center, Portland, Oregon

Heidi D. Nelson, MD, MPH

Mark Helfand, MD, MS

September 2002

Preface

The Agency for Healthcare Research and Quality (AHRQ) sponsors the development of Systematic Evidence Reviews (SERs) through its Evidence-based Practice Program. With guidance from the third U.S. Preventive Services Task Force* (USPSTF) and input from Federal partners and primary care specialty societies, two Evidence-based Practice Centers—one at the Oregon Health Sciences University and the other at Research Triangle Institute-University of North Carolina—systematically review the evidence of the effectiveness of a wide range of clinical preventive services, including screening, counseling, immunizations, and chemoprevention, in the primary care setting. The SERs—comprehensive reviews of the scientific evidence on the effectiveness of particular clinical preventive services—serve as the foundation for the recommendations of the third USPSTF, which provide age- and risk-factor-specific recommendations for the delivery of these services in the primary care setting. Details of the process of identifying and evaluating relevant scientific evidence are described in the “Methods” section of each SER.

The SERs document the evidence regarding the benefits, limitations, and cost-effectiveness of a broad range of clinical preventive services and will help to further awareness, delivery, and coverage of preventive care as an integral part of quality primary health care.

AHRQ also disseminates the SERs on the AHRQ Web site (<http://www.ahrq.gov/uspstfix.htm>) and disseminates summaries of the evidence (summaries of the SERs) and recommendations of the third USPSTF in print and on the Web. These are available through the AHRQ Web site (<http://www.ahrq.gov/uspstfix.htm>), through the National Guideline Clearinghouse (<http://www.ncg.gov>), and in print through the AHRQ Publications Clearinghouse (1-800-358-9295).

We welcome written comments on this SER. Comments may be sent to: Director, Center for Practice and Technology Assessment, Agency for Healthcare Research and Quality, 6010 Executive Blvd., Suite 300, Rockville, MD 20852.

Carolyn M. Clancy, M.D.
Acting Director
Agency for Healthcare Research and Quality

Robert Graham, M.D.
Director, Center for Practice and
Technology Assessment
Agency for Healthcare Research and Quality

* The USPSTF is an independent panel of experts in primary care and prevention first convened by the U.S. Public Health Service in 1984. The USPSTF systematically reviews the evidence on the effectiveness of providing clinical preventive services—including screening, counseling, immunization, and chemoprevention—in the primary care setting. AHRQ convened the third USPSTF in November 1998 to update existing Task Force recommendations and to address new topics.

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services of a particular drug, device, test, treatment, or other clinical service.

Structured Abstract

Context: The incidence of osteoporotic fractures increases with age and is associated with a significant health burden.

Objective: To examine evidence on the benefits and harms of screening asymptomatic postmenopausal women for osteoporosis.

Data Sources: MEDLINE (1966 to May 2001), HealthSTAR (1975 to May 2001), and Cochrane databases, reference lists of systematic reviews, and experts.

Study Selection: We included English-language abstracts with original data about postmenopausal women and osteoporosis that addressed the effectiveness of risk factor assessment, bone measurement tests, or treatment. Two reviewers read each abstract to determine its eligibility.

Data Extraction: We extracted selected information about the patient population, interventions, clinical endpoints, and study design, and applied a set of criteria to evaluate study quality.

Data Synthesis: Although many studies have been published about osteoporosis in postmenopausal women, there have been no trials of screening and, therefore, no direct evidence that screening improves outcomes. Instruments developed to assess clinical risk factors for low bone density or fractures generally have moderate-to-high sensitivity and low specificity, many have not been validated, and none have been widely tested in a practice setting. Among different bone density tests measured at various sites, bone density measured at the femoral neck by dual-energy x-ray absorptiometry is the best predictor of hip fracture and is comparable to forearm measurements for predicting fractures at other sites. Women with low bone density have a 40% to 50% reduction in fracture risk when treated with raloxifene (vertebral fractures) or bisphosphonates (both vertebral and nonvertebral fractures). Trials of estrogen are inconclusive because of methodologic limitations.

Conclusions: Although there is no direct evidence that screening prevents fractures, there is evidence that the prevalences of osteoporosis and fractures increase with age, that the short-term risk of fracture can be estimated by bone measurement tests and risk factor

assessment, and that treatment may reduce fracture risk among women with low bone density.

Contents

Systematic Evidence Review

Chapter 1. Introduction	1
Burden of Suffering/Epidemiology.....	2
Health Care Interventions	3
Prior Recommendations.....	5
Analytic Framework and Key Questions.....	5
Chapter 2. Methods.....	7
Literature Search Strategy.....	7
Inclusion/Exclusion Criteria	7
Size of Literature Reviewed.....	8
Literature Synthesis	8
Chapter 3. Results	10
Arrow 1. Does screening using risk factor assessment and/or bone density testing reduce fractures?	10
Arrow 2. Does risk factor assessment accurately identify women who may benefit from bone density testing?	10
Risk Assessment Based on the Likelihood of Low Bone Density.....	10
Risk Assessment Based on Risk of Fracture	12
Arrow 3. Do bone density measurements accurately identify women who may benefit from treatment?.....	13
Accuracy and Reliability of Tests.....	14
Prediction of Short-term Risk of Fractures	15
Arrow 4. What are the harms of screening?	17
Arrow 5. Does treatment reduce the risk of fractures in women identified by screening?	19
Estrogen Replacement Therapy	19
Selective Estrogen Receptor Modulators.....	21
Biphosphonates	22
Generalizability of Randomized Trials.....	24
Arrow 6. What are the harms of treatment?.....	26
Chapter 4. Discussion	27
Summary of Evidence Quality.....	27
Outcomes Table	27
Conclusions.....	29
Limitations of the Literature	30
Future Research	31
Acknowledgements.....	33
References.....	34

Figures

Figure 1. Osteoporosis Screening Analytic Framework.....	55
Figure 2. Number Needed to Screen to Prevent 1 Hip Fracture in 5 Years.....	56

Tables

Table 1. Recent Systematic Reviews of Osteoporosis.....	57
Table 2. Prevalence of Low Femoral Neck Bone Density in U.S. Women Over Age 50.....	60
Table 3. Osteoporosis Prevalence (%) in White Women in the U.S. Over Age 50 by Decade.....	61
Table 4. 5-Year Hip Fracture Rates for Women Without Previous Hip Fracture by Age.....	62
Table 5. Studies of Risk Factor Assessment.....	63
Table 6. Characteristics of Bone Density and Quantitative Ultrasound Tests.....	67
Table 7. Prospective Studies of DXA of the Hip Reporting Hip Fractures.....	68
Table 8. Randomized Controlled Trials of Estrogen and SERMs with Fracture Outcomes.....	69
Table 9. Randomized Controlled Trials of Alendronate with Fracture Outcomes.....	72
Table 10. Summary of Evidence Quality.....	74
Table 11. Screening for Osteoporosis in 10,000 Postmenopausal Women.....	75

Appendices

Appendix 1. Search Strategies for Individual Topics.....	76
Appendix 2. Criteria for Grading the Internal Validity of Individual Studies.....	81

Chapter 1. Introduction

Half of all postmenopausal women will have an osteoporosis-related fracture during their lives, including one-quarter who will develop a vertebral deformity,¹ and 15% who will suffer a hip fracture.² As early as 1975, it was acknowledged that bone density measurements were related to future fracture risk. In the 1990s, the magnitude of this risk in relation to age and bone density was carefully measured in several well-designed longitudinal studies. Recently, data indicating that newer therapies can prevent fractures in asymptomatic postmenopausal women with osteoporosis have become available as well.

Despite the high prevalence of osteoporosis and the impact of fractures on mortality, independence, and quality of life, whether it is appropriate to screen asymptomatic postmenopausal women is unclear. Recent systematic reviews and guidelines disagree about which women should be screened and when (Table 1).³⁻¹¹ This disagreement reflects, in part, gaps in the evidence. For example, most guidelines recommend using risk factors to select patients for bone density testing, but because of inadequate data there is no consensus on what risk factors to use.

As part of the U.S. Preventive Services Task Force update of its recommendations,¹² we examined evidence on the benefits and harms of screening asymptomatic postmenopausal women for osteoporosis. Specifically, we addressed the role of risk factors in identifying high-risk women, techniques of bone measurements to identify risk of fractures, effectiveness of treatment in reducing risk for fractures, and harms of screening and treatment.

Burden of Suffering/Epidemiology

Osteoporosis affects a large proportion of American women over the age of 50. The third National Health and Nutrition Examination Survey (NHANES III) estimated that 12 million (41%) white women over age 50 met WHO criterion for osteopenia and 5 million (15%) for osteoporosis (Table 2). The prevalence of osteoporosis in Mexican-American women is similar to white women. While rates in black women are approximately half that of the other groups, they are still substantial (8%). Including all races, an estimated 14 million women over age 50 have osteopenia and over 5 million have osteoporosis.¹³

The prevalence of osteoporosis increases with age for all sites measured. By the WHO definition (low bone density at the hip, spine, or forearm), up to 70% of women over age 80 have osteoporosis (Table 3). Percentages almost double across all sites during the eighth decade and again during the ninth.¹⁴

Using data from the 5% U.S. Medicare sample, the actuarial (life table) risk of a 65-year-old white woman sustaining a fracture by age 90 is 16% for the hip, 9% for distal forearm, and 5% for proximal humerus.² Age is an important factor in the relationship between bone density and the absolute risk of fracture. An increase in age of 13 years increases the risk of hip fracture by the same amount as a one standard deviation decrease in bone density. As illustrated in Table 4, the 5-year risk of hip fracture for a 90-year-old woman with a T-score of -1.0 is 2 %, equivalent to that of a 70-year-old woman with a T-score of -2.0.⁶ Older women have a much higher fracture rate than younger women with the same bone density because of increasing risk from other factors such as bone quality and tendency to fall.¹⁵ Hip fractures are associated with high rates of mortality and loss of independence.^{16, 17}

Sixteen percent of postmenopausal women have osteoporosis of the lumbar spine.¹⁴ Five percent of 50-year-old white women and 25% of 80-year-old women have had at least one vertebral fracture.¹ Although many vertebral fractures are only incidentally detected on x-rays, some cause severe pain leading to 150,000 hospital admissions per year in those over age 65, 161,000 physician office visits, and more than 5 million days of restricted activity in those age 45 or older.¹⁸ The functional impact of vertebral fracture on quality of life can be substantial.¹⁹

The burden of osteoporosis extends beyond the consequences of fracture. The process of diagnosis and treatment can also affect quality of life. An osteoporosis-targeted quality of life questionnaire was developed to assess the impact of the disease among women in the community, specifically focusing on physical difficulty with activities of daily living, necessary adaptations, and fears.²⁰ Using this questionnaire, women with osteoporosis indicated significantly more difficulties with routine daily activities compared to women with osteopenia or normal bone density.²¹ Also, women who had osteoporosis had significantly more fears than women who had normal bone density. It is not clear how comorbidities influence these differences in quality of life.

Health Care Interventions

Low bone density has been used to predict risk for fractures as well as to diagnose osteoporosis. Osteoporosis has been defined as “a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk.”^{22,23} This definition emphasizes that, in addition to bone mass, the structure of bone is also an important factor in the mechanism of fractures. In

practice, however, the diagnosis is most often made in a patient who has had an osteoporotic fracture or who has been found to have low bone density.

In 1994, a World Health Organization (WHO) working group²⁴ noted that there was a wide range in estimates of the prevalence of osteoporosis (2.5% to 95%), depending on what value of bone density was classified as abnormal. Therefore, while it endorsed the earlier definition of osteoporosis, the working group proposed that, *in epidemiologic studies*, osteoporosis should be diagnosed when bone mineral density (BMD) is 2.5 standard deviations below the mean for healthy young adult women at the spine, hip, or wrist, or when a history of an atraumatic fracture is present.²⁵

The number of standard deviation units above or below the young healthy mean is called the T-score. For example, if a 60-year-old woman has a T-score of minus one, her bone density is one standard deviation below the average bone density of healthy women in their twenties. A Z-score is the number of standard deviation units above or below the mean for one's own age group. The woman in the example has a Z-score of zero, meaning that a T-score of minus one is average for a 60-year-old woman.

Ideally, the value of bone density used to define the disease osteoporosis would be selected in order to identify a group of patients known to have a high risk of complications (fractures) and likely to benefit from identification and treatment. The WHO working group chose a T-score of minus 2.5 or less as the criterion. By screening 3 sites, twice as many women would be diagnosed with osteoporosis than by screening at the hip alone. The working group also proposed that low bone mass or osteopenia be diagnosed when bone density was 1.0 to 2.5 standard deviations below the young healthy mean. These diagnostic criteria have been incorporated into bone density reports and in the inclusion criteria for recent randomized

controlled trials of therapies for osteoporosis. Although they were not intended for use as a clinical treatment threshold, they are being used as such.

Recently, several participants in the WHO working group recommended that the diagnosis of osteoporosis should be based only on the T-score obtained at the hip measured by dual-energy x-ray absorptiometry (DXA).²⁶ They proposed that measurements at other sites and with other technologies may be useful for assessing risk for fracture, but should not be used for diagnosis of osteoporosis.

Prior Recommendations

The Second Task Force stated that there was insufficient evidence to recommend for or against routine screening for osteoporosis with bone densitometry in postmenopausal women.¹² The Task Force recommended that all postmenopausal women should be counseled about hormone prophylaxis and be advised on the importance of modifying certain risk factors such as smoking, exercise, and calcium intake. They felt that screening may be appropriate for high-risk women considering treatment decisions.

Analytic Framework and Key Questions

The analytic framework (Figure 1) depicts a screening strategy that includes an assessment of clinical risk factors, then measurement of bone density on a high-risk group.²⁷ We addressed these key questions for the target population of asymptomatic postmenopausal women:

Arrow 1: Does screening using risk factor assessment and/or bone density testing reduce fractures?

Arrow 2: Does risk factor assessment accurately identify women who may benefit from bone density testing?

Arrow 3: Do bone density measurements accurately identify women who may benefit from treatment?

Arrow 4: What are the harms of screening?

Arrow 5: Does treatment reduce the risk of fractures in women identified by screening?

Arrow 6: What are the harms of treatment?

It is important to note that our review has a limited perspective: that of selecting asymptomatic postmenopausal women from the general population for testing and treatment. Many experts would argue that public health efforts and provider attention should be focused on adolescents and should promote measures to increase peak bone mass. Peak bone mass, which is achieved in a woman's twenties, is an important determinant of bone strength throughout life. This review does not address primary prevention of osteoporosis or the impact screening can have relative to the potential impact of primary prevention of osteoporosis.

Chapter 2. Methods

Literature Search Strategy

Relevant studies were identified from multiple searches of MEDLINE (1966 to May 2001), HealthSTAR (1975 to May 2001), and Cochrane databases, reference lists of systematic reviews, and experts. The search strategy is described in Appendix 1. We reviewed a set of Cochrane meta-analyses of treatment trials presented at the National Institutes of Health Consensus Development Conference on Osteoporosis in March 2000.²⁸ In addition, we sent letters to manufacturers of bone measurement devices requesting additional information about the performance of their instruments, but we received no new data. This report was reviewed by 6 content reviewers.

Inclusion/Exclusion Criteria

Two reviewers read each abstract to determine its eligibility. We included English-language abstracts that had original data about postmenopausal women and osteoporosis and that addressed screening, or the effectiveness of risk factor assessment, bone measurement testing, or treatment. Postmenopausal women were those who had experienced surgical or natural menopause, regardless of age. Women with pre-existing atraumatic fractures were not considered in the screening population because they already meet the WHO definition of osteoporosis. We did not include studies of primary prevention of osteoporosis such as the role of nutrition, calcium consumption, and physical activity. We did not review known secondary causes of osteoporosis such as corticosteroid use and certain chronic diseases because these are beyond the scope of population screening. We also did not systematically review data describing

the link between fractures and morbidity and mortality because this relationship has been previously established.

For studies of prediction, we selected articles if they reported the relationship between risk factor assessment methods or bone measurement tests and bone density, bone loss, or fractures. To address treatment issues, we reviewed studies of hormone replacement therapy, selective estrogen receptor modulators (SERMs), and bisphosphonates. We focused on randomized controlled trials of current therapies reporting radiographically verified, nontraumatic fracture outcomes, because fractures are a stronger measure of effectiveness than bone density. Investigators read the full-text version of the retrieved papers and re-applied the initial eligibility criteria. We excluded articles if they did not provide sufficient information to determine the methods for selecting subjects and for analyzing data.

Size of Literature Reviewed

The initial literature search included 6,194 titles and abstracts about risk factors. Of these, 230 were reviewed, and 18 studies about risk factor assessment were included. For bone measurement tests, 2,125 titles and abstracts were initially found, and 85 studies were reviewed. Studies of treatment were initially identified from a Cochrane meta-analysis²⁸ supplemented by MEDLINE searches for alendronate, risedronate, estrogen, and raloxifene. Periodic updates for all topics continued through May 2001.

Literature Synthesis

To assess the internal validity of individual studies, we applied a set of criteria developed by the current U.S. Preventive Services Task Force (Appendix 2).²⁷

We created an outcomes table to summarize the number of hip and vertebral fractures prevented based on age-specific prevalence rates, and treatment effects obtained from results of the reviewed studies. We conducted a sensitivity analysis to determine the influence of risk factors on the number needed to screen.

Chapter 3. Results

Arrow 1: Does screening using risk factor assessment and/or bone density testing reduce fractures?

No studies have evaluated the effect of screening in reducing fractures.

Arrow 2: Does risk factor assessment accurately identify women who may benefit from bone density testing?

Hundreds of studies have reported associations between clinical risk factors and low bone density and fractures in postmenopausal women. Several clinical risk factors are consistently associated with increased risks of low bone density and fractures (advancing age, white race, low weight or weight loss, nonuse of estrogen replacement, history of previous fracture, family history of fracture, history of falls, and low scores on one or more measures of physical activity or function), while others are important in some, but not all, studies (smoking, alcohol use, caffeine use, low calcium and vitamin D intake, and use of certain drugs).²⁹ Fewer studies, however, evaluate how to use these risk factors to identify individual women at risk for low bone density or fracture. Ideally, risk factor assessment would aid clinicians in selecting who should and should not undergo bone density testing, reduce modifiable risk factors, and consider treatment.

Risk Assessment Based on the Likelihood of Low Bone Density

One approach to risk assessment is to identify women at high risk for osteoporosis for bone density testing. This approach assumes that those with low bone density could benefit from identification and treatment.

We identified 10 cross-sectional studies that describe methods of determining risk for individual women based on selected clinical risk factors (Table 5, bone density outcomes).³⁰⁻³⁹ The most common methodologic limitations of these studies were lack of generalizability because of small numbers of subjects or nonrepresentative subjects, and lack of validation.

One risk assessment instrument from a study with a good-quality rating assigned points to selected risk factors for low femoral neck bone density (age, weight, race, estrogen use, presence of rheumatoid arthritis, history of fractures) to create a summary measure referred to as the Simple Calculated Osteoporosis Risk Estimation (SCORE).³⁷ These risk factors were obtained from over 1,200 women from the community and were subsequently tested in a validation group of 259 women. SCORE had an area under the ROC curve of 0.81 in the development group, and a sensitivity of 91% and specificity of 40% in the validation group.

SCORE performed poorly, however, in a cohort from the Rancho Bernardo Study.⁴⁰ In this study, women had a mean age approximately 10 years above that of women in the original SCORE cohorts. A total of 1,013 postmenopausal white women age 44 to 98 years underwent assessment with SCORE protocol and DXA of the femoral neck. Using the recommended SCORE cutpoint of 6, all but 5.5% of the women would have been recommended for bone density testing.

The Osteoporosis Risk Assessment Instrument (ORAI) uses 3 items--age, weight, and current use of hormone replacement therapy--to identify women with low bone density.³⁹ It was developed in a community-based sample of 1,376 women over 45 years of age who did not

report a previous diagnosis of osteoporosis. The prevalence of osteoporosis in the development cohort (n=926), defined in the study as a T-score of -2.5 or lower at the hip or lumbar spine, was 11%. In the validation set (n=450), the three items identified 94.4% of the women who had osteoporosis. The specificity was 41.4%.

A recent study compared the performance of 5 clinical decision rules for bone density testing among 2,365 postmenopausal women aged 45 years or older enrolled in a community-based study of osteoporosis in Canada.⁴¹ These included guidelines from the National Osteoporosis Foundation (NOF),⁴² SCORE,³⁷ ORAI,³⁹ Age, Body Size, No Estrogen (ABONE),⁴³ and body weight criterion (weight <70 kg).³³ In this study, SCORE and ORAI had the highest area under the ROC curves (SCORE: sensitivity 99.6%, specificity 17.9%, ROC 0.80; ORAI: sensitivity 97.5%, specificity 27.8%, ROC 0.79).

Risk Assessment Based on Risk of Fracture

The assumption underlying this approach is that women who have a higher overall risk of fracture have a higher benefit from identification and treatment--i.e. the number needed to treat is lower, and the cost-effectiveness of treatment is higher, for higher-risk patients.

We identified 8 studies of clinical prediction models of fracture risk (Table 5, fracture outcomes).⁴⁴⁻⁵¹ None of these studies received a good rating for internal validity. Four studies evaluated hip fracture outcomes, 2 vertebral fractures, and 2 all types of fractures. These studies determined how well risk factors were associated with fractures known to have occurred already (4 case-control studies), or how well they would predict fractures in the future (4 prospective cohort studies).

The best-performing model for fracture outcomes reported an area under the ROC curve of 0.83.⁴⁶ This model was based on a prospective study of 5,208 subjects whose risk factors were determined at baseline. Hip fracture outcomes were determined 3.8 years later, and risk points were assigned from beta coefficients of regression models. Variables in the model included age, gender, height, use of a walking aid, current smoking, and weight. Adding femoral neck bone density to the model and remodeling all of the parameters improved its performance only slightly. This model has not been tested prospectively in a separate population.

Arrow 3: Do bone density measurements accurately identify women who may benefit from treatment?

Several technologies are available to measure bone (Table 6).⁵²⁻⁵⁵ Single photon absorptiometry (SPA), first described in 1963, was based on the principle of measuring photons absorbed by mineral in the tissues.⁵⁶ More recent densitometry techniques, such as dual-energy x-ray absorptiometry (DXA), are based on the same principle, but have the advantages of access to axial sites, better precision, lower radiation exposure, shorter examination time, reduced influence of soft tissue thickness, and a more convenient or reliable source of photons. Results of these densitometry tests are expressed as grams of mineral in a projected area (g/cm^2). Quantitative computed tomography (QCT), an alternative method of measuring axial bone density, is expressed as the grams of mineral in a volume of bone (g/cm^3). QCT also provides a computed tomography image that radiologists can use to assess bone architecture and structural integrity.

Densitometry devices that measure peripheral bone density are considerably less expensive to buy and use than axial DXA. Radiographic absorptiometry and quantitative

microdensitometry use computer software to estimate bone density from conventional radiographs of the hand. Other devices for measuring bone density in the arm or heel include single-energy x-ray absorptiometry (SXA), peripheral dual-energy x-ray absorptiometry (pDXA), and peripheral quantitative computed tomography (pQCT).

Quantitative ultrasound (QUS) devices report the way that bone attenuates sound waves and/or the speed with which sound travels through the bone.⁵⁷ Commercial devices measure the "broadband ultrasonic attenuation" (BUA, expressed as dB/MHz), the speed of sound (SOS), and the "stiffness," a measure derived from the BUA and SOS. Quantitative ultrasound does not measure bone mineral content and is categorized separately from the other technologies. While these measures are not highly correlated with measures of bone density made by DXA, some *in vitro* studies⁵⁸⁻⁶¹ but not all⁶² suggest that QUS might reflect other aspects of bone structure that could be associated with fragility.

Accuracy and Reliability of Tests

The accuracy of densitometry tests in everyday practice has not been studied outside research protocols and major referral centers. Patient factors such as obesity, handedness, and edema (for ultrasound),⁶³⁻⁶⁸ and osteoarthritis (for DXA of the spine)⁶⁹ affect estimates of bone density.

Correlations among different bone density devices are low (0.35 to 0.60).⁷⁰⁻⁹⁰ When used in the same patients, DXA machines from different makers differ by 6 to 15%⁹¹⁻⁹⁶ in the proportion of patients diagnosed to have osteoporosis.

Published studies consistently show that the probability of being diagnosed with osteoporosis depends on the choice of test and site.⁹⁷⁻¹⁰¹ One analytical study, for example, found that, in the NHANES III sample, 6% of women over 60 years old would be diagnosed to have

osteoporosis if DXA of the total hip were used as the only test, versus 14% for DXA of the lumbar spine, 3% for QUS, and 50% for QCT.⁹⁸

The likelihood of being diagnosed with osteoporosis also depends on the number of sites tested. Testing in the forearm, hip, spine, or heel will generally identify different groups of patients. A physician cannot say, based only on a forearm test, that the patient "does not have osteoporosis." Conversely, although the results of a test at any site are associated to some degree with fractures at other sites, the physician may not be able to assess whether the patient who has a low T-score on a hand or forearm test has significant bone loss at other sites.

Prediction of the Short-term Risk of Fractures

The literature describing the performance of bone measurement tests to predict fractures is extensive. We focused our review on a meta-analysis that provided a summary of older studies, and on prospective cohort studies not included in the meta-analysis.

The meta-analysis assessed 23 publications from 11 separate prospective cohort studies published before 1996.¹⁰² Studies were pooled to estimate the age-adjusted relative risk of various types of fractures for a one standard deviation decrease in bone density. Nearly all available data included in these studies were from women in their late 60s or older and all tests used densitometry techniques. Results of the meta-analysis indicated that DXA measured at the femoral neck predicted hip fracture better than measurements made at other sites, and was comparable to forearm measurements for predicting fractures at other sites.¹⁰³⁻¹⁰⁵ For bone density measurements made at the femoral neck, the pooled relative risk per one standard deviation decrease in bone density was 2.6 (2.0-3.5).

To update these results, we identified 15 cohort studies either not included in the meta-analysis or published since 1996 of peri- or postmenopausal women who had DXA, QUS, radiographic absorptiometry (including quantitative microdensitometry), or other bone measurement tests that reported fracture outcomes.^{103, 105-118} The cohorts include the Study for Osteoporotic Fractures (SOF), Epidemiologie de L'Osteoporose (EPIDOS), Hawaii Osteoporosis Study (HOS), Diagnostisch Onderzoek Mammacarcinoom (DOM), two projects from Aberdeen, Scotland, and one each from Rotterdam, Netherlands, Kuopio, Finland, and Modena, Italy. Details of these studies are reported elsewhere.²⁹

For this review, we included studies that reported DXA of the hip for the prediction of hip fractures (Table 7).^{103, 105-110, 112, 114, 116, 117} Rather than reporting details of individual studies, we reported results from the cohorts according to the age groups addressed. We confined all analyses of prospective studies to the time horizon actually observed, rather than extrapolating to 5-year or lifetime risk. We described the probability of a fracture in subjects classified as "high risk" or "low risk" according to their hip DXA result. Classification of "high" and "low" risk varied across studies, and was based on the cutoff value for the test. In one analysis of the SOF cohort, for example, the overall risk of hip fracture during approximately 2 years of followup was about 1% (9 per 1,000). If the hip DXA indicated osteoporosis, the probability of a hip fracture was 2.3%; if not, the probability was 0.46% (about 5 per 1,000). For the average patient in this cohort, then, low bone density increased the chance of fracture within 2 years from 9 to 23 per 1,000, while an osteopenic or normal result decreased the probability of fracture from 9 to 5 per 1,000. The limitations of measuring these differences in probabilities are that they depend on the length of followup and the pretest probability of fracture.

DXA of the hip predicted hip fracture best in the SOF and Rotterdam studies (relative risk per one standard deviation decrease 2.6 [CI 1.9-3.8] and 2.5 [CI 1.8-3.6]). Prediction was not as good among women over 75 years of age. Two recent studies included younger, perimenopausal women between the ages of 45 and 56, but these subjects had too few hip fractures to estimate the relative risk. The adjusted relative risk for all nonvertebral fractures was 1.39 (CI 1.2-1.6) in one study¹¹² and 1.4 (1.3-2.4) in the other,¹¹⁷ which is comparable to the relative risks for all nonvertebral fractures in older women. The 2-year probability of fracture in the group with high bone density was 2.4%, versus 1.7% in the group with low bone density. The low difference (7 per 1,000) reflects the fact that bone density testing provides less information about the short-term risk of fracture in younger, lower-risk individuals.

Arrow 4: What are the harms of screening?

There are several potential harms of screening for women, however, very few studies have been published. Receiving a test result indicating osteoporosis could produce anxiety and perceived vulnerability¹¹⁹ that may be unwarranted. Women with osteoporosis voiced significantly more fears than women who had normal bone density on a quality of life questionnaire.²¹

Some clinicians believe that the results of bone measurement tests motivate patients to exercise, adhere to medication regimens, and change other behaviors to reduce their risk of fracture. Conversely, women informed that they do not have osteoporosis may interpret this to mean that they do not need to engage in preventive measures such as exercise, calcium intake, and smoking cessation. Some women may be falsely reassured that abnormal results from a hip DXA scan from last year appear "improved" by a normal calcaneal ultrasound report this year.

The potential time, effort, expense, and radiation exposure of repeated scans over a lifetime have not yet been determined.

A limited number of studies have evaluated the impact bone density test results have on women including adherence to therapy. Of 1,335 women from the SOF cohort who completed a questionnaire about estrogen therapy and were taking estrogen, 34% reported their primary reason was prevention or treatment of osteoporosis.¹²⁰ In a randomized trial of 141 women within 3 years of menopause who were referred from 3 private practices, those who underwent bone density testing were more likely to fill a prescription for HRT than women who received an educational message about osteoporosis (63% vs. 20%, $p < 0.05$).¹²¹ Another study found that test results appear to influence physicians' decisions to prescribe HRT.¹²²

A randomized trial of screening for osteoporosis measured quality-of-life and use of hormone replacement therapy 2 years later.¹²³ A total of 1,600 women aged 45 to 54 years were randomly selected from a population-based registry in Aberdeen, Scotland. In the group assigned to screening, 576 (72%) responded to an invitation to have bone density testing. Two years later, use of hormone replacement therapy was higher in the screened group (30% vs. 24%), and was highest (43%) among screened women who had bone density in the lowest quartile. There were no differences between screened and unscreened women in any aspect of quality of life, menopausal symptoms, anxiety, or frequency of falls.

A survey of 261 women who received densitometry studied their perception of risk of fracture and degree of worry, and whether they made changes based on their test results.¹²⁴ Of the 53% who reported that their test was below normal, virtually all of them (94%) reported initiating preventive measures. Fifty-six percent of those who reported their test was normal also initiated prevention measures. Women who reported below normal test results were more likely

to start HRT than those with normal or above results. Of the entire group, 24% who reported below normal results said they began limiting activities to avoid falling; for women 65 years and older, this proportion rose to 31%. About one-third of the sample (86 women) had a second bone measurement test. About 26 women reported losing bone at an accelerated rate; of these, 22 initiated additional prevention.

Potential harms may also arise from inaccuracies and misinterpretations of bone measurement technology. The variation between techniques, along with the lack of methods to integrate bone density results with clinical predictors, makes it difficult for clinicians to provide accurate information to patients about their test results. One randomized trial examined the effect of densitometry on the practice patterns of 57 primary care physicians who ordered DXA tests for their patients, and also their understanding of the reports they received.¹²⁵ Physicians found densitometry reports confusing and were not confident that their interpretations of T-scores were accurate. This was especially true among those who received a short technical report compared to a longer written report.

Arrow 5: Does treatment reduce the risk of fractures in women identified by screening?

Although several forms of treatment have been studied, we focused on recent trials of three types of agents: estrogen, SERMs, and bisphosphonates.

Estrogen Replacement Therapy

A recent meta-analysis of 22 trials of estrogen reported an overall 27% reduction in nonvertebral fractures (RR 0.73, 0.56-0.94).¹²⁶ Several trials included in the meta-analysis,

however, do not meet inclusion criteria for our review because they used unpublished data, did not verify fractures radiographically, included traumatic fractures, or included women who were hospitalized or had secondary causes of osteoporosis. Five randomized controlled trials of estrogen with vertebral or nonvertebral fracture outcomes met our inclusion criteria (Table 8).¹²⁷⁻

131

A trial of 78 postmenopausal women age 47 to 75 years with one or more pre-existing vertebral fractures evaluated incident vertebral fractures.¹²⁷ The treatment group was provided with a cyclic regimen of transdermal estrogen and progesterone for one year and was compared to an untreated placebo group. The estrogen group experienced 8 new vertebral fractures in 7 women, while the placebo group had 20 fractures in 12 women. Despite a lower vertebral fracture rate in the estrogen group (RR 0.39; CI 0.16-0.95), the number of women experiencing new vertebral fractures was not significantly different between groups. A smaller trial of 4 years duration comparing 18 women using a cyclic oral estrogen regimen to 18 women in a comparison group found no significant difference in vertebral fractures.¹²⁸

Three trials reported nonvertebral fracture outcomes. A primary prevention trial enrolled a subgroup of a large prospective osteoporosis study based in Finland.¹²⁹ In this study, 464 early postmenopausal women without osteoporosis were randomly assigned to one of four groups: cyclic oral estradiol with progestin, vitamin D alone, estradiol with progestin and vitamin D, or placebo. New, symptomatic, radiographically confirmed nonvertebral fractures were recorded during a mean 4.3 years of follow-up. The risk for fracture was significantly lower for the estrogen/progestin alone group (RR 0.29; CI 0.10-0.90), but not for the estrogen/progestin and vitamin D group, or the vitamin D alone group, compared to placebo when adjusted for baseline bone density and prior fractures. Another primary prevention trial randomized 1,006 early

postmenopausal women (age 45 to 58 years) in Denmark to oral estradiol/norethisterone or placebo. After 5 years, the relative risk for all types of fractures was 0.82 (0.53-1.29), and for forearm fractures 0.40 (0.16-1.01).¹³⁰

The Heart and Estrogen/progestin Replacement Study (HERS) is a secondary prevention trial of the effects of estrogen on cardiovascular outcomes.¹³² This study enrolled 2,763 postmenopausal women with pre-existing coronary disease under 80 years old (mean 66.7 years). A subgroup of women underwent bone densitometry, and 15% of them had osteoporosis. Fractures at various sites were secondary outcomes. The treatment group was given a continuous combined regimen of conjugated estrogen with medroxyprogesterone per day and compared with an equal-sized placebo group. After 4 years, this study found no difference between groups for all fractures combined, or hip, wrist, spine, or other types of fractures specifically.¹³¹

These trials did not meet USPSTF criteria to be ranked as good-quality studies because they did not assemble or maintain comparable groups,^{130, 131} were not blinded,^{129, 130} were small,¹²⁷⁻¹²⁹ or used inappropriate analyses.^{127, 128} The largest trial, HERS, did not monitor for asymptomatic incident vertebral fractures, potentially missing as many as 2/3 of vertebral fractures that would be diagnosed solely by radiographic morphometric criteria.¹³³ These limitations in study design and generalizability led to inconclusive evidence about the effectiveness of estrogen in fracture prevention.

Selective Estrogen Receptor Modulators

Two good-quality randomized controlled trials of raloxifene with fracture outcomes have been published (Table 8). The largest study, the Multiple Outcomes of Raloxifene Evaluation

(MORE) study, included 7,705 postmenopausal women aged 31 to 80 years.¹³⁴ These women met WHO criteria for osteoporosis based on low bone density or presence of vertebral fractures. Incident vertebral fractures were determined using radiographic criteria at the 24- and 36- month visits and at other times if new symptoms of vertebral fractures developed. Nonvertebral fractures were determined by interviewing subjects at 6-month visits.

After 3 years of treatment, women in the raloxifene group had a significantly reduced risk for vertebral fractures compared with women in the placebo group (RR 0.59; CI 0.50-0.70). The risk for nonvertebral fractures was not significantly reduced. A smaller trial of one-year duration also evaluated the effects of raloxifene but found no differences in fracture outcomes compared to placebo.¹³⁵

Women enrolled in the National Surgical Adjuvant Breast and Bowel Project Breast Cancer Prevention Trial (P-1) of tamoxifen were also monitored for fractures, although this was a secondary outcome.¹³⁶ In this study, 13,388 women were randomized to tamoxifen or placebo and followed for 5 years. Incident fractures of the hip, wrist, and spine were confirmed by x-rays. Relative risks for total fractures, hip, wrist, and spine were not significantly reduced.

Bisphosphonates

A recent unpublished meta-analysis²⁸ of 11 randomized trials¹³⁷⁻¹⁴⁷ including 12,855 women found that alendronate significantly reduced vertebral fractures (RR 0.52, 95% CI 0.43-0.65), forearm fractures (RR 0.48, 95% CI 0.29-0.78), and all nonvertebral fractures (RR 0.50, 95% CI 0.38-0.69). There was a non-significant trend toward reduction in hip fractures.

We evaluated data from these trials to determine if women who have a similar overall risk of fracture, but different bone densities, have a similar benefit from treatment. The published

studies are summarized in Table 9. This question is clinically important because the lack of accepted criteria for initiating treatment remains a problem. The meta-analysis review found no relation between the average bone density and the effect size of alendronate, but the researchers did not have individual patient data. None of the studies stratified patients by overall risk, or published subgroup analyses of the effect size in relation to overall fracture risk.

The Fracture Intervention Trial (FIT) of alendronate was conducted with 2 different groups of participants and provides some information about levels of risk. One group (FIT-I) included a higher-risk group of 2,027 women who had T-scores of -1.6 or lower and pre-existing vertebral fractures.¹³⁷ The 3-year risk of hip fracture in the placebo group was 2.2% (1.1% in the alendronate group, relative hazard 0.49 [0.23-0.99]), and the 3-year risk of any clinical fracture was 18.2% (13.6% in the alendronate group, RH 0.72 [0.58-0.90]). Within FIT-I, the highest-risk women—those who were oldest, had the most vertebral fractures at baseline, or had the lowest bone density—had the largest absolute benefit from treatment.¹⁴⁸ The numbers needed to treat with alendronate for 5 years to prevent one new vertebral fracture were 8 for women aged 75 years or older compared with 9 for women younger than 75 years, and 4 for women with 2 or more existing vertebral fractures compared with 16 for women with one existing vertebral fracture.

A second study from FIT (FIT-II) included a lower-risk sample of 4,432 women who also had T-scores of -1.6 or lower, but did not have pre-existing vertebral fractures.¹⁴⁹ The 4-year incidences of hip fracture (1.1%) and any clinical fractures (14.1%) in the placebo group were fewer than those of the women in the placebo group of the FIT-I study. In FIT-II, only the subgroup of treated patients (n=1,627) who had a T-score under -2.5 had a significant risk

reduction for all clinical fractures, from 19.6% to 13.1% (RR 0.64; 0.50-0.82). There was no risk reduction for fractures for patients who had T-scores between -1.6 and -2.5.

The results from FIT suggest that women with more risk factors for fracture, such as those who are older, have very low bone density, or have pre-existing vertebral fractures, have the greatest absolute benefit with treatment. However, FIT did not examine other nonskeletal risk factors, such as psychomotor impairment, poor gait, and other factors that increase the risk of falling.

One recent, fair-quality randomized trial examined the effect of these factors on the benefit of treatment with another bisphosphonate, risedronate.¹⁵⁰ Risedronate had no effect on hip fracture rates among women 80 years of age or older who had one or more risk factors for falls. These women did not necessarily have osteoporosis. As the authors note, many of these women had only one risk factor, so the result does not exclude the possibility that women who have multiple risk factors for falls would benefit from treatment. In the same report, in women aged 70 to 79 with severe osteoporosis (T score < -3 or worse), risedronate reduced hip fractures by 40% (0.6; CI 0.4-0.9; NNT 77). Many of these women had risk factors for falls, but the report did not say whether risedronate was more or less effective for them than for women who had the same degree of osteoporosis but did not have risk factors.

Generalizability of Randomized Trials

For the results of trials to be applicable to a screening program, the trials must include patients who would be identified by screening the general population. We examined recruitment and eligibility characteristics of the 11 randomized trials of alendronate to assess whether selection biases or other biases might affect their generalizability. Overall, the trials included

relatively healthy women who were not using estrogen. Except for the 2 trials of early postmenopausal women who were not osteoporotic, most of the subjects were older than 65 years of age.

FIT-II is the largest study and provided the most detailed description of recruitment and results. In FIT-II, recruitment of the sample of 4,432 women began with a query that was mailed to over a million women selected from the general population in 11 cities. Women who had medical problems, including dyspepsia, or who used estrogen were excluded. Fifty-four thousand (about 5.4%) responded by telephone, of whom 26,137 (52%) had a screening visit. A higher than expected proportion of these (65%) had sufficiently low bone density to enroll in the study. Of these, 57% were classified as "ineligible, did not wish to continue, or screened after recruitment to this arm"; it is not clear from this description how many patients did not meet the eligibility criteria. In addition, an unspecified number of patients (up to 28,000) were found to be ineligible at the initial stage of recruitment. The demographic characteristics of eligible and screened, but excluded, subjects were not reported. None of the other randomized trials disclosed any details of how the sample was recruited or how many respondents were found to be ineligible before randomization.

In other clinical areas, the results of industry-sponsored trials were significantly more favorable to newer therapies than trials funded by nonprofit organizations.^{151, 152} Because all 11 trials of alendronate were funded wholly or in part by the maker, we were unable to assess the influence of sponsorship on effect size. If effectiveness of treatments is smaller than estimated in these trials, the efficiency of screening to identify candidates for treatment will be reduced and the number needed to screen will increase.

Arrow 6: What are the harms of treatment?

Potential risks of estrogen use include thromboembolic events,¹⁵³ cholecystitis,¹⁵⁴ endometrial cancer in those with a uterus taking unopposed estrogen,¹⁵⁵ and possibly ovarian cancer,¹⁵⁶ breast cancer (particularly in long-term users),¹⁵⁷ and thromboembolic stroke.¹⁵⁸ Some women may experience effects such as breast tenderness, vaginal bleeding, and mood changes, among others. Both raloxifene and tamoxifen are associated with thromboembolic events, leg cramps, and hot flashes.¹⁵⁹

Overall, gastrointestinal side effects occur in about 25% of patients taking alendronate, but these rates were usually not higher, or only a little higher, than those for placebo. Higher rates have been observed among Medicare enrollees taking alendronate.¹⁶⁰ In randomized trials, rates of ulcers are higher for patients taking alendronate; in the FIT II trial, 2.2% of alendronate patients developed ulcer disease, versus 1.2% in the placebo group ($p < 0.05$). The long-term adverse effects of alendronate are unknown.

Chapter 4. Discussion

Summary of Evidence Quality

Table 10 summarizes the evidence obtained for this systematic review by indicating the type of study design and quality of evidence for each key question, using criteria developed by the USPSTF.²⁷ Quality ratings include scores for both internal validity (the strength of individual studies) and external validity (the extent to which studies are generalizable to a primary care population).

Outcomes Table

To estimate the effect of screening for osteoporosis on reducing hip and vertebral fractures in 10,000 postmenopausal women, we created an outcomes table based on assumptions from the reviewed studies (Table 11). These estimates include age-specific prevalence rates expressed in 5-year age intervals¹⁶¹ and treatment effects based on trials (37% risk reduction for hip fracture, 50% for vertebral fracture).^{150,28,137,162} We estimated an adherence rate of 70% based on reports of adherence and side effects from treatment trials and allowing for less optimal adherence in the general population.

Using the assumptions in the table, if 10,000 65-to-69-year old women with an osteoporosis prevalence rate of 0.120 underwent bone densitometry (DXA of the femoral neck), 1,200 would be identified as high-risk (T-score <-2.5). If these women were offered treatment that resulted in a 37% reduction of hip fracture risk and 50% reduction of vertebral fracture risk, and 70% of them adhered to therapy, then 14 hip fractures and 40 vertebral fractures would be prevented over 5 years. The number of women in this age group needed to screen to prevent 1 hip fracture in 5 years is 731, and the number of women with low bone density needed to treat is

88. The number needed to screen to prevent 1 vertebral fracture is 248, and the number needed to treat is 30. These numbers become more favorable in the older age groups because the prevalence of osteoporosis increases steadily with age. For women age 75 to 79, the number needed to screen is 143, and the number needed to treat is 41 for hip fractures.

The literature review indicated that the prevalence of osteoporosis, the predictability of densitometry, and the effectiveness of treatment are less, or may be less, for younger than for older postmenopausal women. To determine if considering clinical risk factors when screening younger postmenopausal women is useful, we also included risk estimates for clinical risk factors in a sensitivity analysis. Because studies of risk assessment instruments that we reviewed included predominantly older women, we determined risk estimates by reviewing 7 observational studies of risk factors and fractures specifically conducted in populations with at least 50% of subjects under age 65.¹⁶³⁻¹⁶⁹

The 3 most consistent predictors of fracture in these studies were increasing age, low weight or body mass index, and nonuse of hormone replacement therapy (defined by current, ever, or certain durations of use). These are also the 3 variables used in the Osteoporosis Risk Assessment Instrument (ORAI) to identify women with low bone density,³⁹ and the variables most strongly associated with low bone density in a study enrolling mostly younger postmenopausal women in the U.S.¹⁷⁰ Based on these studies, we estimated that having one of these risk factors increased the probability of having osteoporosis by up to 100% and the risk of fracture by 70% (relative risk 1.7).

For the younger age groups, the presence of clinical risk factors influences the outcomes. For example, only 5 hip fractures are prevented over 5 years when screening all women age 60 to 64 years, but this increases to 9 if a risk factor increasing the risk of fracture by 70% is present.

For women 60 to 64 years old with a risk factor, the number needed to screen (1092) and number needed to treat (72) to prevent hip fractures approach those of the 65-to-69-age group (Figure 2).

Conclusions

Although many studies have been published about osteoporosis in postmenopausal women, there have been no trials of screening and, therefore, no direct evidence that screening improves outcomes. Instruments developed to assess clinical risk factors for low bone density or fractures generally have moderate-to-high sensitivity and low specificity, many have not been validated, and none have been widely tested in a practice setting. Among different bone density tests measured at various sites, bone density measured at the femoral neck by dual-energy x-ray absorptiometry is the best predictor of hip fracture and is comparable to forearm measurements for predicting fractures at other sites. Women with low bone density have a 40 to 50% reduction in fracture risk when treated with raloxifene (vertebral fractures) or bisphosphonates (both vertebral and nonvertebral fractures). Trials of estrogen are inconclusive because of methodologic limitations.

Support for population screening would be based on evidence that the prevalences of osteoporosis and fractures increase with age, that the short-term risk of fracture can be estimated by bone measurement tests and risk factors, and that the fracture risk among women with low bone density can be significantly reduced with treatment. When applying these data to outcomes tables of screening strategies, estimates of the numbers of women needed to screen and treat to prevent fractures can be determined. Age-based screening is supported by prevalence data; i.e., the number needed to screen to prevent fractures decreases sharply as age and prevalence increase. Use of risk factors to screen younger women may identify additional high-risk women and provide number needed to screen estimates comparable to screening older women without

risk factors. These findings relate to screening asymptomatic women only, and do not apply to women who would be considered for testing based on pre-existing or incident fractures or presence of secondary causes of osteoporosis.

There are several limitations to this approach, however, and results from a well-designed trial of screening strategies will supercede our estimations based on indirect evidence. Our estimates in the outcomes table are limited by assumptions that are arguable or highly variable by patient and setting. Our treatment effect and compliance assumptions are especially optimistic and reflect results of clinical trials, not clinical practice. We chose a 5-year time horizon based on the short-term predictability of bone measurement tests as well on results of short-term treatment trials. Long-term outcomes may provide a more accurate estimate of benefits. Also, we cannot exclude the possibility that harms outweigh benefits, particularly since the long-term effects of bisphosphonates are not yet known.

Limitations of the Literature

Evidence upon which we based our conclusions is also limited. Overall, evidence is generally stronger for women older than 65 years than younger because more research has been done in older age groups. Bone loss in the peri- and early postmenopausal years is important to long-term bone health, but few published studies address screening and treatment for younger postmenopausal women. Also, younger women are more likely to experience fractures of the radius, vertebrae, and other sites than of the hip. Most studies currently focus on hip fracture outcomes.

Similarly, studies of non-white women are limited. Prevalence data indicate that white and Asian women have the highest rates of osteoporosis,¹⁷¹ and white women are generally 2 to 3

times more likely than non-white women to suffer a hip fracture.¹⁷² Rates in African American women are roughly half of other groups.^{13, 173} Although limited in number, studies of risk factors for fracture indicate similarities across racial groups.^{164, 174, 175} However, no bone measurement studies or treatment trials include large numbers of non-white women. It may be difficult to provide race-specific screening recommendations in the absence of more evidence.

The role of clinical risk factors is still unclear. Although many risk factors are associated with osteoporosis and fractures, how to use them to select women to test or treat is uncertain. The risk factors identified by our literature review and used in the outcomes table are only best estimates. It may be that other risk factors will prove to be equally predictive when used for screening purposes.

Future Research

Further validation of existing risk assessment instruments or development of new ones would be useful. Few studies have evaluated the effect of altering modifiable risk factors such as smoking cessation, strength and balance training, and visual correction. These interventions may prove to be as effective as drug therapy in preventing fractures, and may also be important effect modifiers that would confound the observed effectiveness of the treatments.

The use of peripheral bone measurement tests in screening, either alone or as part of a sequential approach, has not been well studied. Most treatment trials use hip DXA as entry criteria and results may not apply to women diagnosed by other tests. Further research is needed to define the appropriate use of these technologies.

How frequently to screen has also not been specifically studied, but data are needed to determine optimal screening intervals. Estimations can be made based on the age-specific

prevalence of osteoporosis and precision of bone measurement devices. Less frequent testing for younger postmenopausal women when prevalence is lower (e.g., 5-year intervals), and more frequent testing for older women (e.g., 2-year intervals) might be reasonable. Screening intervals less than 2 years seem unwarranted because the precision error of densitometry would likely exceed the estimated bone loss in such a brief time interval.¹⁷⁶ Once a woman is screened and determined to have osteoporosis, future bone measurement testing for screening purposes would not be necessary.

Many gaps remain in the evidence for osteoporosis screening. Although the effectiveness of population screening has yet to be demonstrated, current literature provides support for screening based on evidence that osteoporosis and fractures increase with age, that short-term risk of fracture can be estimated by bone measurement tests and risk factor assessment, and that treatment may reduce fracture risk among women with low bone density. Osteoporotic fractures present an enormous health burden on an expanding elderly population, underscoring the importance of further research to more accurately determine the benefits and harms of screening.

Acknowledgements

This systematic evidence review was prepared for the Agency for Healthcare Research and Quality (contract #290-97-0018, task order no. 2) to be used by the U.S. Preventive Services Task Force. Task Force members Steven Woolf, MD, MPH and Janet Allen, PhD, RN, served as liaisons. Oregon Health & Science University, Evidence-based Practice Center staff who contributed to this project included Peggy Nygren MA, research associate, Susan Wingenfeld, administrative assistant, and Patty Davies, MA, librarian.

References

1. Melton LJ, 3rd, Kan SH, Frye MA, Wahner HW, O'Fallon WM, Riggs BL. Epidemiology of vertebral fractures in women. *Am J Epidemiol*. 1989;129:1000-11.
2. Barrett JA, Baron JA, Karagas MR, Beach ML. Fracture risk in the US Medicare population. *J Clin Epidemiol*. 1999;52:243-9.
3. Swedish Council on Technology Assessment in Healthcare. Bone density measurement-a systematic review. *J Intern Med*. 1997;241:1-60.
4. Hailey D, Sampietro-Colom L, Marshall D, et al. INAHTA project on the effectiveness of bone density measurement and associated treatments for prevention of fractures. Statement of findings. Edmonton, Alberta, Canada: Alberta Heritage Foundation for Medical Research, on behalf of the International Network of Agencies for Health Technology Assessment; 1996.
5. Green CJ, Bassett K, Foerster V, Smith LL. Bone mineral density testing: does the evidence support its selective use in well women? Vancouver, British Columbia, Canada: British Columbia Office of Health Technology Assessment; 1997.
6. National Osteoporosis Foundation. Osteoporosis: Review of the evidence for prevention, diagnosis, and treatment and cost-effectiveness analysis. *Osteoporos Int*. 1998;Suppl 4:S7-S88.
7. Homik J, Hailey D. Quantitative ultrasound for bone density measurement. Edmonton, Alberta, Canada: Alberta Heritage Foundation for Medical Research; 1998.
8. Homik J, Hailey D. Selective testing with bone density measurement. Edmonton, Alberta, Canada: Alberta Heritage Foundation for Medical Research; 1999.

9. Royal College of Physicians of London. Osteoporosis: Clinical guidelines for prevention and treatment. London; 1999.
10. Blue Cross and Blue Shield Association. Ultrasonography of the heel for diagnostic osteoporosis and selecting patients for pharmacologic treatment. Chicago; 1999.
11. Siebzehner MI. Consensus statement on prevention and treatment of osteoporosis. Israel Center for Technology Assessment, Israel Medical Association. 1999.
12. US Preventive Services Task Force. Guide to clinical preventive services. Report of the US Preventive Services, 2nd Ed. Baltimore: Williams & Wilkins; 1996.
13. Looker AC, Wahner LW, Dunn WL, Smith, et al. Updated data on proximal femur bone mineral levels of US adults. *Osteoporos Int*. 1998;8:468-89.
14. Melton LJ, 3rd. How many women have osteoporosis now? *J Bone Miner Res*. 1995;10:175-7.
15. Heaney RP. Bone mass, bone loss, and osteoporosis prophylaxis. *Ann Intern Med*. 1998;128:313-4.
16. White B, Fisher W, Lacerin C. Rate of mortality for elderly patients after fracture of hip in the 1980s. *J Bone Joint Surg*. 1987;69A:1335-40.
17. Cauley J, Thompson D, Ensrud K, Scott J, Black D. Risk of mortality following clinical fractures. *Osteo Intern*. 2000;11:556-61.
18. Melton LJ, 3rd. Epidemiology of spinal osteoporosis. *Spine*. 1997;22:2S-11S.
19. Gold DT. The clinical impact of vertebral fractures: quality of life in women with osteoporosis. *Bone*. 1996;18:S185-S189.
20. Chander JM, Martin AR, Girman C, et al. Reliability of an osteoporosis-targeted quality of life survey instrument for use in the community: OPT-QOL. *Osteoporos Int*. 1998;8:127-35.

21. Lyles KW, Gold DW, Shipp KM, Smith KK. Association of osteoporotic vertebral fractures with impaired functional status. *Am J Med.* 1993;94:595-601.
22. Consensus development conference: prophylaxis and treatment of osteoporosis. *Am J Med.* 1991;90:107-110.
23. Consensus development conference: diagnosis and treatment of osteoporosis. *Am J Med.* 1993;94:646-50.
24. World Health Organization Technical Report Series. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. 1994;843:1-129.
25. Kanis JA. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. *Osteoporos Int.* 1994;4:368-81.
26. Kanis JA, Gluer CC, Advisors. An update on the diagnosis and assessment of osteoporosis with densitometry. *Osteoporos Int.* 2000;11:192-202.
27. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force. *Am J Prev Med.* 2001;20:21-35.
28. Guyatt G. Systematic reviews of osteoporosis therapies: NIH Consensus Development Conference on Osteoporosis Prevention, Diagnosis and Therapy. Bethesda, MD: National Institutes of Health; 2000.
29. Nelson H, Morris C, Kraemer D, et al. Osteoporosis in postmenopausal women: Diagnosis and monitoring. Rockville, MD: Agency for Healthcare Research and Quality; 2001.
30. Falch JA, Sandvik L, Van Beresteijn EC. Development and evaluation of an index to predict early postmenopausal bone loss. *Bone.* 1992;13:337-41.

31. Verhaar HJ, Koele JJ, Neijzen T, Dessens JA, Duursma SA. Are arm span measurements useful in the prediction of osteoporosis in postmenopausal women? *Osteoporos Int.* 1998;8:174-6.
32. Elliot JR, Gilchrist NL, Wells JE, Ayling E, Turner J, Sainsbury R. Historical assessment of risk factors in screening for osteopenia in a normal Caucasian population. *Aust N Z J Med.* 1993;23:458-62.
33. Michaelsson K, Bergstrom R, Mallmin H, Holmberg L, Wolk A, Ljunghall S. Screening for osteopenia and osteoporosis: selection by body composition. *Osteoporos Int.* 1996;6:120-6.
34. Ballard PA, Purdie DW, Langton CM, Steel SA, Mussurakis S. Prevalence of osteoporosis and related risk factors in UK women in the seventh decade: osteoporosis case finding by clinical referral criteria or predictive model? *Osteoporos Int.* 1998;8:535-9.
35. Ribot C, Pouilles JM, Bonneau M, Tremollieres F. Assessment of the risk of post-menopausal osteoporosis using clinical factors. *Clin Endocrinol.* 1992;36:225-8.
36. Slemenda CW, Hui SL, Longcope C, Wellman H, Johnston CC, Jr. Predictors of bone mass in perimenopausal women. A prospective study of clinical data using photon absorptiometry. *Ann Intern Med.* 1990;112:96-101.
37. Lydick E, Cook K, Turpin J, Melton M, Stine R, Byrnes C. Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. *Am J Manag Care.* 1998;4:37-48.
38. Goemaere S, Zegels B, Toye K, et al. Limited clinical utility of a self-evaluating risk assessment scale for postmenopausal osteoporosis: lack of predictive value of lifestyle-related factors. *Calcif Tissue Int.* 1999;65:354-8.

39. Cadarette SM, Jaglal SB, Kreiger N, McIsaac WJ, Darlington GA, Tu JV. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *Can Med Assoc J.* 2000;162:1289-94.
40. Von Muhlen D, Visby Lunde A, Barrett-Connor E, Bettencourt R. Evaluation of the simple calculated osteoporosis risk estimation (SCORE) in older Caucasian women: the Rancho Bernardo study. *Osteoporos Int.* 1999;10:79-84.
41. Cadarette S, Jaglal S, Murray T, McIsaac W, Joesph L, Brown J. Evaluation of decision rules for referring women for bone densitometry by dual-energy x-ray absorptiometry. *JAMA.* 2001;286:57-63.
42. National Osteoporosis Foundation. Physician's guide to prevention and treatment of osteoporosis. Washington, D.C.: Excerpta Medica, Inc.; 1999.
43. Weinstein L, Ullery B. Identification of at-risk women for osteoporosis screening. *Am J Obstet Gynecol.* 2000;183:547-549.
44. van Hemert AM, Vandenbroucke JP, Birkenhager JC, Valkenburg HA. Prediction of osteoporotic fractures in the general population by a fracture risk score. A 9-year follow-up among middle-aged women. *Am J Epidemiol.* 1990;132:123-35.
45. Ranstam J, Elffors L, Kanis JA. A mental-functional risk score for prediction of hip fracture. *Age Ageing.* 1996;25:439-42.
46. Burger H, de Laet CE, Weel AE, Hofman A, Pols HA. Added value of bone mineral density in hip fracture risk scores. *Bone.* 1999;25:369-74.
47. Tromp AM, Smit JH, Deeg DJ, Bouter LM, Lips P. Predictors for falls and fractures in the Longitudinal Aging Study Amsterdam. *J Bone Miner Res.* 1998;13:1932-9.

48. Johnell O, Gullberg B, Kanis JA, et al. Risk factors for hip fracture in European women: the MEDOS Study. Mediterranean Osteoporosis Study. *J Bone Miner Res.* 1995;10:1802-15.
49. Wolinsky FD, Fitzgerald JF. The risk of hip fracture among noninstitutionalized older adults. *J Gerontol.* 1994;49:S165-75.
50. Kleerekoper M, Peterson E, Nelson D, et al. Identification of women at risk for developing postmenopausal osteoporosis with vertebral fractures: role of history and single photon absorptiometry. *Bone Miner.* 1989;7:171-86.
51. Cooper C, Shah S, Hand DJ, et al. Screening for vertebral osteoporosis using individual risk factors. *Osteoporos Int.* 1991;2:48-53.
52. Blake GM, Gluer CC, Fogelman I. Bone densitometry: current status and future prospects. *Br J Radiol.* 1997;70 Spec No:S177-86.
53. Blake GM, Fogelman I. Applications of bone densitometry for osteoporosis. *Endocrinol Metab Clin North Am.* 1998;27:267-88.
54. Genant HK, Engelke K, Fuerst T, et al. Noninvasive assessment of bone mineral and structure: state of the art. *J Bone Miner Res.* 1996;11:707-30.
55. Jergas M, Genant HK. Spinal and femoral DXA for the assessment of spinal osteoporosis. *Calcif Tissue Int.* 1997;61:351-7.
56. Cameron JR, Sorenson J. Measurement of bone mineral in vivo: an improved method. *Science.* 1963;142:230-2.
57. Langton CM, Palmer SB, Porter RW. The measurement of broadband ultrasonic attenuation in cancellous bone. *Eng Med.* 1984;13:89-91.

58. Lochmuller EM, Eckstein F, Kaiser D, et al. Prediction of vertebral failure loads from spinal and femoral dual-energy X-ray absorptiometry, and calcaneal ultrasound: an in situ analysis with intact soft tissues. *Bone*. 1998;23:417-24.
59. Bouxsein ML, Radloff SE. Quantitative ultrasound of the calcaneus reflects the mechanical properties of calcaneal trabecular bone. *J Bone Miner Res*. 1997;12:839-46.
60. Bouxsein ML, Coan BS, Lee SC. Prediction of the strength of the elderly proximal femur by bone mineral density and quantitative ultrasound measurements of the heel and tibia. *Bone*. 1999;25:49-54.
61. Lochmuller EM, Zeller JB, Kaiser D, et al. Correlation of femoral and lumbar DXA and calcaneal ultrasound, measured in situ with intact soft tissues, with the in vitro failure loads of the proximal femur. *Osteoporos Int*. 1998;8:591-8.
62. Nicholson PHF, Muller R, Lowet G, et al. Do quantitative ultrasound measurements reflect structure independently of density in human vertebral cancellous bone? *Bone*. 1998;23:425-31.
63. Kotzki PO, Buyck D, Hans D, et al. Influence of fat on ultrasound measurements of the os calcis. *Calcif Tissue Int*. 1994;54:91-5.
64. Njeh CF, Boivin CM, Langton CM. The role of ultrasound in the assessment of osteoporosis: a review. *Osteoporos Int*. 1997;7:7-22.
65. Rico H, Gomez M, Aguado F, Villa LF, Hernandez ER, Cortes J. Impact of weight in obese subjects on bone speed of sound. *Invest Radiol*. 1999;34:596-9.
66. Gluer CC, Genant HK. Impact of marrow fat on accuracy of quantitative CT. *J Comput Assist Tomogr*. 1989;13:1023-35.

67. Howard GM, Nguyen TV, Pocock NA, Kelly PJ, Eisman JA. Influence of handedness on calcaneal ultrasound: implications for assessment of osteoporosis and study design. *Osteoporos Int.* 1997;7:190-4.
68. Johansen A, Stone MD. The effect of ankle oedema on bone ultrasound assessment at the heel. *Osteoporos Int.* 1997;7:44-7.
69. Jones G, Nguyen T, Sambrook PN, et al. A longitudinal study of the effect of spinal degenerative disease on bone density in the elderly. *J Rheumatol.* 1995;22:932-6.
70. Agren M, Karellas A, Leahey D, Marks S, Baran D. Ultrasound attenuation of the calcaneus: a sensitive and specific discriminator of osteopenia in postmenopausal women. *Calcif Tissue Int.* 1991;48:240-4.
71. Alenfeld FE, Wuster C, Funck C, et al. Ultrasound measurements at the proximal phalanges in healthy women and patients with hip fractures. *Osteoporos Int.* 1998;8:393-8.
72. Cunningham J, Fordham J, Hewitt T, Speed C. Ultrasound velocity and attenuation at different skeletal sites compared with bone mineral density measured using dual energy X-ray absorptiometry. *Br J Radiol.* 1996;69:25-32.
73. Faulkner KG, McClung MR, Coleman LJ, Kingston-Sandahl E. Quantitative ultrasound of the heel: correlation with densitometric measurements at different skeletal sites. *Osteoporos Int.* 1994;4:42-7.
74. Formica CA, Nieves JW, Cosman F, Garrett P, Lindsay R. Comparative assessment of bone mineral measurements using dual X-ray absorptiometry and peripheral quantitative computed tomography. *Osteoporos Int.* 1998;8:460-7.

75. Graafmans WC, Van Lingen A, Ooms ME, Bezemer PD, Lips P. Ultrasound measurements in the calcaneus: precision and its relation with bone mineral density of the heel, hip, and lumbar spine. *Bone*. 1996;19:97-100.
76. Grampp S, Jergas M, Lang P, et al. Quantitative CT assessment of the lumbar spine and radius in patients with osteoporosis. *Am J Roentgenol*. 1996;167:133-40.
77. Greenspan SL, Bouxsein ML, Melton ME, et al. Precision and discriminatory ability of calcaneal bone assessment technologies. *J Bone Miner Res*. 1997;12:1303-13.
78. Langton CM, Ballard PA, Bennett DK, Purdie DW. A comparison of the sensitivity and specificity of calcaneal ultrasound measurements with clinical criteria for bone densitometry (DEXA) referral. *Clin Rheumatol*. 1997;16:117-8.
79. Laval-Jeantet AM, Bergot C, Williams M, Davidson K, Laval-Jeantet M. Dual-energy X-ray absorptiometry of the calcaneus: comparison with vertebral dual-energy X-ray absorptiometry and quantitative computed tomography. *Calcif Tissue Int*. 1995;56:14-8.
80. Martin J, Reid D. Appendicular measurements in screening women for low axial bone mineral density. *Br J Radiol*. 1996;69:234-40.
81. Massie A, Reid DM, Porter RW. Screening for osteoporosis: comparison between dual energy X-ray absorptiometry and broadband ultrasound attenuation in 1000 perimenopausal women. *Osteoporos Int*. 1993;3:107-10.
82. Naganathan V, March L, Hunter D, Pocock NA, Markovey J, Sambrook PN. Quantitative heel ultrasound as a predictor for osteoporosis. *Med J Aust*. 1999;171:297-300.
83. Pocock NA, Noakes KA, Howard GM, et al. Screening for osteoporosis: what is the role of heel ultrasound? *Med J Aust*. 1996;164:367-70.

84. Rosenthall L, Tenenhouse A, Caminis J. A correlative study of ultrasound calcaneal and dual-energy X-ray absorptiometry bone measurements of the lumbar spine and femur in 1000 women. *Eur J Nucl Med.* 1995;22:402-6.
85. Roux C, Lemonnier E, Lolta S, et al. Broadband ultrasound attenuation of the calcaneus and bone density measurements. *Rev Rhum Engl Ed.* 1993;60:771-80.
86. Salamone L, Krall E, Dawson-Hughes B. Comparison of broadband ultrasound attenuation to single X-ray absorptiometry measurements at the calcaneus in postmenopausal women. *Calcif Tissue Int.* 1994;54:87-90.
87. Schott AM, Weill-Engerer S, Hans D, Duboeuf F, Delmas PD, Meunier PJ. Ultrasound discriminates patients with hip fracture equally well as dual energy X-ray absorptiometry and independently of bone mineral density. *J Bone Miner Res.* 1995;10:243-9.
88. Tromp AM, Smit JH, Deeg DJ, Lips P. Quantitative ultrasound measurements of the tibia and calcaneus in comparison with DXA measurements at various skeletal sites. *Osteoporos Int.* 1999;9:230-5.
89. Turner C, Peacock M, Timmerman L, Neal J, Johnston C. Calcaneal ultrasonic measurements discriminate hip fracture independently of bone mass. *Osteoporos Int.* 1995;5:130-5.
90. Young H, Howey S, Purdie DW. Broadband ultrasound attenuation compared with dual-energy X-ray absorptiometry in screening for postmenopausal low bone density. *Osteoporos Int.* 1993;3:160-4.
91. Ahmed AIH, Blake GM, Rymer JM, Fogelman I. Screening for osteopenia and osteoporosis: do the accepted normal ranges lead to overdiagnosis? *Osteoporos Int.* 1997;7:432-38.

92. Petley GW, Cotton AM, Murrills AJ, et al. Reference ranges of bone mineral density for women in southern England: the impact of local data on the diagnosis of osteoporosis. *Br J Radiol.* 1996;69:655-60.
93. Chen Z, Maricic M, Lund P, Tesser J, Gluck O. How the new Hologic hip normal reference values affect the densitometric diagnosis of osteoporosis. *Osteoporos Int.* 1998;8:423-27.
94. Simmons A, Barrington S, O'Doherty M, Coakley A. Dual energy X-ray absorptiometry normal reference range use within the UK and the effect of different normal ranges on the assessment of bone density. *Br J Radiol.* 1995;68:903-9.
95. Lunt M, Felsenberg D, Reeve J, et al. Bone density variation and its effects on risk of vertebral deformity in men and women studied in thirteen European centers: the EVOS Study. *J Bone Miner Res.* 1997;12:1883-94.
96. Laskey M, Crisp A, Cole T, Compston J. Comparison of the effect of different reference data on Lunar DPX and Hologic QDR-1000 dual-energy X-ray absorptiometers. *Br J Radiol.* 1992;65:1124-29.
97. Arlot ME, Sornay-Rendu E, Garnero P, Vey-Marty B, Delmas PD. Apparent pre- and postmenopausal bone loss evaluated by DXA at different skeletal sites in women: the OFELY cohort. *J Bone Miner Res.* 1997;12:683-90.
98. Faulkner KG, von Stetten E, Miller P. Discordance in patient classification using T-scores. *J Clin Densitomet.* 1999;2:343-50.
99. Grampp S, Henk CB, Fuerst TP, et al. Diagnostic agreement of quantitative sonography of the calcaneus with dual X-ray absorptiometry of the spine and femur. *Am J Roentgenol.* 1999;173:329-34.

100. Kroger H, Lunt M, Reeve J, et al. Bone density reduction in various measurement sites in men and women with osteoporotic fractures of spine and hip: the European quantitation of osteoporosis study. *Calcif Tissue Int.* 1999;64:191-9.
101. Varney LF, Parker RA, Vincelette A, Greenspan SL. Classification of osteoporosis and osteopenia in postmenopausal women is dependent on site-specific analysis. *J Clin Densitomet.* 1999;2:275-83.
102. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ.* 1996;312:1254-59.
103. Cummings SR, Black DM, Nevitt MC, et al. Bone density at various sites for prediction of hip fractures. *Lancet.* 1993;341:72-5.
104. Melton LJ, Atkinson E, O'Fallon W, Wahner H, Riggs B. Long-term fracture prediction by bone mineral assessed at different skeletal sites. *J Bone Miner Res.* 1993;8:1227-33.
105. Black DM, Cummings SR, Genant HK, Nevitt MC, Palermo L, Browner W. Axial and appendicular bone density predict fractures in older women. *J Bone Miner Res.* 1992;7:633-38.
106. Bauer DC, Gluer CC, Cauley JA, et al. Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women. A prospective study. Study of Osteoporotic Fractures Research Group. *Arch Intern Med.* 1997;157:629-34.
107. de Laet CE, Van Hout BA, Burger H, Weel AE, Hofman A, Pols HA. Hip fracture prediction in elderly men and women: validation in the Rotterdam study. *J Bone Miner Res.* 1998;13:1587-93.

108. Duboeuf F, Hans D, Schott AM, et al. Different morphometric and densitometric parameters predict cervical and trochanteric hip fracture: the EPIDOS Study. *J Bone Miner Res.* 1997;12:1895-902.
109. Garnero P, Dargent-Molina P, Hans D, et al. Do markers of bone resorption add to bone mineral density and ultrasonographic heel measurement for the prediction of hip fracture in elderly women? The EPIDOS prospective study. *Osteoporos Int.* 1998;8:563-9.
110. Hans D, Dargent-Molina P, Schott AM, et al. Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. *Lancet.* 1996;348:511-4.
111. Huang C, Ross PD, Yates AJ, et al. Prediction of fracture risk by radiographic absorptiometry and quantitative ultrasound: a prospective study. *Calcif Tissue Int.* 1998;63:380-4.
112. Kroger H, Huopio J, Honkanen R, et al. Prediction of fracture risk using axial bone mineral density in a perimenopausal population: a prospective study. *J Bone Miner Res.* 1995;10:302-6.
113. Mele R, Masci G, Ventura V, de Aloysio D, Bicocchi M, Cadossi R. Three-year longitudinal study with quantitative ultrasound at the hand phalanx in a female population. *Osteoporos Int.* 1997;7:550-7.
114. Nevitt MC, Johnell O, Black DM, Ensrud K, Genant HK, Cummings SR. Bone mineral density predicts non-spine fractures in very elderly women. Study of Osteoporotic Fractures Research Group. *Osteoporos Int.* 1994;4:325-31.
115. Porter RW, Miller CG, Grainger D, Palmer SB. Prediction of hip fracture in elderly women: a prospective study. *BMJ.* 1990;301:638-41.

116. Schott AM, Cormier C, Hans D, et al. How hip and whole-body bone mineral density predict hip fracture in elderly women: the EPIDOS Prospective Study. *Osteoporos Int.* 1998;8:247-54.
117. Stewart A, Torgerson DJ, Reid DM. Prediction of fractures in perimenopausal women: a comparison of dual energy x ray absorptiometry and broadband ultrasound attenuation. *Ann Rheum Dis.* 1996;55:140-2.
118. Vecht-Hart CM, Zwamborn AW, Peeters PH, Collette HJ. Prediction of peripheral fracture risk by quantitative microdensitometry. *Prev Med.* 1997;26:86-91.
119. Rimes KA, Salkovskis PM, Shipman AJ. Psychological and behavioural effects of bone density screening for osteoporosis. *Psychol Health.* 1999;14:585-608.
120. Salamone LM, Pressman AR, Seeley DG, Cauley JA. Estrogen replacement therapy. A survey of older women's attitudes. *Arch Intern Med.* 1996;156:1293-7.
121. Silverman SL, Greenwald M, Klein RA, et al. Effect of bone density information on decisions about hormone replacement therapy: a randomized trial. *Obstet Gynecol.* 1997;89:321-5.
122. Rozenberg S, Kroll M, Vandromme J, Paesmans M, Ham H. Effect of bone density evaluation on hormone replacement therapy prescription. *Maturitas.* 1996;24:57-61.
123. Torgerson DJ, Thomas RE, Campbell MK, Reid DM. Randomized trial of osteoporosis screening. Use of hormone replacement therapy and quality-of-life results. *Arch Intern Med.* 1997;157:2121-5.
124. Rubin SM, Cummings SR. Results of bone densitometry affect women's decisions about taking measures to prevent fractures. *Ann Intern Med.* 1992;116:990-95.

125. Stock JL, Waud CE, Coderre JA, Smith KK. Clinical reporting to primary care physicians leads to increased use and understanding of bone densitometry and affects the management of osteoporosis: a randomized trial. *Ann Intern Med.* 1998;128:996-9.
126. Torgerson D, Bell-Syer S. Hormone replacement therapy and prevention of nonvertebral fractures: A meta-analysis of randomized trials. *JAMA.* 2001;285:2891-2897.
127. Lufkin E, Wahner H, O'Fallon W, et al. Treatment of postmenopausal osteoporosis with transdermal estrogen. *An Intern Med.* 1992;117:1-9.
128. Wimalawansa S. A four-year randomized controlled trial of hormone replacement and bisphosphate, alone or in combination, in women with postmenopausal osteoporosis. *Am J Med.* 1998;104:219-226.
129. Komulainen MH, Kroger H, Tuppurainen MT, et al. HRT and Vit D in prevention of non-vertebral fractures in postmenopausal women; a 5 year randomized trial. *Maturitas.* 1998;31:45-54.
130. Mosekilde L, Beck-Nielsen H, Sorensen OH, et al. Hormonal replacement therapy reduces forearm fracture incidence in recent postmenopausal women—results of the Danish Osteoporosis Prevention Study. *Maturitas.* 2000;36:181-193.
131. Cauley J, Black D, et al. Effects of hormone replacement therapy on clinical fractures and height loss: the heart and Estrogen/progestin Replacement Study (HERS). *Am J Med.* 2001;110:442-450.
132. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *Jama.* 1998;280:605-13.

133. Cooper C, Melton L. Vertebral fractures: How large is the silent epidemic. *BMJ*. 1992;304:793-794.
134. Ettinger B, Black D, Mitlak B, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene. *JAMA*. 1999;282:637-645.
135. Lufkin EG, Whitaker MD, Nickelsen T, et al. Treatment of established postmenopausal osteoporosis with raloxifene: a randomized trial. *J Bone Miner Res*. 1998;13:1747-54.
136. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst*. 1998;90:1371-88.
137. Black DM, Cummings SR, Karpf DB, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. *Lancet*. 1996;1535-41.
138. Adami S, Passeri M, Ortolani S, et al. Effects of oral alendronate and intranasal salmon calcitonin on bone mass and biochemical markers of bone turnover in postmenopausal women with osteoporosis. *Bone*. 1995;383-90.
139. Bone HG, Downs RW, Jr., Tucci JR, et al. Dose-response relationships for alendronate treatment in osteoporotic elderly women. Alendronate Elderly Osteoporosis Study Centers. *J Clin Endocrinol Metab*. 1997;265-74.
140. Chesnut CH, 3rd, McClung MR, Ensrud KE, et al. Alendronate treatment of the postmenopausal osteoporotic woman: effect of multiple dosages on bone mass and bone remodeling. *Am J of Med*. 1995;144-52.

141. Hosking D, Chilvers CE, Christiansen C, et al. Prevention of bone loss with alendronate in postmenopausal women under 60 years of age. Early Postmenopausal Intervention Cohort Study Group. *NEJM*. 1998:485-92.
142. Liberman UA, Weiss SR, Broll J, et al. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. *NEJM*. 1995:1437-43.
143. McClung M, Clemmesen B, Daifotis A, et al. Alendronate prevents postmenopausal bone loss in women without osteoporosis. A double-blind, randomized, controlled trial. Alendronate Osteoporosis Prevention Study Group. *An of Intern Med*. 1998:253-61.
144. Greenspan SL, Parker RA, Ferguson L, Rosen HN, Maitland-Ramsey L, Karpf DB. Early changes in biochemical markers of bone turnover predict the long-term response to alendronate therapy in representative elderly women: A randomized clinical trial. *J of Bone Miner Res*. 1998;13:1431-8.
145. Pols HA, Felsenberg D, Hanley DA, et al. Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Foxamax International Trial Study Group. *Osteo Intern*. 1999:461-8.
146. Cummings SR, Black DM, Thompson DE, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial . *JAMA*. 1998:2077-82.
147. Bonnick S, Rosen C, Mako B, DeLuca P, Byrnes C, Melton M. Alendronate vs calcium for treatment of osteoporosis in postmenopausal women. *Bone* 1998;23(5S):S476.

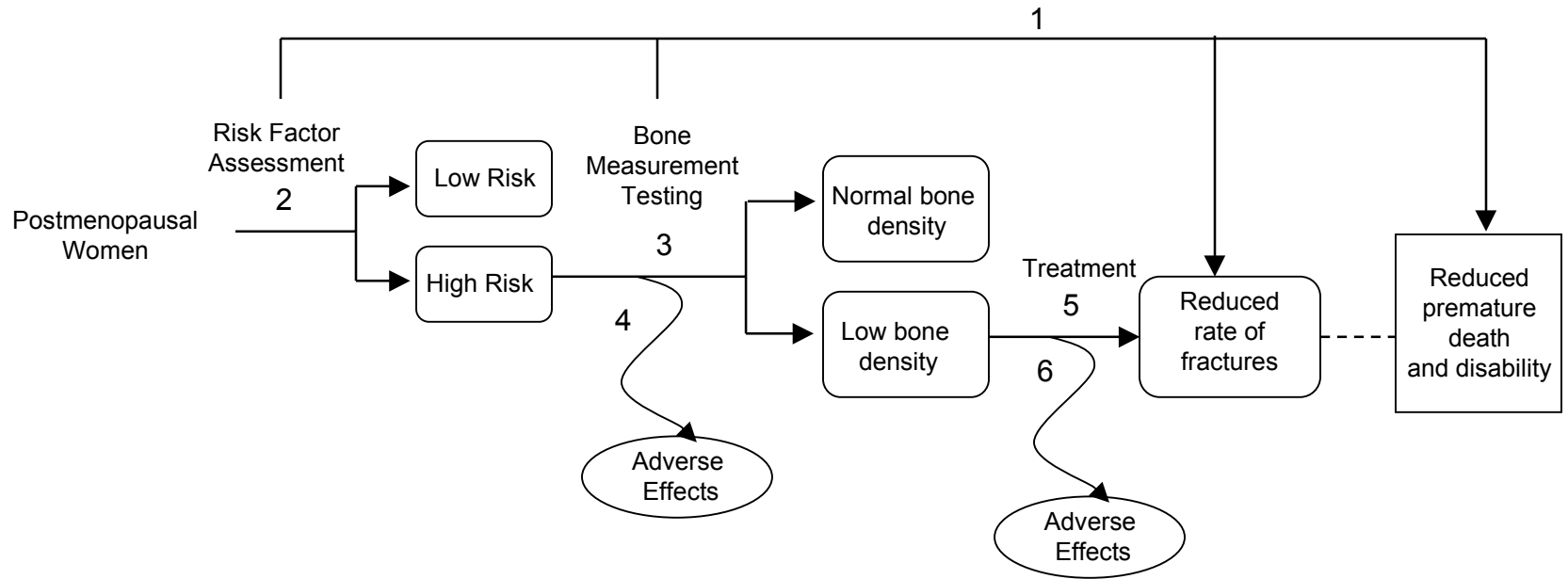
148. Ensrud KE, Black DM, Palermo L, et al. Treatment with alendronate prevents fractures in women at highest risk: results from the Fracture Intervention Trial . *Arch Intern Med*. 1997;2617-24.
149. Cummings SR, Black DM, Thompson DE. Effect of alendronate reduces on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998;280:2077-82.
150. McClung M, Geusens P, Miller P, et al. Effect of risedronate on the risk of hip fracture in elderly women. *NEJM*. 2001;344:333-40.
151. Johansen HK, Gotzsche PC. Problems in the design and reporting of trials of antifungal agents encountered during meta-analysis. *JAMA*. 1999;282:1752-9.
152. Djulbegovic B, Lacevic M, Cantor A, et al. The uncertainty principle and industry-sponsored research. *Lancet*. 2000;356:635-38.
153. Miller J, Chan B, Nelson H. Postmenopausal estrogen replacement and risk of venous thromboembolism: A systematic review and meta-analysis. . Rockville, MD: Agency for Healthcare Research and Quality; 2001.
154. Grodstein F, Colditz G, Stampfer M. Postmenopausal hormone use and cholecystectomy in a large prospective study. *Obstet Gynecol*. 1994;85:5-11.
155. Grady D, Gebretsadik T, Ernster V, Petitti D. Hormone replacement therapy and endometrial cancer risk: A meta-analysis. *Obstet Gynecol*. 1995;85:304-313.
156. Rodriguez I, Kilborn MJ, Liu XK, Pezzullo JC, Woosley RL. Drug-induced QT prolongation in women during the menstrual cycle. *JAMA*. 2001;285:1322-6.
157. Collaborative Group on Hormonal Factors in Breast Cancer: Breast cancer and hormone replacement therapy: Collaborative reanalysis of data from 51 epidemiological studies of

- 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet*. 1997;350:1047-59.
158. Humphrey L, Takana L. Postmenopausal hormone replacement therapy and cardiovascular disease: Systematic evidence review. Rockville, MD: Agency for Healthcare Research and Quality; 2001.
159. Davies G, Huster W, Lu Y, Plouffe LJ, Lakshmanan M. Adverse events reported by postmenopausal women in controlled trials with raloxifene. *Obstet and Gynecol*. 1999;93:558-65.
160. Ettinger B, Pressman A, Schein J. Clinic visits and hospital admissions for care of acid-related upper gastrointestinal disorders in women using alendronate for osteoporosis. *Am J Manag Care*. 1998;4:1377-82.
161. Melton LJ, Chrischilles EA, Cooper C, Lane AW, Riggs BL. Perspective. How many women have osteoporosis? *J Bone Miner Res*. 1992;7:1005-10.
162. Reginster J, Minne H, Sorensen O, et al. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis: Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. *Osteo Intern*. 2000;11:83-91.
163. Meyer HE, Tverdal A, Falch JA. Risk factors for hip fracture in middle-aged Norwegian women and men. *Am J Epidemiol*. 1993;137:1203-11.
164. Fujiwara S, Kasagi F, Yamada M, Kodama K. Risk factors for hip fracture in a Japanese cohort. *J Bone Miner Res*. 1997;12:998-1004.
165. Kreiger N, Gross A, Hunter G. Dietary factors and fracture in postmenopausal women: a case-control study. *Int J Epidemiol*. 1992;21:953-8.

166. Tuppurainen M, Kroger H, Honkanen R, et al. Risks of perimenopausal fractures--a prospective population-based study. *Acta Obstet Gynecol Scand.* 1995;74:624-8.
167. Mallmin H, Ljunghall S, Persson I, Bergstrom R. Risk factors for fractures of the distal forearm: a population-based case-control study. *Osteoporos Int.* 1994;4:298-304.
168. Honkanen R, Tuppurainen M, Kroger H, Alhava E, Saarikoski S. Relationships between risk factors and fractures differ by type of fracture: a population-based study of 12,192 perimenopausal women. *Osteoporos Int.* 1998;8:25-31.
169. Torgerson DJ, Campbell MK, Thomas RE, Reid DM. Prediction of perimenopausal fractures by bone mineral density and other risk factors. *J Bone Miner Res.* 1996;11:293-7.
170. Weinstein L, Ullery B, Bourguignon C. A simple system to determine who needs osteoporosis screening. *Obstet Gyn.* 1999;93:757-60.
171. Tobias JH, Cook DG, Chambers TJ, Dalzell N. A comparison of bone mineral density between Caucasian, Asian and Afro-Caribbean women. *Clin Sci.* 1994;87:587-91.
172. Farmer ME, White LR, Brody JA, Bailey KR. Race and sex differences in hip fracture incidence. *Am J Public Health.* 1984;74:1374-80.
173. Bauer RL. Ethnic differences in hip fracture: a reduced incidence in Mexican Americans. *Am J Epidemiol.* 1988;127:145-9.
174. Grisso JA, Kelsey JL, Strom BL, et al. Risk factors for hip fracture in black women. The Northeast Hip Fracture Study Group. *N Engl J Med.* 1994;330:1555-9.
175. Suzuki T, Yoshida H, Hashimoto T, et al. Case-control study of risk factors for hip fractures in the Japanese elderly by a Mediterranean Osteoporosis Study (MEDOS) questionnaire. *Bone.* 1997;21:461-7.

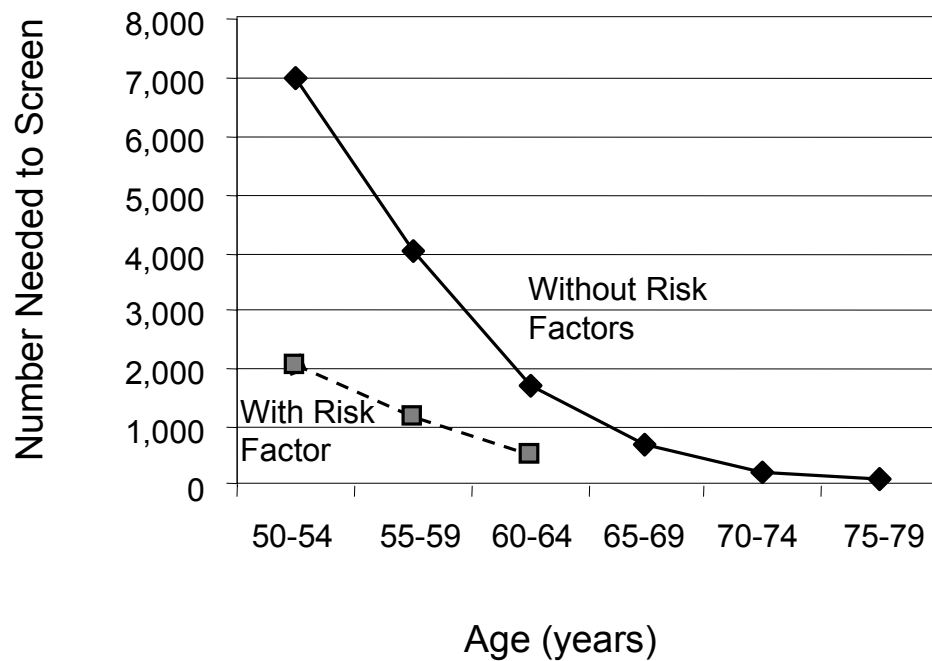
176. Agency for Health Care Policy and Research, U.S. Department of Health and Human Services. Health Technology Assessment. Number 6: Bone Densitometry: Patients with Asymptomatic Hyperparathyroidism. 1995;AHRQ Pub.:96-0004.

Figure 1. Osteoporosis Screening Analytic Framework



Key questions addressed in this review include: (1) Does screening using risk factor assessment and/or bone density testing reduce fractures? (2) Does risk factor assessment accurately identify women who may benefit from bone density testing? (3) Do bone density measurements accurately identify women who may benefit from treatment? (4) What are the harms of screening? (5) Does treatment reduce the risk of fractures in women identified by screening? (6) What are the harms of treatment?

Figure 2. Number Needed to Screen to Prevent 1 Hip Fracture in 5 Years



Number Needed to Screen to Prevent 1 Hip Fracture in 5 Years. The number needed to screen decreases with advancing age and for women under age 65 with risk factors.

Table 1. Recent Systematic Reviews of Osteoporosis

Publication	Year	Organization Producing Report	Conclusions				
			Risk factors	Choice of test	Monitoring	Use of markers	Cost-effectiveness
Bone Density Measurement - A Systematic Review³	1997	Swedish Council on Technology Assessment in Health Care	Must consider risk factors other than bone density alone to make decisions about testing or treatment.	BMD in hip or spine cannot be reliably estimated from measurements in arm or heel.	Measurements at intervals <2 years are unnecessary.	No documentation that repeated measurements of markers influence treatment in a way that improves long-term clinical outcomes.	
Effectiveness of Bone Density Measurement and Associated Treatments for Prevention of Fractures⁴	1996	International Network of Agencies for Health Technology Assessment		The precision and accuracy of all BMD tests in community settings are unknown. Accuracy of ultrasound still not proven.	BMD would require minimum followup of 1 to 1.5 years to detect bone loss of 2 to 3%.		
Bone Mineral Density Testing: Does the Evidence Support its Selective Use in Well Women?⁵	1997	British Columbia Office of Health Technology Assessment	Currently there are no validated risk assessment tools to select patients for BMD testing. In the general population, clinical assessment was no worse than BMD measurement in assessment of fracture risk.	Result of BMD test by any current technology is an unsuitable measure upon which to base clinical decisions.			Available economic evaluations are not adequate evidence that BMD testing is more cost-effective than universal hormone therapy or no intervention.
Osteoporosis: Review of the Evidence for Prevention, Diagnosis, and Treatment and Cost-effectiveness Analysis⁶	1998	National Osteoporosis Foundation	Appropriateness of measuring BMD depends on fracture risk, determined by age and other risk factors, and treatment being considered.	Given the better predictive value of hip measurements for hip fractures, hip DXA should be the primary measurement.	The longer the interval between measurements, the more precise the estimate of changes in bone mass; effect of monitoring on treatment is unknown.	Biochemical markers are promising but their role in patient management is not yet known.	
Note: BMD indicates bone mineral density; DXA, dual-energy X-ray absorptiometry							

Table 1. Recent Systematic Reviews of Osteoporosis (continued)

Publication	Year	Organization Producing Report	Conclusions				
			Risk factors	Choice of test	Monitoring	Use of markers	Cost-effectiveness
Quantitative Ultrasound for Bone Density Measurement⁷	1998	Alberta Heritage Foundation for Medical Research		Quantitative calcaneal ultrasound is a promising diagnostic technology, but its role in osteoporosis diagnosis is unclear. Good evidence that ultrasound can identify increased risk of fracture in populations but not individuals.			
Selective Testing with Bone Density Measurement⁸	1999	Alberta Heritage Foundation for Medical Research	There is potential for selective use of BMD in association with appraisal of other risk factors. Assessment protocols for such an approach have promise as a useful tool for selecting whom to test. Advice on treatment options should consider evidence of efficacy and effectiveness in terms of absolute reduction in risk of fracture, long term compliance, and adverse effects.	Substantial uncertainty with the performance of BMD in correctly classifying an individual as osteoporotic. Ultrasound is less precise than DXA.	Minimum acceptable interval between measurements may be as long as 2 years.		
Osteoporosis: Clinical Guidelines for Prevention and Treatment⁹	1999	Royal College of Physicians	Recommends selective testing in women with risk factors (based not on evidence but on expert opinion).	DXA at the hip is preferred because of higher predictive value for fracture risk.	Optimal use of BMD measurements in monitoring response to treatment is uncertain, recommend future research.	Until biochemical markers become more widely established and supported by evidence, their use in clinical practice will remain limited.	The cost-effectiveness of BMD measurements improves as the expense of the therapy goes up.
Note: BMD indicates bone mineral density; DXA, dual-energy X-ray absorptiometry							

Table 1. Recent Systematic Reviews of Osteoporosis (continued)

Publication	Year	Organization Producing Report	Conclusions				
			Risk factors	Choice of test	Monitoring	Use of markers	Cost-effectiveness
Ultrasonography of the Heel for Diagnostic Osteoporosis and Selecting Patients for Pharmacologic Treatment¹⁰	1999	Blue Cross and Blue Shield Association		Use of ultrasound to direct treatment may result in a substantially smaller health outcome benefit compared to DXA. 43 to 76% of patients benefiting from treatment would be identified by ultrasound (sensitivity); 75 to 90% of patients not benefiting by treatment would be identified by ultrasound (specificity).			
Consensus Statement on Prevention and Treatment of Osteoporosis¹¹	1999	Israel Center for Technology Assessment in Health Care/Israel Ministry of Health/Israel Medical Association	Physician's responsibility to estimate risk of osteoporosis and fractures and to consider performing additional tests; based on some risk factors, report recommends BMD every 2 years	Recommends DXA done at facilities with quality control and following regulations on operations and interpretation of results; additional technologies to be considered if efficiency proven compared to DXA.	Calls for publicly funded BMD in women over 65 every 5 years; in postmenopausal women over 50 every 2 years if certain risk factors present; in those with disease entailing increased risk of osteoporosis, no limit on age or frequency of test.		
Note: BMD indicates bone mineral density; DXA, dual-energy X-ray absorptiometry							

Table 2. Prevalence of Low Femoral Neck Bone Density in U.S. Women Over Age 50 (DXA)

Race	Osteopenia (-2.5<T-score<-1.0)		Osteoporosis (T-score < -2.5)	
	Prevalence (%)*	Millions**	Prevalence (%)*	Millions**
All	40	14	15	5
NHW	41	12	15	5
NHB	28	0.9	8	0.3
MA	38	0.3	16	0.1

*Age adjusted to 1980 U.S. Census

** Undercount adjusted estimates from March 1990 & 1993 Current Population Surveys

Note: Data taken from NHANES III Looker (1998).¹³

DXA indicates dual-energy X-ray Absorptiometry; MA, Mexican American; NHB, non-Hispanic black; NHW, non-Hispanic white.

Table 3. Osteoporosis Prevalence (%) in White Women in the U.S. Over Age 50 by Decade (DPA and SPA)*

Age	Spine	Hip	Wrist	At spine, hip, or wrist
50 – 59	7.6	3.9	3.7	14.8
60 – 69	11.8	8.0	11.8	21.6
70 – 79	25.0	24.5	23.1	38.5
≥80	32.0	47.5	50.0	70.0
All	16.5	16.2	17.4	30.3

*Data taken from Melton (1995).¹⁴

Note: DPA indicates dual photon absorptiometry; SPA, single photon absorptiometry.

Table 4 5-Year Hip Fracture Rates for Women Without Previous Hip Fracture by Age*

Age	T-score							
	-5	-4	-3	-2	-1	0	1	2
50 yrs			0.024	0.0095	0.0038	0.0015	0.0006	0.00023
60 yrs		0.069	0.029	0.011	0.0047	0.0018	0.0007	0.00025
70 yrs		0.127	0.055	0.023	0.0096	0.0039	0.001	
80 yrs	0.35	0.2	0.09	0.042	0.018	0.007	0.0028	
90 yrs	0.29	0.19	0.097	0.046	0.02	0.006		

*Data derived from decision model presented in the National Osteoporosis Foundation Report (1998)⁶

Table 5. Studies of Risk Factor Assessment

Author	Design	N	Validated	Risk Factors Included	Outcome	Performance	Quality Rating*
<i>Bone Density Outcomes</i>							
Stemenda, 1990 ³⁶	Cross-sectional	124	No	Age, height, weight, calcium intake, caffeine intake, alcohol and tobacco use, urinary markers of bone turnover.	Correct classification of high or low BMD (lowest third of subjects)	Midshaft radius: 68% low, 77% high. Lumbar spine: 61% low, 45% high. Femoral neck: 66% low, 53% high.	Poor
Falch, 1992 ³⁰	Cross-sectional	73	Yes	Low body weight, reduced renal phosphate reabsorption, smoking.	Bone loss	Sensitivity 36%, specificity 89%, PPV 74%.	Poor

Table 5. Studies of Risk Factor Assessment (continued)

Ribot, 1992 ³⁵	Cross-sectional	1565	No	Weight, menopause, duration of menopause.	Vertebral BMD < -2 SD	Sensitivity 73%, specificity 66%.	Fair
Elliot, 1993 ³²	Cross-sectional	320	Yes	Spine BMD: age, weight, smoking status, age at menarche. Femoral neck BMD: age, weight, family history, activity, smoking status.	Low lumbar spine and femoral neck BMD (lowest third of age matched normal	Lumbar spine: sensitivity 86%, specificity 32%. Femoral neck: sensitivity 89%, specificity 25%.	Fair
Michaelsson, 1996 ³³	Cross-sectional	175	No	Weight > 70kg.	Femoral neck BMD < -2.5 SD	Sensitivity 94%, specificity 36%, PPV 21%, NPV 97%.	Fair

Table 5. Studies of Risk Factor Assessment (continued)

Ribot, 1992 ³⁵	Cross-sectional	1565	No	Weight, menopause, duration of menopause.	Vertebral BMD < -2 SD	Sensitivity 73%, specificity 66%.	Fair
	Cross-sectional	61	No	1. Arm span-height difference of at least 3 cm. 2. Arm span-height difference, age above or below 70, and whether arm span was below or above 160 cm.	BMD <= -2.5 SD and vertebral fracture	Arm span only: sensitivity 58%, specificity 56%. Arm span, age, arm span length: sensitivity 81%, specificity 64%.	Poor
Ballard, 1998 ³⁴	Cross-sectional	1158	No	Age, age at menopause, height, weight, gravidity, parity, current use of steroids, current HRT.	Osteoporosis of femoral neck and/or spine	ROC area 0.73.	Fair
Lydick, 1998 ³⁷	Cross-sectional	1279	Yes	Age, weight, race, estrogen use, rheumatoid arthritis, history of fractures.	Femoral neck BMD <= -2 SD	Sensitivity 89%, specificity 50%; ROC area 0.81.	Good

Table 5. Studies of Risk Factor Assessment (continued)

Ribot, 1992 ³⁵	Cross-sectional	1565	No	Weight, menopause, duration of menopause.	Vertebral BMD < -2 SD	Sensitivity 73%, specificity 66%.	Fair
Goemaere, 1999 ³⁸	Cross-sectional	300	No	18-item questionnaire of risk factors for osteoporosis (race, height loss, age, weight, smoking, coffee, alcohol, dairy product use, activity, family history, existence of comorbidities, history of wrist fracture, menopause before 45 years, corticosteroid use).	Lumbar spine, femoral neck, and hip BMD	Lumbar spine: ROC area 0.66; Femoral neck: ROC area 0.69; Hip: ROC area 0.76.	Fair
Cadarette, 2000 ³⁹	Cross-sectional	926	Yes	Age, weight, current use of HRT.	Hip or lumbar spine BMD ≤ -2.5	Sensitivity 95%, specificity 41%.	Good

Table 5. Studies of Risk Factor Assessment (continued)

Fracture Outcomes

Kleerekoper, 1989 ⁵⁰	Case-control	663	No	Model 1; total months of lactation, family history of osteoporosis, years post menopause, weight. Model 2: breast fed, surgical menopause, age at menarche, age, smoking status.	Vertebral fractures	Model 1: ROC area (SE) 0.55 (0.07); sensitivity 56%; specificity 54%. Model 2: ROC 0.51 (0.042); sensitivity 63% specificity 39%.	Fair
van Hemert, 1990 ⁴⁴	Cohort	1014	No	Age, metacarpal cortical area, relative cortical area, BMI, height, diameter of forearm, diameter of knee, age at menarche, age at menopause, smoking, number of children, period of lactation.	Osteoporotic fractures	Sensitivity 48%, specificity 82%.	Fair
Cooper, 1991 ⁵¹	Case-control	1012	No	Age, height, vertebral fracture after age 45, age of last menstrual period, number of children, ever use oral corticosteroid.	Vertebral fractures	Sensitivity 51%, specificity 69%.	Fair

Table 5. Studies of Risk Factor Assessment (continued)

Wolinsky, 1994 ⁴⁹	Cohort	368	No	White race, female gender, living in southern U.S., age, having been hospitalized in the previous year, previous fall, body mass.	Hip fractures	ROC 0.71; sensitivity 64.7%, specificity 65.7%.	Fair
Johnell, 1995 ⁴⁸	Case-control	5618	No	Late menarche, poor mental score, low BMI, low physical activity, low exposure to sunlight, and low consumption of calcium and tea.	Hip fractures	Sensitivity 55%, specificity 65%.	Fair
Ranstam, 1996 ⁴⁵	Case-control	7474	No	Mental-functional risk score: knowledge of the day of week, knowledge of age, ability to wash, ability to dress.	Hip fractures	A less than perfect score had a sensitivity 46%, specificity 79%.	Fair
Tromp, 1998 ⁴⁷	Cohort	1469	No	Female gender, living alone, past fractures, inactivity, height, use of analgesics	Probability of fractures	No predictors = 0%; 4 predictors 0.129	Fair

Table 5. Studies of Risk Factor Assessment (continued)

Burger, 1999 ⁴⁶	Cohort	5208	No	Model with BMD: age, gender, height, use of a walking aid, current smoking, BMD of femoral neck. Model without BMD: age, gender, height, use of a walking aid, current smoking, weight.	Hip fractures	Model with BMD: ROC area 0.88; sensitivity 70%, specificity 84%. Model without BMD: ROC area 0.83; sensitivity 70%, specificity 83%.	Fair
----------------------------	--------	------	----	---	---------------	---	------

*Harris²⁷

Note: BMD indicates bone mass density; NPV, negative predictive value; PPV, positive predictive value; SD, standard deviation. ROC indicates receiver operating characteristic (values ≥ 0.80 are usually required to consider a test to be effective).

Table 6

Characteristics of Bone Density and Quantitative Ultrasound Tests								
Test	Sites	Examination time, operator skill needed	Radiation exposure	Capital Costs-Technology Purchase	Precision	Cost	Charges	Comments
		<i>minutes, skill level</i>		<i>\$U.S.</i>	<i>%</i>		<i>\$US</i>	
Single photon absorptiometry (SPA)	Wrist, heel	5 - 15, low	low				\$50 - 150	Uses isotopes
Dual-energy photon absorptiometry (DPA)	Spine, proximal femur, whole body	20-40	low	\$20,000	3 - 10	inexpensive	\$150 -300	Uses isotopes
Single x-ray absorptiometry (SXA)	Peripheral sites		0.08-4.6 uSv	\$20,000	0.5 - 2		\$50 - 150	
Dual-energy X-ray absorptiometry (DEXA or DXA)	Lumbar spine, proximal femur	5 - 10, high	0.08-4.6 uSv (pencil beam) or 60 uSv (fan beam)	\$100,000 - 200,000	1 - 5	Fairly expensive	\$136*	Influenced by osteoarthritis
	Total body				0.5	Expensive		
Peripheral dual-energy X-ray absorptiometry (pDXA)	peripheral (wrist, heel)				2 - 5	Inexpensive		

Table 6

Characteristics of Bone Density and Quantitative Ultrasound Tests								
Test	Sites	Examination time, operator skill needed	Radiation exposure	Capital Costs-Technology Purchase	Precision	Cost	Charges	Comments
		<i>minutes, skill level</i>		<i>\$U.S.</i>	<i>%</i>		<i>\$US</i>	
Quantitative computed tomography (QCT)	Spine	10- 30, high	25-360 uSv	\$5,000 - 15,000 [†]	2 - 5		\$150 -300	Higher radiation exposure; Measures the true volumetric density
Quantitative ultrasonography (QUS)	Heel, fingers, tibia, patella	5 - 10, low	none	\$10,000 - 100,000	0.4 - 4		\$35*	Low cost, portable, no radiation
Radiographic absorptiometry (RA) & Quantitative microdensitometry (QMD)	Hand	5 - 10, high	0.08-4.6 uSv	†	1 - 2		\$90 - 160	Low cost, portable
* Average Medicare reimbursement								
† uses conventional CT or radiographic equipment								
Note: uSv indicates microSieverts.								

Table 7. Prospective Studies of DXA of the Hip Reporting Hip Fractures

Cohort	Population	Age range	Follow up (years)	N	Probability of hip fracture	Mean bone density	Cutoff values	Relative risk per 1 SD (CI)	Probability of hip fracture*	
									"low risk"	"high risk"
Study of Osteoporotic Fractures (SOF)103, 105, 106, 114	Community-dwelling white women from 4 areas in the U.S. recruited from lists.	>=65	1.8-2.9	5236	0.009	0.63	lowest quartile	2.6 (1.9-3.8)	0.005	0.023
		65-79	2.9			0.66		2.9 (2.2-3.9)		
		>=80	2.9			0.59		2.1 (1.4-3.2)		
		65-69	1.8	2371	0.003		age-specific		0.0028	0.005
		70-74	1.8	3013	0.0076		quartiles		0.005	0.016
		75-79	1.8	1728	0.007				0.003	0.019
		80-84	1.8	731	0.018				0.007	0.049
		>=85	1.8	291	0.024			0.014	0.028	
Epidemiologie de L'Osteoporose	Women from 5 cities in France recruited from voting lists and health insurance companies.	>=75	2	5656	0.02	0.69-0.72	<0.703 g/cm ²	1.9 (1.6-2.4)	0.033	0.008
		<80	2	3982	0.013	0.73	T-score <-2.5		0.002	0.025
(EPIDOS)108-110,116		>=80	2	3616	0.028	0.67-0.71	T-score <-2.5		0.006	0.04

Table 7. Prospective Studies of DXA of the Hip Reporting Hip Fractures (continued)

Rotterdam107	People aged 55 and over	>=55	3.8	3078	0.13	risk function†	2.5 (1.8-3.6)	0.007	0.069
--------------	-------------------------	------	-----	------	------	----------------	---------------	-------	-------

living in Rotterdam,
Netherlands.

*Probability of hip fracture if bone density was classified as high or low risk.

†Regression equation using age and bone density.

Note: CI indicates confidence interval; DXA, dual energy x-ray absorptiometry; SD, standard deviation.

Table 8. Randomized Controlled Trials of Estrogen and SERMs with Fracture Outcomes

Author	Drug/dose	N	Duration	Age	Population	Relative risk (95% CI) for fractures		Quality
			(years)	(years)		Vertebral	Nonvertebral	Rating*
Lufkin, 1992 ¹²⁷	Cyclic estradiol transdermal patch (0.1 mg) with oral medroxyprogesterone (10 mg) or placebo.	78	1	45-75	1 or more vertebral fractures; Mayo Clinic, Minnesota	0.39 (0.16-0.95)	NA	Fair
Wimalawansa, 1998 ¹²⁸	Cyclic premarin (0.625 mg/day) with norgestrel (150 microgm/12 days each month) or placebo; all subjects received calcium and vitamin D.	36	4	Mean 65	Established osteoporosis; attending metabolic bone disease outpatient clinics	HRT: 33.3/1,000 patient-yrs; calcium/vit D: 89.3/1,000 patient -yrs; No significant difference between groups	NA	Poor
Koulainen, 1998 ¹²	1. Sequential estradiol (2 mg/day) with cyproterone acetate (1 mg/day); 2. Vit D (300IU/day); 3. HRT & vit D; 4. Placebo.	464	4.3	Early post-menopause	No osteoporosis; a subgroup of the Kuopio Osteoporosis Study (n=13,100) based in Finland	NA	HRT alone: 0.29 (0.10-0.90); HRT & Vit D: 0.44 (0.17-1.15)	Fair

Table 8. Randomized Controlled Trials of Estrogen and SERMs with Fracture Outcomes (continued)

Mosekilde, 2000 ¹³⁰	Sequential estradiol (1-2 mg/day with norethisterone acetate (1 mg/day for 10 days each month; continuous estradiol (2 mg/day) if hysterectomy or non use.	1006	5	45-52	Postmenopausal Danish women recruited by mailed questionnaire.	2.0 (0.62-6.49)	All:0.82 (0.53-1.29); forearm: 0.40 (0.16-1.01); other: 0.96 (0.57-1.64)	Poor
Cauley, 2001 ¹³¹	Continuous combined conjugated equine estrogen (0.625 mg) with medroxyprogesterone (2.5 mg) or placebo.	2763	4.1	<80 (mean 67)	Heart and Estrogen/progestin Replacement Study (HERS); with coronary disease, intact uterus; fractures a secondary outcome.	0.69 (0.3-1.4)	Any: 0.94 (0.8-1.2); hip: 1.09 (0.5-2.3); wrist: 1.01 (0.3-1.4); other: 0.91 (0.7-1.2)	Fair
Lufkin, 1998 ¹³⁵	Raloxifene (60 or 120 mg/day) or placebo; all received calcium and vitamin D.	143	1	45-75 (mean 65)	At least 1 vertebral fracture and low BMD; Mayo Clinic, Minnesota and Arizona.	1.15 (0.75-1.75)	0.51 (0.12-2.16)	Good

Table 8. Randomized Controlled Trials of Estrogen and SERMs with Fracture Outcomes (continued)

Ettinger, 1999 ¹³⁴	Raloxifene (60 or 120 mg/day) or placebo; all received calcium and vitamin D.	7705	3	31-80 (mean 67)	Multiple Outcomes of Raloxifene Evaluation (MORE); 25 countries; met WHO criteria for osteoporosis.	0.59 (0.05-0.70)	0.91 (0.79-1.06)	Good
Fisher, 1998 ¹³⁶	Tamoxifen (20 mg/day) or placebo.	13,388	5	35 and over	National Surgical Adjuvant Breast and Bowel Project Breast Cancer Prevention Trial (P-1); high risk for breast cancer in US and Canada; fractures a secondary outcome.	0.74 (0.41-1.32)	Hip: 0.55 (0.25-1.15); Colles' 0.61 (0.29-1.23)	Fair

*Harris, 2001²⁷

Note: CI indicates confidence interval; SERMs, selective estrogen receptor modulators.

BMD indicates bone mass density; CI, confidence interval

Table 9. Randomized Controlled Trials of Alendronate with Fracture Outcomes

Author	Duration (years)	Age (years)	Population	Population	Exclusion Criteria*	Lost to Followup	Quality Rating†
Adami, 1995 ¹³⁸	2	48-76	9 Italian Centers, T-score < -2 (0.67g/cm ²); 5% vertebral fractures	9 Italian Centers, T-score < -2 (0.67g/cm ²); 5% vertebral fracture	Narrow	32/211 (15.2 %)	Fair to Good
Black, 1996 ¹³⁷	3	55-81	11 U.S. cities, BMD < 0.68 g/cm ² ; no previous vertebral fractures	11 U.S. cities, BMD < 0.68 g/cm ² ; no previous vertebral fracture	Broad (medical illness, dyspepsia, etc)	81/2027 (4%)	Good
Bone, 1997 ¹³⁹	2	>60	15 U.S. sites, BMD < 0.84 g/cm ² ; average 20 yrs since menopause; 30.7% vertebral fractures	15 US sites, BMD < .84 g/cm ² , average 20 years since menopause; 30.7% vertebral fractures	Broad (medical illness, NSAIDs, GI drugs)	19/359 (5.3%)	Fair to Good

Table 9. Randomized Controlled Trials of Alendronate with Fracture Outcomes (continued)

Chesnut, 1995 ¹⁴⁰	2	42-75 (avg. 63)	7 centers, spine BMD<0.88, average hip BMD 0.7; at least 5 years since menopause	7 centers, LS BMD<.88, average hip BMD .7; at least 5 years since menopause	Broad	26/157 (16.6%)	Fair
Cummings, 1998 ¹⁴⁶	4	55-81	11 U.S. cities, BMD<0.68 g/cm ² (aver. 0.59); no previous vertebral fractures	11 U.S. cities, BMD<0.68 g/cm ² (aver. 0.59), no previous vertebral fracture	Broad (medical illness, dyspepsia)	179/4432 (4%)	Good
Greenspan, 1998 ¹⁴⁴	2.5	Over 65	1 Boston center, no BMD entry criteria	Over 65	Narrow ("good health")	33/120 (27.5%)	Fair

Table 9. Randomized Controlled Trials of Alendronate with Fracture Outcomes (continued)

†Harris, 2001²⁷

*In general, "narrow" criteria excluded estrogen users and patients with illnesses affecting bone metabolism

GI=Gastrointestinal

BMI= Bone mass density

NSAID= Nonsteroidal anti-inflammatory drug

Table 10. Summary of Evidence Quality

Key Questions	Evidence Code*	Quality of Evidence†	
		Internal Validity	External Validity
Arrow 1: Does screening using risk factor assessment and/or bone density testing reduce fractures?	None		
Arrow 2: Does risk factor assessment accurately identify women who may benefit from bone density testing?	II-2	Poor-good: small studies, risk assessment instruments often not validated	Poor-fair: no instruments used widely for screening purposes although some were developed from
Arrow 3: Do bone density measurements accurately identify women who may benefit from treatment?	II-2	Fair-good: studies indicate the short-term predictability for fracture.	Fair: not known how well results of studies translate to practice.
Arrow 4: What are the harms of screening?	II-2, III	Poor-fair: small studies, descriptive.	Poor: small studies, selected subjects.
Arrow 5: Does treatment reduce the risk of fractures in women identified by screening?	I	Poor-good: no good-quality trials for estrogen. Good trials indicate fractures prevention for raloxifene and bisphosphonates.	Poor-fair: subjects of trials may be different than primary care patients.
Arrow 6: What are the harms of treatment?	I, II-2	Poor-good: long-term effects of newer agents not known. †Quality of evidence based on criteria developed by the USPSTF27	Poor-fair: difficult to know how risks affect individual patients.

*Evidence codes based on study design categories27:
 I = randomized, controlled trials
 II-1 = controlled trials without randomization
 II-2 = cohort or case-control analytic studies,
 II-3 = multiple time series, dramatic uncontrolled experiments
 III = opinions of respected authorities, descriptive studies

Table 11. Screening for Osteoporosis in 10,000 Postmenopausal Women

Hip and Vertebral Fracture Outcomes by 5-year Age Intervals

	Age (years)					
	50-54	55-59	60-64	65-69	70-74	75-79
<i>Base Case Assumptions†</i>						
Prevalence of osteoporosis	0.0305	0.0445	0.065	0.120	0.2025	0.285
Relative risk for hip fracture with treatment	0.63	0.63	0.63	0.63	0.63	0.63
Relative risk for vertebral fracture with treatment	0.52	0.52	0.52	0.52	0.52	0.52
Adherence to treatment	0.7	0.7	0.7	0.7	0.7	0.7
<i>Results, n</i>						
Identified as high-risk (osteoporotic)	305	445	650	1200	2025	2850
Hip fractures prevented	1	2	5	14	39	70
NNS to prevent 1 hip fracture	7446	4338	1856	731	254	143
NNT to prevent 1 hip fracture	227	193	121	88	51	41
Vertebral fractures prevented	5	7	22	40	95	134
NNS to prevent 1 vertebral fracture	1952	1338	458	248	105	75
NNT to prevent 1 vertebral fracture	60	60	30	30	21	21

*NNS = number needed to screen for benefit; NNT = number needed to treat. reduction for hip fractures143, and 50% risk reduction for vertebral fractures28; treatment compliance of 70% based on results of treatment trials and allowing less optimal compliance.

Appendix 1. Search Strategies for Individual Topics

Risk Factors

- 1 exp osteoporosis
osteoporosis, postmenopausal
- 2 bone density
- 3 1 or 2
- 4 exp risk
logistic models risk assessment
risk factors
- 5 3 and 4
- 6 exp transplantation
cell transplantation transplantation, autologous
organ transplantation transplantation, heterologous
replantation transplantation, heterotopic
tissue transplantation transplantation, homologous
- 7 exp kidney failure
kidney failure, acute diabetes insipidus, nephrogenic
kidney failure, chronic
- 8 su.fs. (Surgery as a subheading anywhere in the article)
- 9 athletic injuries
- 10 exp sports
baseball basketball bicycling
boxing golf football
gymnastics hockey mountaineering
racquetball running martial arts
skating skiing track and field
soccer swimming weight lifting
walking wrestling
- 11 exp fractures/dt,su,th (limited to surgery and other therapies)
femoral fractures fractures, closed
fractures, comminuted fractures, malunited
fractures, open fractures, spontaneous
fractures, stress fractures, ununited
humeral fractures radius fractures
rib fractures shoulder fractures
skull fractures spinal fractures
tibial fractures ulna fractures
- 12 orthopedic\$.mp. (As a textword anywhere)
- 13 6 or 7 or 8 or 9 or 10 or 11 or 12
- 14 5 not 13 (statements 6 through 12 were excluded from the study)
- 15 limit 13 to female

Risk Factors (continued)

- 16 limit 14 to human
- 17 limit 15 to english language
- 18 looked at english abstracts of foreign articles
- 19 exp fractures or exp osteoporosis or bone density

- 20 exp risk (terms as in 4)

- 21 19 and 20

- 22 exp cohort studies
 - longitudinal studies
 - follow-up studies
 - prospective studies

- 23 meta-analysis
- 24 exp case control studies
 - retrospective studies
- 25 predictive value of tests
- 26 evidence-based medicine

- 27 22 or 23 or 24 or 25 or 26

- 28 21 and 27

- 29 limit 28 to human
- 30 limit 29 to english language

Bone Density and Quantitative Ultrasound Testing

- 1 exp osteoporosis
osteoporosis, postmenopausal
- 2 bone density
- 3 1 or 2
- 4 densitometry, x-ray
- 5 exp ultrasonography
echocardiography ultrasonography, doppler
echoencephalography ultrasonography, interventional
endosonography ultrasonography, mammary
- 6 calcaneous/us (us = ultrasonics)
- 7 (dxa or sxa or bua or qct or qus or mxa or mrx or ra or dip or sos or ubps or spa or dpa).tw.
- 8 exp osteoporosis/us (us = ultrasonics)
- 9 4 or 5 or 6 or 7 or 8
- 10 3 and 9
- 11 limit 10 to human
- 12 limit 11 to english language
- 13 looked at english abstracts of foreign articles

Hormone Replacement Therapy

- 1 exp hormone replacement therapy
estrogen replacement therapy
- 1 hormone replacement.tw. (text word taken from title and abstract of article)
- 2 estrogen replacement.tw.
- 4 exp estrogens/ad,tu (ad = administration & dosage; tu = therapeutic use)
equilenin estrogens, catechol
equilin estrogens, conjugated
estradiol estrogens, non-steroidal
estriol estrone
- 5 exp estrogens, synthetic/ad,tu
estrogens, non-steroidal epimestrol
chlorotrianisene ethinyl estradiol
coumestrol mestranol
dienestrol quinestrol
diethylstilbestrol hexestrol
zearalenone zeranol
- 6 1 or 2 or 3 or 4 or 5
- 7 exp osteoporosis
osteoporosis, postmenopausal
- 8 exp fractures
femoral fractures fractures, closed
fractures, comminuted fractures, malunited
fractures, open fractures, spontaneous
fractures, stress fractures, ununited
humeral fractures radius fractures
rib fractures shoulder fractures
skull fractures spinal fractures
tibial fractures ulna fractures
- 6 fracture\$.tw.
- 10 bone density
- 11 7 or 8 or 9 or 10
- 12 6 and 11
- 13 limit 12 to human
- 14 limit 13 to english language
- 15 looked at english abstracts of foreign articles

Randomized Controlled Trials of Estrogen and SERMs

Estrogen

- 1 exp hormone replacement therapy
estrogen replacement therapy
- 2 hormone replacement.tw. (text word taken from title and abstract of article)
- 3 estrogen replacement.tw.
- 4 exp estrogens/ad,tu (ad = administration & dosage; tu = therapeutic use)
equilenin estrogens, catechol
equilin estrogens, conjugated
estradiol estrogens, non-steroidal
estriol estrone
- 5 exp estrogens, synthetic/ad,tu
estrogens, non-steroidal epimestrol
chlorotrianisene ethinyl estradiol
coumestrol mestranol
dienestrol quinestrol
diethylstilbestrol hexestrol
zearalenone zeranol
- 6 1 or 2 or 3 or 4 or 5
- 7 limit 6 to randomized controlled trials (check for document type)
- 8 randomized controlled trials
- 9 randomized.tw.
- 10 8 or 9
- 11 6 and 10
- 12 7 or 11
- 13 limit 12 to human
- 14 limit 13 to english language
- 15 looked at english abstracts of foreign articles

Tamoxifen and raloxifene

- 1 (tamoxifen or raloxifene).mp.
- 2 Bone density/ or "bone density".mp
- 3 exp osteoporosis/ or "osteoporosis".mp
- 4 exp fractures/ or fracture\$.mp.
- 5 exp hormone replacement therapy
- 6 estrogen replacement.mp.
- 7 2 or 3 or 4 or 5 or 6
- 8 1 and 7
- 9 limit 8 to (human and english language)
- 10 exp breast neoplasms/
- 11 9 not 10
- 12 from 11 keep 1-145

Appendix 2: Criteria for Grading the Internal Validity of Individual Studies

Design-Specific Criteria and Quality Category Definitions

Presented below are a set of minimal criteria for each study design and then a general definition of three categories-- “good,” “fair,” and “poor” --based on those criteria. These specifications are not meant to be rigid rules but rather are intended to be general guidelines, and individual exceptions, when explicitly explained and justified, can be made. In general, a “good” study is one that meets all criteria well. A “fair” study is one that does not meet (or it is not clear that it meets) at least one criterion but has no known important limitations. “Poor” studies have at least one important limitation.

Systematic Reviews

Criteria:

- Comprehensiveness of sources considered/search strategy used
- Standard appraisal of included studies
- Validity of conclusions
- Recency and relevance are especially important for systematic reviews

Definition of ratings from above criteria:

Good: Recent, relevant review with comprehensive sources and search strategies; explicit and relevant selection criteria; standard appraisal of included studies; and valid conclusions.

Fair: Recent, relevant review that is not clearly biased but lacks comprehensive sources and search strategies.

Poor: Outdated, irrelevant, or biased review without systematic search for studies, explicit selection criteria, or standard appraisal of studies.

Case Control Studies

Criteria:

- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both
- Response rate
- Diagnostic testing procedures applied equally to each group
- Measurement of exposure accurate and applied equally to each group
- Appropriate attention to potential confounding variable

Definition of ratings based on criteria above:

Good: Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.

Fair: Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80 percent or attention to some but not all important confounding variables.

Poor: Major selection or diagnostic work-up biases, response rates less than 50 percent, or inattention to confounding variables.

Randomized Controlled Trials and Cohort Studies

Criteria:

- Initial assembly of comparable groups
 - for RCTs: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups
 - for cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination)
- Important differential loss to followup or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention to treat analysis for RCTs.

Definition of ratings based on above criteria:

Good: Meets all criteria: comparable groups are assembled initially and maintained throughout the study (followup at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention to treat analysis is used.

Fair: Studies will be graded “fair” if any or all of the following problems occur, without the fatal flaws noted in the “poor” category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in followup; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention-to-treat analysis is done for RCTS.

Poor: Studies will be graded “poor” if any of the following fatal flaws exists: groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention-to-treat analysis is lacking.

Diagnostic Accuracy Studies

Criteria:

- Screening test relevant, available for primary care, adequately described
- Study uses a credible reference standard, performed regardless of test results
- Reference standard interpreted independently of screening test
- Handles indeterminate results in a reasonable manner
- Spectrum of patients included in study
- Sample size
- Administration of reliable screening test

Definition of ratings based on above criteria:

Good: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (more than 100) broad-spectrum patients with and without disease.

Fair: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50 to 100 subjects) and a “medium” spectrum of patients.

Poor: Has fatal flaw such as: uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size of very narrow selected spectrum of patients.