

# Screening and Monitoring Tests for Osteopenia/Osteoporosis

---

## Final Report

October 20, 2014

**Health Technology Assessment Program (HTA)**

Washington State Health Care Authority

P.O. Box 42712

Olympia, WA 98504-2712

(360) 725-5126

[www.hca.wa.gov/hta](http://www.hca.wa.gov/hta)

[shtap@hca.wa.gov](mailto:shtap@hca.wa.gov)



# ***Screening and Monitoring Tests for Osteopenia/Osteoporosis***

## **A Health Technology Assessment**

**Prepared for Washington State Healthcare Authority**

### **FINAL REPORT**

**October 20, 2014**

#### **Acknowledgement**

This report was prepared by:

Hayes, Inc.  
157 S. Broad Street Suite 200  
Lansdale, PA 19446  
P: 215.855.0615 F: 215.855.5218

This report is intended to provide research assistance and general information only. It is not intended to be used as the sole basis for determining coverage policy or defining treatment protocols or medical modalities, nor should it be construed as providing medical advice regarding treatment of an individual's specific case. Any decision regarding claims eligibility or benefits, or acquisition or use of a health technology is solely within the discretion of your organization. Hayes, Inc. assumes no responsibility or liability for such decisions. Hayes employees and contractors do not have material, professional, familial, or financial affiliations that create actual or potential conflicts of interest related to the preparation of this report

## Table of Contents

EVIDENCE SUMMARY.....	1
Summary of Background.....	1
Summary of Technical Aspects of DXA .....	7
Policy Context .....	8
Summary of Review Objectives and Methods.....	9
Summary of Search Results.....	13
Findings .....	16
Key Question #1: Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 1a: For individual patients, and do these outcomes vary according to age, sex, or other risk factors for BMD or fracture? 1b: In populations, and do these outcomes vary by population characteristics?.....	16
Key Question #2: Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 2a: For individual patients, and do these outcomes vary according to age, sex, other risk factors (including previous BMD measurements), treatment status, or testing interval? 2b: In populations, and do these outcomes vary by population characteristics or testing interval? .....	22
Key Question #2c: What is the minimum interval required to detect transition from normal or low BMD to osteoporosis or to assess treatment effect? .....	22
Key Question #3: What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?.....	27
Key Question #4: Are bone density tests safe and what are the potential downstream adverse effects? .....	30
Key Question #5: What are the costs and cost-effectiveness of osteoporosis screening and monitoring? .....	32
Practice Guidelines.....	34
Selected Payer Policies.....	38
Overall Summary and Discussion.....	39
TECHNICAL REPORT .....	42
Clinical Background.....	42
Technical Aspects of DXA.....	58
Review Objectives and Analytic Framework.....	60
Methods.....	63
Search Results .....	66
Literature Review .....	67

KQ#1: Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 1a: For individual patients, and do these outcomes vary according to age, sex, or other risk factors for BMD or fracture? 1b: In populations, and do these outcomes vary by population characteristics?.....	67
KQ#2: Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 2a: For individual patients, and do these outcomes vary according to age, sex, other risk factors (including previous BMD measurements), treatment status, or testing interval? 2b: In populations, and do these outcomes vary by population characteristics or testing interval? .....	71
KQ #2c: What is the minimum interval required to detect transition from normal or low BMD to osteoporosis or to assess treatment effect? .....	71
KQ #3: What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?.....	77
KQ #4: Are bone density tests safe and what are the potential downstream adverse effects?.....	81
KQ #5: What are the costs and cost-effectiveness of osteoporosis screening and monitoring? .....	82
Practice Guidelines.....	86
Selected Payer Policies.....	88
References .....	92
Washington State Agency Utilization Data .....	99
APPENDICES .....	114
APPENDIX I. Search Strategy .....	114
APPENDIX II. Overview of Evidence Quality Assessment Methods .....	117
APPENDIX III. Evidence Tables .....	121
Evidence Table IIIa. Controlled Studies Evaluating the Effectiveness of Osteoporosis Screening by Dual X-Ray Absorptiometry (DXA) in Individuals (Key Question #1a) .....	121
Evidence Table IIIb. Longitudinal Studies Designed to Estimate Optimal Screening Intervals (Key Question #2c) .....	124
Evidence Table IIIc. Studies Reporting Calculations of Numbers Needed to Screen (NNS) (Key Question #3) .....	127
Evidence Table IIId. Economic Evaluations (Key Question #5).....	129
Appendix IV. Summary of Practice Guidelines.....	133

## List of Tables

[Table 1.](#) Interpretation of DXA Results

[Table 2.](#) Summary of Search Results

[Table 3.](#) Summary of Findings, Key Questions #1a and #1b

[Table 4.](#) Summary of Findings, Key Question #2c.

[Table 5.](#) Summary of Findings, Key Question #3

[Table 6:](#) Summary of Findings, Key Question #4

[Table 7.](#) Summary of Findings, Key Question #5

[Table 8.](#) Summary of Practice Guideline Recommendations

[Table 9.](#) Components of the WHO FRAX Tool

[Table 10.](#) Prescription Medications for Osteoporosis

## EVIDENCE SUMMARY

The **EVIDENCE SUMMARY** summarizes background information, the methods and search results for this report, findings with respect to the Key Questions, and payer policies and practice guidelines. The **EVIDENCE SUMMARY** also includes conclusions and an assessment of the quality of the evidence for each Key Question. In general, references are not cited in the **EVIDENCE SUMMARY**. The **EVIDENCE SUMMARY** ends with an **Overall Summary and Discussion**. The **TECHNICAL REPORT** provides additional detail, with full citations, regarding background information, study results, and payer policies and guidelines.

### Summary of Background

#### Osteoporosis and Dual X-Ray Absorptiometry (DXA)

Osteoporosis is the most common bone disease in humans. It is a systemic skeletal disease involving low bone mass and microarchitectural deterioration, both of which lead to fragility and increased risk of fracture. According to the latest available (2005 to 2006) data from the National Health and Nutrition Examination Survey (NHANES III), osteoporosis of the hip is prevalent in 4% of American men older than age 50 and in 16% of American women older than age 50. Other sources suggest that as many as 50% of Americans who are older than age 50 will be at risk for osteoporotic fracture during their lifetime. Prevalence is expected to increase as the proportion of the population older than age 65 increases. A large economic burden due to osteoporotic fractures is demonstrated by recent findings that these events cause more than 432,000 hospital admissions, 2.5 million medical office visits, and approximately 180,000 nursing home admissions per year in the United States (U.S.). Hip fractures are associated with considerable excess mortality, estimated at 8.4% to 36% for 1 year. Mortality related to hip fracture is higher in men than in women. *Peak* bone mass, which occurs around the age of 30, is largely determined by genetics. Loss of bone mass occurs thereafter in general populations as the result of age-related hormonal changes. In women, bone loss usually occurs more rapidly for several years after menopause and then slows down again so that men and women age 65 to 70 years and older lose bone mass at about the same rate. The World Health Organization (WHO) has defined osteopenia and osteoporosis in terms of bone mineral density (BMD) at the hip or lumbar spine, as measured by dual x-ray absorptiometry (DXA). (See additional details regarding the WHO definitions in the following paragraph.) Alternatively, a diagnosis of osteoporosis is considered valid on the basis of adulthood hip or vertebral fracture in the absence of major trauma. Examples of major trauma are an automobile accident or a fall from a multiple-story height.

In clinical practice, the standard technology for measuring BMD and diagnosing osteopenia or osteoporosis is DXA. BMD can be expressed as grams of mineral per square centimeter ( $\text{g}/\text{cm}^2$ ) scanned or as a score that expresses the relationship to normal values. A DXA T-score represents a comparison to

a *young adult* reference population of the same sex and is used to express the relative BMD status of older adults. A DXA Z-score represents a comparison to the BMD of an age-, sex-, and ethnicity-matched reference population, and is commonly used to express the BMD status of premenopausal women, men younger than 50 years old, and children. Z-score cutoff values have been defined by the International Society for Clinical Densitometry (ISCD). The ISCD cautions that *osteoporosis cannot be diagnosed in men under the age of 50 on the basis of BMD status alone*, but advises that the *WHO definition of osteoporosis is appropriate for women in menopausal transition*. See **Table 1** for cutoff values.

**Table 1. Interpretation of DXA Results**

WHO Definitions for Postmenopausal Women and Men Older Than 50 Years	
T-score –1.0 and above:	Normal
T-score above –2.5 but below –1.0:	Low bone mass (osteopenia)
T-score at or below –2.5:	Osteoporosis
T-score at or below –2.5 with $\geq 1$ fractures:	Severe or established osteoporosis
ISCD Definitions for Premenopausal Women, Men Younger Than 50 Years, and Children	
Z-score above –2.0:	BMD within the expected range for age
Z-score at or below –2.0:	Low BMD for chronological age

### The Rationale for Screening and Monitoring

Several expert sources point to the underdiagnosis and undertreatment of osteoporosis. The potential utility of screening for osteoporosis relates to the opportunity to identify individuals for whom treatment is appropriate. Research suggests that, among older women with an osteoporotic fracture, fewer than 25% receive either a BMD test or a prescription for an osteoporotic drug in the 6 months following fracture. In patients who are being treated for osteoporosis or low bone mass, the objectives of monitoring BMD are to determine whether treatment is working and to assess the appropriateness of treatment cessation. There is no consensus on how frequently BMD should be tested in order to detect a meaningful change in fracture risk.

### Osteoporosis and Fracture

The chief clinical concern associated with osteoporosis is risk of fracture. Otherwise, osteoporosis does not produce symptoms. Fractures that are thought to be attributable to osteoporosis are variously referred to as *osteoporotic fractures*, *fragility fractures*, *low-stress fractures*, and *nontraumatic fractures*. The lifetime risk of osteoporotic fracture among white women is approximately 50%. Chronic pain, disability, and even death can occur because of fracture in an older population. Compared with other types of fragility fractures, hip fractures tend to have the greatest impact on mortality, function, and quality of life. The excess 1-year mortality rate associated with hip fractures has been estimated at 8.4% to 36%.

## Risk Factors for Osteoporosis

A wide range of factors have may cause or contribute to osteoporosis. These include reversible lifestyle factors such as alcohol abuse, vitamin D and calcium intake, physical activity, body mass index (BMI), and smoking; hormonal disorders; type 1 diabetes mellitus; malnutrition or conditions that cause malabsorption; rheumatoid arthritis; and a variety of medications such as glucocorticoids and androgen-deprivation therapy (ADT).

The current recommendations of the U.S. Preventive Services Task Force (USPSTF) regarding screening for osteoporosis are stated primarily in terms of sex and age but also refer to “additional risk factors.” Specifically, screening is recommended for women age  $\geq 65$  years without previous known fractures or secondary causes of osteoporosis, and in women age  $< 65$  years whose 10-year fracture risk is equal to or greater than that of a 65-year-old white woman without additional risk factors. Those risk factors are not defined in either the published or online versions of the USPSTF recommendation statement. Nor does the recommendation specify the tool that should be used to calculate risk. However, the USPSTF chose the Fracture Risk Assessment Tool (FRAX) as the best instrument for its own estimates of 10-year fracture risk for different risk profiles. The evidence review (referred to throughout the present report as the *2010 Nelson* review) sponsored by the Agency for Healthcare Research and Quality (AHRQ) to support the updated USPSTF recommendation found that the FRAX tool estimated the 10-year fracture risk of a 65-year-old white woman with no more than 1 additional factor to be 9.3% for any osteoporotic fracture and 1.2% for hip fracture. The authors of the report then used the FRAX tool to identify risk factor and age combinations for which fracture risk would exceed that of the index case of a 65-year-old white woman.<sup>1</sup> See [Table 9](#) in the **TECHNICAL REPORT** for a list of the risk factors for osteoporosis and/or fracture that are included in the FRAX tool and are most commonly used in clinical practice.

A wide range of risk factors include reversible lifestyle factors such as alcohol abuse, vitamin intake, physical activity, and smoking; genetic diseases; hormonal disorders; diabetes mellitus; hyperparathyroidism; gastrointestinal disorders; hematologic disorders; rheumatologic and autoimmune diseases; numerous other conditions; and a variety of medications. Neither the National Osteoporosis Foundation (NOF) guidelines nor those of other organizations provide much detail regarding the evidence or biologic rationale for most associations. A search for systematic reviews published in the last 10 years identified no comprehensive and systematic assessment of the direction of causality and strength of association for the many factors linked to osteoporosis. However, some

---

<sup>1</sup> See Figure 3 in [Screening for Osteoporosis: Systematic Review to Update the 2002 U.S. Preventive Services Task Force Recommendation \[Internet\]](#). Under the heading *Risk for Osteoporotic or Hip Fracture - >one risk factor*, gray shading is used to identify women *younger than age 65* whose risk profile would make them eligible for screening according to the USPSTF recommendation. For example, a 55-year-old woman with low body mass index and a parent who had a hip fracture has an 11% 10-year risk of any osteoporotic fracture and a 0.7% 10-year risk of a hip fracture.



relevant systematic reviews and large observational studies have evaluated the comparative contribution of particular sets of factors to the risk of osteoporosis or the prevalence of osteoporosis in certain diseases (see corresponding section in the **TECHNICAL REPORT** for more detail).

### **Tools for Assessing Risk of Osteoporosis and Risk of Fracture**

A meta-analysis published in 1996 and cited by key review articles established a strong association between BMD, as measured by DXA, and fracture, with the best predictor being hip BMD. The 2010 Nelson review evaluated 14 externally validated instruments that can be used to predict a DXA T-score less than or equal to  $-2.50$  (osteoporosis). The tools used various combinations of age, sex, and the factors mentioned in the preceding section. The most commonly included elements were age, weight or BMI, and previous fracture. DXA scores as a measure of BMD can in turn be used to predict fracture risk. Estimates of the relative risk (RR) of hip fracture per decrease of 1 standard deviation (SD) in femoral neck BMD in older adults have been in the range of 2.3 to 2.6. In other words, for every 1 SD in femoral neck BMD lower than the normal value for a healthy young reference population, the risk of hip fracture is more than doubled in older adults. However, BMD does not completely explain the incidence of fractures in older women, and some experts have reported that more than half of individuals with fragility fractures do not meet the BMD criteria for a diagnosis of osteoporosis (although they may have osteopenia, or low bone mass).

Clinical factors, either alone or in combination with DXA score, can also be used to predict fracture risk. The 2010 Nelson review identified 11 externally validated instruments for assessment of fracture risk. Only 3 instruments included BMD (DXA) T-score: Fracture Index, FRAX, and Garvan nomogram. At least 1 of these 3 instruments, FRAX, will also compute risk if BMD is not available. (See **BACKGROUND, Assessment of Fracture Risk, Performance of Clinical Factors With and Without DXA Score for Prediction of Fracture** for more detail on the FRAX tool.) Nelson and colleagues did not comment on the relative performance of tools that did or did not include BMD. Five of the fracture risk tools took into account 1 or more secondary causes of osteoporosis, the most common being glucocorticoid use. Nelson and colleagues stated similar conclusions for both DXA/BMD prediction tools and fracture risk prediction tools: (1) simple tools performed as well as the more complex tools both for predicting BMD and for predicting fracture; and (2) tools may have limited generalizability to clinical populations since they were validated in general populations.

### **Technologies Other Than DXA for Assessment of Bone Health and Predicting Fracture**

DXA is a successor to dual-photon absorptiometry (DPA). Other older technologies include single-photon absorptiometry (SPA) and single-energy x-ray absorptiometry (SXA). Quantitative computed tomography (QCT) produces volumetric measures of BMD and predicts fracture, but neither the performance of QCT as a screening tool nor the relationship of QCT measurements with standard T-scores has been very well studied. Quantitative ultrasound densitometry (QUS) measures the speed of sound and/or broadband ultrasound attenuation at the heel, tibia, patella, and other peripheral skeletal sites. These 2 measurements are used to compute a composite clinical score that represents features of bone quality other than BMD. There is high but not perfect concordance between QUS and central DXA

measurements, but QUS is not considered a standard tool for diagnosing osteoporosis. Biochemical markers of bone remodeling are associated with risk of fracture in untreated patients (independent of BMD), predict rapidity of bone loss in untreated patients, predict the extent of fracture risk reduction when repeated after 3 to 6 months of treatment, predict the magnitude of BMD increases after initiation of treatment, and reflect patient adherence to treatment. However, bone turnover markers are not considered an alternative to BMD testing or screening; they are used primarily in research settings and have limited clinical application.

## Treatment

Nonpharmacologic treatment of osteoporosis includes increasing the intake of calcium and vitamin D, as well as exercise. A large number of medications may be used to prevent or treat osteoporosis. Bisphosphonates are the first-line drugs for treating diagnosed osteoporosis in postmenopausal women, but no first-line recommendation is made for men.

### *Threshold for Medical Treatment*

According to major practice guidelines, 1 or more of the following factors is considered by expert sources to be sufficient to offer medical treatment:

- A clinical or radiographic fracture of the spine or hip
- A hip DXA T-score  $\leq -2.5$
- Osteopenia *and* a 10-year WHO probability of a hip fracture that is  $\geq 3\%$  or a osteopenia and a 10-year risk of major osteoporotic fracture  $\geq 20\%$

### *Treatment Efficacy*

Treatment efficacy is relevant to the question of whether screening for osteopenia or osteoporosis is effective. In the absence of an effective management strategy or treatment, a screening program cannot be effective. Even though adherence to medical treatment is poor, efficacy has been demonstrated. The ability of Food and Drug Administration (FDA)-approved drugs to reduce fracture risk has been studied primarily in postmenopausal women without secondary causes of low bone mass, and the overwhelming majority of trials have involved bisphosphonates. The 2010 Nelson review concluded that bisphosphonates, parathyroid hormone (PTH), raloxifene, and estrogen reduce primary vertebral fractures in postmenopausal women, and that according to sensitivity analyses, bisphosphonates also reduce nonvertebral fractures in this population. The authors considered the evidence to be poor to good, depending on the medication. They noted that because of strict patient enrollment criteria (e.g., exclusion of comorbid conditions and use of other medications), the evidence may not be generalizable to typical clinical populations.

Less is known about the effectiveness of medications for treatment of glucocorticoid-induced osteoporosis, for *prevention* of osteoporosis in patients at high risk because of secondary causes, or for treatment of men. The authors of 3 systematic reviews, including the 2010 Nelson review, considered the evidence of medical osteoporosis treatment in men to be inconclusive. Systematic reviews have

shown osteoporosis medication to have a possible effect on fracture risk in patients with biliary cirrhosis (nonsignificant RR) and to have a significant effect on fracture risk in patients with rheumatoid arthritis or spinal cord injury. The review of osteoporosis treatment and rheumatoid arthritis found the effect to be especially strong in patients taking glucocorticoids. Other systematic reviews have shown *no* effect on fracture risk in patients with chronic obstructive pulmonary disease (COPD), patients with Crohn's disease plus a diagnosis of osteopenia or osteoporosis, or patients who are taking glucocorticoids for a variety of other inflammatory disorders. The reviews of treatment in populations with special medical conditions provide little information on the differential effectiveness of osteoporosis medication according to different combinations of risk factors.

Also relevant to the question of whether screening is effective is whether treatment efficacy is dependent on baseline BMD. According to the Nelson 2010 review, data were too sparse to allow a pooled analysis, and the only single trial that provided a stratified analysis of baseline BMD and efficacy was the large and pivotal Fracture Intervention Trial (FIT) of alendronate in postmenopausal women. In FIT, alendronate was shown to have a statistically significant effect on fracture incidence (any, vertebral, and hip) only in patients with a baseline femoral neck BMD at or below the threshold for a BMD-based diagnosis of osteoporosis ( $T\text{-score} \leq -2.5$ ). Women with prior fragility fracture were excluded from FIT. In contrast, an AHRQ evidence review conducted in 2012 concluded that moderate evidence from post hoc analyses demonstrated that low femoral neck BMD did *not* predict the effect of alendronate on vertebral fracture or nonvertebral fracture risk.

### *Safety of Bone Drugs*

A wide range of adverse events have been reported in conjunction with use of osteoporosis medications. The 2010 Nelson review found that reliable estimates of the incidence of *serious* events are not easily computed; the evidence pertaining to serious adverse events is summarized in the following way (Nelson et al., 2010a):

- The evidence of serious upper gastrointestinal adverse events, atrial fibrillation, and osteonecrosis of the jaw in otherwise healthy patients taking bisphosphonates for fracture prevention has not been consistent.
- The evidence regarding the harms associated with calcitonin and PTH is limited.
- According to meta-analysis, raloxifene increases the risk of thromboembolic events (RR, 1.60; 95% confidence interval [CI], 1.15 to 2.23; 2 trials), but is not associated with coronary heart disease or stroke.

Other systematic reviews have reported no association of bisphosphonates with gastrointestinal cancer, esophageal cancer, gastric cancer, or colorectal cancer, but have reported an increased risk of atypical femur fracture attributable to bisphosphonates and an increased risk of osteonecrosis associated with bisphosphonate use. The absolute risk of atypical femur fracture and that of osteonecrosis are both very small (fractions of a percent).

## Summary of Technical Aspects of DXA

### How DXA Scanning Works

A DXA system consists of an x-ray source underneath the examining table on which the patient lies, and a detection system that moves over the patient. Both high-energy and low-energy photons are produced by the x-ray tube; attenuation of these beams is measured by the detection system. The attenuation values of soft tissues are subtracted from total attenuation values according to an algorithm, and the remaining value represents the degree to which the energy is diminished as it passes through bone. These resulting values are compared with standard values in phantoms of known density to produce bone mineral content in grams. Dividing the number of grams by the scanned area in centimeters squared ( $\text{cm}^2$ ) yields the BMD value. In the U.S., the NHANES III survey data serve as the basis for a reference standard for total hip and femoral neck T-scores, and manufacturers use their own data to define reference standards for lumbar spine T-scores.

BMD measurements can be made at central sites, i.e., the spine and femur, or at peripheral sites such as the calcaneus (heel), proximal phalanges of the hand, the tibial shaft, and the radius. For purposes of diagnosing osteoporosis and assessing fracture risk, DXA scans of the lumbar spine and proximal femur are emphasized since fractures at these sites are the most severe.

### FDA Approval

As of July 21, 2014, numerous bone densitometers, including DXA scanning devices, have been cleared for marketing as devices intended for medical purposes to measure bone density and mineral content by x-ray or gamma ray transmission measurements through the bone and adjacent tissues.

### Precision

An important area of concern with DXA scanning is the *precision*, or reproducibility, of the measurements. The reason precision is of particular concern with DXA scanning is that changes in BMD occur very slowly, so it is possible that the difference in BMD values between repeated tests in an individual may primarily reflect the imprecision of the scanning process rather than a true biological change. Precision varies across machines, facilities, and operators, and is also affected by patient-related factors. Furthermore, it varies across anatomical scanning sites. To quantify precision error, the ISCD recommends that facilities calculate a *least significant change* (LSC) value for each machine and operator. The combined LSC values for machine and operator(s) at each anatomical site define the site-specific precision errors for a particular facility. If a patient is repeatedly scanned on the same machine and by the same operator, and the magnitude of BMD change exceeds the relevant LSC value, the BMD change is interpreted as representing a statistically significant change due at least in part to true biological change. LSC values are typically in the range of 3% to 6% at the hip and 2% to 4% at the spine. It is problematic when the facility, machine, operator, or number of sites scanned does not remain the same as a patient is monitored over time.

## Safety of DXA

The radiation dose of a modern DXA scan is small, but it is sufficient to be taken into account in large-scale population screenings. The effective radiation dose from a routine DXA scan of the lumbar spine and hip is typically 10 microsievert ( $\mu\text{Sv}$ ), while radiation from cosmic rays and naturally occurring radioactive materials in the earth and human bodies amounts to a daily background dose of about 8  $\mu\text{Sv}$ . A conventional chest x-ray consisting of posterior-anterior and lateral views delivers an effective dose of 60  $\mu\text{Sv}$ . A conventional mammogram delivers about 130  $\mu\text{Sv}$ .

## Policy Context

In 2011, the U.S. Preventive Services Task Force (USPSTF) issued updated recommendations on screening for osteoporosis, taking into account 2 systematic evidence reviews prepared by or for the Agency for Healthcare Research and Quality (AHRQ) in 2002 and 2010.

The following are current USPSTF recommendations:

- Screening for osteoporosis should be conducted in women age  $\geq 65$  years without previous known fractures or secondary causes of osteoporosis. (Grade B recommendation)
- Screening for osteoporosis should be conducted in women age  $< 65$  years whose 10-year fracture risk is equal to or greater than that of a 65-year-old white woman without additional risk factors. (Grade B recommendation) [NOTE: The USPSTF chose the FRAX tool for assessment of fracture risk. See **Tools for Assessing Risk of Osteoporosis and Risk of Fracture** under **Summary of Background** in the **EVIDENCE SUMMARY**, or the corresponding section in the **TECHNICAL REPORT** for detail on the FRAX tool.]
- No recommendation may be made for men who have no previous known fractures or secondary causes of osteoporosis. (Grade I for insufficient evidence)

The evidence reviews found no randomized controlled trials (RCTs) of screening. However, the authors of the review contended that support for population screening could be derived from evidence regarding the reliability of tools used to assess fracture risk and from evidence regarding the ability to reduce fracture risk through treatment.

Regarding the optimal interval for screening, the USPSTF has no formal recommendation but advises that a minimum of 2 years may be needed to reliably measure a change in BMD, and that longer intervals may be necessary to improve fracture risk prediction.

In 2013, the Oregon Health Evidence Review Commission (HERC) issued guidance that mirrors the USPSTF recommendations for initial screening in men and women. The guidance statement also recommends that screening should not be repeated more often than every 15 years for women with normal BMD, every 4 years for women with moderate osteopenia, and every 2 years for women with advanced osteopenia or osteoporosis. The recommendations for repeat screening are based on a recently published population-based study designed to address this issue. This particular study is reviewed in the present report as evidence pertaining to Key Question #2c. The Oregon HERC guidance

document did not cite any new evidence pertaining to screening in men and recommended against routine screening in men.

This health technology assessment (HTA) was commissioned on the basis that an analysis of the evidence supporting current public health and policy statements and an analysis of the most recently published evidence can help promote the most appropriate use of osteoporosis screening for beneficiaries of Washington Health Care Authority (HCA) plans. Additionally, this report analyzes the evidence for questions not addressed by the USPSTF statement and corresponding review, which will permit a more comprehensive policy. Additional questions investigated here include screening in individuals with risk factors other than age and sex, and serial testing for purposes of treatment monitoring in patients taking osteoporosis medications.

## Summary of Review Objectives and Methods

### Review Objectives

**Population:** Adult men and women.

**Interventions:** Bone mineral density (BMD) testing with dual x-ray absorptiometry (DXA).

**Comparisons:** Clinical assessment of fracture risk or treatment success without BMD testing.

**Outcomes:** Health outcomes such as fractures, fracture-related morbidity, fracture-related mortality; intermediate outcomes such as clinical management decisions and patient behavior; harms associated with screening, including potential harms resulting from osteoporosis treatment; cost and cost-effectiveness.

### Key Questions

1. Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices?
  - 1a. For individual patients—and do these outcomes vary according to age, sex, or other risk factors for BMD or fracture?
  - 1b. In populations—and do these outcomes vary by population characteristics?
2. Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices?
  - 2a. For individual patients—and do these outcomes vary according to age, sex, other risk factors (including previous BMD measurements), treatment status, or testing interval?
  - 2b. In populations—and do these outcomes vary according to population characteristics or testing interval?

- 2c. What is the minimum interval required to detect transition from normal or low BMD to osteoporosis, or to assess treatment effect?
3. What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?
4. Are bone density tests safe and what are the potential downstream adverse effects?
5. What are the costs and cost-effectiveness of osteoporosis screening and monitoring?

NOTE: Improvement of outcomes “in populations” (Key Questions #1b and #2b) was assumed to refer to an assessment based on either individual- or group-level data for an entire community or region and were analyzed for the purposes of assessing the effect of a public health program, as opposed to data from a clinic setting or from a community sample.

### **Analytic Framework**

See **TECHNICAL REPORT, Review Objectives and Analytic Framework**.

### **Methods**

See the **Methods** section of the **TECHNICAL REPORT**, [Appendix I](#), and [Appendix II](#) for additional details.

#### *Search Strategy and Selection Criteria*

Core databases, PubMed, and the websites of relevant specialty societies were searched for systematic reviews, meta-analyses, economic evaluations, and practice guidelines published in the last 10 years. Systematic reviews were selected if they reviewed studies considered eligible for answering the Key Questions or if they provided useful background information. A variety of search strategies were used to identify primary studies that have been published since the 2010 evidence review conducted for AHRQ to support the USPSTF recommendation (Nelson et al., 2010a; Nelson et al., 2010b) and to identify studies published before that date but excluded by the AHRQ review. Initial searches for primary studies were conducted in PubMed on June 12, 2014, except for search #3, which was conducted on July 17, 2014. An update search was conducted on August 4, 2014. PubMed searches were restricted to articles published in the English language.

#### *Inclusion Criteria*

- To answer Key Questions #1a, #1b, #2a, #2b, and #4: Randomized controlled trials (RCTs) or quasi-RCTs, cohort studies, or case-control studies that assessed the impact of osteoporosis screening or testing on the outcomes of interest were considered. In other words, any study that involved a comparison between a screened group and an unscreened group was considered. The search was not limited to RCTs since the evidence review supporting the current USPSTF recommendation on screening for osteoporosis identified no RCTs of screening strategies.



- To answer Key Question #2c, the following studies were considered: (1) trials or cohort studies that compared different strategies for the timing of screening or serial monitoring; (2) longitudinal studies that serially measured BMD and assessed some measure of time to change in osteoporosis status (normal, osteoporosis, or osteopenia) and/or fracture.
- To answer Key Question #3: Any relevant published analyses of NNS were included. In addition, event rates in studies selected for Key Questions #1 and #2 were to be used to calculate NNS estimates.
- To answer Key Question #5: Any cost studies or economic evaluations published within the last 10 years were included.
- Systematic reviews of any of the above were also included.

#### *Exclusion Criteria*

No exclusion criteria were defined a priori other than publication of economic evaluations earlier than 10 years prior to the search date.

#### *Number-Needed-to-Screen (NNS) Calculations*

Where fracture rates were reported, NNS was calculated from the results of screening studies selected for Key Question #1 or #2. NNS was assumed to be equivalent to number-needed-to-treat (NNT). The following formula for NNT was used (Gordis, 2000):

$$\text{NNT} = \frac{1}{(\text{rate in untreated [unscreened] group}) - (\text{rate in treated [screened] group})}$$

Screening studies reported fracture incident rates in terms of cumulative incidence per person-year. NNS values were calculated according to the preceding formula by using the 1-year cumulative incidence of fractures for each group (screened and control). These values were then adjusted to represent the NNS to prevent 1 fracture over the time frame represented by the study's mean follow-up interval. In other words, in a study with a mean follow-up of 5 years, the NNS to prevent fracture over 1 year, as calculated by the formula, was then divided by 5 to estimate the NNS to prevent 1 fracture over a 5-year period. We adopted this adjustment to be consistent with the analysis reported in the 2010 evidence review conducted for AHRQ to support the USPSTF recommendation (Nelson et al., 2010a; Nelson et al., 2010b). The analysis by Nelson and colleagues expressed NNS in terms of 5 years, which was the mean follow-up interval of the bisphosphonate trial that served as the basis for their assumed fracture rates.

#### *Quality Assessment*

The process used by Hayes for assessing the quality of primary studies and bodies of evidence is in alignment with the methods recommended by the GRADE Working Group. Like the GRADE Working Group, Hayes uses the phrase *quality of evidence* to describe bodies of evidence in the same manner that other groups, such as AHRQ, use the phrase *strength of evidence*. A tool created for internal use at



Hayes was used to guide interpretation and critical appraisal of economic evaluations. The tool for economic evaluations was based on best practices as identified in the literature and addresses issues such as the reliability of effectiveness estimates, transparency of the report, quality of analysis (e.g., the inclusion of all relevant costs, benefits, and harms), generalizability/applicability, and conflicts of interest. The Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool, along with a consideration of commercial funding and conflicts of interest among the guideline authors, was used to assess the quality of practice guidelines. See the **Methods** section of the **TECHNICAL REPORT** and [Appendix II](#) for details on quality assessment methods.

## Summary of Search Results

One systematic review (Nelson et al., 2010a; Nelson et al., 2010b) and 8 primary studies were selected for detailed analysis as evidence pertaining to the Key Questions. **Table 2** identifies by Key Question the evidence that met selection criteria. See the **Search Results** section of the **TECHNICAL REPORT** for a description of 10 studies that were excluded from analysis after full text review. For some questions, additional studies that did not meet inclusion criteria or were omitted after full text review and did not contribute to conclusions were reviewed briefly to provide additional policy-relevant information, but were not considered in analysis and are not reflected in **Table 2**. The additional policy-relevant evidence is discussed primarily in the **TECHNICAL REPORT**.

**Table 2. Summary of Search Results**

**Key: ADT, androgen-deprivation therapy; NNS, number needed to screen; RCT, randomized controlled trial**

Systematic Reviews	RCTs	Nonrandomized Controlled or Quasirandomized Trials; Observational Studies With Controls; Eligible Longitudinal Studies	Other
<b>KQ #1 – Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices?</b>			
<b>1a. For individual patients</b>			
0*	General populations: Sedlack 2007 (women age 50-65 yrs); Barr 2010 (women age 45-54 yrs)  2	General populations: Kern 2005 (age ≥65 yrs; men and women); Doheny 2011 (older men) Individuals taking medications known to be associated with osteoporosis: Khan 2014 (men, ulcerative colitis, corticosteroids); Zhumkhawala 2013 (ADT for prostate cancer)  4	0
<b>1b. In populations</b>			
0	0	0	0

Systematic Reviews	RCTs	Nonrandomized Controlled or Quasirandomized Trials; Observational Studies With Controls; Eligible Longitudinal Studies	Other
<b>KQ #2 – Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices?</b>			
<b>2a. For individual patients</b>			
0	0	0	0
<b>2b. In populations</b>			
0	0	0	0
<b>2c. Minimum interval</b>			
0+	0	Screening: Frost 2009 (older men and women in Australia); Gourlay 2012 (postmenopausal women in the U.S.) Treatment monitoring: None (indirect data for both indications discussed but not included in analysis)	2 0
<b>KQ #3 –What is the NNS to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?</b>			
Nelson 2010a; 2010b (special analysis in systematic review based on assumptions regarding prevalence and efficacy of treatment) 1	0	Kern 2005, Zhumkhawala 2013, Khan 2014 (calculations based on difference in fracture incidence rates) 3	0
<b>KQ #4 – Are bone density tests safe and what are the potential downstream adverse effects?</b>			
No studies that directly assessed the harms of screening or monitoring. (Review articles, systematic reviews cited in Background section for information on radiation safety and harms associated with treatment.)			
			0

Systematic Reviews	RCTs	Nonrandomized Controlled or Quasirandomized Trials; Observational Studies With Controls; Eligible Longitudinal Studies	Other
<b>KQ #5 – What are the costs and cost-effectiveness of osteoporosis screening and monitoring?</b>			
0	0	0	Nayak 2011, Nshimyumukiza 2013 2

\*The question was addressed in the AHRQ evidence review (2010 Nelson review) that served as the basis of the latest USPSTF recommendation, but the review searched only for RCTs and identified none. The review was not able to answer the question.

†The question of optimal interval for treatment monitoring was addressed in a different AHRQ evidence review of treatments for osteoporosis (Crandall et al., 2012), but no trials comparing monitoring protocols were identified. The other related evidence cited by the treatment review did not meet inclusion criteria for the present report but is summarized as *Other Potentially Policy-Relevant Information* in the **TECHNICAL REPORT**.

## Practice Guidelines

Fourteen relevant practice guidelines published in the last 10 years were identified.

## Findings

Summary of Findings tables follow each Key Question. See **EVIDENCE SUMMARY, Methods, Quality Assessment** and the corresponding section in the **TECHNICAL REPORT**, as well as [Appendix II](#), for details regarding the assessment of bodies of evidence. See [Appendix III](#) for full evidence tables.

**Key Question #1: Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 1a: For individual patients, and do these outcomes vary according to age, sex, or other risk factors for BMD or fracture? 1b: In populations, and do these outcomes vary by population characteristics?**

See [Table 3](#) for a summary of findings.

### For Individual Patients (Key Question #1a)

Two RCTs, 2 nonrandomized or quasirandomized trials, and 2 cohort studies met inclusion criteria. These studies addressed different outcomes and drew from different populations. The non-RCTs had generally good control for confounders, at least in calculations of relative risk. Sample sizes ranged from 196 to 5736. Studies followed patients for 1 to 9 years. Patients with known osteoporosis or a history of fragility fracture and patients currently using osteoporosis medications or hormone replacement therapy were excluded.

#### *Fracture Incidence (4 studies)*

Two fair-quality studies (total n=7907), an RCT and a nonrandomized trial, reported results suggesting a possible effect of screening on reduction in fracture risk in healthy populations of middle-aged women (9 years of follow-up) and older men and women (5 years of follow-up). However, hazard ratios (HRs) were nonsignificant or had confidence intervals (CIs) that approached the null value. The non-RCT recruited older men and women and was well controlled, with the addition of a propensity score to clinical covariates, but unmeasured confounders remain possible. The RCT recruited middle-aged women. The evidence suggesting a reduction of fracture risk through screening in middle-aged to older adults was of low quality due to study quality as well as imprecise estimates. In these studies, participants were recruited from community advertising; results may not be generalizable to clinical populations.

Two fair-quality studies (total n=7168) of patients being treated with certain medications suggested that screening in these clinical populations is effective. Both studies were retrospective analyses of databases and both included men only. In 1 study, men were taking androgen-deprivation therapy (ADT) for prostate cancer, and in the other study, men were taking corticosteroids for ulcerative colitis. Control for known confounders was good; however, these 2 studies may have been subject to unknown

confounders. The evidence suggesting a reduction of fracture risk through screening in individuals taking medications known to be associated with osteoporosis is considered to be of *low quality* because of study quality, unknown applicability to the full spectrum of medications that are thought to be associated with osteoporosis risk, the availability of only 1 study each for the 2 medications addressed, and the exclusion of women as study participants. The evidence might also be considered to be of very low to low quality with respect to ADT for prostate cancer, of very low to low quality with respect to corticosteroids for ulcerative colitis, and insufficient for all other situations in which osteoporosis-inducing medications are administered. It was assumed that fracture risk, as a biological phenomenon, would vary according to medication and possibly, the underlying condition.

#### *Clinical Management Decisions (3 studies)*

One community-based RCT conducted in Scotland evaluated the impact of screening on clinical decisions in general populations. Depending on screening results and current use of corticosteroids, lifestyle changes were advised and prophylactic treatment advice was sent to participants' general practitioners. The study was conducted in Scotland. After 9 years, use of hormone replacement therapy (HRT) and vitamin D and calcium supplementation were greater in the screened group. The 2 fair-quality studies (total n=7168) of men being treated with osteoporosis-inducing medications suggested that screening was associated with much higher frequency of prescriptions for medications to treat osteoporosis. Among men taking ADT for prostate cancer, clinical records showed that during the follow-up period, 29% of screened men had received a prescription for osteoporosis medication, while only 3% ( $P<0.0001$ ) of unscreened men had a prescription. The study of men with ulcerative colitis found that, according to clinical records, men in the screening group were substantially more likely to receive prescriptions for osteoporosis medication and for vitamin D and calcium. Evidence regarding a positive impact on clinical management decisions was considered to be of *low quality* because of study quality and unknown group differences in treatment appropriateness. It was assumed that although clinician behavior might not vary according to underlying disease or osteoporosis-inducing medication, patients with different diseases and healthy populations might vary in their acceptance of medication recommendations that resulted from osteoporosis screening. Furthermore, impact on clinical management decisions should be considered an *intermediate* outcome.

#### *Osteoporosis-Preventing Behavior (2 studies)*

Two other poor- or good-quality studies (total n=399) recruited participants through community advertising and evaluated the impact of screening on change in patient behavior. One study was an RCT (good quality), and the other was a quasi-RCT that was considered to be of poor quality because of deficiencies in study conduct. Results suggested that DXA scanning has minimal or no effect on calcium intake, exercise, alcohol use, or smoking in postmenopausal women and/or older men over the short term (1 year). Given the small sample sizes and the possibility that even the small observed effects could diminish over the long term, evidence of minimal positive effect of screening on osteoporosis-preventing behavior was of *low quality*. Results may not be generalizable to a clinical population, and impact on patient behavior should be considered an *intermediate* outcome.

*Differential Effectiveness According to Risk Factors (3 studies)*

Subgroup analysis in 1 of the 2 studies of general screening in older adults and indirect age comparison across both the studies suggest that the effectiveness of screening for prevention of fractures increases with advanced age; however, the quantity of evidence does not allow a precise age cutoff or a sex-specific cutoff. Subgroup analysis in 1 of the general screening studies suggested possibly greater effectiveness in older women than in older men, but the estimated HR for each subgroup was nonsignificant. Therefore, evidence suggesting greater effectiveness with female sex and advanced age was very low quality due to study quality and quantity, imprecision, and/or indirect evidence.

In the study of screening in men with ulcerative colitis, results according to the extent of corticosteroid exposure during the study period suggested that screening had no effect at low levels, a nonsignificant effect at moderate levels, and a substantial as well as significant effect at high levels. However, a single observational study does not permit conclusions regarding differential effectiveness according to corticosteroid exposure. No other data regarding the differential effectiveness of screening were available. Evidence pertaining to differential effectiveness according to corticosteroid exposure was therefore of low quality due to the availability of only a single, fair-quality study.

**In Populations (Key Question #2b)**

Evidence pertaining to effectiveness in terms of population-level outcomes was insufficient due to a lack of studies.

**Table 3. Summary of Findings, Key Questions #1a and #1b**

Key: ADT, androgen-deprivation therapy (for prostate cancer); CI, confidence interval; CS, corticosteroids; DXA, dual x-ray absorptiometry; f/u, follow-up; HR, hazard ratio; HRT, hormone replacement therapy; IRR, incidence rate ratio; ITT, intention-to-treat; IU, international units; NA, not applicable; NS, not statistically significant; OP, osteoporosis; RCT, randomized controlled trial

Number, Size, and Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results (statistically significant results bolded)
<b>#1a. Effectiveness of Screening for Prevention of Fracture in Middle-Aged or Older Adults</b>				
2 fair-quality studies (n=7907)  Kern 2005 (nonrandomized trial), Barr 2010 (RCT)	<b>Quantity and precision:</b> Wide CIs and/or marginally significant results <b>Consistency:</b> ✓ <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> Unknown  Uncertain generalizability to clinical populations	<b>Low:</b> Study quality, imprecision	Favored screening	<b>Kern 2005 (men and women, age ≥65 yrs; f/u 5 yrs):</b> Overall adjusted HR: <b>0.64 (CI, 0.4-0.99)</b> Unadjusted cumulative incidence: 4.8% vs 8.2% <b>Barr 2010 (women; age 45-54 yrs; f/u 9 yrs):</b> Overall adjusted HR: ITT: 0.791 (CI, 0.600-1.042) Treatment-received: 0.759 (NS) Per-protocol HR: <b>0.734 (CI, 0.546-0.988; P=0.041)</b> Unadjusted cumulative incidence: 8.9% vs 9.4%  All results represent a comparison of screened with unscreened individuals.
<b>#1a. Effectiveness of Screening for Prevention of Fracture in Individuals Taking Medications Known to Be Associated with OP</b>				
2 fair-quality screening studies (n=7168)  Zhumkhawala 2013, Khan 2014 (both, retrospective cohort)	<b>Quantity and precision:</b> (✓). Only 1 study per medication. <b>Consistency:</b> ✓ <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> NA	<b>Low:</b> Unknown applicability to full spectrum of medications, only 1 study for each of 2 medications, exclusion of women	Screening reduces fracture risk and may depend on medication use intensity	<b>Zhumkhawala 2013 (1432 men undergoing ADT; f/u 2-3 yrs):</b> Adjusted HR (control vs screened): <b>4.19 (CI, 1.92-9.13)</b> Unadjusted cumulative incidence rates (screened vs control): 5.1% vs 18.1%  <b>Khan 2014 (5736 men with ulcerative colitis with varying intensities of CS; f/u 3 yrs):</b> Adjusted HR (screening vs no screening) for fragility fracture: <b>0.5 (CI, 0.3-0.9; P=0.03)</b> Unadjusted cumulative incidence rates (screened vs no screening): 1.6% vs 2.8%



Number, Size, and Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results (statistically significant results bolded)
<b>KQ #1a. Impact of Screening on Clinical Management Decisions in Any Population (intermediate outcome)</b>				
3 fair-quality screening studies (n=7168)  Barr 2010 (RCT), Zhumkhawala 2013 (retrospective cohort), Khan 2014 (retrospective cohort)	<b>Quantity and precision:</b> (✓). <b>Consistency:</b> ✓ <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> NA	<b>Low:</b> Study quality, unknown group differences in treatment appropriateness	Screening associated with greater OP medication use	<b>Zhumkhawala 2013 (1432 men undergoing ADT; f/u 2-3 yrs):</b> Screened men more likely to be taking OP drugs during f/u period (29% screened vs 3% unscreened; $P<0.0001$ ) <b>Khan 2014 (5736 men; ulcerative colitis with varying intensities of CS; f/u 3 yrs):</b> <i>Medication use (DXA screen, no screen) (% pts) :</i> Some type of OP medication, excluding HRT: 36.6%, 21.6% ( $P<0.001$ ) Vitamin D and calcium: 32.9%, 13.4% ( $P<0.001$ ) <b>Barr 2010 (2604 women, ages 45-54 yrs; f/u 9 yrs):</b> <i>Medication or supplement use (screened, control) (% pts):</i> HRT: 52.4%, 44.5% ( $P<0.01$ ) VitD: 24.2%, 12.5% ( $P<0.01$ ) Calcium: 20.0%, 14.1% ( $P<0.01$ )
<b>KQ #1a. Impact of Screening in Older Adults for Encouraging OP-Preventing Behavior (intermediate outcome)</b>				
2 poor- or good-quality studies (n=399)  Sedlak 2007 (RCT, good), Doheny 2011 (quasi-RCT, poor)	<b>Quantity and precision:</b> Few data, NS findings <b>Consistency:</b> ✓ <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> NA  Uncertain generalizability to a clinical population	<b>Low:</b> Small quantity and size of studies, study quality, short f/u (1 yr)	Minimal or no effect	<b>Sedlak 2007 (203 postmenopausal women; f/u 1 yr):</b> <i>Total calcium intake over 1 yr (screened, wait-list) (units unclear; assumed to be IUs per day):</i> 786, 668 ( <b>global <math>P&lt;0.001</math></b> ) <i>Exercise:</i> No change over time in either group <b>Doheny 2011 (men age <math>\geq 50</math> yrs):</b> <i>Mean # mins of vigorous activity:</i> 22, 19 (NS) <i>Mean # mins walking:</i> 15.3, 13 (NS) <i>Calcium:</i> No group or knowledge effect
<b>KQ #1a. Differential Effectiveness According to Age or Sex</b>				
2 fair-quality studies (n=7907)  Kern 2005 (nonrandomized trial),	<b>Quantity and precision:</b> Wide CIs and/or NS results; only 1 study per sex-age subgroup	<b>Very low:</b> Study quality, imprecision, indirect comparisons	Effectiveness may be greater with advanced age	<b>Kern 2005 (men and women, age <math>\geq 65</math> yrs; f/u 5 yrs) (HR):</b> <i>Women:</i> 0.61 (CI, 0.35-1.06); <i>Men:</i> 0.68 (CI, 0.32-1.42) <i>Age 65-74 yrs:</i> 0.73 (CI, 0.29-1.87) <i>Age 75-84 yrs:</i> 0.82 (CI, 0.47-1.44) <i>Age <math>\geq 85</math> yrs:</i> <b>0.22 (CI, 0.06-0.79)</b>

Number, Size, and Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results (statistically significant results bolded)
Barr 2010 (RCT)	<b>Consistency:</b> ✓ <b>Applicability to PICO:</b> ✓ <b>Applicability to KQ:</b> Indirect comparisons <b>Publication Bias:</b> Unknown  Uncertain generalizability to clinical populations			<b>Barr 2010 (women, age 45-54; 9 yrs):</b> <i>ITT: 0.791 (CI, 0.600-1.042)</i>  All results represent an adjusted comparison of screened with unscreened individuals.
<b>KQ #1a. Differential Effectiveness According to CS Exposure</b>				
1 fair-quality screening study (n=5736)  Khan 2014 (retrospective cohort)	<b>Quantity and precision:</b> (✓): Only 1 study <b>Consistency:</b> Unknown <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> NA	<b>Low:</b> Quality and quantity of studies	Screening effectiveness may depend on medication use intensity	<b>Khan 2014 (5736 men with ulcerative colitis with varying intensities of CS; f/u 3 yrs):</b> <i>Interaction between DXA screen and CS exposure (IRR, screen vs no screen):</i> <u>Low corticosteroid exposure:</u> No difference <u>Moderate:</u> 0.44 (NS) <u>High:</u> 0.38 ( <b>P=0.02</b> )
<b>KQs #1b. Population-Wide Effectiveness: Insufficient (no evidence)</b>				

**Key Question #2: Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 2a: For individual patients, and do these outcomes vary according to age, sex, other risk factors (including previous BMD measurements), treatment status, or testing interval? 2b: In populations, and do these outcomes vary by population characteristics or testing interval?**

Previous systematic reviews did not address these questions, and the literature search conducted for the present report did not identify any studies designed to answer questions about the effectiveness of repeat screening in untreated individuals or of serial testing in patients who are undergoing treatment for osteoporosis. The evidence for the effectiveness of monitoring (serial screening/testing) for osteoporosis and low bone density was *insufficient* because of the lack of eligible studies.

**Key Question #2c: What is the minimum interval required to detect transition from normal or low BMD to osteoporosis or to assess treatment effect?**

Two eligible studies were identified for this subquestion, both having to do with screening in older adults. See [Table 4](#) for a summary of findings. At the end of this section, the results of other types of analysis are briefly described to provide additional information that may be policy relevant.

#### *Repeat Screening in Older Adults*

Two large and well-controlled prospective longitudinal cohort studies provided consistent evidence that for adults older than age 60 without osteoporosis at the last screening and without risk factors other than age, repeat screening generally does *not* improve the estimation of fracture risk, or by implication identify the need to start treatment, for several years after initial screening. Exceptions are individuals who are very elderly and have at least moderate osteopenia at the time of the previous screen.

A study that followed 1758 Australian men and women who were age  $\geq 60$  years and without osteoporosis (defined as T-score  $\leq -2.5$ ) for a median of 7 years used data collected during the study to construct a clinical fracture prediction model. The model included age, sex, initial BMD, and the competing risk of death; it did not take into account secondary causes of osteoporosis or clinical history. Model predictions suggested that repeat screening every 2 years or less was not useful to identify individuals transitioning to a 20%, 10-year risk of major fracture or osteoporosis. The following results are illustrative of the findings of the Australian study and assume that the utility of repeat screening is defined by the individual reaching the U.S. treatment threshold of a 20%, 10-year risk of fracture:

- Repeat screening at  $< 2$  years would have utility for no individuals.
- Repeat screening at  $< 3$  years would have utility only in elderly adults with substantial osteopenia:
  - Men who were 80 years of age with T-score  $\leq -2.2$  at the time of the last screening. (Younger men or men with higher T-scores would not reach the treatment threshold for an average of  $\geq 3$  years.)

- Women who were 75 years of age with T-score  $\leq -2.0$  at the time of the last screening. (Younger women or women with higher T-scores would not reach the treatment threshold for an average of  $\geq 3$  years.)
- Women who were 80 years of age with T-score  $\leq -1.5$  at the time of the last screening. (Younger women or women with higher T-scores would not reach the treatment threshold for an average of  $\geq 3$  years.)
- For men age 70 at the time of initial screening (70 is the typical initial screening age in the U.S.) with normal BMD, repeat screening would not be necessary for another 9 years.
- For women age 65 at the time of initial screening (65 is the age recommended by the USPSTF for women without other risk factors) and normal BMD, repeat screening would not be necessary for another 12 years.

The Australian study is somewhat limited by substantial discrepancy between fracture risk prediction in the oldest participants and fracture risk prediction for the same age group according to the U.S. version of the FRAX tool. The Australian model would suggest that an 80-year-old woman with a T-score of  $-2.2$  should be screened again in about 2.5 years, whereas using the FRAX tool to estimate risk would suggest screening would not be necessary for another 10 years. However, concordance was found to be good for all but the oldest age increments.

Another eligible study was conducted in a U.S. population of women who were age  $\geq 67$  years and without osteoporosis (defined as T-score  $\leq -2.5$ , or with previous hip or clinical vertebral fracture). The optimal screening interval was defined a priori as the time it took for 10% of women to transition to osteoporosis from different baseline T-score cutoffs. Osteoporosis was defined as the occurrence of a hip or vertebral fracture or attainment of a BMD score that was  $\leq -2.5$ . Observations over a *mean* follow-up of 8 years suggested that repeat screening to identify women who had transitioned to osteoporosis would not be necessary for approximately 17 years if the T-score at baseline was  $\geq -1.5$  (characterized by the authors as normal BMD to mild osteopenia). The data suggested a 5-year interval if the T-score at baseline was  $< -1.50$  to  $-1.99$  (characterized as moderate osteopenia), and a 1-year interval if the T-score at baseline was  $-2.00$  to  $-2.49$  (characterized as advanced osteopenia). It is important to bear in mind that these time estimates were adjusted for most of the covariates in the FRAX model. Thus, the estimates are applicable when risks other than age, sex, and baseline osteoporosis status are not present. However, unlike the Australian study, the U.S. study did not evaluate whether testing intervals should shorten with advancing age beyond 67 years.

The overall evidence for older adults is considered to be of *moderate quality*. The moderate rating reflects good-quality studies with large sample sizes and general consistency of findings but lack of corroboration for either model. It also reflects the probable lack of precision in the estimates for the individuals for whom the estimated repeat screening intervals were very long (15+ years in the Australian study, 17 years in the U.S. study) since only a small proportion of participants in each study were actually followed this long. In other words, the estimated repeat screening intervals following a normal or near-normal DXA scan were imprecise. Other limitations of this evidence relate to generalizability: screening intervals for men in the U.S. have not been studied, the Australian model may

not apply to the very elderly in the U.S., and optimal intervals have not been investigated in clinical (nonvolunteer) populations.

#### *Adults Younger Than 60 Years of Age and Perimenopausal Women*

Evidence of optimal screening intervals in adults younger than age 60 and in perimenopausal women is *insufficient* due to a lack of studies that met inclusion criteria.

#### *Treatment Monitoring*

Evidence regarding optimal screening intervals for patients being treated for osteoporosis is *insufficient* due to a lack of eligible studies. An AHRQ evidence review of treatments to prevent osteoporotic fracture in men and women found no RCTs comparing different schedules of serial BMD monitoring. No controlled, comparative, or longitudinal studies of testing intervals in patients being treated for osteoporosis were identified by the searches conducted for the present report.

#### *Repeat Testing Based on Factors Other Than Age or Treatment Status*

Evidence regarding optimal screening intervals for individuals with risk factors other than age or treatment status is *insufficient* due to a lack of studies.

#### *Other Potentially Policy-Relevant Information*

Numerous studies that did not meet inclusion criteria for this report because they were not designed to estimate screening intervals nevertheless provide substantial evidence that change in BMD over time is not a strong predictor of change in fracture risk unless an individual was at high risk at the previous screen.

A study cited in the 2010 Nelson review for the USPSTF recommendation determined that the accuracy for prediction of fracture was nearly the same, whether it was based on the initial BMD, a second BMD measurement 8 years later, the change in BMD between the 2 measurements, or initial BMD plus change. Other studies have shown that change in BMD accounts for < 50% of the reduction in fracture risk in individuals being treated for osteoporosis. A small quantity of studies of BMD change in patients who have discontinued osteoporosis treatment or who have risk factors other than age or treatment status were also identified. These studies are all discussed in more detail in the **Literature Review** of the **TECHNICAL REPORT**.

**Table 4. Summary of Findings, Key Question #2c**

Key: CI, confidence interval; f/u, follow-up; NA, not applicable; OP, osteoporosis

Number, Size, and Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results
<b>Repeat Screening in Older Adults</b>				
<p>2 good-quality longitudinal cohort studies (n=5707)</p> <p>Frost 2009, Gourlay 2012</p> <p>No OP in either study group at baseline</p>	<p><b>Quantity and precision:</b> Some concern about precision</p> <p><b>Consistency:</b> ✓</p> <p><b>Applicability to PICO:</b> ✓</p> <p><b>Publication Bias:</b> NA</p> <p>Uncertain generalizability to clinical populations, to screening starting at age 50 for men, or to very elderly U.S. populations</p>	<p><b>Moderate:</b> Heterogeneity in models, lack of corroboration or validation of the specific prediction models in other study groups, and imprecision for individuals with normal or near-normal BMD</p>	<p>Transition to OP or occurrence of fracture did not occur for at least several yrs after last screening with normal findings</p>	<p><b>*Frost 2009 (750 men and 1003 women, age ≥60 yrs and no OP; Australia):</b> <i>Shortest/longest time in yrs to reach 20%, 10-yr risk of OP and/or clinical fracture:</i></p> <p><u>Men:</u> Longest: Screened at age 60, T-score 0: 15.0+ (90% CI, <b>14.3</b>-15.0+) (f/u did not go beyond 15 yrs) Shortest: Screened at age 80, T-score -2.2: 2.9 (90% CI, <b>2.6</b>-3.8)</p> <p><u>Women:</u> Longest: Screened at age 60, T-score 0: 14.1 (90% CI, <b>12.7</b>-15.0+) (f/u did not go beyond 15 yrs) Shortest: Screened at age 80, T-score -2.2: 2.4 (90% CI, <b>2.2</b>-2.6)</p> <p><i>Time in yrs to reach 20%, 10-yr risk of OP and/or fracture if initial screen was at typical U.S. screening age:</i></p> <p><u>Men:</u> Age 70, T-score 0: 10.7 (90% CI, <b>9.0</b>-12.2) Age 70, T-score -1.0: 8.9 (90% CI, <b>7.8</b>-9.8) Age 70, T-score -1.5: 8.1 (90% CI, <b>7.2</b>-9.0) Age 70, T-score -2.0: 7 (9.40% CI, <b>6.5</b>-8.7) Age 70, T-score -2.2: 7.3 (90% CI, <b>7.1</b>-8.4)</p> <p><u>Women:</u> Age 65, T-score 0: 12.3 (90% CI, <b>10.6</b>-13.4) Age 65, T-score -1.0: 8.3 (90% CI, <b>7.2</b>-9.8) Age 65, T-score -1.5: 7/5 (90% CI, <b>5.5</b>-7.3) Age 65, T-score -2.0: 4.9 (90% CI, <b>4.4</b>-5.9) Age 65, T-score -2.2: 4.6 (90% CI, <b>3.8</b>-5.4)</p> <p><i>Age and T-score at initial screen in which time to reach 20%,</i></p>

Number, Size, and Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results
				<p>10-year risk was:</p> <p><u>Men:</u></p> <p>&lt;2 yrs: No individuals</p> <p>&lt;3 yrs: Age 80, T-score <math>\leq -2.2</math></p> <p>&lt;5 yrs: Age 75, T-score <math>\leq -2.2</math>; age 80, any T-score</p> <p><u>Women:</u></p> <p>&lt;2 yrs: No individuals</p> <p>&lt;3 yrs: Age 75, T-score <math>\leq -2.0</math>; age 80, T-score <math>\leq -1.5</math></p> <p>&lt;5 yrs: Age 75, T-score <math>\leq -1.5</math>; age 80, any T-score</p> <p><b>Gourlay 2012 (4957 women, age <math>\geq 67</math> yrs and no OP; U.S.):</b>  Adjusted time in yrs for 10% of women to make the transition to OP before fracture or treatment, by OP status at last screen:</p> <p><u>Normal BMD (T-score <math>\geq -1.00</math>):</u> 16.8 (CI, 11.5-24.6)</p> <p><u>Mild osteopenia (T-score <math>-1.01</math> to <math>-1.49</math>):</u> 17.3 (CI, 13.9-21.5)</p> <p><u>Moderate osteopenia (T-score <math>-1.50</math> to <math>-1.99</math>):</u> 4.7 (CI, 4.2-5.2)</p> <p><u>Advanced osteopenia (T-score <math>-2.00</math> to <math>-2.49</math>):</u> 1.1 (CI, 1.0-1.3)</p>
<b>Repeat Screening in Perimenopausal Women:</b> <i>Insufficient (lack of studies)</i>				
<b>Repeat Screening in Adults Younger Than 60:</b> <i>Insufficient (lack of studies)</i>				
<b>Treatment Monitoring:</b> <i>Insufficient (lack of studies)</i>				
<b>Repeat Screening Based on Factors Other Than Age or Treatment Status:</b> <i>Insufficient (lack of studies)</i>				

\*Values for time to fracture in the study by Frost et al. (2009) were calculated for 5-year increments of age and initial BMD values of 0,  $-1.0$ ,  $-1.5$ ,  $-2.0$ , and  $-2.2$ . The authors considered the **lower bound of the 90% CI** rather than the point estimate to be the better benchmark, to take into account sampling variation. See Table 4 in published study report for full set of estimates.

**Key Question #3: What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?**

For Key Question #3, the following analyses were considered or performed:

- A special analysis in the 2010 Nelson review (older women)
- Conversion of event rates from a screening study in older men and women
- Conversion of event rates from 2 studies of individuals taking osteoporosis-inducing medications

See **Table 5** for a summary of findings.

*NNS to Prevent 1 Fracture in Older Adults*

NNS values for postmenopausal women were reported in a good-quality systematic review, the 2010 Nelson review, but these calculations provide indirect evidence since they are not based on empirical evidence from trials designed to assess the effect of screening on fracture risk. Using data from a population study and a pivotal trial of the bisphosphonate alendronate (the Fracture Intervention Trial [FIT]), the 2010 Nelson review estimated that, to prevent 1 osteoporotic fracture over a period of 5 years, the following numbers of postmenopausal women would have to be screened: age 55 to 59 years (n=278), age 60 to 64 years (n=187), age 65 to 69 years (n=103), age 70 to 74 years (n=61), age 75 to 79 years (n=43). The corresponding NNS estimates for prevention of hip fracture were 1667, 1000, 556, 323, and 238. The prevalence assumptions of these calculations may not apply to clinical settings. Furthermore, the treatment assumptions may not apply to real-world settings. Treatment assumptions were derived from an RCT in which adherence was > 80%, while a 2012 AHRQ review of treatment to prevent osteoporotic fractures found that in observational studies in the U.S., only about half of patients demonstrated treatment adherence at 1 year following initiation. The Nelson analysis also assumes that all individuals diagnosed with osteoporosis will be offered and will accept treatment. Thus, the analysis very likely underestimates the true NNS, which the authors acknowledge in their technical report (the full version available from the AHRQ website).

More direct assessment of NNS was possible with data from a fair-quality nonrandomized trial selected for evidence pertaining to Key Question #1a. Calculations based on results from this screening study suggested that only 7 women age 85 or older would need to be screened to prevent 1 fracture over 5 years and that 46 women or 96 men age 65 or older would have to be screened to prevent 1 hip fracture over a 5-year period. These values for NNS are derived from the study's cumulative fracture incidence rates for screened and unscreened groups after a mean follow-up of 5 years. Comparison of results between the Nelson analysis and the screening study is hampered since the screening study did not provide results by 5-year age group. Furthermore, the 2 sets of data seem very inconsistent. The NNS for hip fracture in women age 65 years or older in the screening study was much smaller than the various hip fracture estimates for women in this age bracket in the Nelson analysis, even though the screening study represents a more direct assessment of actual practice and presumably, a more realistic estimate. The screening study authors used a propensity score to correct for the possibility that individuals with better outcomes were more likely to live in the geographic locations in which screening was performed.



Nevertheless, the unsystematic nature of treatment assignment in this study may have created an unmeasured bias. In the screening study, a statistically significant effect of screening was detected only for individuals age 85 or older, an age group not represented in the Nelson NNS analysis. The inconsistency might derive from different treatment thresholds. The Nelson analysis assumed that the threshold for treatment was T-score < -2.50, whereas screened participants in the screening study may have been offered treatment on the basis of osteopenia (T-score < -2.00) plus risk factors. For reasons described in the **TECHNICAL REPORT**, it was difficult to determine with certainty what the data used in the Nelson analysis represent. Since the cumulative incidence rates, in contrast to the screening study's measures of relative risk, do not reflect any adjustment for risk factors, the corresponding values for NNS are subject to bias.

Taking into account the unexplained inconsistency between the analysis in the Nelson review and the results based on the screening study and the indirect (Nelson review) or possibly confounded (screening study) nature of the data, the evidence concerning the NNS to prevent 1 fracture in older women is of low quality. The evidence concerning NNS to prevent 1 fracture in older men is of very low quality because of the availability of data from only 1 study and the potential confounding of the incidence rates in that study.

#### *NNS to Prevent 1 Fracture in Younger Adults*

The evidence concerning NNS to prevent 1 fracture in younger adults is insufficient due to lack of published analyses or screening studies with useful data.

#### *NNS to Prevent 1 Fracture in Individuals Using Osteoporosis-Inducing Medications*

Two screening studies, which followed a retrospective cohort study design, suggested that screening prevents fractures in men being treated with ADT for prostate cancer or in men with ulcerative colitis, especially with greater use of corticosteroids to treat the ulcerative colitis. However, the utility of screening according to NNS values is unclear. Translation of the first study's findings into NNS values suggested that 26 men being treated for prostate cancer would have to be screened to prevent 1 hip fracture over 3 years. The study did not provide data that would allow calculation of NNS values for different subgroups defined by degree of previous ADT exposure. Translation of the second study's findings suggested that 278 men with ulcerative colitis and taking glucocorticoids would have to be screened to prevent 1 hip fracture over 3 years. Study results suggested a dose-response relationship between glucocorticoid use and the benefits of screening, but dose-specific NNS values could not be computed with the data provided.

Evidence concerning NNS to prevent fracture in screening in individuals taking medications known to be associated with osteoporosis is considered to be of very low quality because of limited applicability to the full spectrum of medications that are thought to be associated with osteoporosis risk, the availability of only 1 study each for the 2 medications addressed, lack of data for computing NNS by dose, lack of long-term data, and lack of data for women. Furthermore, since the fracture incidence rates were not adjusted for risk factors, the NNS values are subject to possible confounding.

*NNS to Prevent 1 Fracture in Individuals with Other Risk Factors*

The evidence with regard to NNS in groups defined by any factor other than age, sex, or osteoporosis-inducing medication treatment was *insufficient*.

**Table 5. Summary of Findings, Key Question #3**

Key: ADT, androgen-deprivation therapy (for prostate cancer); CS, corticosteroids; NA, not applicable; NNS, number needed to screen; OP, osteoporosis; SR, systematic review

Number, Size, Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results (statistically significant results bolded)
<b>Older Adults</b>				
1 fair-quality nonrandomized screening trial (Kern 2005)  1 good-quality SR with calculations based on data from multiple sources, older women (Nelson 2010a/2010b)	<b>Quantity and precision:</b> Serious concern regarding precision of underlying effectiveness estimates (Kern 2005) <b>Consistency:</b> Inconsistent <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> NA  Uncertain generalizability to clinical populations (Nelson 2010) and to starting screening at typical age for men (50 yrs) (Kern 2006)	<b>Women: Low</b> for study quantity, study quality and imprecision (Kern 2006), indirect evidence (Nelson 2010a/2010b), inconsistency  <b>Men: Very low</b> due to potentially confounded data from a single study	NNS diminishes with age; considerably greater for prevention of hip fracture than for prevention of any OP fracture	<b>Kern 2005 (older men and women, age ≥65 yrs):</b> <i>NNS to prevent 1 hip fracture over 5 yrs:</i> Overall: 59 Men: 96 Women: 46 Adults age ≥85 yrs: 7  <b>Nelson 2010 (women, age ≥55 yrs):</b> <i>NNS to prevent 1 fracture over 5 yrs in women (any fracture, hip fracture):</i> Age 55-59 yrs: 278, 1667 Age 60-64 yrs: 187, 1000 Age 65-69 yrs: 103, 556 Age 70-74 yrs: 61, 323 Age 75-79 yrs: 43, 238
<b>Younger Adults: Insufficient (no evidence)</b>				
<b>Individuals Taking Medications Known to Be Associated With OP</b>				
2 fair-quality screening studies  Zhumkhawala 2013, Khan 2014 (both retrospective cohort)	<b>Quantity and precision:</b> Some concerns <b>Consistency:</b> Uncertain <b>Applicability to PICO:</b> ✓ <b>Publication Bias:</b> NA	<b>Very low:</b> Unknown applicability to full spectrum of medications, only 1 study for each of 2 medications, NNS by dose not possible, possible	Lower NNS values for ADT-prostate cancer than for glucocorticoids-ulcerative colitis	<b>Zhumkhawala 2013 (1432 young to old men undergoing ADT; f/u 2-3 yrs, mean 3.2):</b> <i>NNS for prevention of 1 hip fracture over 3 yrs:</i> 26 <b>Khan 2014 (5736 young to old men with ulcerative colitis with varying intensities of CS; f/u mean 3 yrs):</b> <i>NNS for prevention of 1 hip fracture over 3 yrs:</i> 278

		confounding, no data for women		
<b>Any Group Defined by Factors Other Than Age, Sex, or OP-Inducing Medication: <i>Insufficient (no studies)</i></b>				

#### **Key Question #4: Are bone density tests safe and what are the potential downstream adverse effects?**

No studies designed to assess harms associated with DXA scanning or the consequences of DXA scanning were identified in systematic reviews or in the searches conducted for this report. See **Table 6** for a summary of findings, including a summary of the following supplemental information from sources that did not meet inclusion criteria.

DXA scanning is a reasonably safe technology, with radiation exposure much less than that of other common radiation-based diagnostic and screening technologies such as standard x-ray and mammography. However, the safety of repeated DXA scanning over a long time frame has not been studied. Patients may suffer harm in the form of inappropriate treatment if DXA scans produce false-positive results or from missed treatment opportunities if results are false-negatives, but the rate of false results is unknown. Harms may also occur if prescreening tools and tests or DXA scan results are not interpreted correctly. Given the wide number of risk factors associated with both osteoporosis and fracture risk, there remains some clinical uncertainty in selecting patients both for screening and for treatment. Thus, unnecessary screening and unnecessary treatment are possibilities, as is a missed opportunity to appropriately screen and treat. However, actual data regarding inappropriate use of DXA scanning were not identified in the literature.

#### Downstream Adverse Effects of Screening:

Serious adverse events have been reported in conjunction with osteoporosis medications, including the following:

- Serious gastrointestinal adverse events with use of bisphosphonates (pooled quantitative data were not available in the literature reviewed for this report; inconsistent evidence according to the 2010 Nelson review).
- Atrial fibrillation with use of bisphosphonates (pooled quantitative data were not available in the literature reviewed for this report; inconsistent evidence according to the 2010 Nelson review).
- Osteonecrosis and bisphosphonates, results of recent meta-analyses:
- Noncancer patients: Odds ratio (OR), 2.91 (95% CI, 1.62 to 5.22; high heterogeneity) (8 studies with adjustment for risk factors). Across all 12 studies with a total of 574,649 participants, 2642 individuals (0.46% of all participants) developed osteonecrosis. The odds were much higher with intravenous delivery (OR, 47.8) than with oral delivery (OR, 3.15).
- Cancer patients: OR, 4.22 (95% CI, 3.21 to 5.54; no heterogeneity) (4 studies with adjustment for risk factors). Across all 8 studies with a total of 571,009 participants, 1389 individuals (0.24% of

all participants) developed osteonecrosis. The odds were much higher with intravenous delivery (OR, 4.27) than with oral delivery (OR, 1.18).

- Thromboembolic events and raloxifene: Pooled RR of 1.60 (95% CI, 1.15 to 2.3) (2 trials) (meta-analysis in 2010 Nelson review).
- Subtrochanteric, femoral shaft, or atypical femur fracture and bisphosphonates: Adjusted RR, 1.7 (95% CI, 1.22 to 2.37) (11 observational studies). Four studies evaluated  $\geq 5$  years of bisphosphonate use; the RR based on these studies was 1.62.

**Table 6. Summary of Findings, Key Question #4**

Key: CI, confidence interval; GI, gastrointestinal; IV, intravenous; MA, meta-analysis; OP, osteoporosis; OR, odds ratio; RR, relative risk; SR, systematic review

Number, Size, Quality of Studies	Other Quality Considerations; Generalizability	Quality Rating	Direction of Findings	Study Results (statistically significant results bolded)
<b>Direct Evidence of the Safety of Screening: <i>Insufficient (no studies)</i></b>				
<b>Safety of DXA Technology</b>				
Review articles	---	---	Safe	Less radiation exposure per scan than with chest x-ray or mammogram. Some concern about repeated scanning over a lifetime.
<b>Harms Associated With False-Positive Results, False-Negative Results, or Misinterpretation</b>				
Commentary in SR. Inference from uncertainty in risk assessment. No actual data.	---	---	Some concern	<i>Potential for inappropriate treatment or missed opportunities for treatment if:</i> False-positive or false-negative scan results Incorrect risk assessment for OP prior to screening or for fracture after screening. (Actual data not available.)
<b>Serious Adverse Events Associated With OP Medication Use</b>				
1 good-quality SR/MA (Nelson 2010a/2010b)  Other SRs (unrated) cited in Background	---	---	Some concern	<i>GI events and atrial fibrillation:</i> Inconsistent evidence (Nelson 2010a/2010b) <i>Osteonecrosis, noncancer patients:</i> Adjusted <b>OR, 2.91 (95% CI, 1.62-5.22;</b> high heterogeneity) (8 studies); absolute risk, 0.46% with or without bisphosphonates. <i>Osteonecrosis, cancer patients:</i> Adjusted <b>OR, 4.22 (95% CI, 3.21-5.54;</b> no heterogeneity) (4 studies); absolute risk, 0.24% with or without bisphosphonates NOTE: Risk of osteonecrosis much higher with IV delivery. <i>Thromboembolic events and raloxifene:</i> <b>RR, 1.60 (95% CI, 1.15-2.3)</b> (2 trials) (Nelson 2010a/2010b) <i>Subtrochanteric, femoral shaft, or atypical femur fracture and bisphosphonates:</i> Adjusted <b>RR, 1.7 (95% CI, 1.22-2.37)</b> (11 observational studies). 4 studies evaluated $\geq 5$ years of bisphosphonate use: <b>RR, 1.62</b>

**Key Question #5: What are the costs and cost-effectiveness of osteoporosis screening and monitoring?**

See **Table 7** for a summary of findings.

*Cost*

The U.S. economic evaluation reviewed in the following discussion assumed the cost of a central DXA scan to be \$98 in 2010 (approximately \$144 if converted to 2014 dollars), based on median Medicare reimbursement. [However, public comment offered in response to the draft version of this report stated that Medicare reimbursement has been markedly reduced since 2010.] The Washington State Agency Utilization Data added to this report indicate that over the past 3 years, the State has paid the following average dollar amounts per DXA scan: \$104 for all beneficiaries of Public Employee Benefits (PEBB) and the Uniform Med Plan (UMP), \$124 per non-Medicare PEBB and UMP beneficiaries, and \$59 for Medicaid fee-for-service (FFS) beneficiaries. A search of the Internet suggests that out-of-pocket costs for patients without insurance are in the range of \$150 to \$250 for a standard set of DXA scans.

*Cost-Effectiveness of Screening in Women*

Two economic evaluations based on modeling suggested that screening in women older than age 65 is cost-effective, but came to conflicting conclusions regarding screening in women younger than age 65. Neither economic evaluation addressed screening in men.

One evaluation was from a U.S. perspective and appeared shortly after publication of the 2011 USPSTF recommendations for screening. The other one was designed for application to the Canadian population. Both were modeling studies, both addressed cost-effectiveness in women, and both assumed a lifetime horizon. It is important to keep in mind that these evaluations were based on numerous assumptions derived from evidence and information from various sources; they were not based on empirical evidence of the effectiveness of screening. The Canadian models differed from the U.S. models in that the Canadian models assumed: (1) use of a formal fracture risk assessment tool after DXA results were known; and (2) that women who were not found to be at sufficiently high risk of fracture to warrant treatment would be encouraged to participate in a national prevention program. The Canadian study assumed that patients judged to be at high risk of fracture according to the assessment tool would be treated, while the U.S. study assumed that patients with a T-score of  $\leq -2.5$  would be treated. The Canadian study also did not take drug-related events into account or assess the effectiveness of different screening intervals, and it assumed an initial screening with follow-up DXA at 2 or 5 years, depending on fracture risk. The U.S. study modeled adverse events and included various screening intervals in its models.

With similar assumptions about the threshold for cost-effectiveness (\$50,000 per quality-adjusted life-year [QALY]), both studies led to the conclusion that BMD screening by DXA scanning would be cost-effective in older women (age  $\geq 55$  years in the U.S. study; age  $\geq 65$  years in the Canadian study). The 2 studies differed in their findings with respect to screening in women younger than age 65. The U.S. study found screening to be cost-effective even in women as young as 55 (and did not evaluate

screening in women younger than 55), while the Canadian study found screening *not* to be cost-effective in women age 40 to 64. The Canadian authors presented a single set of findings for all women age 40 to 64 because results according to 5-year age ranges were very similar. The authors of the U.S. study pointed out that the predicted differences between their screening strategies—some of which involved prescreening for risk of osteoporosis before deciding to perform a DXA scan—were very small. They concluded that initiating screening at age 55 years and continuing every 5 years to age 80 would be effective and within typical cost-effectiveness limits, *regardless of screening strategy*, as long as the treatment threshold was T-score  $\leq -2.5$ . No cost-effectiveness studies in populations other than older women were identified.

#### *Cost-Effectiveness of Screening in Men*

There were no economic evaluations of screening in men that appeared to appropriately model usual practice. The evidence was therefore *insufficient*.

#### *Cost-Effectiveness of Serial Monitoring*

There were no economic evaluations of repeat screening or treatment monitoring. The evidence was therefore *insufficient*.

**Table 7. Summary of Findings, Key Question #5**

Key: AE, adverse event; BMD, bone mineral density; CAROC; Canadian Association of Radiologists and Osteoporosis Canada; DXA, dual x-ray absorptiometry; EE, economic evaluation; ICER, incremental cost-effectiveness ratio; ICUR, incremental cost-utility ratio; OP, osteoporosis; QALY, quality-adjusted life-year; WTP, willingness-to-pay

Number and Type of Studies	Limitations	Direction of Findings	Study Results* (statistically significant results bolded)
2 EEs  Nayak 2011 (U.S.) Nshimyumukiza 2013 (Canada) (both modeling studies)	<b>Nayak 2011:</b> Possible overestimation of effectiveness of bisphosphonate treatment.  Analysis assumes all women at a T-score threshold will receive treatment.  <b>Nshimyumukiza 2013:</b> No consideration of medication AEs.  Findings may not be generalizable to screening programs without a prevention program or to screening followed by treatment based on T-score rather than a multifactorial	Cost-effective at WTP threshold of \$50,000/QALY	<b>Nayak 2011 (women age <math>\geq 55</math> yrs; 7 screening strategies compared with usual care [treat only after OP fracture]):</b>  <i>Best strategy overall (most effective and still within WTP threshold):</i> Initiate at age 55; DXA screen; treat if T-score $\leq -2.5$ ; screen every 5 yrs; \$45,450/QALY (\$48,581 in 2014 USD).  <b>Nshimyumukiza 2013 (women age <math>\geq 40</math> yrs; 12 programs combining universal screening† with universal primary prevention programs, compared with no program [possible DXA scan and treatment after fracture]):</b>  <i>BMD/CAROC plus universal primary prevention with physical activity + Vitamin D + calcium:</i> Would avert the greatest number of fractures and add the most QALYs. ICER of \$60,205 (\$55,019 in 2014 USD) and ICUR of

Number and Type of Studies	Limitations	Direction of Findings	Study Results* (statistically significant results bolded)
	risk tool.		\$55,300 (\$50,537 in 2014 USD).
2 EEs  Nayak 2011 (U.S.) Nshimyumukiza 2013 (Canada) (both modeling studies)	---	Conflicting findings at threshold of \$50,000/QALY	<b>Nayak 2011:</b> Cost-effective for initiation as young as at age 55 <b>Nshimyumukiza 2013:</b> BMD/CAROC plus universal primary prevention with physical activity + Vitamin D + calcium: Would avert the greatest number of fractures but unacceptable ICER of \$346,776 and ICUR of \$239,573 (2007 CAD).

\*In both studies, ICERs and ICURs were computed by comparing each nondominated strategy with the next less expensive nondominated strategy. *Nondominated* refers to a finding that the strategy is not both more expensive and less effective than any other strategy.

†Initial screen with 1 follow-up DXA at 2 or 5 yrs, depending on risk.

## Practice Guidelines

The search of the core sources and relevant specialty group Web sites identified 14 guidelines with relevant recommendations and published within the past 10 years. The general recommendations provided by the guidelines are summarized in **Table 8**. Additional details, by guideline, are presented in [Appendix IV](#). See also **Practice Guidelines** in the **TECHNICAL REPORT** for additional background information on some guidelines and a description of guidelines that were reviewed but found not to contain relevant recommendations.

### *Populations Defined Primarily by Age, Sex, and Age-Related Hormonal Status*

Eleven (11) guidelines addressed the screening and monitoring of osteoporosis in generally healthy populations. These included guidelines from the American Association of Clinical Endocrinologists (AACE) (Watts et al., 2010), American College of Physicians (ACP) (Qaseem et al., 2008), American College of Obstetricians and Gynecologists (ACOG) (ACOG, 2012), American College of Preventive Medicine (Lim et al., 2009), American College of Radiology (ACR) (ACR, 2010), Endocrine Society (Watts et al., 2012), International Society for Clinical Densitometry (ISCD) (ISCD, 2013), Institute for Clinical Systems Improvement (ICSI) (Florence et al., 2013), National Osteoporosis Foundation (NOF, 2014), North American Menopause Society (NAMS, 2010), and the USPSTF (USPSTF, 2011a). Most guidelines were of fair to good quality. The guidelines were generally in good agreement and recommended BMD screening for all women  $\geq 65$  years of age and for those women  $< 65$  years of age who were at a high risk of fracture. Three of the guidelines had recommendations specific to women in menopausal transition, which is a period of accelerated bone loss. The ACR guideline (poor quality) recommends screening all women during menopausal transition. The ISCD guideline (not rated) and NOF guideline (poor) recommend screening during menopausal transition if there are additional risk factors; example risk factors provided by the NOF document are low body weight, prior fracture, and high-risk medication.



There was some variability in recommendations for men. Five guidelines recommended BMD screening in all men  $\geq 70$  years of age and 1 guideline recommended screening in men  $> 50$  years of age who had a fracture risk. In contrast, the USPSTF guidelines stated that the current evidence was insufficient to assess the benefit and harm of osteoporosis screening in men and provided the following considerations for physicians regarding screening in men (USPSTF, 2011a):

- BMD determination may potentially detect osteoporosis in a large number of men and prevent substantial burden of fractures and fracture-related illnesses in this group.
- The potential harms of osteoporosis screening are likely to be small.
- Routine osteoporosis screening is not common practice in men.
- The men most likely to benefit from screening would be those who have a 10-year fracture risk equal to or greater than that of a 65-year-old woman who has no additional risk factors.

Recommendations predicated on risk factors, i.e., recommendations for women younger than 65 or men younger than 70, refer to varying but overlapping lists of appropriate risk factors. The ICSI and the USPSTF recommendations for women younger than 65 directly or indirectly refer to risk as assessed by the FRAX tool. ICSI specifies that in postmenopausal women younger than 65 years, screening is recommended when 10-year fracture risk is  $\geq 9.3\%$  according to the FRAX tool, but the guideline allows for screening when there are “other” indications of increased risk. The current USPSTF policy recommends screening for women (postmenopausal or not) whose 10-year fracture risk is  $\geq$  that of a 65-year-old white woman. The USPSTF does not prescribe a risk assessment tool but used the FRAX tool to develop its recommendations. The ICSI and USPSTF recommendations for women younger than 65 are equivalent since the FRAX-derived 10-year fracture risk for a 65-year-old white woman with  $\leq 1$  risk factor is  $9.3\%$ .<sup>2</sup> The USPSTF statement acknowledges the lack of RCTs designed to directly measure the benefits and harms of screening, but the authors of the supporting systematic review (2010 Nelson review) stated that population screening can be justified if there is evidence that individual risk for fracture can be estimated and fractures can be significantly reduced for persons at risk. Thus, the 2010 Nelson review included key questions about the accuracy of tools for assessing osteoporosis risk and fracture risk and about the benefits and harms of osteoporosis medications. (See **SUMMARY OF BACKGROUND, Tools for Assessing Risk of Osteoporosis and Risk of Fracture and Treatment.**)

#### *Populations Defined by Medical Conditions and Treatment*

Additionally, 5 guidelines addressing screening and/or monitoring of BMD in patients with particular medical conditions were identified. Three guidelines were produced by the American College of

---

<sup>2</sup>See Figure 3 in [Screening for Osteoporosis: Systematic Review to Update the 2002 U.S. Preventive Services Task Force Recommendation \[Internet\]](#). Under the heading *Risk for Osteoporotic or Hip Fracture - >one risk factor*, gray shading is used to identify women *younger than age 65* whose risk profile would make them eligible for screening according to the USPSTF recommendation. For example, a 55-year-old woman with low body mass index and a parent who had a hip fracture has an 11% 10-year risk of any osteoporotic fracture and a 0.7% 10-year risk of a hip fracture.



Rheumatology (Grossman et al., 2010), the American College of Gastroenterology (Kornbluth et al., 2010), and the European Urologic Association (Dohle et al., 2012). Universal osteoporosis screening is recommended in these guidelines for *severe* late-onset male hypogonadism (European Urologic Association) and for any patient starting glucocorticoid therapy that is expected to last for at least 3 months (American College of Rheumatology). In addition to screening, the American College of Rheumatology also recommends serial testing in patients taking glucocorticoids for at least 3 months (the guideline recommends treatment with osteoporosis medication for most individuals who are taking glucocorticoids for  $\geq 3$  months). Screening is recommended for patients with inflammatory bowel disease if they have additional risk factors (American College of Gastroenterology). The ICSI guidelines also strongly recommend BMD screening for individuals using glucocorticoids (at a dose equivalent to  $> 5$  milligrams [mg] prednisone for  $\geq 3$  months) (Florence et al., 2013). The ACP guideline mentioned in discussion of its recommendations that single factors such as ADT may be sufficient justification for screening (Qaseem et al., 2008).

**Table 8. Summary of Practice Guideline Recommendations**

Key: AACE, American Association of Clinical Endocrinologists; ACG, American College of Gastroenterology; ACOG, American College of Obstetricians and Gynecologists; ACP, American College of Physicians; ACPM, American College of Preventive Medicine; ACR, American College of Radiology; ADT, androgen-deprivation therapy; BMD, bone mineral density; BMI, body mass index; DXA, dual x-ray absorptiometry; EUA, European Urological Association; GL, guideline; IBD, inflammatory bowel disease; ICSI, Institute for Clinical Systems Improvement; ISCD, International Society for Clinical Densitometry; NAMS, North American Menopause Society; NOF, National Osteoporosis Foundation; OP, osteoporosis; USPSTF, United States Preventive Services Task Force

Quantity of Individual GLs	Individual GL Quality*	Recommendations
<b>Screening in Postmenopausal Women &lt;65 Yrs of Age</b>		
<b>9</b> (AACE, ACOG, ACPM, ACR, ICSI, ISCD, NAMS, NOF*, USPSTF)	2 Good 4 Fair 1 Poor 2 Not rated	Postmenopausal women age <65 yrs should have BMD screening if they have risk factors for fracture. For example, ICSI and USPSTF recommend screening if 10-yr fracture risk exceeds 9.3% (risk for 65-year-old white woman with $\leq 1$ additional risk factor). <i>Exceptions:</i> ACR (poor) policy applies to women in menopausal transition and does not require risk factors other than menopause. ISCD (not rated) and NOF (poor) also advise screening during menopausal transition if risk factors are present.
<b>Screening in Women <math>\geq 65</math> Yrs of Age</b>		
<b>9</b> (AACE, ACOG, ACPM, ACR, ICSI, ISCD, NAMS, NOF*, USPSTF)	2 Good 4 Fair 2 Poor 1 Not rated	All women age $\geq 65$ yrs should have BMD screening. (ACR recommendation applies to all women age $\geq 50$ yrs.)

Quantity of Individual GLs	Individual GL Quality*	Recommendations
<b>Screening in Men &lt;70 Yrs of Age</b>		
<b>7</b> (ACP, ACPM, ACR, Endocrine Society, ISCD, ICSI, NOF*)	3 Good 1 Fair 2 Poor 1 Not rated	Men age 50-69 yrs should have BMD screening if they have risk factors for fracture. <i>(Presented as a consideration, not a recommendation, by ICSI; and as a weak recommendation by the Endocrine Society.)</i> Risk factors identified by ≥1 GL: Low BMI, weight loss, physical inactivity, corticosteroid use, ADT, fragility fracture.
<b>1</b> (USPSTF)	1 Good	Evidence is insufficient to support a recommendation.
<b>Screening in Men ≥70 Years of Age</b>		
<b>7</b> (ACP, ACPM, ACR, Endocrine Society, ISCD, ICSI, NOF*)	3 Good 1 Fair 2 Poor 1 Not rated	All men age ≥70 yrs should have BMD screening. <i>(Presented as a consideration, not a recommendation, by ICSI; and as a weak recommendation by the Endocrine Society.)</i>
<b>1</b> (USPSTF)	1 Good	Evidence is insufficient to support a recommendation.
<b>Follow-Up Testing After Initial Screen</b>		
<b>5</b> (AACE, ACR, NAMS, ISCD, USPSTF)	2 Good 2 Fair 1 Not rated	AACE, ACR, NAMS: Every 1-5 yrs, depending on risk factors and T-score in patients with risk factors or low BMD (osteopenia) at last scan. ISCD: To monitor BMD if evidence of bone loss would result in treatment. USPSTF: Lack of evidence regarding appropriate intervals.
<b>Treatment Monitoring</b>		
<b>7</b> (AACE, ACOG, ACR, Endocrine Society, ISCD, ICSI, NAMS, NOF)	1 Good 3 Fair 2 Poor 1 Not rated	Typical: BMD monitoring recommended every 2 yrs (3 GLs), every 1-2 yrs (3 GLs), or without specification of interval (1 position statement). Some GLs add that DXA can be discontinued or performed less frequently if BMD improves or stabilizes and there are no new risk factors. The ISCD recommends more frequent monitoring for conditions associated with rapid bone loss, e.g., glucocorticoid therapy.
<b>Screening in Special Situations</b>		
<b>1</b> (ACG: Ulcerative Colitis)	1 Poor	DXA screening should be considered in IBD patients: (1) with risk factors for OP such as smoking, low body mass, sedentary lifestyle, hypogonadism, family history, and nutritional deficiencies; (2) age ≥60 yrs; (3) using corticosteroids >3 months consecutively or recurrently.
<b>1</b> (EUA: Male Hypogonadism)	1 Fair	Adult men with established severe hypogonadism (late-onset) should be screened for concomitant OP. <i>(Severe was not defined.)</i>
<b>2</b> (American College of Rheumatology, ICSI: Patients Taking Glucocorticoids)	2 Good	Baseline DXA <i>recommended</i> for patients before starting glucocorticoid for an anticipated ≥3 months. <i>(Considered a consensus-based recommendation by American College of Rheumatology but a strong recommendation with moderate-quality evidence by ICSI.)</i>

Quantity of Individual GLs	Individual GL Quality*	Recommendations
<b>1 (ACP)</b>	1 Good	In certain situations, a single risk factor, e.g., ADT, may be sufficient reason to screen ( <i>not a formal recommendation</i> ).
<b>Treatment Monitoring in Special Situations</b>		
<b>1 (American College of Rheumatology: Patients Taking Glucocorticoids)</b>	1 Good	Serial BMD testing should be <i>considered</i> for patients receiving glucocorticoid therapy for ≥3 months. As often as 6 months for treatment of OP, yearly for prevention of OP.

\*Recommendations on BMD measurements for diagnosis and monitoring are supported by the American College of Rheumatology.

## Selected Payer Policies

At the direction of Washington State HTA Program, the coverage policies for the following organizations were reviewed: Aetna, Centers for Medicare & Medicaid Services (CMS), Oregon Health Evidence Review Commission (HERC), GroupHealth, and Regence Blue Cross/Blue Shield. The following highlights existing policies:

- A [CMS National Coverage Determination \(NCD\) for Bone \(Mineral \) Density Studies \(150.3\)](#), which was issued in January 2007, documented the transfer of conditions for coverage of bone mass measurements to the CMS Manual System. A document on [Bone Mass Measures](#) in the Manual System states that effective January 1, 2007, bone mass measurement is covered, generally every 2 years but subject to certain conditions. Neither the NCD nor the Manual System provides the rationale or evidence base for these policies.
- Aetna considers bone mass measurements using established techniques to be medically necessary for members who have risk factors such as certain causes of secondary osteoporosis, age > 70 years (men), and history of fragility fracture. Age is not identified as a risk factor for women, but estrogen deficiency is listed as a factor that qualifies coverage for testing.
- Oregon HERC has concluded that osteoporosis screening by DXA should be covered for women ≥ 65 years of age and for men or younger women whose 10-year risk of major osteoporotic fracture is ≥ 9.3%.
- Aetna and CMS recognize several technologies, including, but not limited to, DXA, as established methods of bone mass measurement. However, CMS places some limits on technologies other than DXA.
- Generally, Aetna and CMS cover repeat bone mass measurement no more often than every 2 years; exceptions are made when follow-up testing is to be performed with a technology different from the first test, and during long-term glucocorticoid therapy. The Oregon HERC guidance statement recommends that screening should not be repeated more often than every 15 years for women with normal BMD, every 4 years for women with moderate osteopenia, and every 2 years for women with advanced osteopenia or osteoporosis.
- GroupHealth and Regence do not have coverage policies related to general screening for osteoporosis.

See **Selected Payer Policies** in the **TECHNICAL REPORT** for additional details and links to policy documents.

## Overall Summary and Discussion

### Evidence-Based Summary Statement

The effectiveness of screening programs for preventing osteoporotic fracture has not been established, although some positive evidence for screening older adults has been published. Results of screening studies and NNS analyses suggest that the effectiveness of screening increases with age and that screening may be somewhat more effective in women than in men. Consistent with this evidence were the 2 fairly well-designed modeling studies supporting the cost-effectiveness of screening in older women. As noted in the Literature Review (but not considered for forming evidence-based conclusions), the Geisinger Health Plan observed both an increase in screening and a decrease in fracture incidence after instituting a comprehensive osteoporosis disease management program for women over the age of 55. However, neither the optimal age to start screening nor the type and number of risk factors that must be present to make screening effective have been empirically determined. The large number of practice guidelines on osteoporosis in the U.S. population are almost universally in agreement with the current recommendation of the USPSTF that screening begin at age 65 for women and that screening in women younger than age 65 be guided by the presence of risk factors, with the threshold for screening being a risk comparable to that of a 65-year-old white woman without additional risk factors. (According to the most commonly used fracture prediction tool, the WHO's FRAX, a 65-year-old woman with no more than 1 risk factor has a 9.3% risk of any osteoporotic fracture in the next 10 years. This tool was used to support the current USPSTF recommendations.) In the absence of evidence from controlled studies designed to assess the benefits and harms of screening, the USPSTF recommendation is predicated on evidence that accurate risk assessment tools exist and that osteoporosis medications are effective and safe, at least in postmenopausal women. All organizations agree that neither the effectiveness of osteoporotic treatment in men nor the value of screening in men is clear, but practice guidelines commonly advise screening men who are older than 70 years and men age 50 to 69 years who have additional risk factors.

Two large longitudinal studies have shown that changes in BMD are slow until individuals are approximately 75 or 80 years old, and that a younger individual with a normal DXA scan is not likely to become osteoporotic for many years. No studies specifically pertaining to the rate of BMD change in perimenopausal women were identified for any of the Key Questions. This is noteworthy since bone loss accelerates for a period following menopause. The presence of moderate to advanced osteopenia in an individual of any age may justify repeat screening within 1 to several years. The 2 studies used different analytic models, neither of which has been validated in other study groups. Thus, the evidence does not definitively support a particular screening schedule.

Two large observational studies reported positive outcomes attributable to screening in men taking ADT for prostate cancer and in men taking corticosteroids for ulcerative colitis. Practice guidelines support osteoporosis screening in the presence of ADT and prolonged use of glucocorticoids even without

consideration of any other factors. However, no evidence with regard to screening outcomes or relevant practice guidelines are available for most of the wide range of medical conditions and treatments that can contribute to osteoporosis or fracture risk. It is noteworthy that glucocorticoids and rheumatoid arthritis are standard factors in the FRAX calculator, but other secondary medical causes of osteoporosis are not taken into account unless a BMD score is unavailable. No studies designed to estimate optimal screening and treatment monitoring intervals in these special populations were identified. The effectiveness of osteoporosis medications in patients also taking *osteoporosis-inducing* medications or in patients with various diagnoses of chronic disease has not been well-studied.

Empirical evidence of the effectiveness of serial BMD testing in patients taking osteoporosis medications is also lacking, as is any type of study to address the optimal interval for testing in this population. Although practice guidelines advise testing to monitor treatment effect every 1 to 3 years, or until BMD values stabilize, there is substantial evidence that change in BMD explains less than half of the reduction in fracture risk caused by these medications. Standard practice is for osteoporosis medications to be taken for 5 years and then discontinued; there is no consensus on when individuals who have finished a 5-year course of treatment should be tested for possible resumption of therapy.

DXA, the standard technology for osteoporosis screening, is a safe technology, although cumulative radiation exposure over a lifetime of frequent screening would be of concern. In limited circumstances, there is a small risk of serious adverse effects associated with treatment for osteoporosis. The uncertainty associated with optimal screening intervals and the precise definition of high risk for osteoporosis and for fracture has the potential to result in both unnecessary screening and unnecessary treatment, as well as missed opportunities to screen and/or treat.

The evidence pertaining to osteoporosis screening is limited by an overall issue of generalizability. Fracture risk tools, estimates of fracture prevalence, and most studies evaluating the effect of screening strategies in general populations have been community-based. In other words, study participants have been recruited through advertising rather than through medical facilities. Thus, the evidence regarding the utility of screening may not be entirely generalizable to a clinical population.

### **Gaps in the Evidence**

The following evidence is needed to better answer the Key Questions of this report:

- Very large observational studies (cohort design) and pragmatic RCTs designed to measure the reduction in fracture risk attributable to screening in patients without obvious evidence of osteoporosis.
- Treatment and screening trials, as well as validation of risk prediction tools, in non-white populations.
- Treatment and screening trials in men.
- Additional longitudinal studies in U.S. populations that use a standardized approach to determine the screening interval required to detect transition to osteoporosis treatment threshold, given age and T-score at the time of the last screening. Studies that might allow a

more precise age cutoff in postmenopausal women and older men and studies of perimenopausal women are needed.

- Pragmatic RCTs designed to measure the reduction in fracture risk attributable to BMD monitoring versus no monitoring and to compare fracture outcomes between different monitoring strategies.
- Additional cost-effectiveness studies, preferably trial-based, designed to assess the cost-effectiveness of screening and treatment monitoring combined with BMD-guided management, compared with no screening/monitoring combined with management based on clinical assessment alone

## TECHNICAL REPORT

### Clinical Background

This section covers the following topics:

[Osteoporosis: Prevalence, Consequences, and Definition](#)

[Dual X-Ray Absorptiometry \(DXA\)](#)

[The Rationale for Screening and BMD Monitoring](#)

[Osteoporosis and Fracture](#)

[Risk Factors for Osteoporosis](#)

[Tools for Assessing Risk of Osteoporosis](#)

[Tools for Assessing Risk of Fracture](#)

[Technologies Other Than DXA for Assessing Bone Health and Predicting Fracture](#)

[Treatment](#)

#### **Osteoporosis: Prevalence, Consequences, and Definition**

Osteoporosis is the most common bone disease in humans. It is a systemic skeletal disease involving low bone mass and microarchitectural deterioration, both of which lead to fragility and increased risk of fracture (NOF, 2014). According to data collected through the National Health and Nutrition Examination Survey (NHANES) for the years 2005 to 2008, osteoporosis is prevalent in 4% of American men who are 50 years of age or older and in 16% of American women who are 50 years of age or older. Osteopenia (low bone mass) is prevalent in 38% of American men and 61% of American women (Looker et al., 2012). Other sources from the National Center for Health Statistics (NCHS) sources suggest that as many as 50% of Americans older than age 50 will be at risk for osteoporotic fracture during their lifetime (CDC, 2013). Prevalence is expected to increase as the proportion of the population age > 65 years increases (Lane, 2006).

A large economic burden due to osteoporotic fractures is demonstrated by recent findings that these events cause more than 432,000 hospital admissions, 2.5 million medical office visits, and approximately 180,000 nursing home admissions per year in the United States (U.S.) (Lane, 2006). Hip fractures are associated with considerable excess mortality, estimated at 8.4% to 36% for 1 year. Mortality related to hip fracture is higher in men than in women (NOF, 2014).

Peak bone mass, which occurs around the age of 30, is largely determined by genetics. Loss of bone mass thereafter occurs in general populations as the result of aging. Hormonal changes cause the natural process of removing older bone and replacing it with new bone to become imbalanced, with loss exceeding replacement. In women, bone loss usually occurs more rapidly for several years after

menopause and then slows down again so that men and women age 65 to 70 years and older lose bone mass at about the same rate. The World Health Organization (WHO) has defined osteopenia and osteoporosis in terms of bone mineral density (BMD) at the hip or lumbar spine, as measured by dual x-ray absorptiometry (DXA). (See additional details regarding the WHO definitions in the following section.) Alternatively, a diagnosis of osteoporosis is considered valid on the basis of adulthood hip or vertebral fracture in the absence of major trauma. Examples of major trauma are an automobile accident or a fall from a multiple-story height. There is some evidence that in an adult age > 50 years, any fracture, whether trauma-related or not, should raise the suspicion of low BMD. Compared with standards for measuring BMD, standards for measuring the *rate* of bone loss and the quality of bone are not as well defined (Nelson et al., 2010b; Warriner and Saag, 2013; Lo et al., 2011; NOF, 2014).

### Dual X-Ray Absorptiometry (DXA)

In clinical practice, the standard technology for measuring BMD and diagnosing osteopenia or osteoporosis is DXA. BMD can be expressed as grams of mineral per square centimeter ( $\text{g}/\text{cm}^2$ ) scanned or as a score that expresses the relationship to normal values. A DXA T-score represents a comparison to a *young adult* reference population of the same sex and is used to express the relative BMD status of older adults. A *T-score* is calculated by subtracting the individual's BMD (in  $\text{g}/\text{cm}^2$ ) from the reference population mean and then dividing by the standard deviation (SD) of the reference population. The resulting score represents the number of SDs above or below normal. T-score cutoff points defined by WHO are universally recognized for diagnosing osteopenia and osteoporosis. The International Society for Clinical Densitometry (ISCD) prefers the terms *low bone mass* or *low bone density* to *osteopenia*. Currently, the NHANES III data are considered the appropriate reference for calculating total hip and femoral neck T-scores, but the ISCD recommends that manufacturers use their own databases for the reference ranges for lumbar spine T-scores. A DXA *Z-score* represents a comparison to the BMD of an age-, sex-, and ethnicity-matched reference population and is commonly used to express the BMD status of premenopausal women, men younger than age 50 years, and children. Z-score cutoff values have been defined by the ISCD. The ISCD cautions that *osteoporosis cannot be diagnosed in men under the age of 50 on the basis of BMD status alone*, but advises that the *WHO definition of osteoporosis is appropriate for women in menopausal transition*. See **Table 1** for a summary of the cutoff values (Lane, 2006; Nelson et al., 2010b; ISCD, 2013; Warriner and Saag, 2013; NOF, 2014).

DXA measurement at the hip is considered the best predictor of future hip fracture, but for diagnosis according to WHO criteria, DXA measurements are made at both the lumbar spine and femoral neck. DXA measurements at the hip and spine are considered *central DXA*. Radial measurements may be substituted if hip and lumbar spine measurements cannot be made or are unusable). Peripheral DXA (pDXA) measures bone density of the forearm, finger, or heel, and can predict vertebral and overall fracture incidence in postmenopausal women, with an accuracy in the range of 59% to 66%; the ability of pDXA to assess fracture prediction in men is unknown. pDXA is not appropriate for monitoring BMD after treatment (Nelson et al., 2010b; Warriner and Saag, 2013; NOF, 2014).



**Table 1. Interpretation of DXA Results (repeated from Evidence Summary)**

WHO Definitions for Postmenopausal Women and Men Older Than 50 years	
T-score –1.0 and above:	Normal
T-score above –2.5 but below –1.0:	Low bone mass (osteopenia)
T-score at or below –2.5:	Osteoporosis
T-score at or below –2.5 with $\geq 1$ fractures:	Severe or established osteoporosis
ISCD Definitions for Premenopausal Women, Men Younger Than 50 Years, and Children	
Z-score above –2.0:	BMD within the expected range for age
Z-score at or below –2.0:	Low BMD for chronological age

### The Rationale for Screening and BMD Monitoring

Several expert sources point to the underdiagnosis and undertreatment of osteoporosis (Lim et al., 2009; Ciaschini et al., 2010; Nelson et al., 2010b; Lim et al., 2013). Research suggests that, among older women with an osteoporotic fracture, fewer than 25% receive either a BMD test or a prescription for an osteoporotic drug in the 6 months following fracture (NOF, 2014). The potential utility of screening for osteoporosis relates to the opportunity to identify individuals for whom treatment is appropriate. Utility further depends on the effectiveness of the treatment for osteoporosis. In patients who are already being treated for osteoporosis or low bone mass, the objectives of monitoring BMD are to determine that treatment is working and to assess the appropriateness of treatment cessation.

Studies have suggested that although initial BMD is the stronger predictor, the rate of DXA-measured BMD *loss over time* is a statistically independent predictor of fracture risk, after adjusting for initial BMD (Hillier et al., 2007). There is no consensus on how frequently BMD should be tested in order to detect a meaningful change in fracture risk.

### Osteoporosis and Fracture

The chief clinical concern associated with osteoporosis is risk of fracture. Otherwise, osteoporosis does not produce symptoms. Fractures that are thought to be attributable to osteoporosis are variously referred to as *osteoporotic fractures*, *fragility fractures*, *low-stress fractures*, and *nontraumatic fractures*. They occur with little or no trauma and can occur even in the presence of bone mass that does not meet the definition of osteoporosis but is considered below normal. In fact, most fragility fractures occur in individuals with low bone mass rather than in individuals with osteoporosis, because low bone mass is more common. The lifetime risk of osteoporotic fracture among white women is approximately 50%. Osteoporosis is less common in African Americans compared with whites, but the elevated risk of fracture associated with osteoporosis is the same. A fracture that occurs at a major skeletal site in any adult older than 50 years of age is considered to be a probable indicator of osteoporosis, even if the fracture occurs as the result of trauma. The risk of fracture in older populations is caused not only by lower bone mass but also by an increased propensity to fall. Heightened fall risk occurs not only because of deteriorating musculoskeletal condition, but also because of a greater prevalence of such medical

problems as orthostatic hypotension, arrhythmia, poor vision, diminished cognitive skills, and urge urinary incontinence (NOF, 2014).

Chronic pain, disability, and even death can occur because of fracture in an older population. Compared with other types of fragility fractures, hip fractures tend to have the greatest impact on mortality, function, and quality of life. The excess 1-year mortality rate associated with hip fractures has been estimated at 8.4% to 36%. Compared with women, men have a higher rate of mortality associated with osteoporotic fracture. A substantial proportion of individuals with hip fracture require a period of long-term care. Although vertebral fractures are usually initially clinically silent, the resulting chronic pain and postural changes contribute to considerable disability. Multiple thoracic fractures can cause respiratory and digestive problems. Wrist fractures can interfere with some activities of daily living (ADL). Pelvic and humeral fractures can also lead to morbidity and mortality (Lane, 2006; Nelson et al., 2010b; NOF, 2014).

### **Risk Factors for Osteoporosis**

A wide range of factors have been identified as causing or contributing to osteoporosis. These include reversible lifestyle factors such as alcohol abuse, vitamin D and calcium intake, physical activity, BMI, and smoking; hormonal disorders; type 1 diabetes mellitus; malnutrition or conditions that cause malabsorption; rheumatoid arthritis; and a variety of medications such as glucocorticoids and ADT (Grossman et al., 2010; Watts et al., 2012; Warriner and Saag, 2013; WHO, 2014).

The current recommendations of the USPSTF regarding screening for osteoporosis are stated primarily in terms of sex and age but also refer to “additional risk factors.” Specifically, screening is recommended for women age  $\geq 65$  years without previous known fractures or secondary causes of osteoporosis, and in women  $< 65$  years whose 10-year fracture risk is equal to or greater than that of a 65-year-old white woman without additional risk factors. Those factors are not defined in either the published or online version of the USPSTF recommendation statement. Nor does the recommendation specify the tool that should be used to calculate risk. However, the USPSTF chose the Fracture Risk Assessment Tool (FRAX) as the best instrument for its own estimates of 10-year fracture risk for different risk profiles (USPSTF, 2011a; USPSTF, 2011b). The evidence review sponsored by the Agency for Healthcare Research and Quality (AHRQ) to support the updated USPSTF recommendation (Nelson et al., 2010a) found that the FRAX tool estimated the 10-year fracture risk of a 65-year-old white woman with no more than 1 additional factor to be 9.3% for any osteoporotic fracture and 1.2% for hip fracture. The authors of the report then used the FRAX tool to identify risk factor and age combinations for which fracture risk would exceed that of the index case of a 65-year-old white woman.<sup>3</sup> See [Table 9](#) for a list of the risk factors for

---

<sup>3</sup> See Figure 3 in [Screening for Osteoporosis: Systematic Review to Update the 2002 U.S. Preventive Services Task Force Recommendation \[Internet\]](#). Under the heading *Risk for Osteoporotic or Hip Fracture - >one risk factor*, gray shading is used to identify women younger than age 65 whose risk profile would make them eligible for screening according to the USPSTF recommendation. For example, a 55-year-old woman with low body mass index (BMI) and

osteoporosis and/or fracture that are included in the FRAX tool and most commonly used in clinical practice. (See **Assessment of Fracture Risk, Performance of Clinical Factors With and Without DXA Score for Prediction of Fracture, FRAX.**) Many lifestyle factors, medication conditions, and medications that are not included in the FRAX tool are also considered to have a potential relationship with osteoporosis and fracture risk (NOF, 2014).

A search for systematic reviews published in the last 10 years identified no comprehensive and systematic assessment of the many factors linked to osteoporosis. However, the following relevant systematic reviews and large observational studies have evaluated the comparative contribution of particular sets of factors to the risk of osteoporosis or the prevalence of osteoporosis in certain diseases:

- A systematic evaluation of risk factors for low BMD in healthy women age 40 to 60 years was able to confirm only low body weight and postmenopausal status as risk factors for low BMD (good evidence). Alcohol and caffeine intake and reproductive history were found *not* to be risk factors (fair or good evidence). The evidence was inconsistent or insufficient for the effect on BMD of calcium intake, physical activity, smoking, age at menarche, history of amenorrhea, family history of osteoporosis, race/ethnicity, and current age. The authors concluded that in healthy white women age 40 to 60 years, only those with a low body weight (< 70 kilograms [kg]) be selected for BMD testing (Waugh et al., 2009).
- The 2008 American College of Physicians guidelines (Qaseem et al., 2008) referred to an earlier meta-analysis (Espallargues et al., 2001) that considered 80 risk factors and identified only the following 12 as being associated with a *high* risk of osteoporosis: age (> 70 to 80 years), low body weight, loss of weight, physical inactivity, use of corticosteroids, use of anticonvulsants, primary hyperparathyroidism, diabetes mellitus type 1, anorexia nervosa, gastrectomy, pernicious anemia, and prior osteoporotic fracture. Participants in the studies included in the meta-analysis were primarily women.
- A systematic review of 11 randomized controlled trials (RCTs) found that aromatase inhibitors (AIs) result in low bone density and high fracture risk in women using AIs for early-stage hormone receptor–positive breast cancer. Trials were selected if they evaluated the adverse effects of AIs compared with placebo and/or tamoxifen. The review authors did not conduct a meta-analysis, but study results were very consistent. However, the evidence was insufficient for assessing the interaction between AI adverse effects and age or other baseline fracture risk factors (Becker et al., 2012).
- A meta-analysis calculated high pooled prevalence estimates for young adults with cystic fibrosis; the results were as follows: osteoporosis, 23.5%; osteopenia, 38%; radiological vertebral fracture, 14%; nonvertebral fracture, 19.7%. Median age of the patients was 28.2

---

a parent who had a hip fracture has an 11% 10-year risk of any osteoporotic fracture and a 0.7% 10-year risk of a hip fracture.

years. The median proportion of patients using corticosteroids across studies was 30.5%, but in sensitivity analysis, corticosteroid use was not found to be explanatory for the variation in prevalence (Paccou et al., 2010).

- A systematic review to assess osteoporosis and osteopenia in older adults (mean age 63 years) with chronic obstructive pulmonary disease (COPD) found a relatively high prevalence of abnormal BMD, ranging from 9% to 69% for osteoporosis and 27% to 67% for osteopenia. Osteoporosis was associated with measures of body composition, COPD severity, and the use of corticosteroids, but conclusions regarding causality were not possible (Graat-Verboom et al., 2009). NOTE: COPD is not listed in the NOF table of contributing factors but is mentioned as a risk factor for men in the Endocrine Society guidelines (Watts et al., 2012).
- A systematic qualitative review of osteoporosis in spinal cord injury found evidence of a bone metabolism imbalance and accelerated early bone resorption, but a statistically significant decrease in BMD according to DXA scanning did not appear until approximately 12 months after injury (Charmetant et al., 2010).
- A prospective cohort study (total n=38,812) followed patients who were  $\geq 50$  years old and found that inflammatory bowel disease (IBD) doubled the risk of *hip* fracture, with similar hazard ratios (HRs) after adjustment of age and sex only, adjustment for FRAX score computed with BMD, and adjustment for FRAX score without knowledge of BMD (Targownik et al., 2013). (IBD was considered an umbrella term for Crohn's disease and ulcerative colitis.) IBD was not statistically associated with an increased overall risk of *major* osteoporotic fracture. A comparison of the HRs for IBD adjusted for FRAX with BMD (HR, 2.14; 95% confidence interval [CI], 1.26 to 3.64) and for IBD adjusted for FRAX minus BMD information (HR, 2.10; 95% CI, 1.23 to 3.57) suggests that *BMD adds negligible information* to an assessment of fracture risk in individuals with IBD. NOTE: FRAX is a common tool used to quantify fracture risk. See **Assessment of Fracture Risk, Performance of Clinical Factors With and Without DXA Score for Prediction of Fracture, FRAX**.
- A systematic literature review estimated that, at diagnosis, approximately one-third of patients with celiac disease have osteoporosis and another one-third have osteopenia (Fouda et al., 2012). However, the authors argued that a gluten-free diet is the most important treatment for bone loss in celiac disease and that *evidence does not support routine screening for low BMD* at the time of diagnosis.

One group of researchers has categorized medication associations with osteoporosis in the following manner (Zhumkhawala et al., 2013):

Drugs that cause osteoporosis: Oral corticosteroids, testosterone, anticonvulsants, heparin, methotrexate.

Drugs associated with hip fracture: Proton pump inhibitors (PPI), histamine-2 blockers, sedative hypnotics, antidepressants, antipsychotics, antihypertensives.

### Tools for Assessing Risk of Osteoporosis

A 2010 review by Nelson and colleagues (referred to as the 2010 Nelson review in this report) evaluated 14 externally validated instruments that can be used to predict a DXA T-score of  $\leq -2.50$  (osteoporosis) (Nelson et al., 2010b). The tools used various combinations of age, sex, and the factors mentioned in the preceding section. Overall accuracy, expressed as the area under the receiver operating characteristics curve (AUC), ranged from 0.60 to 0.80 in most studies. In other words, these tools were 60% to 80% accurate at optimal cutoff values for predicting a diagnosis of osteoporosis. The most commonly included elements were age, weight or body mass index (BMI), and previous fracture. Simple tools (using few factors) performed similarly to more complex tools. The authors also noted that the tools have been validated in general populations and their applicability to clinical populations is somewhat uncertain. Other systematic review authors have come to similar conclusions about tools for predicting low BMD in postmenopausal women and using such tools to identify women who should be screened (McLeod and Johnson, 2009; Rud et al., 2009).

### Tools for Assessing Risk of Fracture

#### *Performance of DXA for Prediction of Fracture*

A meta-analysis published in 1996 and cited by key review articles established a strong association between BMD, as measured by DXA, and fracture, with the best predictor being hip BMD (Marshall et al., 1996). The Marshall meta-analysis pooled data from 11 cohort studies for women in their late 60s or older. The pooled relative risk (RR) of hip fracture per decrease of 1 SD in femoral neck BMD was 2.6 (95% CI, 2.0 to 3.5). According to Marshall and colleagues, this association is stronger than the association between a 1-SD increase in blood pressure with stroke (RR, 1.5) or a 1-SD increase in serum cholesterol concentration for cardiovascular disease (RR, 1.4). Ten-year results from the Study of Osteoporotic Fractures (SOF), which originally enrolled 9704 white women age 65 or older, suggested a very similar age-adjusted RR of 2.37 (95% CI, 2.12 to 2.66) for the relationship between femoral neck BMD and hip fracture. Evidence from the SOF suggested, however, that BMD does not completely explain the incidence of fractures in older women.

The proportion of hip fractures that were attributable to a diagnosis of osteoporosis (T-score  $\leq -2.5$ ), was only 21% based on total spine BMD, and 28% based on total hip BMD (Stone et al., 2003). Other authors have reported that more than half of individuals with fragility fractures do not meet the BMD criteria for a diagnosis of osteoporosis (although they may have osteopenia, or low bone mass) (Silverman and Calderon, 2010).

A Rapid Response Report conducted by the Canadian Agency for Drugs and Technologies in Health (CADTH) sought to determine whether T-score or Z-score was more accurate in assessing the 10-year fracture risk of patients younger than age 50 years who are at increased risk because of systemic steroid use. The report found no studies designed to answer this question (CADTH, 2011a).

*Performance of Clinical Factors With and Without DXA Score for Prediction of Fracture*

The 2010 Nelson review identified 11 externally validated instruments for assessment of fracture risk (Nelson et al., 2010b). Only 3 instruments included BMD (DXA) T-score: Fracture Index (SOF), FRAX, and Garvan nomogram. At least 1 of these 3 instruments, the FRAX, will also compute risk if BMD is not available (WHO, 2014). Nelson and colleagues did not comment on the relative performance of tools that did and did not include BMD. Five of the tools took into account 1 or more secondary causes of osteoporosis, the most common being glucocorticoid use. The FRAX tools and another called the QFracture include a large number of variables related to lifestyle, medical conditions, and medication use. The content of these tools is detailed in Table 1 of the summary of the 2010 AHRQ evidence review (Nelson et al., 2010b).

Nelson and colleagues stated conclusions regarding fracture prediction tools similar to their conclusion regarding tools for predicting BMD-defined osteoporosis: (1) simple tools using combinations such as age plus BMD or BMD plus fracture history performed as well as the more complex tools; and (2) these tools may have limited generalizability to clinical populations since they were validated in general populations (Nelson et al., 2010b). The FRAX tool may have an advantage over other tools since it takes into account the competing risk of death from other causes in older adults, and 1 group of researchers has demonstrated that failure to adjust for the competing risk of death can result in inflated estimates of 10-year fracture probability (Leslie et al., 2013). Nevertheless, data compiled by Nelson and colleagues showed that different studies' estimates of overall accuracy for prediction of hip fracture by the FRAX tool fell within a very similar range (equivalent to 65% to 81%) as corresponding estimates by other studies of the accuracy of other tools (71% to 84%).

FRAX

The most widely used instrument for assessing fracture risk does not require knowledge of BMD status but takes it into account if it is supplied (WHO, 2014). This tool, called FRAX, was developed by WHO and is specifically recommended for fracture assessment by the NOF (NOF, 2014). The FRAX tool is also featured in guidelines published by other groups (Lim et al., 2009; NAMS, 2010; Watts et al., 2012). WHO provides an [online calculator](#). It was devised from an analysis of 46,340 individuals and was then validated in 230,486 individuals. The 10 risk factors included in the tool were chosen on the basis of globally available data, independent (of BMD) association with fracture, clinical practicality, and responsiveness to pharmaceutical intervention. Many of the derivation cohorts but only 1 of the validation cohorts included men. FRAX calculates the 10-year probability of a hip fracture and the 10-year probability of a major osteoporotic fracture (clinical vertebral, hip, forearm, or proximal humerus fracture), based on fracture and mortality rates in the relevant geographic region, i.e., an adjustment for nationality. [Table 9](#) outlines the components of the tool, which are weighted. FRAX underestimates risk in these situations: recent fracture, multiple osteoporosis-related fractures, increased risk for falls, and low BMD at the femoral neck but normal BMD at the lumbar spine. In fact, the supporting algorithm has not been validated for the use of lumbar spine BMD (NAMS, 2010; Nelson et al., 2010b; Silverman and Calderon, 2010; NOF, 2014).

FRAX was intended to be used in untreated postmenopausal women and has not been validated for the following population groups: younger adults, children, and patients currently or previously treated with pharmacotherapy for osteoporosis. Another limitation of the tool is that it assumes a constant relationship between body mass index (BMI) and mortality across racial and ethnic groups, whereas there is some evidence that fracture risk independent of BMD is different in Hispanics and African Americans, compared with Caucasians, and the accuracy of the tool for Asian Americans is unknown. For many secondary causes of osteoporosis, fracture risk has been shown to be mediated primarily through the condition's impact on BMD. Therefore, when femoral neck BMD is inserted into FRAX, the "secondary causes of osteoporosis" button is automatically inactivated. As shown in **Table 9**, several different determinants of osteoporosis are taken into account even when BMD is entered into the calculator: rheumatoid arthritis, current smoking status, alcohol intake, and use of oral glucocorticoids (Silverman and Calderon, 2010; NOF, 2014).

The National Osteoporosis Foundation recommends that clinicians consider additional factors that are not captured in the FRAX model, such as frailty or falls, patient preference, comorbidities, and recent decline in BMD (NOF, 2014).

**Table 9. Components of the WHO FRAX Tool**

- Current age
- Sex
- Weight
- Height
- A prior osteoporotic fracture (including clinical and asymptomatic vertebral fractures\*)
- Parental history of hip fracture
- Current smoking
- Oral glucocorticoids (current exposure or past exposure for >3 months at the equivalent of >5 milligrams of prednisone)
- Rheumatoid arthritis
- Secondary causes of osteoporosis (insulin-dependent type 1 diabetes, osteogenesis imperfecta in adults, untreated longstanding hyperthyroidism, hypogonadism or premature menopause [age <45 years], chronic malnutrition or malabsorption, chronic liver disease)
- Alcohol intake (≥3 drinks/day)†
- Femoral neck BMD if available (total hip BMD may be entered, but femoral neck measurement is preferred; the tool will compute a risk estimate without BMD data)

Source: FRAX Tool for the United States (WHO, 2014)

\*See FAQs

†The "secondary causes of osteoporosis" button is inactivated when femoral neck BMD is inserted into the calculator.



## Technologies Other Than DXA for Assessing Bone Health and Predicting Fracture

### *Older Technologies*

DXA is a successor to dual-photon absorptiometry (DPA). Other older technologies include single-photon absorptiometry (SPA) and single-energy x-ray absorptiometry (SXA) (Pisani et al., 2013).

### *Quantitative Computed Tomography (QCT)*

Quantitative computed tomography (QCT) produces volumetric measures of BMD. Unless otherwise specified, QCT is understood to be applied to the spine or hip. Peripheral QCT (pQCT) measures BMD in the forearm or tibia. High-resolution pQCT (HR-pQCT) at the radius and tibia provides measures of bone structure and microarchitecture as well as volumetric density. In postmenopausal women, QCT measurement of spine trabecular BMD predicts vertebral fractures, and pQCT of the forearm at the ultradistal radius predicts hip fractures. The association of QCT measurements with fracture risk in men is unknown. QCT and pQCT are associated with greater amounts of radiation exposure than patients receive with central DXA or peripheral DXA (pDXA). QCT is also more expensive than DXA. Furthermore, neither the performance of QCT as a screening tool nor the relationship of QCT measurements with standard T-scores has been very well studied (Lim et al., 2009; NOF, 2014). A 2010 Hayes report found that the evidence showed moderate correlation between QCT and DXA measurement of BMD, but that there is a lack of studies designed to measure either the clinical impact of QCT or QCT-based thresholds for a diagnosis of osteoporosis or low bone mass (Hayes, Inc., 2010).

### *Quantitative Ultrasound Densitometry (QUS)*

Quantitative ultrasound densitometry (QUS) measures the speed of sound and/or broadband ultrasound attenuation at the heel, tibia, patella, and other peripheral skeletal sites. These 2 measurements are used to compute a composite clinical score that represents features of bone quality other than BMD. QUS can predict fractures in postmenopausal women and in men 65 years and older. One advantage of this technology is the lack of any radiation exposure. The literature reports at least 1 attempt to compare diagnosis based on QUS measurements with DXA-based BMD, T-score, and Z-score. Preliminary results of this project have shown high (> 80%) but not perfect concordance with DXA measurements. Both pDXA and QUS are often used for community-based screening programs because of the portability of the equipment. The accuracy of QUS for predicting fractures in postmenopausal women is approximately 60%, while the accuracy of pDXA ranges from 60% to 66%. Some investigators have proposed QUS of the calcaneus as a tool for identifying patients who should have a DXA scan (Nelson et al., 2010b; Jiménez-Núñez et al., 2013; Pisani et al., 2013; NOF, 2014).

### *Bone Turnover Markers*

Biochemical markers of bone remodeling are associated with risk of fracture in untreated patients (independent of BMD), predict rapidity of bone loss in untreated patients, predict the extent of fracture risk reduction when repeated after 3 to 6 months of treatment, predict the magnitude of BMD increases after initiation of treatment, and reflect patient adherence to treatment. The ability of bone turnover



markers to predict treatment effects applies only to Food and Drug Administration (FDA)-approved therapies. Their accuracy in determining the appropriate duration of drug holidays and whether medication should be restarted is under investigation. Examples of resorption markers include serum C-telopeptide (CTX) and urinary N-telopeptide (NTX). Examples of formation markers include serum bone-specific alkaline phosphatase (BSAP), osteocalcin (OC), and aminoterminal propeptide of type I procollagen (PINP). Bone turnover markers are not considered an alternative to BMD testing or screening; they are used primarily in research settings and have limited clinical application (Lim et al., 2009; NOF, 2014).

## Treatment

Nonpharmacologic treatment of osteoporosis includes increasing the intake of calcium and vitamin D, as well as increasing exercise (Nelson et al., 2010a; Crandall et al., 2012). **Table 10** outlines the prescription medications that are available for primary and secondary treatment of osteoporosis. Additional detail is available in section 6 of the NOF's [Clinician's Guide to Prevention and Treatment of Osteoporosis](#).

Estrogen is the only pharmaceutical class approved for prevention of osteoporosis. Bisphosphonates are the first-line drugs for treating diagnosed osteoporosis in postmenopausal women (NAMS, 2010), but no first-line recommendation is made for men (Watts et al., 2012).

**Table 10. Prescription Medications for Osteoporosis**

Mechanism of Action	Examples
Antiresorptive agents	<p><u>Estrogen therapy brand names (examples)</u>: Climara, Estrace, Estraderm, Estratab, Ogen, Premarin, Vivelle</p> <p>NOTE: The FDA has approved estrogen only for osteoporosis <i>prevention</i>.</p> <p><u>Bisphosphonates</u>: alendronate (Fosamax), ibandronate (Boniva), risedronate (Actonel, Atelvia), zoledronic acid (Reclast)</p> <p><u>Monoclonal antibody</u>: Denosumab (Prolia)</p> <p><u>Calcitonin brand names</u>: Miacalcin, Fortical</p> <p><u>Estrogen agonist/antagonist, also called selective estrogen receptor modulators (SERMs)</u>: raloxifene (Evista)</p> <p><u>Hormone therapy brand names (examples)</u>: Activella, Femhrt, Premphase, Prempro</p> <p><u>Tissue-selective estrogen complexes</u>: Duavee</p>
Anabolic agent (PTH)	Teriparatide (Forteo)

Sources: Chen et al., 2011; Gourlay et al., 2012; Warriner and Saag, 2013; NOF, 2014

## Threshold for Medical Treatment

According to major practice guidelines, 1 or more of the following factors is considered by expert sources to be sufficient to offer medical treatment (NAMS, 2010; Watts et al., 2012; NOF, 2014):

- A clinical or radiographic fracture of the spine or hip
- A hip DXA T-score  $\leq -2.5$
- Osteopenia *and* a 10-year WHO probability of a hip fracture that is  $\geq 3\%$  or osteopenia and a 10-year risk of major osteoporotic fracture  $\geq 20\%$

NOF guidelines cite 2 studies, both commissioned by NOF, that support these criteria. One study was an economic evaluation (modeling) to determine the 10-year fracture probability above which it would be cost-effective, from a U.S. societal perspective, to treat with pharmacologic agents. The other study adapted the FRAX algorithm to clinical practice. Some experts consider it valid to consider patient preference for treatment even if the 10-year probability thresholds are not met. In addition to these considerations, the risks associated with osteoporosis drugs should also be considered (NOF, 2014).

Some investigators have proposed using absolute risk of future fracture, rather than T-score, as the basis of initiating treatment. For instance, the FRAX tool rather than DXA scanning could be used to identify women to whom treatment should be offered. There is some evidence suggesting this might be a reasonable approach. However, no prospective studies have been published that assess treatment efficacy in patients selected by any means other than DXA-based low BMD (Nayak et al., 2011).

#### *Overview of Treatment Efficacy*

Treatment efficacy is relevant to the question of whether screening for osteopenia or osteoporosis is effective. In the absence of an effective treatment, a screening program cannot be effective. Even though adherence to medical treatment is poor, efficacy has been demonstrated. The ability of FDA-approved drugs to reduce fracture risk has been studied primarily in postmenopausal women without secondary causes of low bone mass, and the overwhelming majority of trials have involved bisphosphonates. Less is known about the effectiveness of medications for treatment of glucocorticoid-induced osteoporosis, for *prevention* of osteoporosis in patients at high risk because of secondary causes, or for treatment of men (NOF, 2014).

#### *Efficacy in Postmenopausal Women*

The 2010 Nelson review concluded that bisphosphonates, PTH, raloxifene, and estrogen reduce primary vertebral fractures in postmenopausal women, and that according to sensitivity analyses, bisphosphonates also reduce nonvertebral fractures in this population (Nelson et al., 2010a). This meta-analysis was based on placebo-controlled RCTs involving women without known previous osteoporosis-related fractures (or, if fracture history was not reported, reporting a mean baseline BMD of  $-3.0$  or more) and without known secondary causes of osteoporosis. Study inclusion criteria did *not* specify that participants have a maximum BMD, e.g., T-scores could be in the normal range. The authors considered the evidence to be poor to good, depending on the medication. They noted that because of the strict patient enrollment criteria (e.g., exclusion of comorbid conditions and use of other medications), the evidence may not be generalizable to typical clinical populations. The following pooled RRs were reported for primary prevention of fractures (Nelson et al., 2010a; Nelson et al., 2010b):

- Bisphosphonates

- Vertebral: RR, 0.66 (95% CI, 0.50 to 0.89); unweighted absolute risks, 2% versus 3% (7 RCTs)
- Nonvertebral: RR, 0.83 (95% CI, 0.64 to 1.08); unweighted absolute risks, 7.4% versus 9.0% (9 RCTs). Statistically significant effects were detected in sensitivity analyses where alternate methods of pooling data were used or that included otherwise excluded trials in which > 40% of patients had a previous vertebral fracture.
- Hip: RR, 0.70 (95% CI, 0.44 to 1.11); unweighted absolute risks, 0.7% versus 0.9% (3 RCTs)
- PTH
  - Vertebral: RR, 0.32 (95% CI, 0.14 to 0.75) (1 RCT)
  - Hip: No evidence
- Raloxifene
  - Vertebral: RR, 0.61 (95% CI, 0.54 to 0.69) (2 RCTs)
  - Hip: RR, 0.97 (95% CI, 0.62 to 1.52)
- Estrogen with progestin
  - Vertebral: RR, 0.66 (95% CI, 0.46 to 0.92) (1 RCT)
  - Hip: RR, 0.67 (95% CI, 0.47 to 0.96) (1 RCT)
- Estrogen alone
  - Vertebral: RR, 0.62 (95% CI, 0.42 to 0.93) (1 RCT)
  - Hip: RR, 0.61 (95% CI, 0.41 to 0.91) (1 RCT)

NOTE: Unweighted absolute risks for drugs classes other than bisphosphonate were not available.

In contrast to the 2010 Nelson review, the more recent AHRQ evidence review on osteoporosis treatment (Crandall et al., 2012) did not exclude studies involving individuals who used glucocorticoids, had a condition associated with the chronic use of glucocorticoids, or had a condition associated with low bone density. Crandall and colleagues concluded that, according to generally high-quality evidence, a variety of antiresorptive agents are effective in reducing vertebral, nonvertebral, and/or hip fractures in women with postmenopausal osteoporosis (Crandall et al., 2012). Menopausal hormone therapy was found to reduce the risk of fracture in postmenopausal women, but not in those who already have established osteoporosis. The evidence was considered insufficient to allow conclusions regarding efficacy in individuals taking osteoporosis-inducing medications or with medical conditions associated with osteoporosis. This review did not include a meta-analysis.

The individual studies and meta-analyses included in the Nelson and Crandall reviews showed statistically significant, positive effects for pharmaceutical products with good consistency. The following study- or meta-analysis-specific data were reported for the effect on *hip or vertebral fracture* in the Crandall review:

- Bisphosphonates: Pooled RR, 0.40 to 0.68
- Denosumab: OR, 0.14 to 0.59
- Raloxifene: OR, 0.53 to 0.86 in most studies
- PTH: Pooled RR, 0.35

- Menopausal estrogen therapy: OR, 0.31 to 0.88
- Calcium: Pooled RR, 0.77 to 0.90 in most studies (but generally nonsignificant)
- Vitamin D: Pooled RR, 0.33 to 2.21 in patients selected because of osteoporosis or postmenopausal status (most RRs were not significant)
- Calcium plus vitamin D: 0.14 to 2.35 (all estimates were nonsignificant)

### *Treatment Efficacy in Men*

The 2010 Nelson review identified a single trial of PTH for men, which showed nonsignificant trends for reduced fractures. The 2012 Crandall review reported that low-quality evidence suggested effectiveness of some agents in men. A more recent systematic review identified 5 RCTs of antiresorptive treatment and 3 RCTs of anabolic treatment in men older than age 50 years (Schwarz et al., 2011). Although all studies demonstrated an increase in BMD comparable to those seen in postmenopausal women, only a nonsignificant trend in the reduction of clinical fractures was observed, either for antiresorptive treatment or anabolic treatment with teriparatide. The authors of all 3 reviews considered the evidence of medical osteoporosis treatment in men to be inconclusive.

### *Treatment Efficacy in Patients Selected by Factors Other Than Age, Sex, or Menopausal Status*

#### Osteoporosis-Inducing Medications and Special Medical Conditions:

Systematic reviews have shown osteoporosis medication to have a possible effect on fracture risk in patients with biliary cirrhosis (nonsignificant RR) and to have a significant effect on fracture risk in patients with rheumatoid arthritis or spinal cord injury. The review of osteoporosis treatment and rheumatoid arthritis found the effect to be especially strong in patients taking glucocorticoids. Other systematic reviews have shown no effect on fracture risk in patients with COPD, with Crohn's disease plus a diagnosis of osteopenia or osteoporosis, or who are taking glucocorticoids for a variety of inflammatory disorders. The reviews provide little information on the differential effectiveness of osteoporosis medication according to different combinations of risk factors.

**COPD:** A Cochrane Review detected no statistically significant reduction in fracture risk across 6 RCTs of oral bisphosphonates or in 1 RCT of intravenous bisphosphonate in adults with COPD (Conwell and Chang, 2014). However, follow-up intervals were short (12 or 24 months).

**Glucocorticoid Use:** A Rapid Response report issued by CADTH identified 3 systematic reviews and 7 RCTs assessing the effectiveness of bisphosphonates for *prevention* of osteoporosis in patients using glucocorticoids (CADTH, 2011b). Bisphosphonates were found to prevent the bone loss associated with the use of corticosteroids to treat a variety of inflammatory conditions, particularly when effectiveness was measured in terms of lumbar spine BMD. However, there was little evidence regarding the ability of bisphosphonates to reduce fracture risk in this population. The review did not provide information on the BMD of trial participants at the start of osteoporosis medication use. No evidence-based guidelines were identified that addressed optimal duration of treatment with bisphosphonates for patients initiating treatment with corticosteroids.

**IBD:** A large cohort study evaluating the contributors to fracture risk in adults older than age 50 who had IBD found that osteoporotic medication reduced fracture risk (Targownik et al., 2013). A systematic review and meta-analysis of 5 RCTs found no effect of bisphosphonate treatment on incident vertebral fracture in patients with a diagnosis of osteopenia or osteoporosis and Crohn's disease (Guo et al., 2013). Mean patient age ranged from 39 to 47 years. Therapy duration ranged from 12 to 24 months in most of the studies and was 42 months in 1 study.

**Primary Biliary Cirrhosis:** A Cochrane Review of 3 RCTs found no effect of bisphosphonates on mortality or BMD and only a nonsignificant and very imprecise association with fracture incidence (RR, 0.87; 95% CI, 0.29 to 2.66) in patients with biliary cirrhosis (Rudic et al., 2011). The review provided no information about baseline BMD or whether patients had been diagnosed with osteoporosis.

**Rheumatoid Arthritis:** A recent systematic review showed bisphosphonates to be effective in preventing vertebral fracture, whether medication was administered for prevention or for treatment of osteoporosis in patients with rheumatoid arthritis, especially in those patients taking glucocorticoids (Feng et al., 2013).

**Spinal Cord Injury:** A systematic qualitative review concluded that bisphosphonate therapy has been shown to be effective in preventing fractures in patients with spinal cord injury (Charmetant et al., 2010).

#### Efficacy According to Baseline BMD:

Also relevant to the question of whether screening is effective is whether treatment efficacy is dependent on baseline BMD. The 2010 Nelson review and the 2012 Crandall review cited evidence from 3 different studies of bisphosphonates with conflicting evidence regarding the dependence of efficacy on baseline BMD. The Nelson 2010 review found too little data to assess across studies whether the efficacy of osteoporosis medications varied by baseline BMD since most studies did not enroll patients with T-scores below  $-2.5$  and only a small number of fractures occurred during the studies. The only trial that provided a stratified analysis of baseline BMD and efficacy was the large and pivotal FIT. In FIT, alendronate was shown to have a statistically significant effect on fracture incidence (any, vertebral, and hip) only in patients with a baseline BMD at or below the threshold for a BMD-based diagnosis of osteoporosis (T-score  $\leq -2.5$ ). In contrast, the Crandall review concluded that moderate evidence from a post hoc analysis of 1 large RCT showed that low femoral neck BMD did not predict the effect of alendronate on vertebral fracture or nonvertebral fracture risk. The Crandall review also referenced another post hoc analysis that suggested risedronate was comparably effective in women with osteopenia or osteoporosis. However, the Crandall review did not identify these 2 studies.

#### *Duration of Treatment*

The optimal duration of bisphosphonate therapy has not been determined and data are sparse with respect to treatment lasting  $> 5$  years. Pharmacologic treatment is typically discontinued after 5 years because of safety concerns. Some persistence of effect beyond 5 years has been demonstrated for

bisphosphonates (NAMS, 2010; NOF, 2014). However, a systematic review and meta-analysis of studies (RCTs or nonrandomized comparative studies) found no statistically significant association between fracture incidence and whether or not patients discontinued bisphosphonate therapy after 5 years (Fraser et al., 2011).

### *Safety of Bone Drugs*

NOF describes these possible adverse effects associated with osteoporosis medications (NOF, 2014):

Bisphosphonates: Gastrointestinal (GI) problems (oral), renal impairment in patients at risk (oral), eye inflammation (oral), osteonecrosis of the jaw (rare, only with long-term use, and typically following high-dose intravenous bisphosphonate treatment in patients with cancer), atypical femur fractures (rare, long-term use)

Calcitonin: Respiratory and allergic reactions (intranasal), malignancy (4.1% versus 2.9% in a meta-analysis; all forms of calcitonin included)

Estrogen/hormone therapy: Increased risk of myocardial infarction or stroke (at  $\geq 10$  years post menopause), invasive breast cancer, pulmonary emboli, and deep vein thrombosis, rapid bone loss after discontinuation.

Raloxifene: Increased risk of deep vein thrombosis, hot flashes, leg cramps

PTH (teriparatide): Leg cramps, nausea, and dizziness; increased risk of osteosarcoma (long-term use in rodents)

Denosumab: Hypocalcemia, serious skin infections, skin rash, osteonecrosis of the jaw (primarily at high doses in cancer patients)

Both the 2010 Nelson and 2012 Crandall reviews acknowledged lists of adverse events similar to those provided by NOF. The Nelson review considered the evidence to be of poor to good quality, depending on the medication. The 2012 review considered the safety evidence to be of high quality. The Nelson review found that a number of serious adverse events have been reported in users of osteoporosis medications but that reliable estimates of incidence are not easily computed (Nelson et al., 2010a). The 2010 Nelson review summarized the evidence pertaining to serious adverse events in the following way:

- The evidence of serious upper GI adverse events, atrial fibrillation, and osteonecrosis of the jaw in otherwise healthy patients taking bisphosphonates for fracture prevention has not been consistent.
- The evidence regarding the harms associated with calcitonin and PTH is limited.
- According to a meta-analysis, raloxifene increases the risk of thromboembolic events (RR, 1.60; 95% CI, 1.15 to 2.23; 2 trials), but is not associated with coronary heart disease or stroke.

GI cancer was investigated in a systematic review and meta-analysis of 6 very large observational studies with at least 2 years of follow-up (Oh et al., 2012). The authors found no association between bisphosphonate use and GI cancer (esophageal, gastric, or colorectal cancers), regardless of whether analysis included all studies or only those studies with long-term follow-up.

A recent meta-analysis concluded that bisphosphonates increase the risk of subtrochanteric, femoral shaft, and atypical femur fracture (Gedmintas et al., 2013). An adjusted RR of 1.7 (95% CI, 1.22 to 2.37) was calculated by pooling data from 11 observational studies. Four studies evaluated  $\geq 5$  years of bisphosphonate use; the RR based on these studies was 1.62.

Two other recent systematic reviews with meta-analysis have demonstrated an increased risk of osteonecrosis associated with bisphosphonate use but have also shown that the absolute risk is extremely small and the risk is greatest in cancer patients or those receiving intravenous bisphosphonates. One review included 12 studies (total  $n=574,649$ ) in *noncancer* patients (Lee et al., 2014a). Pooled data from the 8 studies that adjusted for risk factors yielded an OR of 2.91 (95% CI, 1.62 to 5.22; high heterogeneity). Of the 574,649 patients represented in all 12 selected studies, there were 2642 cases of osteonecrosis; that is, 0.46% of patients developed osteonecrosis. The odds of osteonecrosis was much higher in the 3 studies of intravenous bisphosphonates (OR, 47.8) than in the 9 studies of oral bisphosphonates (OR, 3.15). A second review by the same authors reported an adjusted OR of 4.22 (95% CI, 3.21 to 5.54; no heterogeneity), based on 4 studies, for *cancer* patients (Lee et al., 2014b). Of the 571,009 participants in the 8 studies selected for the second review, there were 1389 cases of osteonecrosis (0.24% of participants). As with noncancer patients, intravenous administration was associated with higher risk (OR, 4.27) than was oral administration (OR, 1.18) in cancer patients.

#### *Treatments/Interventions Other than Bone Drugs*

Whether or not a patient and his or her clinician choose pharmaceutical treatment for a diagnosis of osteoporosis, other measures are generally recommended for individuals at risk of osteoporotic fracture: dietary intake and supplements (calcium, vitamin D), weight-bearing and muscle-strengthening exercise, modifications in the home to prevent falls, balance training, management of orthostatic blood pressure, avoidance of medications that affect the central nervous system, visual correction, smoking cessation, and moderation of alcohol intake. Hip protectors have been shown to have a small protective effect in nursing home populations. Some of these measures are also considered primary treatment for individuals who have not yet developed osteoporosis (Nelson et al., 2002; Warriner and Saag, 2013; NOF, 2014).

The 2012 Crandall review addressed several nonpharmaceutical treatments. Moderate-quality evidence demonstrated a lack of effectiveness from calcium in postmenopausal women, but subgroup analysis in 1 trial suggested that lack of effect may be due to poor compliance. Evidence concerning the effectiveness of vitamin D was mixed. The evidence was insufficient to allow conclusions about the effect of exercise, compared with the other agents.

## **Technical Aspects of DXA**

### **How DXA Scanning Works**

A DXA system consists of an x-ray source underneath the examining table on which the patient lies, and a detection system that moves over the patient. Both high-energy and low-energy photons are produced by the x-ray tube; attenuation of these beams is measured by the detection system. The



attenuation values of soft tissues are subtracted from total attenuation values, according to an algorithm, and the remaining value represents the degree to which the energy is diminished as it passes through bone. These resulting values are compared with standard values in phantoms of known density to produce bone mineral content in grams. Dividing the number of grams by the scanned area in centimeters squared ( $\text{cm}^2$ ) yields the BMD value (Pisani et al., 2013). The ISCD encourages manufacturers to continue using data from the National Health and Nutrition Examination Survey (NHANES) III as the reference standard for femoral neck and total hip T-scores and to use their own databases for the reference standard for lumbar spine T-scores (ISCD, 2013).

BMD measurements can be made at central sites, i.e., the spine and femur, or at peripheral sites such as the calcaneus (heel), proximal phalanges of the hand, the tibial shaft, and the radius. For purposes of diagnosing osteoporosis and assessing fracture risk, DXA scans of the lumbar spine and proximal femur are emphasized since fractures at these sites are the most severe. Certain clinical conditions can interfere with accurate measurement of bone mass by DXA and should be taken into account when interpreting DXA results: osteomalacia, osteoarthritis, soft tissue calcifications, previous fractures, severe scoliosis, and vertebral deformities (Pisani et al., 2013).

### **FDA Approval**

As of July 21, 2014, numerous bone densitometers, including DXA scanning devices, have been cleared for marketing and are listed in the [FDA Premarket Approval Database](#) under Product Code KGI. The corresponding section (Section 892.1170) of the [CRF – Code of Federal Regulations Title 21](#) describes densitometers as devices intended for medical purposes to measure bone density and mineral content by x-ray or gamma ray transmission measurements through the bone and adjacent tissues.

### **Precision**

An important area of concern with DXA scanning is the *precision*, or reproducibility, of the measurements. As explained in a white paper commissioned by ISCD, there is random error associated with any type of quantitative medical testing. For instance, the likelihood of 2 blood pressure measurements for 1 person falling within 5 millimeters of mercury (mm Hg) of each other is < 50%, even when the 2 measurements are made under identical conditions. The reason precision is of particular concern with DXA scanning is that changes in BMD occur very slowly, so it is possible that the difference in BMD values between repeat testing in an individual may primarily reflect the imprecision of the scanning process rather than a true biological change. In healthy adults, BMD diminishes at only 0.5% to 2.0% per year. Scanning facilities are advised to perform phantom scans at least once a week for verification of system calibration and to have every technologist perform an operator-specific in vivo precision assessment with representative patients (Baim et al., 2005; ISCD, 2013).

Precision varies across machines, facilities, and operators, and is also affected by patient-related factors. Furthermore, it varies across scanning sites. To quantify precision error, the ISCD recommends that facilities calculate a *least significant change* (LSC) value for each machine and operator. The combined LSC values for machine and operator(s) at each anatomical site define the site-specific precision errors for a particular facility. If a patient is repeatedly scanned on the same machine and by the same



operator, and the magnitude of BMD change exceeds the relevant LSC value, the BMD change is interpreted as representing a statistically significant change due at least in part to true biological change. The ISCD paper provides the following example (Baim et al., 2005):

Baseline lumbar spine BMD	0.500 g/cm <sup>2</sup>
Scan 2 years later	0.542 g/cm <sup>2</sup>
Difference	0.042 g/cm <sup>2</sup>
LSC for lumbar spine at this DXA center	0.027 g/cm <sup>2</sup>
<b>Conclusion</b>	<b>Since the change in BMD exceeds the LSC, a therapeutic response has occurred</b>

LSC values are typically in the range of 3% to 6% at the hip and 2% to 4% at the spine. In the example, the LSC was 5.4% of the baseline lumbar spine value and the observed change was 8.4%. It is problematic when the facility, machine, operator, or number of sites scanned does not remain the same as a patient is monitored over time. No standards have been defined for addressing precision when a patient is scanned by different operators, but the ISCD suggests either averaging precision estimates between operators or pooling data from the scans used to set both operators' precision values and then calculating a pooled precision value (Nelson et al., 2002; Baim et al., 2005; NOF, 2014).

### Safety of DXA

The radiation dose of a modern DXA scan is small, but it is sufficient to be taken into account in large-scale population screenings (Pisani et al., 2013). The effective radiation dose from a routine DXA scan of the lumbar spine and hip is typically 10 microsieverts (μSv), while radiation from cosmic rays and naturally occurring radioactive materials in the earth and human bodies amounts to a daily background dose of about 8 μSv. A conventional chest x-ray consisting of posterior-anterior and lateral views delivers an effective dose of 60 μSv. A conventional mammogram delivers about 130 μSv (Baim et al., 2005).

## Review Objectives and Analytic Framework

The scope of this report is defined as:

**Population:** Adult men and women.

**Interventions:** Bone mineral density (BMD) testing with dual x-ray absorptiometry (DXA).

**Comparisons:** Clinical assessment of fracture risk or treatment success without BMD testing.

**Outcomes:** Health outcomes such as fractures, fracture-related morbidity, fracture-related mortality; intermediate outcomes such as clinical management decisions and patient behavior; harms associated with screening, including potential harms resulting from osteoporosis treatment; cost and cost-effectiveness.

**Key Questions**

The following key questions will be addressed:

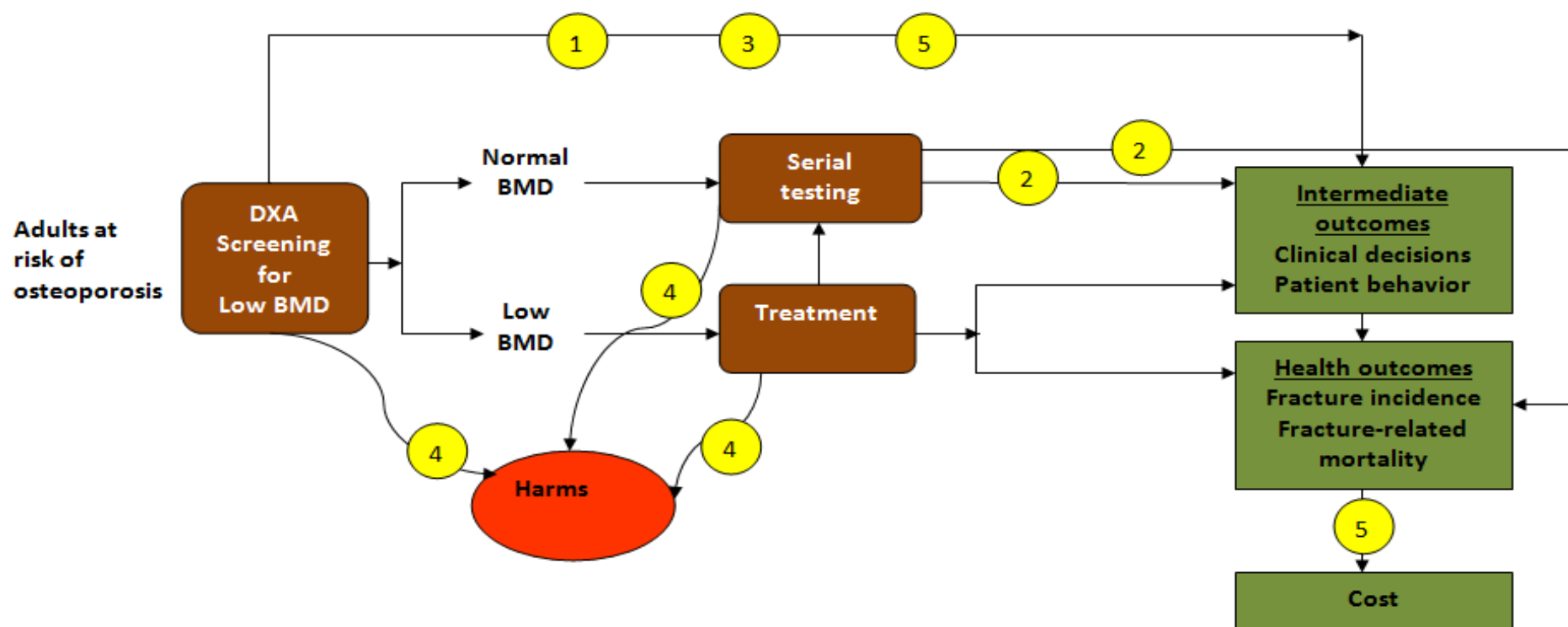
1. Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices?
  - 1a. For individual patients—and do these outcomes vary according to age, sex, or other risk factors for BMD or fracture?
  - 1b. In populations—and do these outcomes vary by population characteristics?
2. Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices?
  - 2a. For individual patients—and do these outcomes vary according to age, sex, other risk factors (including previous BMD measurements), treatment status, or testing interval?
  - 2b. In populations—and do these outcomes vary according to population characteristics or testing interval?
  - 2c. What is the minimum interval required to detect transition from normal or low BMD to osteoporosis or to assess treatment effect?
3. What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?
4. Are bone density tests safe and what are the potential downstream adverse effects?
5. What are the costs and cost-effectiveness of osteoporosis screening and monitoring?

**NOTE:** Improvement of outcomes “in populations” (Key Questions #1b and #2b) was assumed to refer to an assessment based on either individual- or group-level data for an entire community or region and were analyzed for the purposes of assessing the effect of a public health program, as opposed to data from a clinic setting or a community sample.

**Analytic Framework**  
(Key Questions referenced by number in the graphic)

1. Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 1a. In individual patients? 1b. In populations?
2. Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 2a. In individual patients? 2b. In populations?  
2c. What is the minimum interval required to detect transition from normal or low BMD to osteoporosis or to assess treatment effect?
3. What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?
4. Are bone density tests safe and what are the potential downstream adverse effects?
5. What are the costs and cost-effectiveness of osteoporosis screening and monitoring?

Key: BMD, bone mineral density; DXA, dual x-ray absorptiometry



## Methods

### Search Strategy and Selection Criteria

See [Appendix I](#) for additional search details.

#### *Systematic Reviews and Guidelines*

These sources were searched over the time frame July 8, 2014, to August 1, 2014, for systematic reviews, meta-analyses, economic evaluations, and practice guidelines published in the last 10 years:

- Core online databases such as the Agency for Healthcare Research and Quality (AHRQ), Cochrane Library, and National Guidelines Clearinghouse (NGC)
- Websites of relevant professional societies
- PubMed, using filters for Practice Guidelines, Guidelines, Meta-analyses, and Systematic Reviews

Systematic reviews were selected if they reviewed studies considered eligible for answering the Key Questions or if they provided useful background information.

#### *Primary Studies*

A variety of search strategies were used to identify primary studies that have been published since the 2010 evidence review conducted for AHRQ to support the USPSTF recommendation (Nelson et al., 2010b) and to identify studies published before that date but excluded by the AHRQ review. Initial searches were conducted in PubMed on June 12, 2014, except for search #3, which was conducted on July 17, 2014. An update search was conducted on August 4. Specific search strings are documented in [Appendix I](#). PubMed searches were restricted to articles published in the English language.

#### *Inclusion Criteria*

- To answer Key Questions #1a, #1b, #2a, #2b, and #4: RCTs or quasi-RCTs, cohort study, or case-control studies that assessed the impact of osteoporosis screening or testing on the outcomes of interest. In other words, any study that involved a comparison between a screened group and an unscreened group. The search was not limited to RCTs since the evidence review supporting the current USPSTF recommendation on screening for osteoporosis identified no RCTs of screening strategies.
- To answer Key Question #2c: (1) Trial or cohort studies comparing different strategies for the timing of screening or treatment monitoring; (2) longitudinal studies that serially measured BMD and assessed some measure of time to change in osteoporosis status (normal, osteoporosis, or osteopenia) and/or fracture.
- To answer Key Question #3: Any relevant published analysis of NNS. In addition, event rates in studies selected for Key Questions #1 and #2 were to be used to calculate NNS estimates.
- To answer Key Question #5: Any cost studies or economic evaluations published within the last 10 years.

- Systematic reviews of any of the above.

#### *Exclusion Criteria*

No a priori exclusion criteria were observed.

### **Number-Needed-to-Screen (NNS) Calculations**

Where fracture rates were reported, NNS was calculated from the results of screening studies selected for Key Question #1 or #2. NNS was assumed to be equivalent to number-needed-to-treat (NNT). The following formula for NNT was used (Gordis, 2000):

$$\text{NNT} = \frac{1}{(\text{rate in untreated [unscreened] group}) - (\text{rate in treated [screened] group})}$$

Screening studies reported fracture incident rates in terms of cumulative incidence per person-year. NNS values were calculated according to the preceding formula by using the 1-year cumulative incidence of fractures for each group (screened and control). These values were then adjusted to represent the NNS to prevent 1 fracture over the time frame represented by the study's mean follow-up interval. In other words, in a study with a mean follow-up of 5 years, the NNS to prevent fracture over 1 year, as calculated by the formula, was then divided by 5 to estimate the NNS to prevent 1 fracture over a 5-year period. We adopted this adjustment to be consistent with the analysis reported in the 2010 evidence review conducted for AHRQ to support the USPSTF recommendation (Nelson et al., 2010a; Nelson et al., 2010b). The analysis by Nelson and colleagues expressed NNS in terms of 5 years, which was the mean follow-up interval of the bisphosphonate trial that served as the basis for their assumed fracture rates.

### **Quality Assessment**

#### *Clinical Studies*

[Appendix II](#) outlines the process used by Hayes for assessing the quality of individual primary studies and the quality of bodies of evidence. This process is in alignment with the methods recommended by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group. Quality checklists for individual studies address study design, integrity of execution, completeness of reporting, and the appropriateness of the data analysis approach. Individual studies are labeled as *good*, *fair*, *poor*, or *very poor*. For individual studies included in systematic reviews, this report relies on the quality assessment by review authors. To aid in interpreting the assessment by review authors, a systematic review quality checklist, the Assessment of Multiple Systematic Reviews (AMSTAR) tool (Shea et al., 2007), was used.

Like the GRADE Working Group, Hayes uses the phrase *quality of evidence* to describe bodies of evidence in the same manner that other groups, such as AHRQ, use the phrase *strength of evidence*. The

Hayes Evidence-Grading Guides assure that assessment of the quality of bodies of evidence takes into account the following considerations:

- Methodological quality of individual studies, with an emphasis on the risk of bias within studies.
- Applicability to the population(s), intervention(s), comparator(s), and outcome(s) of interest, i.e., applicability to the PICO statement.
- Consistency of the results across studies.
- Quantity of data (number of studies and sample sizes).
- Publication bias, if relevant information or analysis is available.

**NOTE:** Two terms related to applicability are *directness* and *generalizability*. *Directness* refers to how applicable the evidence is to the outcomes of interest (i.e., health outcomes versus surrogate or intermediate outcomes) or to the comparator of interest (indirect comparison of 2 treatments versus head-to-head trials). *Generalizability* usually refers to whether study results are applicable to real-world practice. If the setting is not specified in a PICO (population-interventions-comparator-outcomes) statement, the issue of generalizability to real-world settings is not typically treated as an evidence quality issue. Another term used by some organizations is *imprecision*, which refers to findings based on such a small quantity of data that the CI surrounding a pooled estimate includes both clinically important benefits and clinically important harms, or such a small quantity of data that any results other than large statistically significant effects should be considered unreliable.

Bodies of evidence for particular outcomes are labeled as being of *high*, *moderate*, or *low quality*, or they are deemed to be *insufficient* to permit conclusions. These labels can be interpreted in the following manner:

**High:** Suggests that we can have high confidence that the evidence found is reliable, reflecting the true effect, and is very unlikely to change with the publication of future studies.

**Moderate:** Suggests that we can have reasonable confidence that the results represent the true direction of effect but that the effect estimate might well change with the publication of new studies.

**Low:** We have very little confidence in the results obtained, which often occurs when the quality of the studies is poor, the results are mixed, and/or there are few available studies. Future studies are likely to change the estimates and possibly the direction of the results.

**Very Low/Insufficient:** Suggests no confidence in any result found, which often occurs when there is a paucity of data or the data are such that we cannot make a statement on the findings.

### *Economic Evaluations*

A tool created for internal use at Hayes was used to guide interpretation and critical appraisal of economic evaluations. This tool was based on best practices as identified in the literature and addresses issues such as the reliability of effectiveness estimates, transparency of reports, quality of analysis (e.g., the inclusion of all relevant costs, benefits, and harms), generalizability/applicability, and conflicts of interest. Sources are listed in [Appendix II](#).

### *Guidelines*

The Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool (AGREE Enterprise, 2013), along with a consideration of the items related to commercial funding and conflicts of interest among the guideline authors, was used to assess the quality of practice guidelines. Use of the AGREE tool was limited to these areas because they relate most directly to the link between guideline recommendations and evidence.

## **Search Results**

One systematic review (Nelson et al., 2010a; Nelson et al., 2010b) and 8 primary studies were selected for detailed analysis as evidence pertaining to the Key Questions. For some questions, additional studies that did not meet inclusion criteria or were omitted after full texts were reviewed briefly to provide additional policy-relevant information, but were not considered in analysis, do not contribute to conclusions, and are not reflected in [Table 2](#). Additional policy-relevant evidence is discussed primarily in the **TECHNICAL REPORT**. [Table 2](#) identifies the evidence that met selection criteria, by Key Question.

## **Excluded Studies**

The following studies that were selected on the basis of title and abstract review as potentially eligible were later excluded after full text review:

- An observational study with historical controls involving older patients presenting to an orthopedic clinic because of a fragility fracture (wrist, humeral, vertebral, or hip) (Astrand et al., 2012). The study was excluded from analysis for Key Questions #1a because of very poor quality (incomplete control for confounders and high loss to follow-up) and because the population was not truly a screening population.
- A community-level RCT designed to assess whether a multifaceted intervention that included screening for low BMD increased the likelihood that patients would be managed according to current practice guidelines for Canada (Ciaschini et al., 2010). This trial used a wait-list design, where the control group continued under usual care for 6 months and then entered the intervention program. The intervention group immediately entered a program involving counseling, educational materials, and home nursing visits. Details of the protocol revealed that the specific impact of screening could not be assessed because: (1) most patients in both groups had been referred to the program because of previously ascertained low BMD values or fractures; and (2) DXA scanning was performed only if a DXA scan had not been performed

within the previous year; and (3) results of the DXA scans performed as part of the study were sent to all participants' primary care physicians shortly after randomization. Thus, DXA scanning was performed only to complete baseline data.

- An observational study that was considered as evidence for Key Question #3 but excluded because of the wrong outcome (Sawka et al., 2006). The study was designed to assess NNS to detect 1 case of previously undiagnosed osteoporosis, but did not provide any data on NNS to prevent 1 fracture.
- Two economic evaluations conducted in the context of the healthcare system in Thailand were eliminated because of potentially limited generalizability to a U.S. population (Panichkul et al., 2006; Kingkaew et al., 2012). The decision to exclude was based on the availability of other studies conducted in North America and Europe. These 2 studies were also subject to the limitation described in the following description.
- Five other economic evaluations were excluded from analysis because the reference (control) scenarios did not report all relevant costs associated with usual practice (Kraemer et al., 2006; Mobley et al., 2006; Schousboe et al., 2007; Ito et al., 2009; Mueller and Gandjour, 2009). The model descriptions and cost details did not include any assumptions about the cost of osteoporosis medications and possible confirmatory BMD testing in unscreened individuals who subsequently suffered a fracture; no screening was explicitly or implicitly understood to mean no BMD testing or osteoporosis treatment. In actual practice, an unscreened individual who had an osteoporotic fracture would likely be offered osteoporosis medication and might also be tested to confirm low BMD. Although these 5 studies were not considered as evidence for Key Question #5, their findings are briefly described in the **Literature Review** section of the **TECHNICAL REPORT**.

## Literature Review

In the following discussion, findings are described and synthesized for each Key Question. However, no summary conclusions or assessment of the quality of bodies of evidence are presented here. See the **EVIDENCE SUMMARY** for conclusions and quality assessment.

***KQ#1: Is there direct evidence that screening for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 1a: For individual patients, and do these outcomes vary according to age, sex, or other risk factors for BMD or fracture? 1b: In populations, and do these outcomes vary by population characteristics?***

### For Individual Patients (Key Question #1a)

#### *Study and Participant Characteristics*



Four community-based studies and 2 studies of clinical populations assessed the effect of DXA screening (Kern et al., 2005; Sedlak et al., 2007; Barr et al., 2010; Doheny et al., 2011; Zhumkhawala et al., 2013; Khan et al., 2014). In the 4 earlier studies, community participants were selected on the basis of age, sex, and/or menopausal status. The studies represent a mix of men and women and focused on older adults. The other 2 studies

**KQ #1a. RCTs:** Sedlak 2007, Barr 2010.

Nonrandomized or quasi-randomized trials: Kern 2005, Doheny 2011. Cohort studies: Zhumkhawala 2013, Khan 2014

See [Evidence Table IIIa](#).

addressed screening in men being treated with androgen-deprivation therapy (ADT) for prostate cancer (Zhumkhawala et al., 2013) or screening in patients with ulcerative colitis who were using corticosteroids at various intensities (Khan et al., 2014). The studies represented a mix of study designs, including RCTs and non-RCTs with generally good control for confounders. Sample sizes ranged from 196 to 5736. Studies followed patients for 1 to 9 years. Patients with known osteoporosis or a history of fragility fracture, and patients currently using osteoporosis medications or hormone replacement therapy were excluded.

#### *Fracture Incidence (4 studies)*

Four studies reported mixed results with respect to the effectiveness of screening to reduce the risk of fracture. This inconsistency may be partially explained by differences in the populations studied.

Two fair-quality studies (n=7907) suggested that screening can be effective in reducing fracture risk, but primarily in older adults and perhaps more so in women than in men. In a population-based nonrandomized trial in older men and women (age  $\geq 65$  years), fracture risk at a mean of 5 years after study enrollment was reduced by 36% in the screened group, compared with the unscreened, usual care group (adjusted HR, 0.64) (Kern et al., 2005). However, the upper bound of the CI for this estimate approached the null value (95% CI, 0.4 to 0.99), which suggests that the true RR reduction may range from as little as 1% to as much as 60%. Another community-based RCT detected a possible effect at 9 years on fracture risk in middle-aged women (age 45 to 54 years), but findings do not permit conclusions (Barr et al., 2010). The adjusted HR in the Barr et al. study, according to intention-to-treat analysis, suggested that screening might reduce fractures (HR, 0.791), but results were nonsignificant, with the upper bound of the CI reaching the null value (95% CI, 0.600 to 1.042). The per-protocol HR favored screening and was significant (HR, 0.734).

Two fair-quality studies (n=7168) suggested that DXA screening is effective in patients being treated with medications that are known to be associated with osteoporosis. These 2 studies were retrospective reviews of patient records and used multiple regression analysis to assess independent predictors of fracture, adjusting for time spent in the study. The use of multiple covariates served to adjust for known confounders. The first study enrolled men receiving ADT for prostate cancer (Zhumkhawala et al., 2013). Compared with the men who underwent DXA screening, the men who were *not* screened were 4 times more likely to have a fragility fracture within an average of 2 to 3 years after initiation of ADT. There were also significant differences in the time to fracture between the screened and unscreened groups. The other study was drawn from a database within the Veterans Administration system (Khan et al.,

2014). This study found a 50% reduction (HR, 0.5; 95% CI, 0.3 to 0.9) in fracture risk at approximately 3 years attributable to screening in a group of men being treated for ulcerative colitis.

Summary: The evidence suggesting a reduction of fracture risk through screening in middle-aged to older adults was of low quality due to study quantity and quality as well as imprecise estimates. In these studies, participants were recruited from community advertising; results may not be generalizable to clinical populations.

#### *Clinical Management Decisions (2 studies)*

One fair-quality community-based RCT (Barr et al., 2010) evaluated the impact of screening on clinical decisions in general populations. Depending on screening results and current use of corticosteroids, lifestyle changes were advised and prophylactic treatment advice was sent to participants' general practitioners. The study was conducted in Scotland. After 9 years, use of hormone replacement therapy (HRT) and vitamin D and calcium supplementation were greater in the screened group: HRT, 52.4% versus 44.5%; vitamin D, 24.2% versus 12.5% ; calcium, 20.0% versus 14.1% ( $P<0.01$  for each comparison).

In both of the studies of men being treated with osteoporosis-inducing medications, screening was associated with much higher rates of prescriptions for medications to treat osteoporosis. Among men taking ADT for prostate cancer, the database being reviewed showed that during the follow-up period, 29% of screened men had received a prescription for osteoporosis medication, while only 3% ( $P<0.0001$ ) of unscreened men had a prescription (Zhumkhawala et al., 2013). The investigators in the study of men with ulcerative colitis found that, according to clinical records, men in the screening group were substantially more likely to receive prescriptions for osteoporosis medication (36.6% versus 21.6%;  $P<0.001$ ) and for vitamin D and calcium (32.9% versus 13.4%;  $P<0.001$ ) (Khan et al., 2014).

Since baseline BMD data were not available for the unscreened groups in these 3 studies, it was not possible for the authors of either study to adjust differences between screened and unscreened groups according to appropriateness criteria. In other words, it could not be determined whether more patients in the screened group met the criteria for treatment.

Summary: Evidence regarding a positive impact on clinical management decisions was considered to be of low quality because of unknown group differences in treatment appropriateness, the limited range of health conditions represented by the 2 studies, and the lack of women participants. It was assumed that although clinician behavior might not vary according to underlying disease or osteoporosis-inducing medication, patients with different diseases and healthy populations might vary in their acceptance of medication recommendations that resulted from osteoporosis screening. Furthermore, impact on clinical management decisions should be considered an *intermediate* outcome.

*Osteoporosis-Preventing Behavior (2 studies)*

Two studies (total n=399) suggested that DXA scanning has minimal effect on osteoporosis-preventing behavior in postmenopausal women (Sedlak et al., 2007) and older men (Doheny et al., 2011). Both studies recruited participants through media advertising.

In the Sedlak et al. study, which was a good-quality RCT with wait-list controls, a small, short-term increase in calcium intake was attributed to DXA scanning. Over a 1-year time span after random assignment to a DXA scan or a wait-list, total calcium intake was 786 units (presumably international units [IUs]) per day) in the DXA scanning group and 668 units in the wait-list group. A global comparison between groups and across the 3 measurement times of baseline, 6 months, and 1 year was statistically significant ( $P<0.001$ ). However, no effect on exercise, alcohol use, or smoking was observed. In the Doheny et al. study, a fair-quality quasirandomized trial involving men who were  $\geq 50$  years of age, time spent in exercise after 1 year was slightly greater in the group that received a DXA scan right after enrollment, but the difference was nonsignificant. No difference in calcium intake was observed.

Summary: Given the small sample sizes and the possibility that even the small observed effects could diminish over the long term, evidence of minimal positive effect of screening on osteoporosis-preventing behavior was of low quality. Results may not be generalizable to a clinical population, and impact on patient behavior should be considered an *intermediate* outcome.

*Differential Effectiveness According to Risk Factors (3 studies)*

In stratified analysis conducted by 1 of the general screening studies (Kern et al., 2006), the HR was not statistically significant for either the male or female subgroup, but there was greater imprecision in the estimate for men than for women. In the age subgroups, the HR was significant only for the group who were 85 years of age or older (HR, 0.22; 95% CI, 0.06 to 0.79). The effect was significant in the subgroup defined by white race, but the subgroup defined by black race was too small to permit analysis. In the other general screening study (Barr et al., 2010), the reported HR for middle-aged women (age 40 to 54 years) suggested a smaller effect (HR, 0.791; nonsignificant) than did the HR for older women (age  $\geq 65$  years) reported by Kern et al. (HR, 0.61; nonsignificant), but the quantity of data and the indirectness of this comparison do not permit conclusions.

In the study of screening in men with ulcerative colitis, results according to the extent of corticosteroid exposure during the study period suggested that screening had no effect at low levels, a nonsignificant effect at moderate levels, and a substantial as well as significant effect at high levels.

Summary: Evidence suggesting greater effectiveness with female sex and advanced age was of very low quality due to study quality and quantity, imprecision, and/or indirect evidence. Evidence pertaining to differential effectiveness according to corticosteroid exposure was of low quality due to the availability of only a single, fair-quality study.

## In Populations (Key Question #2b)

No studies that met inclusion criteria assessed population-wide outcomes from a screening program.

Summary: Evidence pertaining to effectiveness in terms of population-level outcomes was insufficient due to a lack of studies.

### *Other Potentially Policy-Relevant Information*

A 5-year observational study evaluating the impact of a comprehensive disease management program was not selected because it was not designed to assess the impact of screening per se but provides useful insight from a payer perspective (Newman et al., 2003). The study analyzed trend data for all women over the age of 55 years who were enrolled in the Geisinger Health Plan (GHP) from 1996 to 2000. A very comprehensive program in clinical pathway distribution, clinician education, and patient education was initiated at the beginning of the study. Osteoporosis screening and prescription treatment of osteoporosis increased significantly over the time frame of the study, and the age-adjusted incidence of hip fractures fell significantly. The 1998 guidelines of the National Osteoporosis Foundation, which added risk factors to recommended treatment criteria, were used to modify the materials being distributed. The study provided no subgroup analysis of fracture outcomes for women who were actually screened during that time, and thus the effect of increased screening versus potentially better adherence to treatment guidelines cannot be assessed. Compared with a predictive model of no intervention, there was the program was judged to be cost-saving.

***KQ#2: Is there direct evidence that monitoring (serial testing) for osteoporosis and low bone density improves health outcomes, clinical management decisions, or patient choices? 2a: For individual patients, and do these outcomes vary according to age, sex, other risk factors (including previous BMD measurements), treatment status, or testing interval? 2b: In populations, and do these outcomes vary by population characteristics or testing interval?***

Previous systematic reviews did not address either of these questions, and the literature search conducted for the present report did not identify any studies designed to answer questions about the effectiveness of repeat screening in untreated individuals or of serial testing in patients who are undergoing treatment for osteoporosis.

Summary: The evidence for the effectiveness of monitoring (serial screening/testing) for osteoporosis and low bone density was insufficient because of the lack of eligible studies.

***KQ #2c: What is the minimum interval required to detect transition from normal or low BMD to osteoporosis or to assess treatment effect?***

Two eligible studies were identified for this subquestion, both having to do with screening in older adults. At the end of this section, the results of other types of analysis are briefly described to provide additional information that may be policy relevant.

*Repeat Screening in Older Adults*

Two large studies took different approaches to estimating the optimal interval for osteoporosis screening (Frost et al., 2009; Gourlay et al., 2012). Both were considered good-quality studies, given their objectives, and both focused on older adults who did not have osteoporosis

**KQ # 2c. 2 longitudinal cohort studies:** Frost et al., 2009; Gourlay et al., 2012  
See [Evidence Table IIIb](#).

at the initial study assessment. These were epidemiological studies employing a cohort design to compare exposure groups defined by age and/or T-score at baseline. Both were community-based studies, as opposed to studies conducted in clinical populations. Neither study was included in the AHRQ evidence review (Nelson et al., 2010b) supporting the current USPSTF recommendation because the publication date did not meet inclusion criteria for the report (Gourlay et al., 2012) or for unknown reasons (Frost et al., 2009).

A study of men and women age  $\geq 60$  years and living near Sydney, Australia, followed 1758 participants (750 men, 1008 women) for a median of 7 years (Frost et al., 2009). Only individuals without osteoporosis, with osteoporosis defined as T-score  $\leq -2.5$ , were included in the study. BMD measurements were made every 2 years. Radiology reports from the community's x-ray centers were used to track fractures, with the nontraumatic nature of fractures confirmed through participant interviews. The study's fracture prediction model was derived from data collected during the study and was designed to predict fracture-related outcomes, taking into account age, sex, initial femoral neck BMD, and the competing risk of death. The model did not take into account secondary causes of osteoporosis or clinical history. The model predicted the risk of osteoporosis, fracture (hip or vertebral), and osteoporosis and/or fracture. These calculations then served as the basis for estimating the mean time it would take an individual to reach a 10%, 5-year risk for hip or vertebral fracture and/or osteoporosis and time to reach a 20%, 10-year risk for one of these events. Representative estimates are displayed in [Evidence Table IIIb](#). The full set of calculations is presented in Table 4 of the published study report.

The authors of the Australian study (Frost et al., 2009) recommended a conservative approach of using the *lower bound of the 90% CI* rather than the point estimate of their time projections to define the interval for subsequent screening. They advocated a 10%, 5-year risk of hip/vertebral fracture as the target for the screening interval. However, in the U.S., the threshold for treatment is typically a 20%, 10-year risk of hip or vertebral fracture, calculated with WHO's FRAX tool (See **BACKGROUND, Treatment**). Frost and colleagues explained that their model for Australians yielded estimates of fracture risk that were either comparable to or slightly higher than estimates from the FRAX-U.S. model for most adults, but that the difference was much greater for individuals at the upper age ranges. Thus, using the Frost model and the 20%/10-year parameters to estimate time to the event of interest would suggest that an 80-year-old woman with a T-score of  $-2.2$  should be screened again in about 2.5 years, whereas using the FRAX tool to estimate risk would suggest screening would not be necessary for another 10 years.

The following results are illustrative of the findings of the Australian study and assume that the utility of repeat screening is defined by the individual reaching the U.S. treatment threshold of a 20%, 10-year risk of fracture:

- Repeat screening at < 2 years would have utility for no individuals.
- Repeat screening at < 3 years would have utility only in elderly adults with moderate to advanced osteopenia:
  - Men who were age 80 years of age with T-score  $\leq -2.2$  at the time of the last screening. (Younger men or 80-year-old men with higher T-scores would not reach the treatment threshold for an average of at least 3 years.)
  - Women who were 75 years of age with T-score  $\leq -2.0$  at the time of the last screening. (Younger women or 75-year-old women with higher T-scores would not reach the treatment threshold for an average of at least 3 years.)
  - Women who were 80 years of age with T-score  $\leq -1.5$  at the time of the last screening
- For men at age 70 at the time of initial screening (typical initial screening age in the U.S.) with normal BMD, repeat screening would not be necessary for another 9 years.
- For women at age 65 at the time of initial screening (the age recommended by the USPSTF for women without other risk factors) and normal BMD, repeat screening would not be necessary for another 12 years.

Additional data from the Australia study are presented in [Evidence Table IIb](#).

The other study of screening intervals that met eligibility criteria was based on the prospective Study of Osteoporotic Fractures (SOF), which recruited older women from the communities of Baltimore, Minneapolis, Monongahela Valley near Pittsburgh, and Portland (Gourlay et al., 2012). The analysis of screening intervals was based on the 4957 SOF participants who were age  $\geq 67$  years and did not have osteoporosis at the time of their initial BMD assessment. The objective of the study was to measure the time it took for participants to have a hip or vertebral fracture or to be diagnosed with osteoporosis before being treated for osteoporosis. The introduction to the published study report specifically references the 2011 USPSTF recommendation for osteoporosis screening and the USPSTF advice that screening intervals longer than 2 years might be needed to accurately estimate fracture risk. As the basis for their approach, Gourlay and colleagues cited earlier literature defining an optimal screening interval for any health condition as the interval needed to identify a predetermined proportion of the total eventual expected number of cases in the population being screened. They chose 10% as the predetermined proportion; the rationale for this choice was unclear.

Mean follow-up in the Gourlay et al. study was 8 years. BMD status was based on the lower of T-scores at the femoral neck and total hip. Participants were classified as having normal BMD, mild osteopenia, moderate osteopenia, or advanced osteopenia for stratified analysis. The T-score cutoff values for these BMD categories were defined as  $\geq -1.00$  (normal BMD),  $-1.01$  to  $-1.49$  (mild osteopenia),  $-1.50$  to  $-1.99$  (moderate osteopenia), and  $-2.00$  to  $-2.49$  (advanced osteopenia). These cutoff values follow WHO standards for defining normal BMD, osteopenia, and osteoporosis, but the authors provided no basis for the cutoffs used to differentiate mild, moderate, and advanced osteopenia. The model adjusted for



baseline factors nearly identical to the covariates in the FRAX model: age, BMI, estrogen use at baseline, any fracture after age 50, current smoking, current or past oral glucocorticoids, and self-reported rheumatoid arthritis. After adjustment of the observed time intervals for baseline factors, 10% of women with normal BMD at baseline were predicted to transition to osteoporosis after 17 years. A 10% transition to osteoporosis also was predicted to take 17 years for women with baseline mild osteopenia (defined as T-score  $-1.1$  to  $-1.49$ ), but the transition was predicted to take 4.7 years for women with moderate osteopenia (defined as T-score  $-1.50$  to  $-1.99$ ) and 1.1 years for women with advanced osteopenia (defined as T-score  $-2.00$  to  $-2.49$ ). Stratification was based on the lower of femoral neck or total hip BMD. Several sensitivity analyses with different target percentages of women transitioning to osteoporosis were conducted and confirmed findings. Stratified analysis showed baseline T-score to be a far bigger determinant of transition time than age, BMI, or current use of estrogen. Transition times diminished with increasing age and with use of estrogen, but differences were small and nonsignificant within each T-score category. It is important to bear in mind that these time estimates were adjusted for most of the covariates in the FRAX model. Thus, the estimates are applicable when risks other than age, sex, and baseline osteoporosis status are not present. However, unlike the study by Frost et al. (2009), the study did not evaluate whether testing intervals should shorten with advancing age beyond 67 years.

Summary: Two large and well-controlled prospective longitudinal cohort studies provided consistent evidence that for adults older than age 60 without osteoporosis at the last screening and without risk factors other than age, repeat screening generally does *not* improve the estimation of fracture risk, or by implication identify the need to start treatment, for several years after initial screening. Exceptions are individuals who are very elderly and have at least moderate osteopenia at the time of the previous screening. The overall evidence for older adults is considered to be of moderate quality. The moderate rating reflects good-quality studies with large sample sizes and general consistency of findings but lack of corroboration for either model. It also reflects the probable lack of precision in the estimates for the individuals for whom the estimated repeat screening intervals were very long (15+ years in the Australian study, 17 years in the U.S. study) since only a small proportion of participants in each study were actually followed this long. In other words, the estimated repeat screening intervals following a normal or near-normal DXA scan were imprecise. Other limitations of this evidence relate to generalizability: screening intervals for men in the U.S. have not been studied, the Australian model may not apply to the very elderly in the U.S., and optimal intervals have not been investigated in clinical (nonvolunteer) populations.

#### *Adults Younger Than 60 Years of Age and Perimenopausal Women*

The study by Frost et al. (2009) estimated screening intervals for 5-year age increments starting at age 60 (men and women), and the study by Gourlay et al. (2012) analyzed data only for women who were  $\geq 67$  years of age. No studies designed to assess optimal screening intervals in adults younger than age 60 or in perimenopausal women met inclusion criteria. Summary: Evidence for optimal screening intervals in adults younger than 60 and perimenopausal women is *insufficient* due to a lack of studies that met inclusion criteria.

*Treatment Monitoring*

An AHRQ evidence review of treatments to prevent osteoporotic fracture in men and women found no RCTs comparing different schedules of serial BMD monitoring (Crandall et al., 2012). No controlled, comparative, or longitudinal studies of testing intervals in patients being treated for osteoporosis were identified by the searches conducted for the present report.

Summary: Evidence regarding optimal screening intervals for patients being treated for osteoporosis is insufficient due to a lack of eligible studies.

*Repeat Testing Based on Risk Factors Other Than Age or Treatment Status*

Searches identified no eligible longitudinal studies of patients with medical conditions associated with osteoporosis, undergoing treatments that are associated with osteoporosis, or selected on the basis of lifestyle risk factors.

Summary: Evidence regarding optimal screening intervals for individuals with risk factors other than age or treatment status is insufficient due to a lack of studies.

*Other Potentially Policy-Relevant Information (Key Question #2c)*Screening in Older Adults:

The only study relevant to screening intervals that was cited in the Nelson evidence review for the USPSTF recommendation was a cohort study that compared the accuracy of initial BMD with that of repeat BMD for predicting nontraumatic fracture in women age  $\geq 65$  years (Hillier et al., 2007). Like the longitudinal study referred to in the previous section (Gourlay et al., 2012), the Hillier et al. study was derived from the SOF. The mean T-score at the time of the first assessment, based on total hip DXA scan, was  $-1.37$ . The study was based on participants' initial BMD measurement and their second BMD measurement, which occurred at a mean of 8 years after the first BMD measurement. Patients were followed for an additional mean 5 years after the second BMD measurement. Accuracy was expressed by area under the [receiver operating characteristic] curve (AUC). An AUC value can be interpreted as the percentage of test results that are correct. The following AUC values were reported for nonspine fracture, hip fracture, and spine fracture:

- Initial BMD (total hip): 65%, 73%, 67%
- BMD at mean 8 years later: 65%, 74%, 68% (nonsignificant comparisons with initial BMD)
- Change in BMD: 61%, 68%, 62% ( $P < 0.05$  for comparison of each with corresponding initial BMD)
- Initial BMD plus change: 65%, 74%, 68% (nonsignificant comparisons with initial BMD)

The similarity of AUC values suggests that initial BMD is as informative as the change in BMD over the next several years. Sensitivity analyses, in which participants were stratified by initial T-score or by estrogen use, yielded similar results. No other attempt to adjust for risk factors was made. The authors



concluded that in a population of older, postmenopausal women, repeat BMD measurement at 8 years did not add information to the initial BMD for prediction of fracture.

#### Patients Undergoing Treatment for Osteoporosis:

The systematic review of osteoporosis treatment by Crandall et al. (2012) identified 2 systematic reviews and 8 studies (primarily post hoc analyses of RCTs) that evaluated the ability of baseline BMD at the time of treatment initiation to predict future fracture. These reviews and studies do not answer the question of optimal testing intervals but were included in the Crandall review because they suggest that change in BMD does not fully answer the question of whether treatment is working. Findings included the following (follow-up times, the type of fracture analyzed, and the sex of participants varied):

For patients undergoing pharmaceutical treatments:

- The RR of fracture predicted by change in BMD was greater (0.80) than the actual RR of fracture according to observed results (0.65) (1 meta-analysis of 12 RCTs). Predicted RR was based on the association of incident vertebral fracture with unit decrease in BMD in the placebo group of FIT, applied to the actual difference in BMD change observed in the 12 trials.
- No association between reduction of fracture risk and change in BMD (2 post hoc analyses).
- Small (4% to 16% for oral antiresorptive agents; 30% to 40% for teriparatide; 23% to 37% for oral or intravenous ibandronate) percentage of fracture risk explained by the magnitude of BMD (5 post hoc analyses).
- Similar reduction of fracture risk between patients with an increase and patients with a decrease in BMD (2 post hoc analyses).
- Between-person variation in the effects of alendronate was small compared with within-person variation. The authors (Bell et al., 2009) concluded that monitoring in the first 3 years is unnecessary and possibly misleading.

For patients taking calcium with or without vitamin D:

- No association between BMD and fracture risk reduction (1 meta-analysis of 15 RCTs).
- Crandall and colleagues commented that fracture risk reduction appears to result from improvements in non-BMD determinants of bone strength.

A study published since the Crandall review also reported similar findings (Bruyère et al., 2012). In this post hoc analysis of 3476 women who had been treated with bazedoxifene (a selective estrogen receptor modulator [SERM]) in an RCT, BMD changes at the femoral neck after 3 years explained only 29% of the reduction in the incidence of hip fracture, and total hip BMD changes explained 44% of the risk reduction. The analysis was adjusted for age, BMI, and number of prevalent vertebral fractures.

#### Patients Who Have Discontinued Osteoporosis Treatment:

In 437 women who were randomized to continued treatment with placebo after 4 to 5 years of alendronate therapy, fracture risk at 5 years after discontinuation was not associated with BMD change

at 1 year after discontinuation (Bauer et al., 2014). The authors concluded on the basis of these findings that BMD measurement at 1 year after discontinuation of medication is not helpful for assessing whether fracture risk has increased. They additionally expressed the opinion that there is no evidence to support testing even at 2 or 3 years after discontinuation.

#### Individuals with Risk Factors Other Than Age or Treatment Status:

A longitudinal study of 44 men treated with highly active antiretroviral therapy (HAART) for human immunodeficiency virus (HIV) found a 5.3% increase in lumbar spine BMD over 6 years of follow-up (Bolland et al., 2012). This increase compared very favorably with a 0.3% increase ( $P < 0.001$ ) over the same time frame in a group of healthy controls, who were selected through workplace advertisements, subject to the same exclusion criteria that were applied to selection of the HIV patients. This observation suggests that men being treated for HIV infection are not at increased risk of bone loss. No difference in change of total hip BMD was observed. The authors concluded that routine monitoring of BMD would not be necessary in HIV-infected men over the short or medium term (Bolland et al., 2012).

A systematic review of RCTs assessing bisphosphonates for prevention or treatment of osteoporosis in patients with rheumatoid arthritis showed a statistically significant reduction in vertebral fracture risk at 12 months if medication was administered for prevention of osteoporosis, but the effect on vertebral fracture did not become significant until 36 months when medication was administered for treatment of osteoporosis (Feng et al., 2013). These findings suggest that the *soonest* BMD testing would need to occur after initiating osteoporosis medical treatment in individuals with rheumatoid arthritis would be 1 year if medication had been prescribed prophylactically and 3 years if medication were being used to treat existing osteoporosis.

In a prospective cohort study with matched controls, ADT-induced loss of BMD in men being treated for prostate cancer occurred almost entirely in the first year (Alibhai et al., 2013), but average relative change was very small.

#### ***KQ #3: What is the number needed to screen (NNS) to prevent 1 fracture in subgroups defined by age, sex, and other risk factors?***

One of the studies selected as evidence for Key Question #1a (Kern et al., 2005) provided event rates in older adults that could be converted to NNS estimates. The 2010 Nelson review (Nelson et al., 2010) provided NNS estimates for postmenopausal women. The data from these studies were derived from community recruits. Two other studies selected as evidence for Key Question #1a provided data that could be converted to NNS estimates for certain patients being treated with medications known to cause osteoporosis (Zhumkhawala et al., 2013; Khan et al., 2014).

**KQ #3: Older Adults:** Kern 2005, Nelson 2010  
**Medications That Cause Osteoporosis:** Zhumkhawala 2013, Khan 2014  
 See [Evidence Table IIIc](#).

*NNS to Prevent 1 Fracture in Older Adults*

The 2010 Nelson review made assumptions about prevalence based on a population-based epidemiological study and made assumptions about treatment effectiveness based on results for women with a baseline T-score  $\leq -2.5$  from a large, pivotal RCT of the bisphosphonate alendronate. The RCT was the Fracture Intervention Trial (FIT) (Cummings et al., 1998). The assumptions of Nelson and colleagues were then used to calculate, by 5-year age group, the NNS to prevent 1 fracture over 5 years. Implied assumptions in this analysis were that all individuals offered osteoporosis treatment would accept treatment and that compliance would be comparable to that in the FIT ( $> 80\%$  [Cummings et al., 1998]). The 2010 Nelson review estimated that 556 postmenopausal women age 65 to 69 years would have to be screened to prevent 1 *hip fracture* over the following 5 years. NNS values ranged from 1667 for women age 50 to 55 years to 238 for women age 75 to 79 years for prevention of hip fracture. NNS values ranged from 278 for women age 50 to 55 years to 43 for women age 75 to 79 years for prevention of *any* fracture. Neither the Nelson report nor the updated recommendation of the USPSTF that resulted from the report made an explicit connection between this calculation and the recommendation.

It was not considered reasonable to update the NNS calculations presented in the 2010 Nelson review with more recent prevalence or effectiveness assumptions, or to extend the model to populations other than postmenopausal women. The prevalence assumptions of the Nelson analysis were based on a 1992 population study, but more recent prevalence data by 5-year age groups were not available from the National Center for Health Statistics (NCHS), and no comparable study was identified in the searches conducted for the present report or mentioned in review articles. The latest prevalence data by 10-year age groups from the NCHS (results of the 2005 to 2008 National Health and Nutrition Examination Survey [NHANES]) appear very consistent with the 1992 estimates used in the Nelson review, at least for individuals age 60 and older (Looker, 2012):

<b>Prevalence of osteoporosis assumed by 2010 Nelson review for postmenopausal women</b>	<b>2005-2008 prevalence of osteoporosis or low bone mass from NHANES, women</b>
Age 50-54: Not included	Age 50-59: 7%
Age 55-59: 4%	
Age 60-64: 7%	Age 60-69: 10%
Age 65-69: 12%	
Age 70-74: 20%	Age 70-79: 27%
Age 75-79: 28%	
Age 80 and older: Not included	Age 80 and older: 35%

It is unclear whether these data are truly comparable. The NNS table in the Nelson review listed prevalence figures for osteoporosis defined explicitly as T-score  $\leq -2.50$ . However, the referenced source for these data presents data for osteoporosis *or low bone mass*. The similarity of the numbers reported in the Nelson review and in the NHANES results suggests that the prevalence data in the

Nelson review may have been erroneously described. The fracture rates were also described as pertaining to women in the FIT with a T-score  $\leq -2.50$ , but the accuracy of this description could not be verified with data reported in the published results of the FIT since the study report did not include data by age category.

The FIT continues to be seen as a representative study for the effectiveness of alendronate, the most effective of osteoporosis medications, in postmenopausal women. The FIT served as the source for effectiveness estimates in recent cost-effectiveness studies of osteoporosis screening (Mobley et al., 2006; Nayak et al., 2011). The most recent meta-analysis of RCTs of osteoporosis medications, also conducted as part of the 2010 Nelson review, did not provide estimates by age category. Nor, apparently, did the authors of the Nelson review consider their own pooled estimate preferable to the results of the FIT for calculating NNS.

The present report also made no attempt to apply the Nelson model to men since definitive estimates for the effectiveness of osteoporosis medications in these populations are lacking according to the National Osteoporosis Foundation (NOF) (NOF, 2014) and systematic reviews (Nelson et al., 2010b; Schwarz et al., 2011; Crandall et al., 2012).

Calculation of NNS estimates based on the screening study by Kern et al. (2005) represents a more direct approach to assessing the utility of a screening program. Using the reported rates for fracture per 1000 person-years, converting that figure to fractures per 1 person-year, and adjusting for the study's mean follow-up of 5 years, calculations would suggest that 46 women age 65 or older and 96 men age 65 or older would have to be screened to prevent 1 hip fracture over a 5-year period. The estimate for women suggested by this study is surprisingly smaller than the corresponding estimates (NNS, 238 to 556) reported in the 2010 Nelson review. Given the real-world context, a larger NNS might be expected in the empirical screening study due to possible inconsistent treatment of low bone mass and poorer adherence to treatment than was observed in the FIT, which was the basis of effectiveness estimates in the analysis of the Nelson review. Kern and colleagues used a propensity score to correct for the possibility that individuals with better outcomes were more likely to live in the geographic locations in which screening was performed. Nevertheless, the unsystematic nature of treatment assignment in this study may have created an unmeasured bias. Another explanation for the inconsistency may stem from the use of data in Nelson analysis only from women with a T-score  $\leq -2.50$ , whereas women with low bone mass (osteopenia) and a clinical fracture are also eligible for treatment in practice. (See the preceding discussion about the uncertainty of how to interpret the prevalence assumptions in the Nelson review.) Regardless of how the inconsistency is interpreted, NNS estimates from the Kern study cannot be considered precise because of the wide CIs around the overall HR and most of the subgroup HRs that served as the basis for NNS values.

Summary: Taking into account the unexplained inconsistency between the analysis in the Nelson review and the results based on the screening study and the indirect (Nelson review) or possibly confounded (screening study) nature of the data, the evidence concerning the NNS to prevent 1 fracture in older women is of low quality. The evidence concerning NNS to prevent 1 fracture in older men is of very low

quality because of the availability of data from only 1 study and the potential confounding of the incidence rates in that study.

#### *NNS to Prevent 1 Fracture in Younger Adults*

No screening studies of younger men were identified. The only study of screening in relatively younger women (Barr et al., 2010) did not provide event rates, so no NNS calculations could be made. The present report also made no attempt to apply the Nelson model to younger adults since definitive estimates for the effectiveness of osteoporosis medications in these populations are lacking according to the NOF (NOF, 2014) and systematic reviews (Nelson et al., 2010b; Crandall et al., 2012).

Summary: The evidence concerning NNS to prevent 1 fracture in younger adults is insufficient due to lack of published analyses or screening studies with useful data.

#### *NNS to Prevent 1 Fracture in Individuals Using Osteoporosis-Inducing Medications*

Two screening studies addressed special populations and provided event rates. Both studies followed patients for a mean of 3 years. The study by Zhumkhawala et al. (2013) was of fair quality and suggested that DXA scanning with subsequent treatment of osteoporosis reduces the risk of hip fracture in men being treated with ADT for prostate cancer. A translation of study results into NNS values suggests that 26 men being treated for prostate cancer would have to be screened to prevent 1 hip fracture over 3 years. The study did not report relative risks or event rates for subgroups defined by duration or total dose of ADT therapy. Thus, NNS values for screening in different exposure subgroups could not be calculated.

A fair-quality study of men with ulcerative colitis and taking corticosteroids suggested that DXA screening would reduce the risk of fracture (Khan et al., 2014). A translation of study results to NNS values suggests that 278 men with ulcerative colitis would have to be screened to prevent 1 hip fracture over 3 years. The study also found that in stratified analysis the effect of screening applied only to individuals with at least moderately intensive use of corticosteroids, and possibly only to those with highly intensive use of corticosteroids. NNS values could not be calculated for the subgroups because subgroup event rates were not reported, but compared with the overall NNS, NNS for the moderate- and high-use subgroups would be smaller.

Summary: Evidence concerning NNS to prevent fracture in screening in individuals taking medications known to be associated with osteoporosis is considered to be of very low quality because of limited applicability to the full spectrum of medications that are thought to be associated with osteoporosis risk, the availability of only 1 study each for the 2 medications addressed, lack of data for computing NNS by dose, lack of long-term data, and lack of data for women. Furthermore, since the fracture incidence rates were not adjusted for risk factors, the NNS values are subject to possible confounding.

#### *NNS to Prevent 1 Fracture in Individuals with Other Risk Factors*

No screening studies in individuals selected on the basis of age, sex, or the 2 medical conditions represented in the preceding discussion were identified. The present report also made no attempt to

apply the Nelson model to individuals with risk factors other than age, sex, or treatment status since definitive estimates for the effectiveness of osteoporosis medications in populations defined by other factors are lacking according to the NOF (NOF, 2014).

Summary: The evidence with regard to NNS in groups defined by any factor other than age, sex, or osteoporosis-inducing medication treatment was *insufficient*.

#### ***KQ #4: Are bone density tests safe and what are the potential downstream adverse effects?***

##### *Direct Evidence from Screening Studies*

No studies designed to assess harms associated with DXA scanning or the consequences of DXA scanning were identified in systematic reviews or in the searches conducted for this report.

##### *Information That May Support Inferences About Adverse Effects from Screening*

The NOF describes the radiation exposure associated with DXA scans to be *trivial* (NOF, 2014; p. 20). The radiation exposure of a typical DXA scan is considerably less (about 8  $\mu$ Sv) than the radiation delivered by a chest x-ray with posterior-anterior and lateral view delivers (about 60  $\mu$ Sv) or a conventional mammogram (about 130  $\mu$ Sv) (Baim et al., 2005). However, experts suggest that the radiation dose of a modern DXA scan is large enough for radiation exposure to be taken into account in large-scale population screenings (Pisani et al., 2014). Furthermore, the safety of frequently repeated scans over a long time frame has not been determined (Nelson et al., 2002).

Patients may suffer harm in the form of inappropriate treatment if DXA scans produce false-positive results, or from missed treatment opportunities if results are false-negatives, but the rate of false results is unknown. Harms may also occur if DXA scan results are not interpreted correctly. The evidence review that supported the 2002 USPSTF screening recommendations (Nelson et al., 2002) cited a study in which physicians reported that they found densitometry reports confusing and lacked confidence in their interpretation of the reports. Given the wide number of risk factors associated with both osteoporosis and fracture risk, there remains some clinical uncertainty in selecting patients both for screening and for treatment. Thus, unnecessary screening and unnecessary treatment are possibilities, as is a missed opportunity to appropriately screen and treat. However, actual data regarding inappropriate use of DXA scanning were not identified in the literature.

Serious GI adverse events, atrial fibrillation, and osteonecrosis of the jaw have been reported in conjunction with bisphosphonates, but according to the 2010 Nelson review, the evidence is inconsistent. As noted in the **CLINICAL BACKGROUND** section of the current report, more recent meta-analyses have found that bisphosphonates are associated with an increased risk of osteonecrosis but have also shown that the absolute risk of osteonecrosis is extremely small. One review included 12 studies (total n=574,649) in *noncancer* patients (Lee et al., 2014a). Pooled data from the 8 studies that adjusted for risk factors yielded an OR for the occurrence of osteonecrosis of 2.91 (95% CI, 1.62 to 5.22; high heterogeneity), comparing patients who were and were not taking bisphosphonates. Of the 574,649 patients represented in all 12 selected studies, there were 2642 cases of osteonecrosis; that is,

0.46% of all patients developed osteonecrosis. The relative odds of osteonecrosis were much higher in the 3 studies of intravenous bisphosphonates (OR, 47.8) than in the 9 studies of oral bisphosphonates (OR, 3.15). A second review by the same authors reported an adjusted OR of 4.22 (95% CI, 3.21 to 5.54; no heterogeneity), based on 4 studies, for *cancer* patients (Lee et al., 2014b). Of the 571,009 participants in the 8 studies selected for the second review, there were 1389 cases of osteonecrosis (0.24% of participants). As with noncancer patients, intravenous administration was associated with higher risk (OR, 4.27) than was oral administration (OR, 1.18) in cancer patients. The 2010 Nelson review also reported a pooled RR of 1.60 (95% CI, 1.15 to 2.3), based on 2 trials, for the association of raloxifene and thromboembolic events.

A recent meta-analysis concluded that bisphosphonates increase the risk of subtrochanteric, femoral shaft, and atypical femur fracture (Gedmintas et al., 2013). An adjusted RR of 1.7 (95% CI, 1.22 to 2.37) was calculated by pooling data from 11 observational studies. Four studies evaluated  $\geq 5$  years of bisphosphonate use; the RR based on these studies was 1.62. Risk difference was not reported, but the authors described these types of fractures as very rare overall.

#### ***KQ #5: What are the costs and cost-effectiveness of osteoporosis screening and monitoring?***

NOTE: For the following currency conversions, The CCEMG-EPPI-Centre web-based cost converter with the International Monetary Fund (IMF) dataset for Purchasing Power Parity (PPP) values was used on July 29, 2014, with 2007 as the price year and 2014 as the target price year: [CCEMG-EPPI-Centre Cost Converter](#) (last updated on January 27, 2014) (Shimelt et al., 2010). These conversions represent an *approximate* translation of the procedural cost and/or product price *values* to current U.S. *values*. These conversions do NOT provide an estimate of the current cost or cost-effectiveness and do not directly reflect the U.S. healthcare system when the original results were expressed in a non-U.S. currency.

#### *Cost*

The U.S. economic evaluation reviewed in the following discussion assumed the cost of a central DXA scan to be \$97.71, based on median Medicare reimbursement for 2010 (Nayak et al., 2011). This translates to approximately \$144 if converted to 2014 dollars. [However, public comment offered in response to the draft version of this report stated that Medicare reimbursement has been markedly reduced since 2010.] The Washington State Agency Utilization Data added to this report indicate that over the past 3 years, the state has paid the following average dollar amounts per DXA scan: \$104 for all beneficiaries of Public Employee Benefits (PEBB) and the Uniform Med Plan (UMP), \$124 per non-Medicare PEBB and UMP beneficiaries, and \$59 for Medicaid fee-for-service (FFS) beneficiaries. A search of the Internet suggests that out-of-pocket costs for patients without insurance are in the range of \$150 to \$250 for a standard set of DXA scans.



*Cost-Effectiveness and Cost-Utility*

Nine economic evaluations published in 2004 or later were identified, and of these, 2 were selected for analysis. The others were omitted from analysis primarily because the reference

Middle-aged to elderly women: Nayak 2011, Nshimyumukiza 2013  
See [Evidence Table III d](#).

scenarios did not report all relevant costs associated with usual practice. The 2 included evaluations support screening in older women but had conflicting findings regarding screening in younger women. One was from a U.S. perspective and appeared shortly after publication of the 2011 USPSTF recommendations for screening (Nayak et al., 2011). The other one was designed for application to the Canadian population (Nshimyumukiza et al., 2013). Following a discussion of findings for each of these studies, a very brief description of the omitted studies is given. No useful evaluations of screening in men and no evaluations of serial screening or treatment monitoring were identified.

U.S. Study (Nayak et al., 2011):

The U.S. study was conducted from a payer perspective and modeled costs and outcomes for postmenopausal women (Nayak et al., 2011). The study considered all direct, lifetime medical costs, including over-the-counter medications (PPIs) required by some individuals using bisphosphonates (Nayak et al., 2011). Seven strategies were ranked in order of increasing cost. The cost-effectiveness of each strategy was computed by comparing effectiveness with the effectiveness of the next less expensive strategy. The strategies included the status quo, 3 strategies where treatment was based on DXA T-score ( $-2.5$ ,  $-2.0$ , or  $-1.5$ ), 2 strategies where DXA was preceded by prescreening with QUS, and 2 strategies where DXA scanning was preceded by prescreening with the Simple Calculated Osteoporosis Risk Estimation (SCORE) tool. It was assumed that individuals in the no screening scenario would be offered treatment if they suffered an osteoporotic fracture. Treatment adherence was assumed to be 50% (for comparison, adherence in the FIT was  $> 80\%$  [Cummings et al., 1998]), and adherence following fracture, as assessed in an insurance plan population, was recently estimated at 56% for women and 61% for men (Balasubramanian et al., 2014). Thus, the rate of treatment adherence assumed by Nayak and colleagues seems to be reasonably conservative for a real-world setting. For all strategies, initiating screening at 65 or older dominated no screening; that is, screening for these older ages was both less expensive and more effective at reducing fracture risk than no screening.

Assuming a cost-effectiveness threshold of \$50,000, the best strategy was:

Initiate screening at age 55; DXA scanning every 5 years; treatment when T-score was  $\leq -2.5$  = \$45,450 in 2010 dollars per quality-adjusted-life-year (QALY) (approximately \$48,581 in 2014 dollars).

Assuming a cost-effectiveness threshold of \$100,000, the best strategy was:

Initiate screening at age 55; DXA scanning every 10 years; treatment when T-score was  $\leq -2.0$  = \$94,210/QALY in 2010 dollars (approximately \$100,107 in 2014 dollars).



Sensitivity analyses that varied fracture risk, adherence rates, costs, and adverse event rates yielded the same conclusions. Nayak and colleagues concluded that multiple screening strategies are cost-effective for postmenopausal women, including strategies that initiate screening at 55 years and that include prescreening tools prior to DXA, and that expansion of osteoporosis screening could improve health outcomes at reasonable cost. They also noted that differences between strategies are small. The analysis may overestimate the benefits of screening by assuming that all women who are offered treatment will start treatment (see [Evidence Table IIIId](#) for additional limitations). The authors will use their model to compare strategies in which treatment is based on FRAX and other tools when relevant efficacy data become available.

In considering the applicability of the analysis by Nayak et al. (2011) to the State of Washington, the assumption of a payer cost of \$98 per DXA scan should be compared with the actual costs (past 3 years) for the State. As summarized in the previous *Cost* section, these costs range from \$59 to more than \$124, depending on the plan and plan subgroup. *Allowable* charges for non-Medicare PEBB/UMP members have exceeded \$150.

Canadian Study (Nshimyumukiza et al., 2013):

The Canadian study was from a national health plan perspective and the time horizon was lifetime. Results for women age 40 years or older were simulated. To compute cost-effectiveness and cost-utility, the authors ranked 12 scenarios in order of increasing costs and then compared outcomes from each scenario with outcomes from the next less expensive scenario. The 12 scenarios included the status quo (in which it was assumed that only individuals who had a fracture would undergo DXA scanning and possibly be treated), 3 variations of a universal prevention promotion program, and 9 strategies representing various combinations of BMD testing versus osteoporosis risk calculation, use of the Canadian Association of Radiologists and Osteoporosis Canada (CAROC) (which incorporates BMD) to estimate fracture risk and guide treatment, and various elements of the universal prevention program. A 1-time screening was assumed. It was assumed that only individuals at high risk of fracture according to the CAROC tool would be offered pharmaceutical treatment and that adherence would be 49%. The costs and quality-of-life impact of drug-related adverse events were not considered. It was assumed in all scenarios that individuals at moderate or low risk of fracture would be encouraged to participate in a national, universal prevention program consisting of physical activity and and/or supplementation with vitamin D and calcium. The model also assumed that a certain proportion of individuals in the status quo scenario would be tested for BMD and offered treatment accordingly in the event of a fracture.

The model made these predictions:

- In *women age  $\geq 65$  years of age*, the following strategy would be cost-effective in terms of averting fracture, and in terms of adding QALYs, when the strategy was compared with the next less expensive scenario: BMD screening with DXA, followed by fracture risk assessment using CAROC, and treatment or encouragement to participate in all elements of the prevention program. The incremental cost-effectiveness ratio (ICER) was \$55,300 in 2007 Canadian dollars (approximately \$50,537 in 2014 USD).

- The same strategy also yielded an incremental cost-utility ratio (ICUR) of \$55,300 (approximately \$50,537 in 2014 USD).
- No strategy involving BMD testing was found to be cost-effective for women who were at the age of 40 to 64 years at the beginning of the program.

Sensitivity analyses yielded the same conclusions. The threshold of \$50,000 per unit of health gain was assumed to be upper limit of what Canadian society would be willing to pay. The authors believe that results would be fairly generalizable to the U.S. because of similar populations and because the CAROC fracture risk tool has a 90% concordance with the FRAX.

#### *Cost-Effectiveness of Screening in Men*

There were no economic evaluations of screening in men that reported appropriate assumptions for the use of DXA scanning and osteoporosis medications in usual practice.

#### *Cost-Effectiveness of Serial Monitoring*

There were no economic evaluations of repeat screening or treatment monitoring.

#### *Findings of Studies Excluded from Analysis*

Several additional economic evaluations, also based on modeling, were considered but were not reviewed in detail or included in analysis for Key Question #5, as previously noted in the **Methods** section. Two studies conducted in the context of the healthcare system in Thailand were eliminated because of potentially limited generalizability to a U.S. population and healthcare system (Panichkul et al., 2006; Kingkaew et al., 2012). The decision to exclude was based on the availability of other studies conducted in North America and Europe. These 2 studies were also subject to the limitations described in the following paragraph.

Five other economic evaluations were excluded from analysis because the descriptions of the reference (control) scenarios did not mirror usual practice (Kraemer et al., 2006; Mobley et al., 2006; Schousboe et al., 2007; Ito et al., 2009; Mueller and Gandjour, 2009). The 2 selected studies (Nayak et al., 2011; Nshimyumukiza et al., 2013) explicitly stated assumptions that some proportion of individuals in the no screening scenarios would subsequently incur the cost of pharmaceutical osteoporosis treatment because of fracture (and in the Nshimyumukiza et al. study, the cost of BMD testing as well, to confirm low BMD). In the excluded studies, the model descriptions and cost details did not include these assumptions; *no screening* was explicitly or implicitly understood to mean no BMD testing or osteoporosis treatment. In actual practice, an unscreened individual who had an osteoporotic fracture would likely be offered osteoporosis medication and might also be tested to confirm low BMD. The excluded studies might be expected to overestimate the cost-effectiveness of screening, but conclusions were fairly somewhat inconsistent, with the more recent studies of postmenopausal women supporting conclusions similar to those of the 2 selected economic evaluations and studies of men possibly reflecting the unknown efficacy of osteoporosis medications in men.

Older Women:

- DXA scanning dominated both no screening and screening with clinical risk factors alone in women at age  $\geq 70$  years and was a cost-effective alternative, in terms of QALYs gained, to no screening or screening with clinical risk factors alone in women at age 60 to 70 years (Mueller and Gandjour, 2009). (This study was conducted in Germany.)
- In women age 65 years, DXA scanning and immediate treatment for 5 years was more expensive and had less utility (QALYs gained) than no screening if the treatment were HRT, resulted in a cost-utility ratio of \$446,315/QALY (2002 USD) compared with no screening if the treatment were raloxifene, and resulted in a cost-utility ratio of \$72,877/QALY (2002 USD) compared with no screening if the treatment were alendronate (Mobley et al., 2006). A DXA scan was assumed to cost \$140, an amount higher than that used in the 2 selected economic evaluations and higher than the costs reported in the Washington State Agency Utilization Data.
- In older women, the sequential use of QUS for prescreening followed by DXA was more effective in reducing fractures and resulted in lower costs when compared with DXA alone. Diagnosis using QUS alone was more expensive than DXA alone under most conditions (Kraemer et al., 2006).

Older Men:

- Assuming a cost-utility threshold of \$50,000/QALY (2004 USD), DXA scanning following by bisphosphonate for those with a diagnosis of osteoporosis had cost-utility under these conditions: men age  $\geq 65$  years with a self-reported prior clinical fracture and men age 80 to 85 years with or without a prior fracture. Assuming a cost-utility threshold of \$100,000/QALY, the same strategy would have cost-utility for men age  $\geq 70$  years with or without a prior fracture. The effectiveness estimate was based on a Bayesian meta-analysis incorporating prior knowledge of the antifracture effects of bisphosphonates in women (Schousboe et al., 2007).
- Compared with no screening, the cost-utility of DXA scanning was \$421,000/QALY (2006 USD) in men age  $\geq 70$  years. A strategy of prescreening with the Osteoporosis Self-Assessment Tool (OST) and then DXA had a cost-utility of \$86,500/QALY (Ito et al., 2009).

**Practice Guidelines**

Fourteen practice guidelines with relevant recommendations for the U.S. population were identified. Eleven guidelines addressed the screening and monitoring of osteoporosis in generally healthy populations. Additionally, 3 guidelines addressing screening and/or monitoring of BMD in patients with particular medical conditions were identified. [Appendix IV](#) presents the recommendations of each guideline.

**Additional Information Not Included in Appendix IV***American Association of Clinical Endocrinologists (AACE)*

In addition to the guidelines described in Appendix IV, the AACE has also published *Medical Guidelines for Clinical Practice for the Diagnosis and Treatment of Menopause*. These guidelines had no recommendations regarding BMD screening or monitoring (Goodman et al., 2011).

*American College of Rheumatology*

The American College of Rheumatology published a position statement on bone density measurement in 2001, which was reviewed in 2008 (American College of Rheumatology, 2008). The following were included in the position statement:

- The use of BMD testing for the diagnosis of osteoporosis or low bone mass is supported and is a critical element in assessment of fracture risk.
- The use of serial BMD testing to monitor treatment response or to monitor progression of osteoporosis or low bone mass near treatment thresholds is supported. The frequency of retesting may be as often as every 6 months, depending upon the rheumatologist's preference.
- The NOF guidelines on the use of BMD measurements for diagnosis and monitoring of osteoporosis are supported.

*International Society for Clinical Densitometry (ISCD)*

In 2013, the ISCD issued or confirmed the following positions regarding technical issues. These statements, which are consistent with the most recent USPSTF recommendations, were approved concerning the reference database for T-scores (ISCD, 2013):

- Use a uniform white (non-race-adjusted) female normative database for women of all ethnic groups.
- Use a uniform white (non-race-adjusted) female reference for men of all ethnic groups.
- Manufacturers should continue to use NHANES III data as the reference standard for femoral neck and total hip T-scores.
- Manufacturers should continue to use their own databases for the lumbar spine as the reference standard for T-scores.
- If local reference data are available they should be used to calculate only Z-scores but not T-scores.

The ISCD endorses the WHO T-score-based definition of osteoporosis and recommends vertebral and hip measurements in all patients, except for these circumstances, in which forearm BMD should be measured: hip and/or spine cannot be measured or interpreted, hyperparathyroidism, and very obese patients who are over the weight limit for the DXA table. The ISCD approves the use of any well-validated technique for fracture risk assessment and prefers the terms *low bone mass* or *low bone density* to *osteopenia*.

*United States Preventive Services Task Force (USPSTF)*

A 2011 update by the USPSTF of its 2002 guidelines on screening for osteoporosis was rated as being of good quality (Nelson et al., 2010b; USPSTF, 2011a; USPSTF, 2011b). The recommendations of the earlier guidelines were limited to osteoporosis screening in women 60 years of age and older. The update was undertaken to broaden the target population by including men and younger women. The USPSTF recommendations apply to older adults who do not have osteoporosis, an osteoporotic fracture, or

other indications for BMD measurement. The USPSTF currently recommends screening in women who are > 65 years of age and in younger women who have a 10-year fracture risk that is equal to or greater than that of a 65-year-old white woman who has no additional risk factors. This recommendation was given a grade of B, meaning that there is a high certainty that the net benefit is moderate or that there is moderate certainty that the net benefit is moderate to substantial. The FRAX tool was the preferred tool to estimate the 10-year fracture risk. Neither the recommendation nor the evidence review explained why a 65-year-old white woman was selected as the reference case for establishing a screening policy. USPSTF did not identify any studies that evaluated the effectiveness and harms from osteoporosis screening.

The current evidence was determined to be insufficient to assess the benefit and harm of osteoporosis screening in men. The USPSTF provided the following considerations for physicians regarding screening in men (USPSTF, 2011):

- BMD determination may potentially detect osteoporosis in a large number of men and prevent substantial burden of fractures and fracture-related illnesses in this group.
- The potential harms of osteoporosis screening are likely to be small.
- Routine osteoporosis screening is not common practice in men.
- The men most likely to benefit from screening would be those who have a 10-year fracture risk equal to or greater than that of a 65-year old woman who has no additional risk factors.
- 

*Guidelines Reviewed and Found Not to Have Recommendations Pertaining to Osteoporosis Screening with DXA*

Endocrine Society guidelines on *Endocrine and Nutritional Management of the Post-Bariatric Surgery Patient* point to the risk of osteoporosis following bariatric surgery due to nutritional deficiencies and metabolic imbalances such as hyperparathyroidism. Bone markers are mentioned as a means of monitoring bone health, but the guidelines make no recommendations concerning osteoporosis screening with DXA or any other technology (Heber et al., 2010). Guidelines on *Diagnosis and Management of Chronic Obstructive Pulmonary Disease (COPD)* produced by the Institute for Clinical Systems Improvement (ICSI) refer briefly to osteoporosis as a possible concomitant disorder in patients with COPD but do not include any recommendations regarding screening (ICSI, 2013).

## Selected Payer Policies

### Aetna

Aetna considers bone mass measurements using established techniques medically necessary for members who meet any of the following criteria:

- Individuals being monitored to assess response to or efficacy of osteoporosis drug therapy; or
- Individuals receiving (or expected to receive) glucocorticoid therapy equivalent to 5 mg of prednisone or greater, per day, for more than 3 months; or

- Individuals with celiac sprue; or
- Individuals with primary hyperparathyroidism; or
- Individuals with vertebral abnormalities as demonstrated by x-ray to be indicative of osteoporosis, osteopenia, or vertebral fracture; or
- Men
  - 50 years of age with specific risk factors for osteoporosis; or
  - With hypogonadism or receiving ADT; or
- Individuals with fragility fractures; or
- Screening of men > 70 years of age; or
- Screening of women who are estrogen-deficient; or
- Women on long-term Depo-Provera contraception injection therapy; or
- Women on long-term anticonvulsant therapy; or
- Women with hyperthyroidism

Aetna considers bone mass measurements experimental and investigational for all other indications.

Repeat bone mass measurements are usually not indicated more frequently than once every 2 years, except:

- For a confirmatory baseline bone mass measurement to permit monitoring of individuals in the future if the initial bone mass test was performed with a technique that is different from the proposed testing method; or
- For monitoring of individuals on long-term glucocorticoid (steroid) therapy or anticonvulsant therapy of more than 3 months' duration; or
- Monitoring of individuals with uncorrected primary hyperparathyroidism.

Aetna recognizes these technologies as established procedures for bone mass measurement of the axial or appendicular (peripheral) skeleton: DXA, QCT, radiographic absorptiometry (photodensitometry), single-energy x-ray absorptiometry (SXA), and ultrasound BMD studies.

See Bone Mass Measurements: [Aetna Clinical Policy Bulletin No. 0134](#).

### Centers for Medicare & Medicaid Services (CMS)

A [CMS National Coverage Determination \(NCD\) for Bone \(Mineral \) Density Studies \(150.3\)](#), which was issued in January 2007, documented the transfer of conditions for coverage of bone mass measurements to the CMS Manual System. A document on [Bone Mass Measures](#) in the Manual System states that effective January 1, 2007, bone mass measurement is covered for Medicare beneficiaries, subject to these conditions:

- Performed to identify bone mass, detect bone loss, or determine bone quality, and performed by either a bone densitometer (other than single-photon or dual-photon absorptiometry) or a bone monometer system that has been cleared or approved for marketing for bone mass measurement by the FDA.

NOTE: This policy would appear to cover not only DXA, but also QUS. However, see the following specific requirements that DXA be used for treatment monitoring. It is unclear whether QCT is

covered. No NCD was identified for DXA, QUS, or QCT on June 16, 2014 (search National Coverage Documents by keywords *dual x-ray absorptiometry*, *quantitative ultrasound*, *quantitative computed tomography* in National Coverage Determinations at: [CMS Advanced Search Database](#)).

- Performed for these indications:
  - Estrogen deficiency and clinical risk for osteoporosis, based on a woman's medical history and other findings.
  - Vertebral abnormalities as demonstrated by an x-ray to be indicative of osteoporosis, osteopenia, or vertebral fracture.
  - Receiving (or expected to receive) glucocorticoid (steroid) therapy equivalent to an average of 5.0 mg of prednisone, or greater, per day, for more than 3 months.
  - Primary hyperparathyroidism.
  - Confirmation of bone mass for beneficiaries who have 1 of the preceding indications to permit monitoring in the future *unless the initial measurement was performed by axial skeletal DXA* (e.g., confirmatory bone mass measurement is not covered if the initial measurement was by axial skeletal DXA).
  - Monitoring to assess the response to or efficacy of an FDA-approved osteoporosis drug therapy. *In this case, must be performed by axial skeletal DXA.*
- Performed every 2 years:
  - Medicare might pay when more frequent measurements are considered medically necessary. Examples include:
    - Monitoring beneficiaries on long-term glucocorticoid (steroid) therapy of more than 3 months.
    - Confirming baseline bone mass measurements to permit monitoring of beneficiaries in the future.

## GroupHealth

No coverage policy for screening for osteoporosis was identified on the GroupHealth website ([GroupHealth Providers](#)) on June 19, 2014 (search by keywords *osteoporosis* or *DXA* or *absorptiometry*).

## Oregon Health Evidence Review Commission (HERC)

The Oregon HERC has concluded that osteoporosis screening by DXA should be covered for women  $\geq 65$  years of age and for men or younger women whose 10-year risk of major osteoporotic fracture is  $\geq 9.3\%$ . HERC recommends the FRAX tool to determine fracture risk. Routine screening of men is not recommended. The frequency of monitoring should not be based upon DXA scores alone. Repeat testing should be covered only if the results will influence clinical management. Testing less frequently than every 2 years to monitor treatment is not recommended for coverage. HERC recommends coverage of repeat DXA screening according to the following plan, unless a patient's risk factors have significantly changed:

- Every 2 years for patients with osteoporosis or advanced osteopenia (T-score of –2 or lower)
- Every 4 years for patients with moderate osteopenia (T-score between –1.5 and –1.99)
- Every 10 years for patients with mild osteopenia (T-score between –1.01 and –1.49)
- Every 15 years for patients with normal bone density

Coverage guidance decisions by HERC are intended to guide public and private purchasers in Oregon in making informed decisions about healthcare services.

See Osteoporosis Screening and Monitoring by Dual-Energy X-Ray Absorptiometry: [HERC Coverage Guidances](#).

### **Regence**

Regence considers screening for vertebral fractures using DXA as a stand-alone procedure or in addition to standard BMD studies investigational.

See Screening for Vertebral Fracture with Dual X-Ray Absorptiometry (DXA): [Regence Medical Policy No. 48](#).



## References

- Alibhai SMH, Mohamedali HZ, Gulamhusein H, et al. Changes in bone mineral density in men starting androgen deprivation therapy and the protective role of vitamin D. *Osteoporos Int*. 2013;24(10):2571-2579. PMID: [23563932](#).
- American College of Obstetricians and Gynecologists (ACOG). ACOG Practice Bulletin N. 129. Osteoporosis. *Obstet Gynecol*. 2012;120(3):718-734. PMID: [22914492](#).
- American College of Radiology (ACR). ACR Appropriateness Criteria: Osteoporosis and Bone Mineral Density. Reviewed 2010. Available at: <http://www.acr.org/~media/289DF665A5D448ADA04F1081172ADF9A.pdf>. Accessed August 20, 2014.
- American College of Rheumatology. American College of Rheumatology Position Statement: Bone Density Measurement. 2008. Available at: <https://www.rheumatology.org/ACR/practice/clinical/position/bonedensitymeasurement.pdf>. Accessed August 20, 2014.
- Appraisal of Guidelines Research & Evaluation (AGREE) Enterprise. Appraisal of Guidelines for Research and Evaluation II Instrument. Updated September 2013. Available at: [http://www.agreetrust.org/wp-content/uploads/2013/10/AGREE-II-Users-Manual-and-23-item-Instrument\\_2009\\_UPDATE\\_2013.pdf](http://www.agreetrust.org/wp-content/uploads/2013/10/AGREE-II-Users-Manual-and-23-item-Instrument_2009_UPDATE_2013.pdf). Accessed August 20, 2014.
- Astrand J, Nilsson J, Thorngren KG. Screening for osteoporosis reduced new fracture incidence by almost half: a 6-year follow-up of 592 fracture patients from an osteoporosis screening program. *Acta Orthop*. 2012;83(6):661-665. PMCID: [PMC3555451](#).
- Baim S, Wilson CR, Lewiecki EM, Luckey MM, Downs RW Jr, Lentle BC. Precision assessment and radiation safety for dual-energy X-ray absorptiometry: position paper of the International Society for Clinical Densitometry. *J Clin Densitom*. 2005;8(4):371-378. PMID: [16311420](#).
- Balasubramanian A, Tosi LL, Lane JM, Dirschl DR, Ho PR, O'Malley CD. Declining rates of osteoporosis management following fragility fractures in the U.S., 2000 through 2009. *J Bone Joint Surg Am*. 2014;96(7):e52. PMID: [24695929](#).
- Barr RJ, Stewart A, Torgerson DJ, Reid DM. Population screening for osteoporosis risk: a randomised control trial of medication use and fracture risk. *Osteoporos Int*. 2010;21(4):561-568. PMID: [19565176](#).
- Bauer DC, Schwartz A, Palermo L, et al. Fracture prediction after discontinuation of 4 to 5 years of alendronate therapy: the FLEX Study. *JAMA Intern Med*. 2014;174(7):1126-1134. PMID: [24798675](#).
- Becker T, Lipscombe L, Narod S, Simmons C, Anderson GM, Rochon PA. Systematic review of bone health in older women treated with aromatase inhibitors for early-stage breast cancer. *J Am Geriatr Soc*. 2012;60(9):1761-1767. PMID: [22985145](#).
- Bell KJ, Hayen A, Macaskill P, et al. Value of routine monitoring of bone mineral density after starting bisphosphonate treatment: secondary analysis of trial data. *BMJ*. 2009;338:b2266. PMCID: [PMC3272655](#).

- Bolland MJ, Grey A, Horne AM, et al. Stable bone mineral density over 6 years in HIV-infected men treated with highly active antiretroviral therapy (HAART). *Clin Endocrinol (Oxf)*. 2012;76(5):643-648. PMID: [22040002](#).
- Bruyère O, Detilleux J, Chines A, Reginster JY. Relationships between changes in bone mineral density or bone turnover markers and vertebral fracture incidence in patients treated with bazedoxifene. *Calcif Tissue Int*. 2012;91(4):244-249. PMID: [22945761](#).
- Canadian Agency for Drugs and Technologies in Health (CADTH). *T-Score as a Measurement to Assess the 10-Year Fracture Risk: A Review of the Clinical Evidence*. Ottawa, Ontario: Canadian Agency for Drugs and Technologies in Health; 2011a. Available at: [http://www.cadth.ca/media/pdf/htis/aug-2011/RC0294\\_T-score\\_final.pdf](http://www.cadth.ca/media/pdf/htis/aug-2011/RC0294_T-score_final.pdf). Accessed August 20, 2014.
- Canadian Agency for Drugs and Technologies in Health (CADTH). *Bisphosphonates for the Prevention of Osteoporosis in Patients Treated with Systematic Corticosteroids: A Review of the Clinical Evidence and Guidelines*. Ottawa, Ontario: Canadian Agency for Drugs and Technologies in Health; 2011b. Available at: [http://www.cadth.ca/media/pdf/htis/aug-2011/RC0293\\_bispho\\_final.pdf](http://www.cadth.ca/media/pdf/htis/aug-2011/RC0293_bispho_final.pdf). Accessed August 20, 2014.
- Centers for Disease Control and Prevention (CDC). FastStats: Osteoporosis. Updated November 21, 2013. Available at: <http://www.cdc.gov/nchs/fastats/osteoporosis.htm>. Accessed August 20, 2014.
- Charmetant C, Phaner V, Condemine A, Calmels P. Diagnosis and treatment of osteoporosis in spinal cord injury patients: a literature review [correction appears in *Ann Phys Rehabil Med*. 2011;54(4):271]. *Ann Phys Rehabil Med*. 2010;53(10):655-668. PMID: [21094110](#).
- Chen P, Miller PD, Delmas PD, Misurski DA, Kregg JH. Change in lumbar spine BMD and vertebral fracture risk reduction in teriparatide-treated postmenopausal women with osteoporosis. *J Bone Miner Res Off J Am Soc Bone Miner Res*. 2006;21(11):1785-1790. PMID: [23548220](#).
- Ciaschini PM, Straus SE, Dolovich LR, et al. Community based intervention to optimize osteoporosis management: randomized controlled trial. *BMC Geriatr*. 2010;10:60. PMCID: [PMC2940796](#).
- Conwell LS, Chang AB. Bisphosphonates for osteoporosis in people with cystic fibrosis. *Cochrane Database Syst Rev*. 2014;3:CD002010. PMID: [24627308](#).
- Crandall CJ, Newberry SJ, Diamant A, et al. *Treatment To Prevent Fractures in Men and Women With Low Bone Density or Osteoporosis: Update of a 2007 Report*. Comparative Effectiveness Review No. 53. (Prepared by Southern California Evidence-based Practice Center under Contract No. HHS-290-2007-10062-I.) Rockville, MD: Agency for Healthcare Research and Quality; 2012. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK92566/>. Accessed August 20, 2014.
- Cummings SR, Black DM, Thompson DE, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998;280(24):2077-2082. PMID: [9875874](#).
- Doheny MO, Sedlak CA, Estok PJ, Zeller RA. Bone density, health beliefs, and osteoporosis preventing behaviors in men. *Orthop Nurs*. 2011;30(4):266-272. PMCID: [PMC3146740](#).
- Dohle G, Arver S, Bettocchi S, Kliesch S, Punab M, de Ronde W. Male Hypogonadism. February 2012. European Association of Urology [website]. Available at: [http://www.uroweb.org/gls/pdf/18%20Male%20Hypogonadism\\_LR.pdf](http://www.uroweb.org/gls/pdf/18%20Male%20Hypogonadism_LR.pdf). Accessed August 20, 2014.

- Espallargues M, Sampietro-Colom L, Estrada MD, et al. Identifying bone-mass-related risk factors for fracture to guide bone densitometry measurements: a systematic review of the literature. *Osteoporos Int*. 2001;12(10):811-822. PMID: [11716183](#).
- Feng Z, Zeng S, Wang Y, Zheng Z, Chen Z. Bisphosphonates for the prevention and treatment of osteoporosis in patients with rheumatic diseases: a systematic review and meta-analysis. *PloS One*. 2013;8(12):e80890. PMID: [24324644](#).
- Florence R, Allen S, Benedict L, et al. Health Care Guideline: Diagnosis and Treatment of Osteoporosis. 2013. Institute for Clinical Systems Improvement [website]. Available at: [https://www.icsi.org/\\_asset/vnw0c3/Osteo.pdf](https://www.icsi.org/_asset/vnw0c3/Osteo.pdf). Accessed August 20, 2014.
- Fouda MA, Khan AA, Sultan MS, Rios LP, McAssey K, Armstrong D. Evaluation and management of skeletal health in celiac disease: position statement. *Can J Gastroenterol*. 2012;26(11):819-829. PMID: [23166906](#).
- Fraser LA, Vogt KN, Adachi JD, Thabane L. Fracture risk associated with continuation versus discontinuation of bisphosphonates after 5 years of therapy in patients with primary osteoporosis: a systematic review and meta-analysis. *Ther Clin Risk Manag*. 2011;7:157-166. PMID: [21691586](#).
- Frost SA, Nguyen ND, Center JR, Eisman JA, Nguyen TV. Timing of repeat BMD measurements: development of an absolute risk-based prognostic model. *J Bone Miner Res*. 2009;24(11):1800-1807. PMID: [19419321](#).
- Gedmintas L, Solomon DH, Kim SC. Bisphosphonates and risk of subtrochanteric, femoral shaft, and atypical femur fracture: a systematic review and meta-analysis. *J Bone Miner Res*. 2013;28(8):1729-1737. PMID: [23408697](#).
- Goodman NF, Cobin RH, Ginzburg SB, Katz IA, Woode DE; American Association of Clinical Endocrinologists. American Association of Clinical Endocrinologists Medical Guidelines for Clinical Practice for the diagnosis and treatment of menopause. *Endocr Pract*. 2011;(17 Suppl 6):1-25. PMID: [22193047](#).
- Gordis L. *Epidemiology*. 2<sup>nd</sup> ed. Philadelphia, PA: W. B. Saunders Company, 2000.
- Gourlay ML, Fine JP, Preisser JS, et al.; Study of Osteoporotic Fractures Research Group. Bone-density testing interval and transition to osteoporosis in older women. *N Engl J Med*. 2012;366(3):225-233. PMID: [PMC3285114](#).
- Graat-Verboom L, Wouters EF, Smeenk FW, van den Borne BE, Lunde R, Spruit MA. Current status of research on osteoporosis in COPD: a systematic review. *Eur Respir J*. 2009;34(1):209-218. PMID: [19567604](#).
- Grossman JM, Gordon R, Ranganath VK, et al. American College of Rheumatology 2010 recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis[correction appears in *Arthritis Care Res (Hoboken)*. 2012;64(3):464]. *Arthritis Care Res (Hoboken)*. 2010;62(11):1515-1526. PMID: [20662044](#).
- Guo Z, Wu R, Gong J, et al. The efficacy and safety of bisphosphonates for osteoporosis or osteopenia in Crohn's disease: a meta-analysis. *Dig Dis Sci*. 2013;58(4):915-922. PMID: [23179146](#).
- Hayes, Inc. Hayes Medical Technology Directory. *Quantitative Computed Tomography for Bone Mineral Density*. Lansdale, PA: Hayes, Inc.; December 8, 2010.

- Heber D, Greenway FL, Kaplan LM, et al.; Endocrine Society. Endocrine and nutritional management of the post-bariatric surgery patient: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab.* 2010;95(11):4823-4843.
- Hillier TA, Stone KL, Bauer DC, et al. Evaluating the value of repeat bone mineral density measurement and prediction of fractures in older women: the study of osteoporotic fractures. *Arch Intern Med.* 2007;167(2):155-160. PMID: [17242316](#).
- Institute for Clinical Systems Improvement (ICSI). Chronic Obstructive Pulmonary Disease (COPD), Diagnosis and Management of. Ninth Edition. Revised March 2013. Available at: [https://www.icsi.org/guidelines\\_more/catalog\\_guidelines\\_and\\_more/catalog\\_guidelines/catalog\\_respiratory\\_guidelines/copd/](https://www.icsi.org/guidelines_more/catalog_guidelines_and_more/catalog_guidelines/catalog_respiratory_guidelines/copd/). Accessed July 22, 2014.
- International Society for Clinical Densitometry (ISCD). 5th ISCD Position Development Conference (Adult). March 2013. Available at: <http://www.iscd.org/official-positions/5th-iscd-position-development-conference-adult/>. Accessed August 20, 2014.
- Ito K, Hollenberg JP, Charlson ME. Using the osteoporosis self-assessment tool for referring older men for bone densitometry: a decision analysis. *J Am Geriatr Soc.* 2009;57(2):218-224. PMID: [19207137](#).
- Jiménez-Núñez FG, Manrique-Arija S, Ureña-Garnica I, et al. Reducing the need for central dual-energy X-ray absorptiometry in postmenopausal women: efficacy of a clinical algorithm including peripheral densitometry. *Calcif Tissue Int.* 2013;93(1):62-68. PMID: [23608922](#).
- Kern LM, Powe NR, Levine MA, et al. Association between screening for osteoporosis and the incidence of hip fracture. *Ann Intern Med.* 2005;142(3):173-181. PMID: [15684205](#).
- Khan N, Abbas AM, Almkhatar RM, Cole EB, Khan AN. Adherence and efficacy of screening for low bone mineral density among ulcerative colitis patients treated with corticosteroids. *Am J Gastroenterol.* 2014;109(4):572-578. PMID: [24469613](#).
- Kingkaew P, Maleewong U, Ngarmukos C, Teerawattananon Y. Evidence to inform decision makers in Thailand: a cost-effectiveness analysis of screening and treatment strategies for postmenopausal osteoporosis. *Value Health.* 2012;15(1 Suppl):S20-28. PMID: [22265062](#).
- Kornbluth A, Sachar DB, Practice Parameters Committee of the American College of Gastroenterology. Ulcerative colitis practice guidelines in adults: American College Of Gastroenterology, Practice Parameters Committee [correction appears in *Am J Gastroenterol.* 2010;105(3):500]. *Am J Gastroenterol.* 2010;105(3):501-523. PMID: [20068560](#).
- Kraemer DF, Nelson HD, Bauer DC, Helfand M. Economic comparison of diagnostic approaches for evaluating osteoporosis in older women. *Osteoporos Int.* 2006;17(1):68-76. PMID: [15889313](#).
- Lane NE. Epidemiology, etiology, and diagnosis of osteoporosis. *Am J Obstet Gynecol.* 2006;194(2 Suppl):S3-11. PMID: [16448873](#).
- Lee SH, Chan RC, Chang SS, et al. Use of bisphosphonates and the risk of osteonecrosis among cancer patients: a systemic review and meta-analysis of the observational studies. *Support Care Cancer.* 2014b;22(2):553-560. PMID: [24203085](#).
- Lee SH, Chang SS, Lee M, Chan RC, Lee CC. Risk of osteonecrosis in patients taking bisphosphonates for prevention of osteoporosis: a systematic review and meta-analysis. *Osteoporos Int.* 2014a;25(3):1131-1139. PMID: [24343364](#).
- Leslie WD, Lix LM, Wu X; Manitoba Bone Density Program. Competing mortality and fracture risk assessment. *Osteoporos Int.* 2013;24(2):681-688. PMID: [22736068](#).

- Lim LS, Hoeksema LJ, Sherin K; ACPM Prevention Practice Committee. Screening for osteoporosis in the adult U.S. population: ACPM position statement on preventive practice. *Am J Prev Med*. 2009;36(4):366-375. PMID: [19285200](#).
- Lim SY, Lim JH, Nguyen D, et al. Screening for osteoporosis in men aged 70 years and older in a primary care setting in the United States. *Am J Mens Health*. 2013;7(4):350-354. PMID: [23440083](#).
- Lo JC, Burnett-Bowie S-AM, Finkelstein JS. Bone and the perimenopause. *Obstet Gynecol Clin North Am*. 2011;38(3):503-517. PMCID: [PMC3920744](#)
- Looker AC, Borrud LG, Dawson-Hughes B, Shepherd JA, Wright NC. *Osteoporosis or Low Bone Mass at the Femur Neck or Lumbar Spine in Older Adults: United States, 2005–2008*. NCHS Data Brief No. 93. Hyattsville, MD: National Center for Health Statistics; 2012. Available at: <http://www.cdc.gov/nchs/data/databriefs/db93.pdf>. Accessed August 20, 2014.
- Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ*. 1996;312(7041):1254-1259. PMID: [8634613](#).
- McLeod KM, Johnson CS. Identifying women with low bone mass: a systematic review of screening tools. *Geriatr Nurs*. 2009;30(3):164-173. PMID: [19520227](#).
- Mobley LR, Hoerger TJ, Wittenborn JS, Galuska DA, Rao JK. Cost-effectiveness of osteoporosis screening and treatment with hormone replacement therapy, raloxifene, or alendronate. *Med Decis Making*. 2006;26(2):194-206. PMID: [16525173](#).
- Mueller D, Gandjour A. Cost-effectiveness of using clinical risk factors with and without DXA for osteoporosis screening in postmenopausal women. *Value Health*. 2009;12(8):1106-1117. PMID: [19706151](#).
- National Osteoporosis Foundation (NOF). *Clinician's Guide to Prevention and Treatment of Osteoporosis*. Washington, DC: National Osteoporosis Foundation; 2014. Available at: [http://redireference.com/support/NOF/Site/NOF\\_Guide\\_Support.html](http://redireference.com/support/NOF/Site/NOF_Guide_Support.html). Accessed August 20, 2014.
- Nayak S, Roberts MS, Greenspan SL. Cost-effectiveness of different screening strategies for osteoporosis in postmenopausal women. *Ann Intern Med*. 2011;155(11):751-761. PMCID: [PMC3318923](#).
- Nelson HD, Haney EM, Chou R, Dana T, Fu R, Bougatsos C. *Screening for Osteoporosis: Systematic Review to Update the 2002 U.S. Preventive Services Task Force Recommendation*. Rockville, MD: Agency for Healthcare Research and Quality; 2010a. AHRQ Publication No. 10-05145-EF-1. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK45201/>. Accessed August 20, 2014.
- Nelson HD, Haney EM, Dana T, Bougatsos C, Chou R. Screening for osteoporosis: an update for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2010b;153(2):99-111. PMID: [20621892](#).
- Nelson HD, Helfand M, Woolf SH, Allan JD. Screening for postmenopausal osteoporosis: a review of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2002;137(6):529-541. PMID: [12230356](#).
- North American Menopause Society (NAMS). Management of osteoporosis in postmenopausal women: 2010 position statement of The North American Menopause Society. *Menopause N Y N*. 2010;17(1):25-54; quiz 55-56. PMID: [20061894](#).
- Nshimyumukiza L, Durand A, Gagnon M, et al. An economic evaluation: simulation of the cost-effectiveness and cost-utility of universal prevention strategies against osteoporosis-related fractures. *J Bone Miner Res*. 2013;28(2):383-394. PMCID: [PMC3580046](#).



- Oh YH, Yoon C, Park SM. Bisphosphonate use and gastrointestinal tract cancer risk: meta-analysis of observational studies. *World J Gastroenterol*. 2012;18(40):5779-5788. PMCID: [PMC3484348](#).
- Paccou J, Zeboulon N, Combescure C, Gossec L, Cortet B. The prevalence of osteoporosis, osteopenia, and fractures among adults with cystic fibrosis: a systematic literature review with meta-analysis. *Calcif Tissue Int*. 2010;86(1):1-7. PMID: [19949942](#).
- Panichkul S, Panichkul P, Sritara C, Tamdee D. Cost-effectiveness analysis of various screening methods for osteoporosis in perimenopausal Thai women. *Gynecol Obstet Invest*. 2006;62(2):89-96. PMID: [16636570](#).
- Pisani P, Renna MD, Conversano F, et al. Screening and early diagnosis of osteoporosis through X-ray and ultrasound based techniques. *World J Radiol*. 2013;5(11):398-410. PMCID: [PMC3856332](#).
- Qaseem A, Snow V, Shekelle P, et al.; Clinical Efficacy Assessment Subcommittee of the American College of Physicians. Screening for osteoporosis in men: a clinical practice guideline from the American College of Physicians [correction appears in *Ann Intern Med*. 2008;148(11):888]. *Ann Intern Med*. 2008;148(9):680-684. PMID: [18458281](#).
- Rud B, Hilden J, Hyldstrup L, Hróbjartsson A. The Osteoporosis Self-Assessment Tool versus alternative tests for selecting postmenopausal women for bone mineral density assessment: a comparative systematic review of accuracy. *Osteoporos Int*. 2009;20(4):599-607. PMID: [18716823](#).
- Rudic JS, Giljaca V, Krstic MN, Bjelakovic G, Gluud C. Bisphosphonates for osteoporosis in primary biliary cirrhosis. *Cochrane Database Syst Rev*. 2011;(12):CD009144. PMID: [22161446](#).
- Sawka AM, Papaioannou A, Josse RG, et al.; CaMos Research Group. What is the number of older Canadians needed to screen by measurement of bone density to detect an undiagnosed case of osteoporosis? a population-based study from CaMos. *J Clin Densitom*. 2006;9(4):413-418. PMID: [17097526](#).
- Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7:10.
- Schousboe JT, Taylor BC, Fink HA, et al. Cost-effectiveness of bone densitometry followed by treatment of osteoporosis in older men. *JAMA*. 2007;298(6):629-637. PMID: [17684185](#).
- Schwarz P, Jorgensen NR, Mosekilde L, Vestergaard P. The evidence for efficacy of osteoporosis treatment in men with primary osteoporosis: a systematic review and meta-analysis of antiresorptive and anabolic treatment in men. *J Osteoporos*. 2011;2011:259818. PMCID: [PMC3138068](#).
- Sedlak CA, Doheny MO, Estok PJ, Zeller RA, Winchell J. DXA, health beliefs, and osteoporosis prevention behaviors. *J Aging Health*. 2007;19(5):742-756. PMID: [17827447](#).
- Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. *Evid Policy*. 2010;6(1):51-59.
- Silverman SL, Calderon AD. The Utility and Limitations of FRAX: A US Perspective. *Curr Osteoporos Rep*. 2010;8(4):192-197. PMCID: [PMC2947011](#)
- Stone KL, Seeley DG, Lui LY, et al.; Osteoporotic Fractures Research Group. BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. *J Bone Miner Res*. 2003;18(11):1947-1954. PMID: [14606506](#).

- Targownik LE, Bernstein CN, Nugent Z, et al. Inflammatory bowel disease and the risk of fracture after controlling for FRAX. *J Bone Miner Res*. 2013;28(5):1007-1013. PMID: [23239264](#).
- United States Preventive Services Task Force (USPSTF). Screening for Osteoporosis: Recommendation Statement. January 2011a. Available at: <http://www.uspreventiveservicestaskforce.org/uspstf10/osteoporosis/osteors.htm>. Accessed August 20, 2014.
- United States Preventive Services Task Force (USPSTF). Screening for osteoporosis: U.S. preventive services task force recommendation statement. *Ann Intern Med*. 2011b;154(5):356-364. PMID: [21242341](#).
- Warriner AH, Saag KG. Osteoporosis diagnosis and medical treatment. *Orthop Clin North Am*. 2013;44(2):125-135. PMID: [23544819](#).
- Watts NB, Bilezikian JP, Camacho PM, et al. American Association of Clinical Endocrinologists Medical Guidelines for Clinical Practice for the diagnosis and treatment of postmenopausal osteoporosis. *Endocr Pract*. 2010;16 Suppl 3:1-37. PMID: [21224201](#).
- Watts NB, Adler RA, Bilezikian JP, et al.; Endocrine Society. Osteoporosis in men: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2012;97(6):1802-1822. PMID: [22675062](#).
- Wagh EJ, Lam MA, Hawker GA, et al.; Perimenopause BMD Guidelines Subcommittee of Osteoporosis Canada. Risk factors for low bone mass in healthy 40-60 year old women: a systematic review of the literature. *Osteoporos Int*. 2009;20(1):1-21. PMID: [18523710](#).
- World Health Organization (WHO) Collaborative Centre for Metabolic Bone Diseases, University of Sheffield, UK. FRAX WHO Fracture Risk Assessment Tool. 2014. Available at: <http://www.shf.ac.uk/FRAX/>. Accessed August 20, 2014.
- Zhumkhawala AA, Gleason JM, Cheetham TC, et al. Osteoporosis management program decreases incidence of hip fracture in patients with prostate cancer receiving androgen deprivation therapy. *Urology*. 2013;81(5):1010-1015. PMID: [23490521](#).

## Washington State Agency Utilization Data

Figure 1. All Agency Dual-energy X-ray absorptiometry (DXA) Bone Mineral Density (BMD) Tests, 2011-2013

Agency/Year Public Employee Benefits (PEBB), Uniform Med Plan (UMP)	2011	2012	2013	3 Yr Overall Total**	Avg Annual % Change	
Average Annual Members	212,596	212,684	222,339		2.3%	
Osteoporosis Member Counts	6032	5601	5604	18948	-5.7%	*
DXA BMD Patients	5933	5102	4658	14058	-13.4%	*
DXA BMD Tests	6067	5242	4799	16108	-13.0%	*
Average DXA Encounters per Patient	1.0	1.0	1.0	1.1	0.4%	
Non DXA Tests (not included in totals)‡	74	83	78	235	1.0%	*
<b>PEBB/UMP Total Paid, All DXA Tests</b>	<b>\$636,180</b>	<b>\$535,862</b>	<b>\$497,900</b>	<b>\$1,669,942</b>	<b>-11.4%</b>	*
Average Paid per DXA Procedure	\$105	\$102	\$104	\$104	-0.5%	
Average Paid/DXA ,Non-Medicare (% of tests)†	\$121 (86%)	\$123 (82%)	\$129 (80%)	\$124 (83%)	3.3%	
<b>Medicaid, Fee For Service (FFS) and Managed Care</b>						
Average Annual Clients, <b>FFS</b>	473,356	477,727	442,698		-3.2%	
Average Annual Clients, <b>Managed Care</b>	695,591	730,250	800,096		7.3%	
Osteoporosis Client Counts, <b>FFS</b>	1174	994	708	2876	-19.6%	*
Osteoporosis Client Counts, <b>Managed Care</b>	339	351	651	1662	34.0%	*
DXA BMD Patients <b>FFS</b>	2696	2033	1136	5582	-32.7%	*
DXA BMD Patients <b>Managed Care</b>	573	814	1655	2951	60.6%	*
DXA BMD Tests <b>FFS</b>	2828	2143	1175	6146	-32.3%	*
DXA BMD Tests <b>Managed Care</b>	595	851	1726	3172	66.7%	*
Average DXA Encounters per Patient (overall)	1.0	1.1	1.0	1.1		
Non DXA Tests (not included in totals)‡	28	23	15	113	-27.8%	*
<b>Medicaid Total Paid , All DXA Tests (FFS Only)</b>	<b>\$171,836</b>	<b>\$130,550</b>	<b>\$62,768</b>	<b>\$365,154</b>	<b>-36.7%</b>	*
Average Paid per Procedure ( <b>FFS</b> only)	\$61	\$61	\$53	\$59	-7.3%	

\*Population adjusted average % change

\*\*3 year total patient counts represent unique patients over 3 years, and may not equal the sum of annual counts.

†PEBB/UMP Non-medicare “allowed amounts” are more representative of test cost. Annual averages: \$158 (2011) to \$172 (2013), \$163 over 3 yrs.



‡Non DXA tests account for less than 2% of bone mineral density tests for PEBB members and less than 1% for Medicaid clients. Other scan types are mainly CT and ultrasound, but use is steady or declining, with 50-70 ultrasounds/year among PEB members the most prevalent of these.

**Note:** L&I BMD scans: During 2011-2013, L&I paid for 286 DXA scans on 164 separate claims. Claims averaged around \$66 per scan, for a total of \$18,000.

**Figure 2: Agency Fee Schedules**

Current pricing by DXA Scan type as available on Agency web sites:

DXA CPT Codes	CPT Code Descriptions	Current Agency Fees		
		PEBB/UMP*	L&I†	Medicaid‡
<b>77080</b>	DXA BONE DENSITY, AXIAL	\$138.48	\$84.36	\$30.29
<b>77081</b>	DXA BONE DENSITY/PERIPHERAL	\$41.01	\$46.86	\$16.83

\*Regence Blue Shield Provider Fee Schedule – effective January 1 2013, MD/DO/DPM Provider rates, Maximum Allowable fee, [http://www.hca.wa.gov/ump/documents/Regence\\_Professional\\_Fee\\_Schedule\\_Jan\\_2013.pdf](http://www.hca.wa.gov/ump/documents/Regence_Professional_Fee_Schedule_Jan_2013.pdf), Accessed 10/13/2014. Payment based on the Regence Fee Schedule is subject to all of the terms and conditions of the applicable Regence BlueShield provider agreement, member benefits, Regence BlueShield policies, and all published Regence BlueShield administrative guidelines. Therefore, the appearance of fees for particular procedure codes does not guarantee coverage. Some providers may have contracted fees at different rates.

†Washington State Labor and Industries Fee Schedules and Payment Policies (MARFS), Fee Schedules and Payment Policies for: 2014, <http://www.lni.wa.gov/apps/FeeSchedules/>, Accessed 10/12/2014.

‡Washington State Medicaid Rates Development Fee Schedule, July 1, 2014 Physician and Related Services Fee Schedule (Updated October 1, 2014), <http://www.hca.wa.gov/medicaid/rbrvs/pages/index.aspx#P>, Accessed 10/12/2014.

Figure 3a. PEBB Osteoporosis Average Patients by Age, 2011-2013

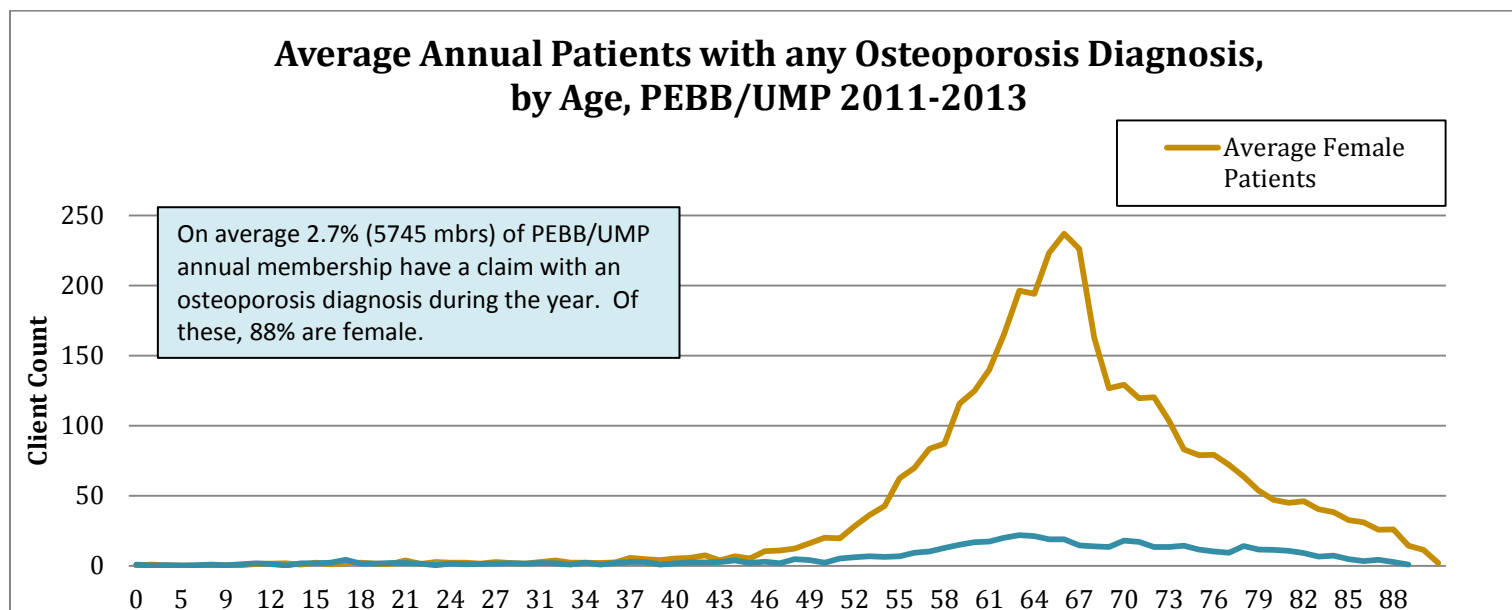


Figure 3b. Medicaid Osteoporosis Average Patients by Age, 2011-2013

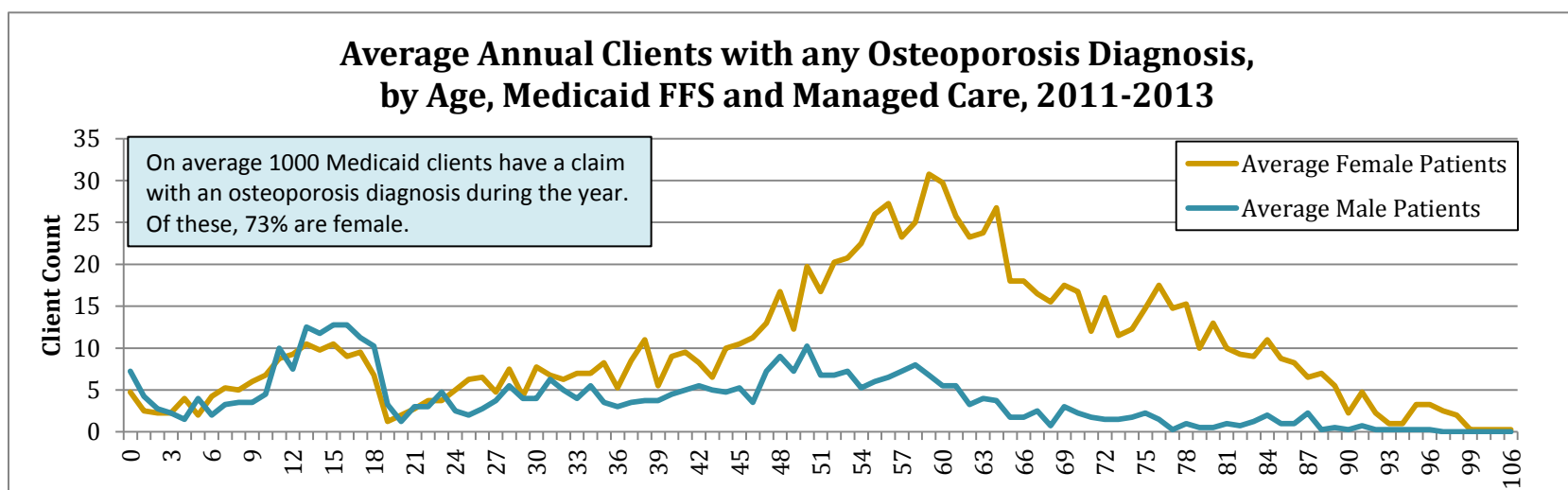


Figure 4a PEBB/UMP Osteoporosis Average Patients by Age, 2011-2013

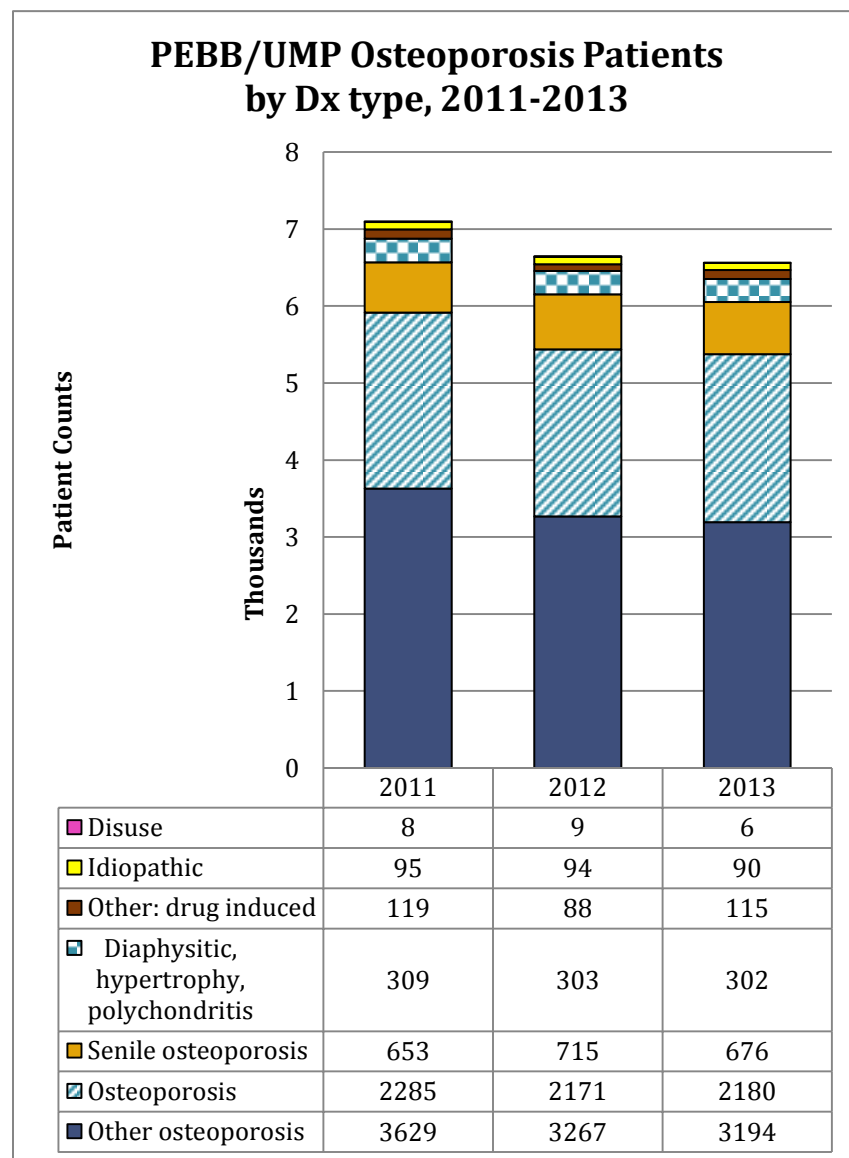
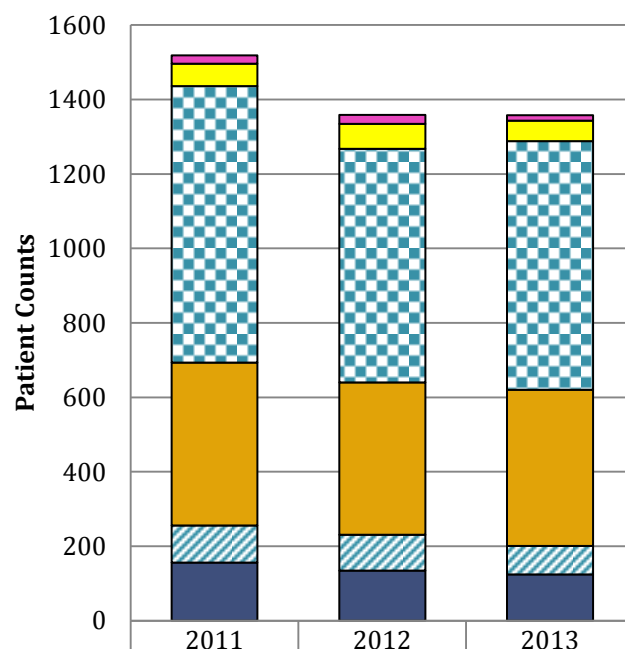


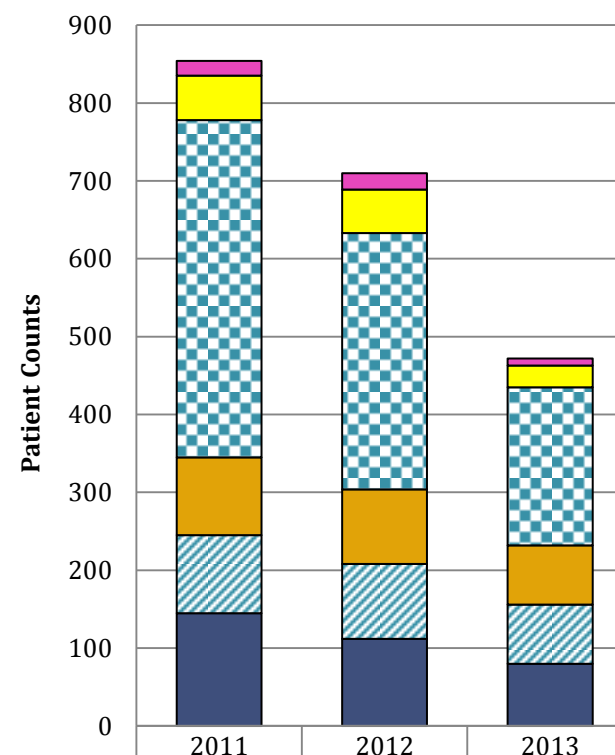
Figure 4b,c Medicaid Osteoporosis Average Patients by Age, 2011-2013

### Medicaid Total Osteoporosis Patients by Diagnosis type, 2011-2013



Disuse	23	24	15
Idiopathic	60	67	55
Diaphysitic, hypertrophy, polychondritis	743	628	668
Senile osteoporosis	437	409	419
Osteoporosis	100	96	77
Other osteoporosis	156	135	124

### Medicaid FFS Osteoporosis Patients by Diagnosis type, 2011-2013



Disuse	19	21	9
Idiopathic	57	56	28
Diaphysitic, hypertrophy, polychondritis	433	329	203
Senile osteoporosis	100	96	76
Osteoporosis	100	96	76
Other osteoporosis	145	112	80

Figure 5a, 5b PEBB/UMP vs Medicaid Progression from Screening DXA BMD to Osteoporosis Diagnosis, 2011-2013

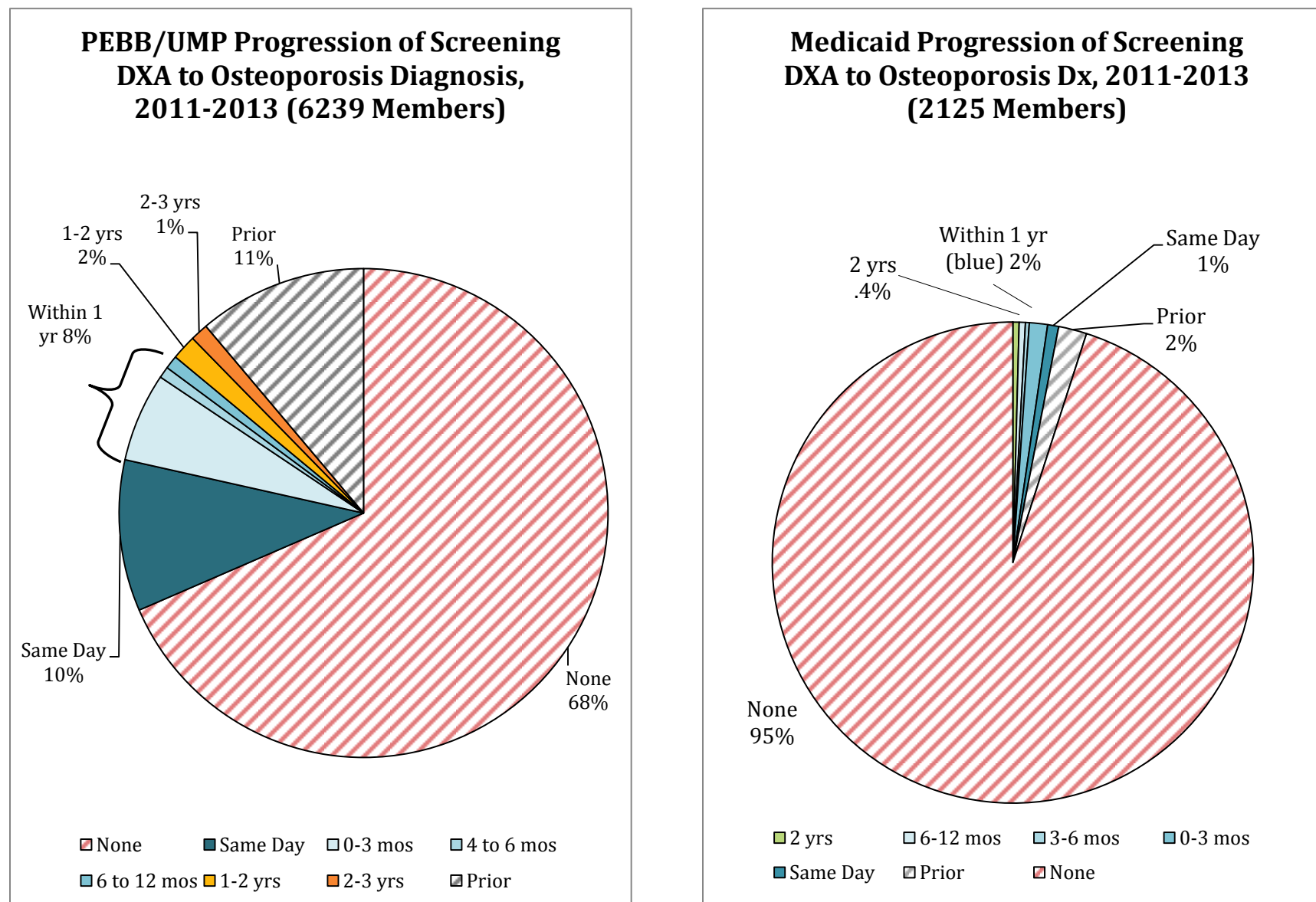
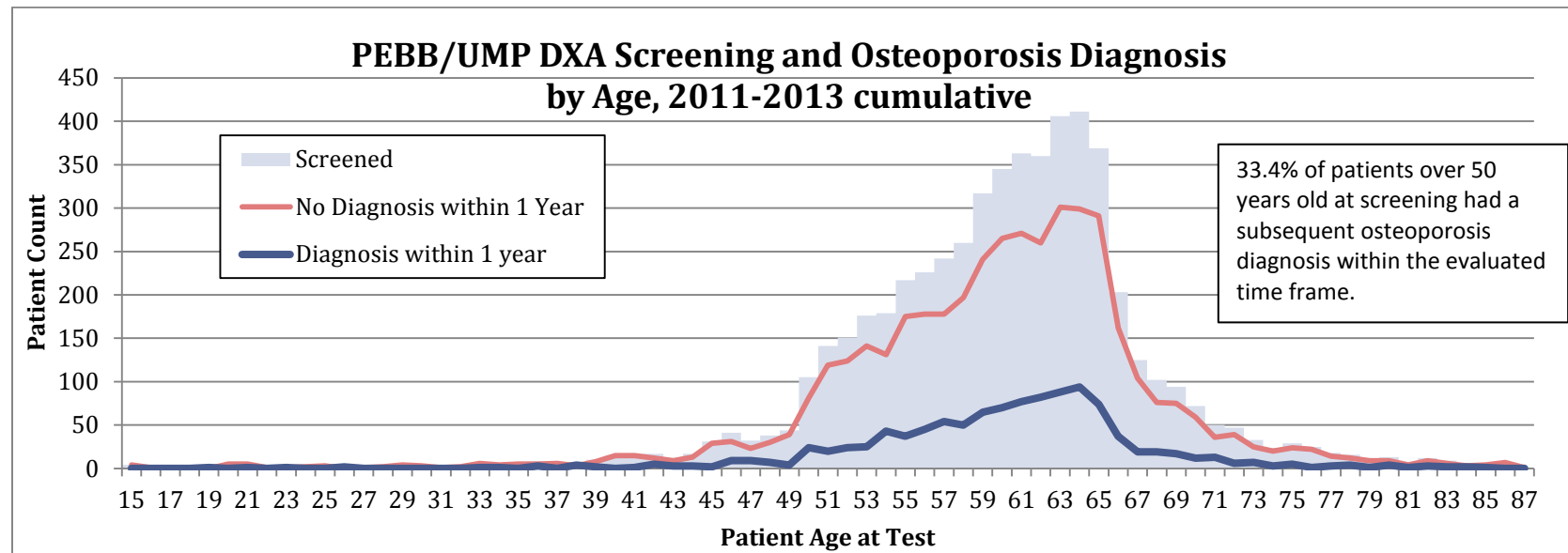
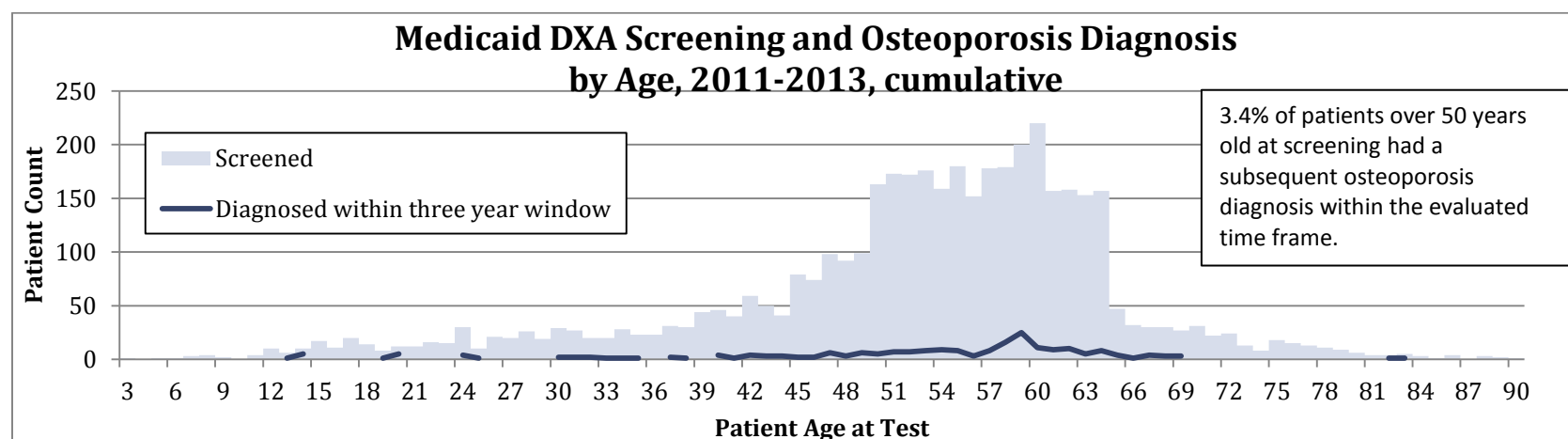


Figure 6a PEBB/UMP Age in Progression from DXA to Osteoporosis Diagnosis within 1 year, 2011-2013



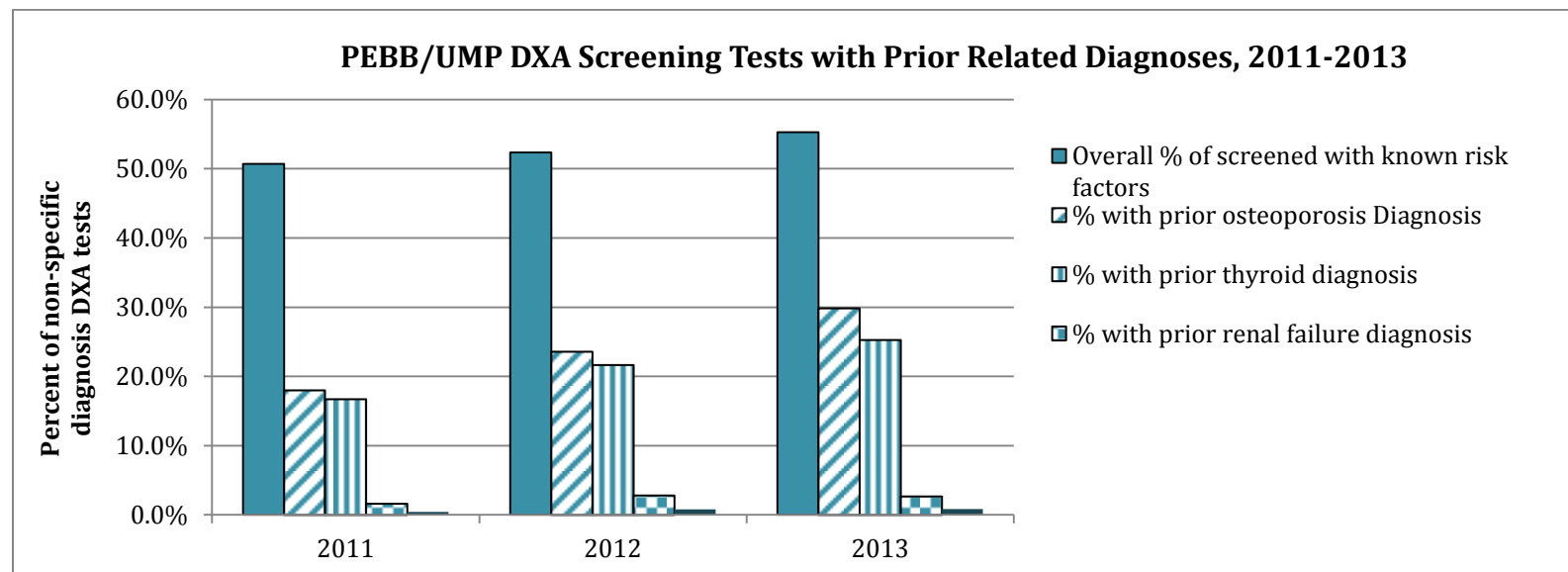
**Screening definition:** Non specific, non-symptomatic or screening diagnosis at BMD. Patients with a prior osteoporosis diagnosis are excluded from 6a &b.

Figure 6b Medicaid Age in Progression from DXA to Osteoporosis Diagnosis within 1 year, 2011-2013



**Note:** Medicaid chart shows client counts diagnosed within 3 years rather than one year due the low diagnosis rate.

Figure 6a. PEBB/UMP Risk Factors in Screening BMD tests, 2011-2013



**Screening definition:** Non specific, non-symptomatic or screening diagnosis for test

Figure 6b. Medicaid Risk Factors in Screening BMD tests, 2011-2013

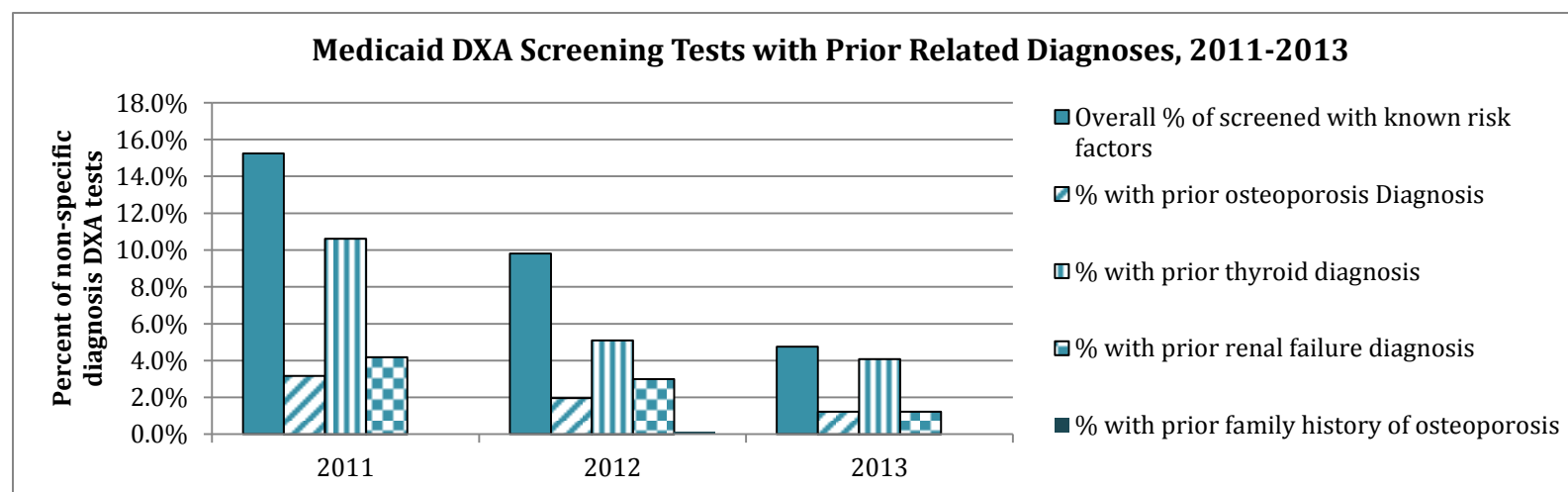


Figure 8a PEBB/UMP Repeated DXA Scans 2010-2013, All scans.

Count of DXA Scans per member in 4 years	2	3	4 or more
<b>Member count with repeated DXA scans in 4 years</b>	<b>3657</b>	<b>511</b>	<b>104</b>
% of Total Members (22030 Members in 4 years)	16.6%	2.3%	0.4%
Average Years from first scan to last repeat	1.8	2.5	3.0
Maximum (days) to repeat	1442	1438	1187
Median (days) to repeat	734	869	1124

Figure 8b. Medicaid Repeated DXA Scans 2010-201, All scans.

Count of DXA Scans per member in 4 years	2	3	4 or more
<b>Member count with repeated DXA scans in 4 years</b>	<b>1159</b>	<b>138</b>	<b>38</b>
% of Total Members (12871 Members in 4 years)	9.0%	1.1%	0.3%
Average Years from first scan to last repeat	1.6	2.5	2.8
Maximum (days) to repeat	1372	1337	1318
Median (days) to repeat	623	879	1095



Figure 9a. PEBB/UMP Repeated DXA Scans by age for members with no osteoporosis diagnosis in 4 years 2010-2013

PEBB/UMP Members with no osteoporosis diagnosis		Scan Count per member			
		1	2	3	4 or more
Age Group/Member count over 4 years		3106	659	49	9
	0-20	18	5	0	0
	21-34	33	6	0	0
	35-49	220	40	3	0
	50-64	2282	426	24	4
	65-84	544	173	19	5
	85+	9	9	3	0
Age Group/Average years from first scan to last repeat (where counts are sufficient)					
	35-49		2.6		
	50-64		1.8	2.5	
	65-84		1.8	2.3	

Figure 9b. Medicaid Repeated DXA Scans by age for members with no osteoporosis diagnosis in 4 years 2010-2013

Medicaid Clients with no osteoporosis diagnosis		Scan Count per member			
		1	2	3	4 or more
Age Group/Client count over 4 years		10613	944	91	31
	0-20	451	78	14	9
	21-34	654	65	7	2
	35-49	1833	167	15	6
	50-64	5639	575	52	14
	65-84	1725	53	3	0
	85+	311	6	0	0
Age Group/Average years from first scan to last repeat (where counts are sufficient)					
	0-20		1.3	1.9	
	21-34		1.6		
	35-49		1.6	2.7	
	50-64		1.6	2.6	2.8
	65-84		1.6		
	85+		1.6		

Figure 10a. PEBB/UMP Top Diagnoses for DXA Testing, Descending by Paid \$, 2011-2013

PEBB/UMP Top Diagnoses for DXA Testing by Paid \$ Descending (of 592 Diagnoses)					Grand Total	% of Total Paid	Trend
Dx Code	Dx Description	2011 \$636,180	2012 \$535,862	2013 \$497,900	\$1,669,942		-11.4%
<b>733.9</b>	DISORDER OF BONE AND CARTILAGE, UNSPECIFIED	\$176,907	\$159,332	\$154,118	\$490,357	29.4%	-6.6%
<b>V82.81</b>	SPECIAL SCREENING FOR OSTEOPOROSIS	\$154,632	\$129,592	\$113,930	\$398,154	23.8%	-14.1%
<b>733</b>	UNSPECIFIED OSTEOPOROSIS	\$58,253	\$53,868	\$61,160	\$173,281	10.4%	3.0%
<b>V49.81</b>	ASYMPT POSTMENO STATUS	\$43,537	\$35,106	\$29,492	\$108,135	6.5%	-17.7%
<b>V76.12</b>	SCREEN MAMMOGRAM NEC	\$42,144	\$26,640	\$27,542	\$96,326	5.8%	-16.7%
<b>627.2</b>	SYMPT FEM CLIMACT STATE	\$29,110	\$24,047	\$13,780	\$66,937	4.0%	-30.0%
<b>733.01</b>	SENILE OSTEOPOROSIS	\$11,958	\$14,219	\$6,854	\$33,031	2.0%	-16.4%
<b>627.9</b>	UNSP MENOPAUSAL/ POSTMENOPAUSAL DISORDER	\$14,796	\$8,794	\$8,226	\$31,816	1.9%	-23.5%
<b>V70.0</b>	ROUTINE GENL MEDICAL EXAM AT HEALTH CARE FACILITY	\$11,401	\$8,795	\$9,030	\$29,226	1.8%	-10.1%
<b>174.9</b>	MAL NEOPLASM OF BREAST (FEMALE), UNSPECIFIED SITE	\$4,457	\$4,283	\$1,629	\$10,369	0.6%	-32.9%
<b>V58.65</b>	LONG-TERM (CURRENT) USE OF STEROIDS	\$2,931	\$2,331	\$3,147	\$8,409	0.5%	7.3%
<b>733.09</b>	OTHER OSTEOPOROSIS	\$1,612	\$1,375	\$4,504	\$7,491	0.4%	106.4%
<b>V76.11</b>	SCREEN MAMMOGRAM HI RISK	\$3,623	\$2,747	\$962	\$7,332	0.4%	-44.6%
<b>781.91</b>	LOSS OF HEIGHT	\$3,311	\$1,787	\$1,758	\$6,856	0.4%	-23.8%
<b>252</b>	HYPERPARATHYROIDISM NOS	\$1,471	\$1,437	\$3,463	\$6,371	0.4%	69.3%
<b>V72.31</b>	ROUTINE GYN EXAMINATION	\$1,328	\$2,556	\$2,069	\$5,953	0.4%	36.7%
<b>627.4</b>	SYMPT STATE W ARTIF MENO	\$2,203	\$1,437	\$1,909	\$5,549	0.3%	-1.0%
<b>V58.69</b>	LONG-TERM USE MEDS NEC	\$1,983	\$1,417	\$2,149	\$5,549	0.3%	11.6%
<b>268.9</b>	UNSPECIFIED VITAMIN D DEFICIENCY	\$2,176	\$1,609	\$1,454	\$5,239	0.3%	-17.8%
<b>V10.3</b>	PERSONAL HIST MALIGNANT NEOPLASM OF BREAST	\$3,190	\$634	\$1,311	\$5,135	0.3%	13.3%
<b>714</b>	RHEUMATOID ARTHRITIS	\$2,131	\$917	\$1,906	\$4,954	0.3%	25.4%
<b>256.39</b>	OVARIAN FAILURE NEC	\$2,087	\$1,878	\$688	\$4,653	0.3%	-36.7%
<b>V17.81</b>	FAMILY HX OSTEOPOROSIS	\$1,832	\$1,435	\$969	\$4,236	0.3%	-27.1%
<b>V58.83</b>	ENCOUNTER FOR THERAPEUTIC DRUG MONITORING	\$1,765	\$1,068	\$1,363	\$4,196	0.3%	-5.9%

Figure 10b, 10c PEBB/UMP Top Diagnoses for First vs Repeat BMD Tests 2011-2013, Descending by Amount Paid

Diag Code	Top Diagnoses for First DXA Tests, Paid \$ Desc.	Total Paid	Diag Code	Top Diagnoses for Repeat DXA Tests, Paid \$ Desc.	Total Paid
<b>V82.81</b>	SPECIAL SCREENING FOR OSTEOPOROSIS	\$314,518	<b>733.9</b>	DISORDER OF BONE AND CARTILAGE, UNSPECIFIED	\$187,550
<b>733.9</b>	DISORDER OF BONE AND CARTILAGE, UNSPECIFIED	\$302,807	<b>733</b>	UNSPECIFIED OSTEOPOROSIS	\$92,953
<b>V49.81</b>	ASYMPT POSTMENO STATUS	\$86,458	<b>V82.81</b>	SPECIAL SCREENING FOR OSTEOPOROSIS	\$83,636
<b>733</b>	UNSPECIFIED OSTEOPOROSIS	\$80,328	<b>V76.12</b>	SCREEN MAMMOGRAM NEC	\$27,215
<b>V76.12</b>	SCREEN MAMMOGRAM NEC	\$69,111	<b>V49.81</b>	ASYMPT POSTMENO STATUS	\$21,677
<b>627.2</b>	SYMPT FEM CLIMACT STATE	\$52,786	<b>733.01</b>	SENILE OSTEOPOROSIS	\$17,478
<b>627.9</b>	UNSPECIFIED MENOPAUSAL AND POSTMENOPAUSAL DISORDER	\$24,504	<b>627.2</b>	SYMPT FEM CLIMACT STATE	\$14,151
<b>V70.0</b>	ROUTINE GENERAL MEDICAL EXAMINATION AT HEALTH CARE FACILITY	\$22,342	<b>627.9</b>	UNSPECIFIED MENOPAUSAL AND POSTMENOPAUSAL DISORDER	\$7,312
<b>733.01</b>	SENILE OSTEOPOROSIS	\$15,553	<b>V70.0</b>	ROUTINE GENERAL MEDICAL EXAMINATION AT HEALTH CARE FACILITY	\$6,884
<b>V58.65</b>	LONG-TERM (CURRENT) USE OF STEROIDS	\$6,376	<b>174.9</b>	MALIGNANT NEOPLASM OF BREAST (FEMALE), UNSPECIFIED SITE	\$5,784
<b>781.91</b>	LOSS OF HEIGHT	\$5,607	<b>733.09</b>	OTHER OSTEOPOROSIS	\$3,824
<b>V76.11</b>	SCREEN MAMMOGRAM HI RISK	\$5,065	<b>V10.3</b>	PERSONAL HISTORY OF MALIGNANT NEOPLASM OF BREAST	\$2,929
<b>252</b>	HYPERPARATHYROIDISM NOS	\$4,663	<b>733.02</b>	IDIOPATHIC OSTEOPOROSIS	\$2,646
<b>174.9</b>	MALIGNANT NEOPLASM OF BREAST (FEMALE), UNSPECIFIED SITE	\$4,585	<b>714</b>	RHEUMATOID ARTHRITIS	\$2,418
<b>627.4</b>	SYMPT STATE W ARTIF MENO	\$4,470	<b>V72.31</b>	ROUTINE GYN EXAMINATION	\$2,319
<b>733.09</b>	OTHER OSTEOPOROSIS	\$3,667	<b>V76.11</b>	SCREEN MAMMOGRAM HI RISK	\$2,267
<b>268.9</b>	UNSPECIFIED VITAMIN D DEFICIENCY	\$3,652	<b>V58.69</b>	LONG-TERM USE MEDS NEC	\$2,254
<b>V72.31</b>	ROUTINE GYN EXAMINATION	\$3,634	<b>V58.65</b>	LONG-TERM (CURRENT) USE OF STEROIDS	\$2,033

Figure 10d. Medicaid Top Diagnoses for DXA Testing, Descending by Paid \$, 2011-2013

Medicaid FFS Top Diagnoses for DXA Testing by Paid \$ Descending (of 706 Diagnoses)		2011	2012	2013	Grand Total	% of Total Paid	Trend
Dx Code	Dx Description	\$184,025	\$165,546	\$139,369	\$488,941		-11.4%
73390	Bone & cartilage dis NOS	\$45,364	\$41,499	\$36,530	\$123,393	25.2%	-10.2%
V8281	Screen - osteoporosis	\$26,619	\$31,079	\$26,420	\$84,118	17.2%	0.9%
73300	Osteoporosis NOS	\$26,493	\$27,538	\$19,976	\$74,008	15.1%	-11.8%
V4981	Asympt postmeno status	\$21,521	\$10,102	\$6,760	\$38,383	7.9%	-43.1%
V7612	Screen mammogram NEC	\$6,726	\$5,575	\$4,015	\$16,316	3.3%	-22.6%
6272	Sympt fem climact state	\$4,606	\$3,323	\$1,774	\$9,703	2.0%	-37.2%
6279	Menopausal disorder NOS	\$4,151	\$2,636	\$1,512	\$8,299	1.7%	-39.6%
V5865	Long-term use steroids	\$2,019	\$1,959	\$1,303	\$5,281	1.1%	-18.2%
1749	Malign neopl breast NOS	\$1,216	\$1,935	\$2,103	\$5,254	1.1%	33.9%
20401	Act lym leuk w rmsion	\$1,966	\$1,629	\$1,015	\$4,610	0.9%	-27.4%
V5869	Therapeutic drug monitor	\$1,715	\$632	\$1,264	\$3,610	0.7%	18.4%
V5883	Long-term use meds NEC	\$1,568	\$521	\$1,022	\$3,111	0.6%	14.8%
6260	Absence of menstruation	\$1,812	\$416	\$503	\$2,731	0.6%	-28.0%
6274	Vitamin D deficiency NOS	\$610	\$1,076	\$756	\$2,443	0.5%	23.3%
7140	Rheumatoid arthritis	\$1,102	\$361	\$974	\$2,437	0.5%	51.4%
2689	Act myl leuk w rmsion	\$861	\$664	\$753	\$2,279	0.5%	-4.8%
25200	Sympt state w artif meno	\$1,321	\$590	\$360	\$2,271	0.5%	-47.2%
5559	Hyperparathyroidism NOS	\$734	\$1,032	\$464	\$2,230	0.5%	-7.2%
V103	Hx of breast malignancy	\$511	\$875	\$537	\$1,923	0.4%	16.3%
75651	Osteogenesis imperfecta	\$681	\$492	\$602	\$1,775	0.4%	-2.6%
20501	Regional enteritis NOS	\$421	\$405	\$914	\$1,741	0.4%	61.0%
7100	Path fx vertebrae	\$694	\$690	\$340	\$1,723	0.4%	-25.6%
7242	Syst lupus erythematosus	\$898	\$661	\$153	\$1,713	0.4%	-51.6%
73313	Lumbago	\$534	\$420	\$440	\$1,394	0.3%	-8.2%

Figure 10e, 10f Medicaid Top Diagnoses for First vs Repeat BMD Tests 2011-2013, Descending by Amount Paid

Diag Code	Top Diagnoses for First DXA Tests, Paid \$ Desc.	Total Paid	Diag Code	Top Diagnoses for Repeat DXA Tests, Paid \$ Desc.	Total Paid
<b>73390</b>	Bone & cartilage dis NOS	\$103,292	<b>73390</b>	Bone & cartilage dis NOS	\$20,101
<b>V8281</b>	Screen - osteoporosis	\$77,021	<b>73300</b>	Osteoporosis NOS	\$14,473
<b>73300</b>	Osteoporosis NOS	\$59,535	<b>V8281</b>	Screen - osteoporosis	\$7,097
<b>V4981</b>	Asympt postmeno status	\$35,548	<b>V4981</b>	Asympt postmeno status	\$2,836
<b>V7612</b>	Screen mammogram NEC	\$14,606	<b>V7612</b>	Screen mammogram NEC	\$1,710
<b>6272</b>	Sympt fem climact state	\$8,726	<b>20501</b>	Act myl leuk w rmsion	\$1,235
<b>6279</b>	Menopausal disorder NOS	\$7,399	<b>20401</b>	Act lym leuk w rmsion	\$1,207
<b>V5865</b>	Long-term use steroids	\$4,546	<b>6272</b>	Sympt fem climact state	\$977
<b>1749</b>	Malign neopl breast NOS	\$4,294	<b>1749</b>	Malign neopl breast NOS	\$960
<b>20401</b>	Act lym leuk w rmsion	\$3,403	<b>20400</b>	Ac lym leuk wo achv rmsn	\$914
<b>V5883</b>	Therapeutic drug monitor	\$2,975	<b>6279</b>	Menopausal disorder NOS	\$900
<b>V5869</b>	Long-term use meds NEC	\$2,917	<b>75651</b>	Osteogenesis imperfecta	\$841
<b>6260</b>	Absence of menstruation	\$2,423	<b>V5865</b>	Long-term use steroids	\$735
<b>2689</b>	Vitamin D deficiency NOS	\$2,335	<b>V5883</b>	Therapeutic drug monitor	\$635
<b>7140</b>	Rheumatoid arthritis	\$2,234	<b>20500</b>	Ac myl leuk wo achv rmsn	\$586
<b>6274</b>	Sympt state w artif meno	\$2,178	<b>82321</b>	Fx shaft fibula-closed	\$510
<b>25200</b>	Hyperparathyroidism NOS	\$2,018	<b>3384</b>	Chronic pain syndrome	\$453
<b>V103</b>	Hx of breast malignancy	\$1,616	<b>6260</b>	Absence of menstruation	\$308

**Related Medical Codes**

Code Type	Codes	Short Description	Code Chgs
Osteoporosis Diagnosis	733	Osteoporosis	
	733.01	Senile osteoporosis; Postmenopausal osteoporosis	
	733.02	Idiopathic osteoporosis	
	733.03	Disuse osteoporosis	
	733.09	Other osteoporosis; Drug-induced osteoporosis	
	733.9	Other osteoporosis & unspecified disorders of bone/cartilage	Code recommended for osteopenia
	733.99	Other osteoporosis: Diaphysitis; Hypertrophy of bone; Relapsing polychondritis	
BMD Test CPT Codes	V82.81	Special screening: osteoporosis	
	V17.81	Family history, chronic disease: osteoporosis	
	G0130	Single energy xray (SEXA) bone density, peripheral	
	76977	US BONE DENSITY MEASURE	
	77078	CT BONE DENSITY, AXIAL	
	77079	CT BONE DENSITY, PERIPHERAL	Del 2012
	77080	DXA BONE DENSITY, AXIAL	
	77081	DXA BONE DENSITY/PERIPHERAL	
	77082	DXA BONE DENSITY/VERTEBRAL FRACTURE ASSESSMT	Not included in utilization data
	77083	RADIO ABSORPTIOMETRY, 1 OR MORE SITES	Del 2012
	78350	BONE MINERAL DENSITY, SINGLE PHOTON	

## APPENDICES

### APPENDIX I. Search Strategy

#### INITIAL SEARCH, SYSTEMATIC REVIEWS AND PRACTICE GUIDELINES (conducted July 8 to August 1, 2014)

Initially, evidence for this report was obtained by searching for systematic reviews and guidelines that had been published in the past 10 years. Searches were conducted in the following databases using the terms *osteoporosis* or *osteoporosis and screening*: Agency for Healthcare Research and Quality (AHRQ), Blue Cross Blue Shield TEC Assessments, Canadian Agency for Drugs and Technology in Health (CADTH), Centre for Reviews and Dissemination (York University), Cochrane Library, Hayes Knowledge Center, Institute for Clinical Systems Improvement (ICSI), National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (UK), United States Preventive Services Task Force (USPSTF), National Guidelines Clearinghouse (NGC), National Institute for Health and Care Excellence (NICE), and Veterans Affairs Technology Assessment Program (VA TAP).

The websites for the American Academy of Family Physicians, the American College of Physicians, the American Geriatrics Society, and the American Academy of Orthopedic Surgeons were also searched since guidelines for both older men and older women were not identified for these groups in a search of the NGC.

Additional systematic reviews were selected from a search of the PubMed database using filters for Practice Guidelines, Guidelines, Meta-Analyses, and Systematic Reviews, according to this search:

1. screen or screening or monitor or monitoring or test or testing or follow-up or interval
2. osteoporosis or osteopenia or "bone density" or "bone mineral density" or BMD
3. 1 and 2

Filters: Meta-Analysis; Systematic Reviews; Publication date from 2009/01/01 to 2014/12/31; English

#### PRIMARY CLINICAL STUDIES NOT INCLUDED IN SYSTEMATIC REVIEWS

Searches were conducted in PubMed on June 12 except for search #3, which was conducted on July 17. An update search was conducted on August 4, 2014.

**Search #1** – Any relevant study published since November 2009 (last search date for the 2010 Evidence Review on *Screening for Osteoporosis: Systematic Review to Update the 2002 U.S. Preventive Services Task Force Recommendation*.)

1. DXA or DEXA or "dual x-ray absorptiometry"
2. screen or screening or monitor or monitoring or test or testing or follow-up or interval
3. osteoporosis or osteopenia or "bone density" or "bone mineral density" or BMD
4. 1 and 2 and 3

Filters: Publication date from 2009/11/01 to 2014/12/31; Humans; English

**Search #2** – This search took into account the fact that the 2002 and 2010 AHRQ reports included only *randomized controlled trials (RCTs)* and the population of interest was *postmenopausal women or men older than age 50*. Some studies that might provide data for Key Questions 1a, 1b, 2a, or 2b for the present report could have been eliminated from the 2010 AHRQ review because of 1 of the following reasons:

- They were not RCTs.
- They selected patients for reasons other than age or menopausal status.
- They addressed strategies for monitoring patients being *treated* for osteoporosis rather than screening strategies.

The following search was expected to identify non-RCTs of either screening or treatment monitoring, regardless of patient selection criteria and regardless of publication date:

1. (osteoporosis or (bone density)) AND fracture AND (screen or screening)
2. "Cohort Studies"[Mesh] OR "Case-Control Studies"[Mesh]
3. 1 and 2

Filters: Publication date from 1966/1/1 to 2014/12/31; Humans; English

Other searching was considered in order to identify studies published prior to November 2009 that were RCTs but that were excluded from the AHRQ reports because the population was defined by risk factors other than age and menopausal status, or because the studies evaluated treatment monitoring strategies. However, several factors suggested that such searches would yield no relevant studies: (1) topic scoping searches were unfruitful; (2) few RCTs were identified for screening strategies in general populations, making RCTs in special risk populations or treatment populations unlikely; (3) practice guidelines and review articles did not cite any RCTs; and (4) the Excluded Studies lists of the 2010 AHRQ report did not identify any RCTs that were excluded because of patient selection criteria. Therefore, no attempt was made to systematically search for RCTs published prior to November 2009 other than those included in the AHRQ reports.

**Search #3** – Studies published since November 2009 that might address Key Question #2c (minimum screening/testing interval) and were not otherwise identified.

("bone density" or "bone mineral density") and "Longitudinal Studies"[Mesh]

Filters: Publication date from 2009/11/01 to 2014/12/31; Humans; English

**Search #4** – Cost studies and economic evaluations

**Search 4a:** National Health Service Economic Evaluation Database (NHS-EED) (on the site for Centre for Reviews and Dissemination, York University)



**Search 4b:**

1. osteoporosis or osteopenia or "bone density" or "bone mineral density"
2. screen or screening or monitor or monitoring or test or testing
3. (((((economic analysis) OR (economic evaluation)))) OR (((((cost AND (analysis OR benefit OR effective\* OR consequence OR minimization)))) OR ((("Costs and Cost Analysis"[MeSH] OR "Cost-Benefit Analysis"[MeSH]))))
4. 1 and 2 and 3

Filters: Publication date from 2005/01/01 to 2014/12/31; English

## APPENDIX II. Overview of Evidence Quality Assessment Methods

### Clinical Studies

Tools used include internally developed Quality Checklists for evaluating the quality (internal validity) of different types of studies, a checklist for judging the adequacy of systematic reviews used instead of de novo analysis, and Hayes Evidence-Grading Guides for evaluating bodies of evidence for different types of technologies. Hayes methodology is in alignment with the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) system, which was developed by the GRADE Working Group, an international collaborative body.

<b>Step 1</b>	<u>Individual study appraisal</u> a. Initial rating according to study design <i>Good</i> : Randomized Controlled Trials <i>Fair</i> : Nonrandomized Trial (controlled, parallel-group, quasirandomized) <i>Poor</i> : Observational Analytic Studies (prospective or retrospective trials involving historical controls, pretest-posttest control trial [patients legitimately serve as their own controls], case-control, registry/chart/database analysis involving a comparison group) <i>Very Poor</i> : Descriptive Uncontrolled Studies (case reports, case series, cross-sectional surveys [individual-level data], correlation studies [group-level data]) b. Consider the methodological rigor of study execution according to items in a proprietary Quality Checklist c. Repeat for each study
<b>Step 2</b>	<u>Evaluation of each body of evidence by outcome, key question, or application</u> a. Initial quality designation according to <i>best</i> study design in a body of evidence b. Downgrade/upgrade <i>Downgrade factors</i> : Study weaknesses (Quality Checklists), small quantity of evidence, lack of applicability, inconsistency of results, publication bias <i>Possible upgrade factors</i> : Strong association, dose-response effect, bias favoring no effect c. Assign final rating: High-Moderate-Low-Insufficient d. Repeat for each outcome/question/application
<b>Step 3</b>	<u>Evaluation of overall evidence</u> a. Rank outcomes by clinical importance b. Consider overall quality of evidence for each <i>critical</i> outcome c. Assign overall rating based on lowest-quality body: High-Moderate-Low-Insufficient
<b>Step 4</b>	<u>Evidence-Based Conclusion</u> Overall quality of evidence + Balance of benefits and harms

**Practice Guidelines (checklist taken from [AGREE Tool](#) and approach to scoring used in this report)**

Rank each item on a scale of 1-7.

Decide on overall quality (1 = lowest to 7 = highest), giving strongest weight to items 7-14 (Rigor of Development Domain) and items 22-23 (Editorial Independence). For qualitative labels:

Very poor = 1

Poor = 2-3

Fair = 4-5

Good = 6-7

1. The overall objective(s) of the guideline is (are) specifically described.
2. The health question(s) covered by the guideline is (are) specifically described.
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.
4. The guideline development group includes individuals from all relevant professional groups.
5. The views and preferences of the target population (patients, public, etc.) have been sought.
6. The target users of the guideline are clearly defined.
7. Systematic methods were used to search for evidence.
8. The criteria for selecting the evidence are clearly described.
9. The strengths and limitations of the body of evidence are clearly described.
10. The methods for formulating the recommendations are clearly described.
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.
12. There is an explicit link between the recommendations and the supporting evidence.
13. The guideline has been externally reviewed by experts prior to its publication.
14. A procedure for updating the guideline is provided.
15. The recommendations are specific and unambiguous.
16. The different options for management of the condition or health issue are clearly presented.
17. Key recommendations are easily identifiable.
18. The guideline describes facilitators and barriers to its application.
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.
20. The potential resource implications of applying the recommendations have been considered.
21. The guideline presents monitoring and/or auditing criteria.
22. The views of the funding body have not influenced the content of the guideline.
23. Competing interests of guideline development group members have been recorded and addressed.

## Economic Evaluations

A tool developed by Hayes for internal use guides interpretation and critical appraisal of economic evaluations. The tool includes a checklist of items addressing issues such as the reliability of effectiveness assumptions, transparency of reporting, quality of analysis, generalizability/applicability, and conflicts of interest. The following publications served as sources of best practice.

### Articles

Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol*. 2013;66(2):140-150.

Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ*. 1996;313(7052):275-283.

Drummond M, Sculpher M. Common methodological flaws in economic evaluations. *Med Care*. 2005;43(7 Suppl):5-14.

Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care*. 2005;21(2):240-245.

Gerken S, Crott R, Cleemput I, et al. Comparison of three instruments assessing the quality of economic evaluations: a practical exercise on economic evaluations of the surgical treatment of obesity. *Int J Technol Assess Health Care*. 2008;24(3):318-325.

Hutubessy R, Chisholm D, Edejer TT. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost Eff Resour Alloc*. 2003;1(1):8.

Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. *Evid Policy*. 2010;6(1):51-59.

Smith KA, Rudmik L. Cost collection and analysis for health economic evaluation. *Otolaryngol Head Neck Surg*. 2013;149(2):192-199.

Ubel PA, Hirth RA, Chernew ME, Fendrick AM. What is the price of life and why doesn't it increase at the rate of inflation? *Arch Intern Med*. 2003;163(14):1637-1641.

### Books

Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. 2nd Edition. Oxford, UK: Oxford University Press; 1997.

Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine*. 1996. Oxford, UK: Oxford University Press; 1996.

*Other*

Canadian Agency for Drugs and Technologies in Health (CADTH). *Guidelines for the Economic Evaluation of Health Technologies*. 3rd Edition. Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health; 2006. Available at: [http://www.cadth.ca/media/pdf/186\\_EconomicGuidelines\\_e.pdf](http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf). Accessed August 20, 2014.

## APPENDIX III. Evidence Tables

**Evidence Table IIIa. Controlled Studies Evaluating the Effectiveness of Osteoporosis Screening by Dual X-Ray Absorptiometry (DXA) in Individuals (Key Question #1a)**

Key: ADT, androgen-deprivation therapy; BMD, bone mineral density; BMI, body mass index; CI, confidence interval; CS, corticosteroid; DXA, dual x-ray absorptiometry; EHR, electronic health record; GP, general practitioner; HR, hazard ratio; HRT, hormone replacement therapy; IR(R), incidence rate (ratio); ITT, intention-to-treat; IU, international unit; NS, not statistically significant; OP, osteoporosis; RCT, randomized controlled trial; SD, standard deviation

Authors/Study Design/Protocol	Patients/Follow-Up from DXA Scan	Main Findings (statistically significant 95% CIs are bolded)	Quality/Comments
<b>Kern et al. (2006)</b>  Community-based nonrandomized trial  Study participants in a Cardiovascular Health Study were assigned to DXA scan or no scan according to the state in which they lived; study protocol did not stipulate treatment	3107 generally healthy, community-dwelling men and women without a history suggesting OP, age ≥65 yrs  Mean 5 yrs	<b>Incidence of hip fracture (screened, usual care) (per 1000 person-yrs):</b> Overall: 4.8, 8.2 Women: 5.3, 9.7; Men: 4.2, 6.3 65-74 yrs: 2.3, 4.1; 75-84 yrs: 6.6, 8.8; ≥85 yrs: 8.1, 38.6 White: 5.6, 9.3; Black: 1.4, 1.7 <b>Adjusted HR for hip fracture (screened vs usual care):</b> <b>Overall: 0.64 (CI, 0.4-0.99)</b> Women: 0.61 (CI, 0.35-1.06); Men: 0.68 (CI, 0.32-1.42) Age 65-74 yrs: 0.73 (CI, 0.29-1.87); 75-84 yrs: 0.82 (CI, 0.47-1.44); ≥85 yrs: <b>0.22 (CI, 0.06-0.79)</b> All participants minus those age ≥85 yrs: 0.85 (CI, 0.50-1.31), but test for group-age interaction was NS. <b>White race: 0.62 (CI, 0.39-0.97); Black race:</b> Too few events to calculate	Fair  HRs were adjusted for sex and propensity score, which reflected 31 potential confounders that could be associated with both treatment effect and likelihood of being in the screened group.  Unknown generalizability to clinical populations
<b>Sedlak et al. (2007)</b>  Community-based RCT w/ wait-list controls  Questionnaires mailed to respondents to media advertising; immediate (screened group) or delayed (control group), free DXA	203 healthy postmenopausal women  1 yr	<b>Total calcium intake over 1 yr (screened, wait-list) (units unclear; assumed to be IUs per day):</b> 786, 668 (global $P<0.001$ )  <b>Exercise:</b> No change over time in either group	Good  Long-term effects unknown. Unknown whether generalizable to a clinical population.

Authors/Study Design/ Protocol	Patients/Follow- Up from DXA Scan	Main Findings (statistically significant 95% CIs are bolded)	Quality/Comments
offered. Follow-up with physician recommended if DXA was abnormal, but no treatment recommendations.			
<b>Barr et al. (2010)</b>  Community-based RCT  Patients were randomized to DXA invitation or no invitation; questionnaires mailed to both groups for baseline and follow-up data. For individuals in screened group with femoral neck or lumbar spine T-scores below lowest quarter of first 1000 women screened, lifestyle changes were advised, and treatment advice (prophylactic HRT after menopause) was mailed to GPs; HRT also recommended for screened women in the top 3 quarters of BMD scores if corticosteroids were prescribed.	2604 evaluable middle-aged (45-54 yrs) women in a primary care registry in a particular locality  9 yrs	<b>Fracture, ITT analysis (screened vs control):</b> % women w/ incident fracture: 8.8%, 9.4% <i>Adjusted HR:</i> 0.791 (CI, 0.600-1.042) Adjusted treatment-received HR was 0.759 (NS); adjusted per-protocol HR was 0.734 (CI, 0.546-0.988; $P=0.041$ ) Global assessment of differences stratified by type of fracture was NS.  <b>Medication or supplement use (screened, control) (% pts):</b> HRT: 52.4%, 44.5% ( $P<0.01$ ) VitD: 24.2%, 12.5% ( $P<0.01$ ) Calcium: 20.0%, 14.1% ( $P<0.01$ )	Fair (high loss to follow-up, i.e., a substantial proportion of the sample was not evaluable)  HR was adjusted for age, weight, and height.  True effect might have been even more negative, depending on the consistency between usual GP practice and the treatment advice mailed to them by investigators. Results may not be generalizable to current practice regarding HRT. No way of knowing whether lack of medication use in controls was appropriate.
<b>Doheny et al. (2011)</b>  Community-based quasi-RCT  Group assignment to free DXA screening or usual care by flip of coin for first respondent, followed by alternate assignment in subsequent respondents. Both groups received educational	196 healthy men ≥50 yrs of age and without history of fracture; recruited through community advertising	<b>Exercise:</b> Mean # mins of vigorous activity: 22, 19 (NS) Mean # mins walking: 15.3, 13 (NS) <b>Calcium:</b> No group or knowledge effect.	Fair

Authors/Study Design/ Protocol	Patients/Follow- Up from DXA Scan	Main Findings (statistically significant 95% CIs are bolded)	Quality/Comments
brochure.	1 yr		
<b>Zhumkhawala et al. (2013)</b>  Retrospective cohort study  Analysis of EHRs and Healthy Bones Program (HBP) registry in a large group-model Managed Care Organization (Kaiser Permanente Southern California).	1432 men with prostate cancer being treated with ADT (leuprolide); young to old  Mean 3.2 yrs (screened) and 2.28 yrs (control)	<b>Hip fracture (screened, control):</b> <u>% of pts:</u> 1.7%, 4.1% <u>IR per 1000 person-yrs:</u> 5.1 (CI, 3.0-8.0), 18.1 (CI, 10.5-29.0). Non-overlapping CIs suggests statistically significant difference. <u>Time to fracture (mean days±SD):</u> 828.7±471.7, 590.3±476 <u>Time to fracture (median days±SD):</u> 800.5, 528 <u>Adjusted HR (control vs screened):</u> <b>4.19 (CI, 1.92-9.13)</b> (# leuprolide doses, obesity, and overweight were also associated with significant HRs <1). <b>Medication use:</b> Screened men more likely to be taking OP drugs during follow-up period (29% vs 3%; $P<0.0001$ )	Fair (upgraded from Poor because of strong association and because confounders created bias against intervention)  HR was adjusted for age, race, BMI, previous or concurrent other fracture, previous disease or diagnosis associated with OP, concurrent drugs causing OP, concurrent drugs associated with hip fracture, propensity score to control for HBP selection bias (possibility that patients with greater risk enrolled in the Healthy Bones Program).
<b>Khan et al. (2014)</b>  Retrospective review of database (cohort design).	5736 veterans with ulcerative colitis in national Veterans Affairs healthcare system and without history of fracture. Even distribution of age from young to elderly  Mean 3 yrs	<b>Adjusted HR (screening vs no screening) for fragility fracture:</b> <b>0.5 (CI, 0.3-0.9; <math>P=0.03</math>).</b> <b>Cumulative incidence (screening, no screening) (fractures per 1000 person-yrs):</b> <i>Overall:</i> 1.6 vs 2.8 <i>CS exposure vs low:</i> <i>Intermediate:</i> 1.9 (CI, 1-3.4; $P=0.04$ ) <i>High:</i> 2.9 (CI, 1.6-5.3; $P<0.001$ ) <b>Interaction between DXA screen and CS exposure (IRR, screen vs no screen):</b> <i>Low CS exposure:</i> No difference <i>Moderate:</i> 0.44 (NS) <i>High:</i> 0.38 ( $P=0.02$ ) <b>Medication use (DXA screen, no screen)(% pts; risk difference):</b> <i>Bisphosphonate (n=727):</i> 32.2%, 7.4% ( $P<0.001$ ); 24.8%. <i>Some type of OP medication, excluding HRT:</i> 36.6%, 21.6% ( $P<0.001$ ); 15%	Fair (upgraded from Poor because of the strong statistical association)  HR was adjusted for age, sex, and race.



Authors/Study Design/ Protocol	Patients/Follow- Up from DXA Scan	Main Findings (statistically significant 95% CIs are bolded)	Quality/Comments
		<i>Vitamin D and calcium</i> : 32.9%, 13.4% ( $P<0.001$ ); 19.5% <i>Calcitonin</i> : 2.3%, 0.8% ( $P<0.001$ ); 1.5%	

### Evidence Table IIIb. Longitudinal Studies Designed to Estimate Optimal Screening Intervals (Key Question #2c)

Key: AUC, area under the curve; BL, baseline; BMD, bone mineral density; BMI, body mass index; CI, confidence interval; DXA, dual x-ray absorptiometry; FRAX, Fracture Risk Assessment Tool; HR, hazard ratio; HRT, hormone replacement therapy; OP, osteoporosis; RA, rheumatoid arthritis; SD, standard deviation

Authors/Study Design/ Protocol	Patients/Follow-Up from DXA Scan	Main Findings (95% CIs unless otherwise noted; statistically significant CIs are bolded)	Quality/Comments
<p><b>Frost et al. (2009)</b> Australia</p> <p>Interim results from the community-based longitudinal Dubbo Osteoporosis Epidemiology study (retrospective cohort study)</p> <p>Participants reported for BMD measurement every 2 yrs. Fracture incidence based on radiology reports, with nontraumatic nature confirmed by participant interview.</p> <p>Risk prediction model (Cox proportional hazards) included age, sex, and initial femoral neck BMD as covariates and took into account the competing risk of death. Separate models for OP, clinical fracture, and OP and/or fracture as dependent variables. Time to reach risk thresholds</p>	<p>1753 men and women with T-score <math>&gt; -2.5</math>, age <math>\geq 60</math> yrs; 99% white; recruitment through invitation (750 men, 1003 women)</p> <p>Median 7.1 yrs</p>	<p><b>Accuracy of the risk prediction model (AUC):</b> Men, 0.74; women, 0.76</p> <p>The following are representative results from the study model. In the second set of data, age 70 has been emphasized for men, since that is closest to the typical screening age for men in the U.S., and age 65 for women, since that is the typical screening age for women in the U.S. Lower bounds of the 90% CIs are bolded since these were the repeat screening intervals recommended by authors.</p> <p><b>Shortest/longest times to reach 10%, 5-yr risk of OP and/or fracture (treatment threshold in Australia) (yrs):</b>  <i>Men at age 60, T-score 0:</i> 12.3 (90% CI, <b>10.6</b>-13.4)  <i>Men at age 80, T-score -2.2:</i> 2.0 (90% CI, <b>1.6</b>-2.4)  <i>Women at age 60, T-score 0:</i> 8.9 (90% CI, <b>6.7</b>-10.6)  <i>Women at age 80, T-score -2.2:</i> 1.9 (90% CI, <b>1.5</b>-2.0)</p>	<p>Good</p> <p>Risk prediction model did not include lifestyle factors, medical conditions, or medical treatments that can contribute to OP (however, Nelson et al. [2010] [AHRQ]) concluded that simple predictive models worked as well as more complex models).</p> <p>Time estimates that substantially exceed the median follow-up are based on small percentages of participants.</p> <p>Results may not be generalizable to a nonvolunteer, clinical population, or to the U.S. Discrepancy between study predictions in the oldest participants and predictions for the</p>

Authors/Study Design/ Protocol	Patients/Follow-Up from DXA Scan	Main Findings (95% CIs unless otherwise noted; statistically significant CIs are bolded)	Quality/Comments
was then calculated by simulation methods for age and T-score categories.		<p><b><i>Shortest/longest times and times for typical initial screening ages to reach 20%, 10-yr risk of OP and/or fracture (treatment threshold in the U.S.) (yrs):</i></b></p> <p><b><i>Men:</i></b></p> <p>Age 60, T-score 0: 15.0+ (90% CI, <b>14.3</b>-15.0+) (follow-up did not go beyond 15 yrs)</p> <p>Age 70, T-score 0: 10.7 (90% CI, <b>9.0</b>-12.2)</p> <p>Age 70, T-score -1.0: 8.9 (90% CI, <b>7.8</b>-9.8)</p> <p>Age 70, T-score -1.5: 8.1 (90% CI, <b>7.2</b>-9.0)</p> <p>Age 70, T-score -2.0: 7 (9.40% CI, <b>6.5</b>-8.7)</p> <p>Age 70, T-score -2.2: 7.3 (90% CI, <b>7.1</b>-8.4)</p> <p>Age 80, T-score -2.2: 2.9 (90% CI, <b>2.6</b>-3.8)</p> <p><b><i>Women:</i></b></p> <p>Age 60, T-score 0: 14.1 (90% CI, <b>12.7</b>-15.0+) (follow-up did not go beyond 15 yrs)</p> <p>Age 65, T-score 0: 12.3 (90% CI, <b>10.6</b>-13.4)</p> <p>Age 65, T-score -1.0: 8.3 (90% CI, <b>7.2</b>-9.8)</p> <p>Age 65, T-score -1.5: 7/5 (90% CI, <b>5.5</b>-7.3)</p> <p>Age 65, T-score -2.0: 4.9 (90% CI, <b>4.4</b>-5.9)</p> <p>Age 65, T-score -2.2: 4.6 (90% CI, <b>3.8</b>-5.4)</p> <p>Age 80, T-score -2.2: 2.4 (90% CI, <b>2.2</b>-2.6)</p> <p><b><i>Age and T-score at initial screen of men in which time to reach 20%, 10-yr risk was</i></b></p> <p>&lt;2 yrs: No individuals</p> <p>&lt;3 yrs: Age 80, T-score ≤ -2.2</p> <p>&lt;5 yrs: Age 75, T-score ≤ -2.2; age 80, any T-score</p> <p><b><i>Age and T-score at initial screen of women in which time to reach 20%, 10-yr risk was</i></b></p> <p>&lt;2 yrs: No individuals</p> <p>&lt;3 yrs: Age 75, T-score ≤ -2.0; age 80, T-score ≤ -1.5</p> <p>&lt;5 yrs: Age 75, T-score ≤ -1.5; age 80, any T-score</p>	<p>same age group based on the U.S. version FRAX tool.</p> <p><i>Clinical application, according to authors:</i> A 60-yr-old man with a T-score of 0 does not need to be screened again for another ~10.5 yrs if the goal is to assess a 10% risk of fracture within 5 yrs at each BMD measurement. But an 80-yr-old man with a T-score of -2.2 should be screened again in about 1.5 yrs with the same clinical goal.</p>

Authors/Study Design/ Protocol	Patients/Follow-Up from DXA Scan	Main Findings (95% CIs unless otherwise noted; statistically significant CIs are bolded)	Quality/Comments
<p><b>Gourlay et al. (2012)</b> Analysis of a subset of the Study of Osteoporotic Fractures (U.S.)</p> <p>Prospective cohort study</p> <p>BMD measurement at 2, 6, 8, and 16 yrs. Stratification based on lower of femoral neck or total hip BMD. Questionnaires at each assessment. Postcard or telephone contact every 4 months to ascertain incident fracture; physician adjudication from review of radiology reports.</p> <p><i>Adjustments in time to event:</i> <u>Normal BMD Group</u>: Continuous BMD T-score, age <u>*Osteopenia Group</u>: Age, BMI, estrogen use at BL, any fracture after 50 yrs of age, current smoking, current or past oral glucocorticoids, and self-reported RA.</p>	<p>4957 women with T-score &gt; -2.5, no previous hip or clinical vertebral fracture, no previous OP treatment, and age ≥67 yrs 17% were currently using HRT</p> <p>Mean 8.2 yrs; range 1.1-14.6 yrs from time of first recorded BMD; &gt;7.9 yrs for 50% of participants</p>	<p><b><i>Adjusted time for 10% of women to make the transition to OP before function or treatment (yrs):</i></b>  <i>Normal BMD (T-score ≥ -1.00):</i> 16.8 (CI, 11.5-24.6)  <i>Mild osteopenia (T-score -1.01 to -1.49):</i> 17.3 (CI, 13.9-21.5)  <i>Moderate osteopenia (T-score -1.50 to -1.99):</i> 4.7 (CI, 4.2-5.2)  <i>Advanced osteopenia (T-score -2.00 to -2.49):</i> 1.1 (CI, 1.0-1.3)</p> <p>No interactions except between BMI and baseline BMD: Increasing BMI associated with longer time to transition only in women with moderate osteopenia at BL (<math>P &lt; 0.001</math> for trend). Transition times diminished with age (0.77 for advanced osteopenia at age 85 to 4.58 yrs for moderate osteopenia at age 67) and with current use of estrogen (0.97 for past/never used and advanced osteopenia to 6.90 for current use and moderate osteopenia), but differences were small within each BL T-score category and CIs overlapped.</p> <p>Sensitivity analyses varied the percentage of women making transition from osteopenia (1%, 2%, 5%) and from normal (20%). Findings were confirmed.</p>	<p>Good</p> <p>Basis of osteopenia categories not reported.</p> <p>Rationale for different covariates for the 2 exposure groups not reported.</p> <p>Missing data for parental history of fracture; no reason given for omitting alcohol use as covariate. Both factors are included in FRAX model.</p> <p>Time estimates &gt;13 yrs are based on small percentage of participants.</p>

**Evidence Table IIIc. Studies Reporting Calculations of Numbers Needed to Screen (NNS) (Key Question #3)**

Key: ADT, androgen-deprivation therapy; CI, confidence interval; FIT, Fracture Intervention Trial; HR, hazard ratio; NNS, number needed to screen; OP, osteoporosis; SR, systematic review

Authors/Study Type/Population	Sources of Data or Assumptions	Assumptions and Results by Age Category	Quality/Comments																																				
<b>Kern et al. (2005)</b>  Nonrandomized trial to detect an association between screening and incidence of hip fracture; 4 states in the U.S.	Empirical findings from this study (3107 older adults, age ≥65 yrs; follow-up mean 5 yrs)  <b>Threshold for treatment:</b> Not defined <b>Prevalence of OP:</b> Not reported. Prevalence of low bone mass (T-score ≤ −2.0) was 1.8% <b>Effectiveness of treatment:</b> Not measured	The following calculations were based on the incidence per 1000 person-yrs reported in the study article. <b>In adults ≥85 yrs of age (the only subgroup with robust HR):</b> Incidence of hip fracture per 1000 person-yrs (screened, usual care): 8.1, 38.6 Incidence per person-yr: 0.0081, 0.0386 Risk difference per person-yr: 0.0305 NNS (inverse of risk difference) to prevent hip fracture over 1 yr: 33 NNS over 5 yrs (33/5 and rounded up): 7 <b>NNS in subgroups defined by sex (nonsignificant point estimates for HR) (over 1 yr, 5 yrs):</b> Men: 477, 96 Women: 228, 46 <b>In all adults ≥65 yrs of age (upper bound of CI for HR approached 1) (over 1 yr, extrapolation to 5 yrs):</b> 295, 59	See <b>Table Evidence Table IIIa</b> for more study details.  NNS values were expected to by the authors of this report to reflect real-world probabilities of treatment and treatment adherence after screening.																																				
<b>Nelson et al. (2010)</b> AHRQ Evidence Review  Older women without previous vertebral fracture, followed ~4 yrs	<b>Threshold for treatment:</b> OP (T-score ≤ −2.5) <b>Prevalence of OP:</b> 1992 population-based study of women living in Rochester, MN (age ≥50 yrs; 100% white) <b>Effectiveness of treatment:</b> Fractures occurring in FIT (alendronate [treatment] vs placebo [no treatment])*	<table><thead><tr><th></th><th>Age 55-59</th><th>60-64</th><th>65-69</th><th>70-74</th><th>75-79</th></tr></thead><tbody><tr><td><b>Prevalence of OP (per 10,000 women)</b></td><td>0.0445</td><td>0.0650</td><td>0.1200</td><td>0.2025</td><td>0.2850</td></tr><tr><td colspan="6"><b>†5-yr risk of any fracture, with/without treatment in FIT subgroup with T-score ≤2.5:</b> 24.5%/ 16.38%</td></tr><tr><td><b>Clinical fractures prevented, any</b></td><td>36</td><td>53</td><td>97</td><td>164</td><td>231</td></tr><tr><td><b>Clinical fractures prevented, vertebral</b></td><td>16</td><td>23</td><td>43</td><td>73</td><td>104</td></tr><tr><td><b>Clinical fractures prevented, hip</b></td><td>6</td><td>10</td><td>18</td><td>31</td><td>42</td></tr></tbody></table>		Age 55-59	60-64	65-69	70-74	75-79	<b>Prevalence of OP (per 10,000 women)</b>	0.0445	0.0650	0.1200	0.2025	0.2850	<b>†5-yr risk of any fracture, with/without treatment in FIT subgroup with T-score ≤2.5:</b> 24.5%/ 16.38%						<b>Clinical fractures prevented, any</b>	36	53	97	164	231	<b>Clinical fractures prevented, vertebral</b>	16	23	43	73	104	<b>Clinical fractures prevented, hip</b>	6	10	18	31	42	Good-quality SR Fair-quality analysis because of indirectness  Probably underestimates the true NNS in real practice
	Age 55-59	60-64	65-69	70-74	75-79																																		
<b>Prevalence of OP (per 10,000 women)</b>	0.0445	0.0650	0.1200	0.2025	0.2850																																		
<b>†5-yr risk of any fracture, with/without treatment in FIT subgroup with T-score ≤2.5:</b> 24.5%/ 16.38%																																							
<b>Clinical fractures prevented, any</b>	36	53	97	164	231																																		
<b>Clinical fractures prevented, vertebral</b>	16	23	43	73	104																																		
<b>Clinical fractures prevented, hip</b>	6	10	18	31	42																																		

Authors/Study Type/Population	Sources of Data or Assumptions	Assumptions and Results by Age Category	Quality/Comments																		
		<b>NNS to prevent:</b> <table><tr><td><b>1 fracture over 5 yrs</b></td><td>278</td><td>187</td><td>103</td><td>61</td><td>43</td></tr><tr><td><b>1 vertebral fracture over 5 yrs</b></td><td>625</td><td>435</td><td>233</td><td>137</td><td>96</td></tr><tr><td><b>1 hip fracture over 5 yrs</b></td><td>1667</td><td>1000</td><td>556</td><td>323</td><td>238</td></tr></table> <b>Calculations:</b> $(10,000 \times \text{prevalence}/10,000) \times (\text{risk}_{\text{no treatment}} - \text{risk}_{\text{treatment}}) = \# \text{ fractures prevented}$ $\text{Inverse of } (\# \text{ fractures prevented}/10,000) = \text{NNS}$	<b>1 fracture over 5 yrs</b>	278	187	103	61	43	<b>1 vertebral fracture over 5 yrs</b>	625	435	233	137	96	<b>1 hip fracture over 5 yrs</b>	1667	1000	556	323	238	
<b>1 fracture over 5 yrs</b>	278	187	103	61	43																
<b>1 vertebral fracture over 5 yrs</b>	625	435	233	137	96																
<b>1 hip fracture over 5 yrs</b>	1667	1000	556	323	238																
<b>Zhumkhawala et al. (2013)</b>  Retrospective analysis of electronic health records and registry  See <a href="#">Evidence Table IIIa</a> for additional study details.	Study results (1432 men being treated for prostate cancer with ADT; follow-up mean 3.2 yrs)	<b>Calculation of NNS from study based on the incidence per 1000 person-yrs reported in the study article:</b> <i>Incidence of hip fracture per 1000 person-yrs (screened, usual care): 5.1, 18.1</i> <i>Incidence per person-yr: 0.0051, 0.0181</i> <i>Risk difference per person-yr: 0.013</i> <i>NNS (inverse of risk difference) to prevent hip fracture over 1 yr: 77</i> <i>NNS over 3 yrs (77/3 and rounded up): 26</i>	Fair																		
<b>Khan et al. (2014)</b>  Retrospective review of database (cohort study)  See <a href="#">Evidence Table IIIa</a> for additional study details.	Study results (5736 veterans with ulcerative colitis in national Veterans Affairs healthcare system; mean follow-up 3 yrs)	<b>Calculation of NNS based on the incidence per 1000 person-yrs reported in the study article:</b> <i>Incidence of hip fracture per 1000 person-yrs (screened, usual care): 1.6, 2.8</i> <i>Incidence per person-yr: 0.0016, 0.0028</i> <i>Risk difference per person-yr: 0.0012</i> <i>NNS (inverse of risk difference) to prevent hip fracture over 1 yr: 834</i> <i>NNS over 3 yrs (834/3 and rounded up): 278</i>	Fair  Stratified analysis by intensity of corticosteroid use showed no effect (low), substantial but nonsignificant effect (moderate), and a statistically significant effect (high); event rates were not available for subgroups.																		

\*Details of FIT trial (Cummings et al., 1998):

4432 postmenopausal women, age 50-80 yrs, without vertebral fracture; selected on the basis of low bone mass; ~1/3 participants in each of 3 T-score categories ( $< -2.5$ ,  $-2.0$  to  $-2.5$ ;  $-1.5$  to  $-2.0$ ).

Recruited through community advertising; 11 medical centers participated.

All participants randomized to alendronate or placebo, regardless of T-score.

Intention-to-treat analysis; mean follow-up 4.2 yrs.

†The technical report (Nelson et al., 2010a) notes that event rates were calculated for 5 years (FIT study followed patients for a mean of 4.2 years). See footnote to Table 11 in the technical report for the Nelson review.

### Evidence Table IIIId. Economic Evaluations (Key Question #5)

Key: AE, adverse effect; BMD, bone mineral density; CAROC, Canadian Association of Radiologists and Osteoporosis Canada; CE, cost-effectiveness; DXA, dual x-ray absorptiometry; FIT, Fracture Intervention Trial; ICER, incremental cost-effectiveness ratio; ICUR, incremental cost-utility ratio; OP, osteoporosis; ORAI, Osteoporosis Risk Assessment Instrument; OST, Osteoporosis Self-Assessment Tool; QALD, quality-adjusted life-day; QALY, quality-adjusted life-year; QUS, quantitative ultrasound; RR, relative risk; SCORE, Simple Calculated Osteoporosis Risk Estimate; SOF, Study of Osteoporotic Fractures; USD, United States dollars; VitD, vitamin D; WTP, willingness-to-pay

Authors/Study Type/Population	Inputs, Assumptions, and Calculations	Findings/Sensitivity Analysis	Author Conclusions/Comments
<b>Nayak et al. (2011)</b> U.S. <b>Design:</b> Individual-level state-transition CE model <b>Perspective:</b> Payer <b>Time horizon:</b> Lifetime <b>Participants:</b> Community-dwelling postmenopausal women age $\geq 55$ yrs <b>Reference:</b> No screening; treatment only after OP fracture <b>7 screening strategies:</b> 3 using DXA with T-score cutoffs for treatment: –2.5, –2.0, and –1.5	<b>Outcomes:</b> Death after OP-based treatment, after fracture, or in hospital with second fracture Nursing home vs community care Medication AE (FIT) Fracture without treatment based on SOF Fracture with treatment, trial-based estimates of RR: 0.44-0.46 for hip, 0.50-0.52 for lumbar <b>Utilities:</b> <i>No fracture:</i> 0.724 (age 85) to 0.837 (age 55). <i>Fracture:</i> 0.797-0.81, depending on site; multiplier $\sim 0.9$ for subsequent yrs. <i>NH residence:</i> 0.4 <i>Esophagitis or esophageal ulcer:</i> Multiplier 0.98 or 0.91 <b>Treatment assumptions:</b>	Nondominated = no other screening strategy is both more effective and less costly. All screening strategies at all initiation ages were effective compared with no screening. <b>Nondominated strategy with smallest ICER:</b> Initiate at age 60; QUS prescreen, then DXA if QUS $\leq -1.0$ ; treatment if T-score $\leq -2.5$ ; no further screening; \$2300/QALY (\$2459 in 2014 USD). <b>Best* strategy at all initiation ages, assuming WTP threshold of \$50,000/QALY:</b> DXA every 5 yrs, treat if T-score $\leq -2.5$ , with or without prescreening. <b>Best strategy overall (most effective and still within WTP threshold):</b> WTP threshold \$50,000/QALY: Initiate at age 55; DXA screen; treat if T-score $\leq -2.5$ ; screen every 5 yrs; \$45,450/QALY (\$48,581 in 2014 USD).	<b>Authors' conclusions:</b> Initiating screening at age 55 yrs, continuing to 80 yrs, is effective and within typical CE thresholds, regardless of screening strategy. Small differences. Expansion of OP screening could improve health outcomes at reasonable cost. <b>Limitations:</b> Possible overestimation of effectiveness of bisphosphonate treatment (larger effects than suggested by

Authors/Study Type/Population	Inputs, Assumptions, and Calculations	Findings/Sensitivity Analysis	Author Conclusions/Comments
<p>2 using QUS as prescreen, 2 QUS cutoffs for DXA screen, treat if T-score &lt; -2.5</p> <p>3 using SCORE as prescreen; DXA if SCORE ≥ 7. 3 different thresholds for treatment: T-score &lt; -2.0, T-score &lt; -1.5, age ≥ 80 yrs</p> <p><b>6 variations according to age of screening initiation:</b> 55, 60, 65, 70, 75, and 80 yrs</p>	<p>Generic oral alendronate if OP diagnosis or fracture</p> <p>50% adherence, 50% of adherent patients switched to more expensive bisphosphonate within first month and remained adherent</p> <p>5 yrs of tx for adherent pts, 5 yrs off, 5 yrs on, etc. if no AE</p> <p><b>Direct Costs:</b> Medications, fracture treatment, nursing home, endoscopy, proton pump inhibitors</p> <p><b>Base yr:</b> 2010</p> <p><b>Discount rate:</b> 3% for costs and QALYs</p> <p><b>Comparisons for ICERS:</b> ICERs were calculated with respect to the next less costly, nondominated strategy.</p>	<p><b>WTP threshold \$100,000/QALY:</b> Initiate at age 55; DXA screen; treat if T-score ≤ -2.0; screen every 10 yrs; \$94,210/QALY (\$100,701 in 2014 USD).</p> <p><b>ICERs for QUS and SCORE strategies:</b> All nondominated strategies had smaller ICERs than ICERS for DXA strategy.</p> <p><b>Sensitivity analyses:</b> Base case and probabilistic. 2 WTP thresholds (\$50,000/QALY and \$100,000/QALY). Fracture risk, ± 50%; adherence at 70%; costs at upper limit of sensitivity analysis range; AEs at 100 times rate in trials. <i>Same strategies were still the most effective with sensitivity analyses.</i></p> <p>In probabilistic analyses, no strategy was ranked first in &gt;11% of analyses at either WTP threshold; consistent with small differences (several hrs to 14 QALDs gained; average lifetime cost differences, \$20-\$1810).</p> <p><b>Model validation:</b> Predictions compared with actual outcomes in U.S. data sources.</p> <p><b>*Best means most effective but still under the WTP threshold.</b></p>	<p>systematic reviews; may be explained by follow-up differences).</p> <p>Analysis assumes all women at a T-score threshold will receive treatment.</p> <p>No analysis of initiation of screening at age &lt;55 yrs or of screening intervals defined by previous T-score.</p> <p>Fracture rates based on SOF may be subject to healthy volunteer bias.</p> <p>Inputs based on white women.</p>
<p><b>Nshimyumukiza et al. (2013)</b> Canada</p> <p><b>Design:</b> Markov decision model</p> <p><b>Perspective:</b> National health plan</p> <p><b>Time horizon:</b> Lifetime</p> <p><b>Participants:</b> Community-dwelling women age ≥ 40</p>	<p><b>Outcomes modeled:</b></p> <p>Death after fracture.</p> <p>Fracture without treatment based on CAROC and population data.</p> <p>Fracture with treatment, trial-based estimates of RR: For risedronate, 0.72 for hip, 0.58 for vertebral, 0.82 for wrist.</p> <p><b>Utilities:</b> No fracture: Unclear. Fracture: 0.30-1.00, depending on fracture site and site of care</p>	<p><b>For women age 40-64 yrs:</b></p> <p><i>Universal primary prevention, physical activity alone:</i> Would save costs, avert fractures, and add QALYs, compared with the status quo.</p> <p><i>BMD/CAROC plus universal primary prevention with physical activity+VitD+calcium:</i> Would avert the greatest number of fractures compared with all other 11 alternatives, but unacceptable ICER of \$346,776 when compared with the next less expensive alternative. This strategy would not add the most QALYs, and the ICUR would be</p>	<p><b>Authors' conclusions:</b></p> <p>A universal program promoting physical activity is the most cost-effective option and has the most cost-utility for women age 40-64 yrs.</p> <p>For women age ≥ 65 yrs, universal BMD screening and pharmacological treatment might be</p>



Authors/Study Type/Population	Inputs, Assumptions, and Calculations	Findings/Sensitivity Analysis	Author Conclusions/Comments
<p>yrs</p> <p><b>Reference:</b> Status quo. Screening and OP treatment according to usual care.</p> <p><b>12 screening scenarios:</b></p> <p>3 universal primary prevention programs: Physical activity, VitD+calcium, physical activity+VitD+calcium</p> <p>12 universal screening programs (3 scenarios each):</p> <p><u>BMD (DXA) screening</u>, followed by fracture risk with CAROC (incorporates BMD); treatment of individuals found to be at high risk of fracture; and promotion of 1 of the 3 prevention programs at low-moderate risk.</p> <p><u>ORAI</u>, followed by CAROC; treatment or prevention promotion as in the BMD screening.</p> <p><u>OST</u>, followed by CAROC; treatment or prevention promotion as in the other scenarios</p> <p><u>SCORE</u>, followed by</p>	<p><b>Miscellaneous assumptions:</b> OP-preventing behavior, participation in universal screening programs, BMD testing and acceptance of medication after fracture; various sources</p> <p><b>Treatment assumptions:</b></p> <p>Risedronate if high risk according to CAROC; 49% adherence. Taken until death (5 and 10 yrs in sensitivity analyses).</p> <p><b>Follow-up assumptions:</b> Repeat DXA at 2 or 5 yrs</p> <p><b>Comparisons for ICERS:</b> ICERs were calculated with respect to the next less costly, nondominated strategy.</p> <p><b>Direct costs:</b> Scans, office visits, medications, fracture treatment, nursing home. <i>One-time screening assumed.</i></p> <p><b>Base yr:</b> 2007</p> <p><b>Discount rate:</b> 3% for costs and QALYs</p>	<p>\$239,573, compared with next less expensive strategy.</p> <p><i>SCORE/CAROC plus universal primary prevention with physical activity+VitD+calcium:</i> ICER was \$105,649. Did not add more QALYs than the next less expensive strategy.</p> <p>No other strategy averted more fractures or added more QALYs than the next less expensive strategy.</p> <p><b>For women age ≥65 yrs:</b></p> <p><i>Universal primary prevention, physical activity alone:</i> Would save costs and avert fractures, compared with the status quo.</p> <p><i>BMD/CAROC plus universal primary prevention with physical activity+VitD+calcium:</i> Would avert the greatest number of fractures and add the most QALYs. ICER, \$60,205 (\$55,019, 2014 USD). ICUR, \$55,300 (\$50,537 2014 USD).</p> <p>No other strategy averted more fractures or added more QALYs than the next less expensive strategy.</p> <p><b>Sensitivity Analyses:</b> Findings remained robust after several forms of one-way and multiway sensitivity analyses. The cost-effectiveness acceptability curve generated for BMD/CAROC plus all prevention programs vs a physical activity program alone suggested a 63% probability of ≤\$50,000 per averted fracture and a 75% probability of ≤\$50,000 per QALY gained.</p>	<p>considered a reasonable alternative to the status quo.</p> <p><b>Limitations:</b></p> <p>No consideration of medication AEs.</p> <p>Findings may not be generalizable to screening programs that do not incorporate a prevention program or to screening followed by treatment based on T-score rather than a multifactorial risk tool.</p>



Authors/Study Type/Population	Inputs, Assumptions, and Calculations	Findings/Sensitivity Analysis	Author Conclusions/Comments
CAROC; treatment or prevention promotion as in the other scenarios			

## APPENDIX IV. Summary of Practice Guidelines

Key: ADT, androgen-deprivation therapy; BMD, bone mineral density; BMI, body mass index; DXA, dual-energy x-ray absorptiometry; FRAX, Fracture Risk Assessment Tool; IBD, inflammatory bowel disease; OP, osteoporosis; QCT, quantitative computed tomography; RCT, randomized controlled trial

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
<b>American Association of Clinical Endocrinologists (AACE)</b> (Watts et al., 2010)  <i>Medical Guidelines for Clinical Practice for the Diagnosis and Treatment of Osteoporosis</i>	<u>Women age <math>\geq 65</math> yrs</u> (Grade B, Best Evidence Label 2) or <u>younger postmenopausal women</u> at increased risk of fracture, based on a list of risk factors (Grade C, Best Evidence Label 3) Screen with DXA (Grade B, Best Evidence Label 3)  4-part scales for both recommendation Grade and Best Evidence Label.	Patients with hip or spine (clinical or radiographic) fracture, T-score $\leq -2.5$ , or T-score between $-1.0$ and $-2.5$ if FRAX score suggests $\geq 20\%$ risk of major OP fracture or $\geq 3\%$ risk of hip fracture.  First-line: alendronate, risedronate, zoledronic acid, denosumab.	<u>Women with normal BMD</u> : In 5 or 10 yrs, or perhaps never, if indications do not change. (Not rated) <u>Women with osteopenia</u> : Every 3-5 yrs, depending on T-score. (Not rated) <u>Women being treated for OP</u> : every 1-2 yrs until findings are stable, every 2 yrs or less frequently thereafter. (Grade B, Best Evidence Label) <u>Women approaching treatment threshold</u> : Every 1-2 yrs (Not rated)	4 (criteria for selecting evidence not described, strengths and limitations of body of evidence not described)
<b>American College of Gastroenterology</b> (Kornbluth et al., 2010)  <i>Ulcerative Colitis Practice Guidelines in Adults</i>	DXA should be considered in IBD patients: (1) with risk factors for OP such as smoking, low body mass, sedentary lifestyle, hypogonadism, family history, and nutritional deficiencies; (2) age $\geq 60$ yrs; (3) using corticosteroids $>3$ months consecutively or recurrently.	Bisphosphonates should be considered in IBD patients using corticosteroids $<3$ months consecutively or recurrently	No recommendations	3 (minimal detail about systematic literature search or quality of evidence; these particular recommendations seem to be offered as expert opinion—no evidence cited and no strength of recommendation)
<b>American College of Physicians (ACP)</b> (Qaseem et al., 2008)  <i>Screening for Osteoporosis in Men: A Clinical Practice Guideline from the American College of Physicians</i>	Clinicians should periodically assess risk factors for OP in older men. (Strong recommendations; moderate-quality evidence) The appropriate age to start risk assessment is uncertain. Clinicians should obtain DXA for	No recommendations	No recommendations	6 (external review not reported)

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
	men who are at increased risk <sup>4</sup> and are candidates for drug therapy. (Strong recommendations; moderate-quality evidence) Also states that in certain situations a single risk factor, e.g., ADT, may be sufficient reason to screen.			
<b>American College of Preventive Medicine (ACPM)</b> (Lim et al., 2009)  <i>Screening for Osteoporosis in the Adult U.S. Population: ACPM Position Statement on Preventive Practice</i>	<u>Women ≥65 yrs</u> and <u>men ≥70 yrs</u> of age should be screened for OP. Screening should be performed with BMD testing by DXA. <u>Adults ≥50 yrs</u> of age should be evaluated for risk factors and undergo testing if they have <u>≥1 major or 2 minor risk factors</u> .	No recommendations	No recommendations	4 (some recommendations were derived from other guidelines and no critical appraisal was offered)
<b>American College of Obstetricians-Gynecologists (ACOG)</b> (ACOG, 2012)  <i>Practice Bulletin: Osteoporosis</i>	OP screening (DXA) recommended for <u>women ≥65 yrs</u> of age and <u>younger postmenopausal women who have OP risk factors</u> according to FRAX.	Treatment recommended in concordance with other major guidelines, when a BMD T-score is ≤ -2.5. Based upon the FRAX calculator, women who have a 10-yr risk of fracture ≥20% or hip fracture risk of ≥3% are candidates for drug therapy. Women who had a low-trauma fracture without evidence of OP are candidates for drug therapy.  Bisphosphonates are first-line	DXA scan 1-2 yrs after treatment initiation; DXA does not need to be repeated if BMD improves or stabilizes or if there are no new risk factors.	4 (criteria for selecting evidence not described, development group and methods for formulating recommendations not described, guideline review and update process not described)

<sup>4</sup>The ACP guideline cites a meta-analysis that considered 12 potential risk factors for osteoporosis in healthy men and identified these as the most important: age > 70 years, low body weight, weight loss compared with normal or in recent years, physical inactivity, oral corticosteroids, and previous fragility fracture.

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
		agents. Raloxifene is a good choice for younger postmenopausal women.		
<b>American College of Rheumatology</b> (Grossman et al., 2010)  <i>Recommendation for the Prevention and Treatment of Glucocorticoid-Induced Osteoporosis</i>  (Also endorsed by American Society for Bone and Mineral Research)	Baseline DXA recommended (level of evidence C) for patients before starting glucocorticoid for an anticipated $\geq 3$ months. (Level C recommendation)  Level C = evidence from consensus, expert opinion, or case series	<u>Postmenopausal women and men age &gt;50 yrs</u> : Different recommendations for low, medium, and high fracture risk <sup>5</sup> and for different doses and duration of glucocorticoids. <i>Risk assessment requires obtaining BMD T-score.</i> <u>Premenopausal women and men age &lt;50 yrs</u> : No recommendation for those without prevalent fragility fracture. Different recommendations depending on dosage and duration of glucocorticoids.	Serial BMD testing should be <i>considered</i> for patients receiving glucocorticoid therapy for $\geq 3$ months. (Level of evidence C) As often as 6 months for treatment of OP, yearly for prevention of OP.	7
<b>American College of Radiology (ACR)</b> (ACR, 2010)	Appropriate imaging studies for women in <u>menopausal transition</u> (late 40s), <u>postmenopausal women &gt;50 yrs</u> of age, and <u>men &gt;50 yrs</u> of	No recommendations	Follow-up imaging every 2 yrs until stabilization to measure efficacy in patients who are receiving treatment and for those who have a	3 (not a systematic or critical assessment of the evidence)

<sup>5</sup>The fracture risk categories that served as the basis for treatment recommendations were based on a consensus process in which an expert panel rated 48 patient examples that were derived by first creating every possible combination of these factors from the FRAX model: sex, age (55, 65, or 75 years), race/ethnicity (white or African American), femoral neck T-score (0.0, -1.0, -1.5, -2.0, or -2.5). Then the corresponding FRAX scores for major osteoporotic and hip fracture, assuming glucocorticoid use was present, was assigned to each of the 48 examples. The other FRAX variables were assumed to be absent but are meant to be taken into account when considering whether an individual taking glucocorticoids is at higher risk than the risk category definitions would otherwise suggest. In addition to the other FRAX variables, the panel recommended that, based on evidence in the literature, these factors be taken into account when assessing risk of osteoporotic fracture: higher daily glucocorticoid dose, higher cumulative glucocorticoid dose, intravenous pulse glucocorticoid usage, declining central BMD measurement that exceeds the least significant change (see **TECHNOLOGY DESCRIPTION** for a discussion of the concept of least significant change).

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
<i>ACR Appropriateness Criteria: Osteoporosis and Bone Mineral Density</i>	<u>age with risk factors</u> : DXA posteroanterior spine; DXA femur and total hip and femoral neck; or QCT spine		risk for fracture or low density and are not undergoing treatment. ACR recommends using the same scanner type for all follow-up BMD measurements.	
<b>The Endocrine Society</b> (Watts et al., 2012)  <i>Osteoporosis in Men: An Endocrine Society Clinical Practice Guideline</i>	BMD testing suggested for <u>men age <math>\geq 70</math> yrs or age 50-69 yrs with risk factors for fracture</u> . <sup>6</sup> (Weak recommendation, low-quality evidence)  DXA of spine and hip is recommended. (Strong recommendation, low-quality evidence)	Drug therapy is recommended for patients with the following: <ul style="list-style-type: none"> <li>• Nontraumatic hip or vertebral fracture</li> <li>• BMD of the spine, femoral neck, and/or total hip is 2.5 standard deviations below that of normal young white males</li> <li>• T-scores from <math>-1</math> to <math>-2.5</math> and a 10-yr risk of major osteoporotic fracture of <math>\geq 20\%</math> or hip fracture of <math>\geq 3\%</math></li> <li>• Long-term glucocorticoid therapy in pharmacologic doses</li> <li>• Current treatment for prostate cancer with ADT</li> </ul>	DXA of the hip and spine every 1 or 2 yrs is recommended to monitor response to treatment.	6
<b>European Urological Association (EUA)</b> (Dohle et al., 2012)  <i>Guidelines on Male Hypogonadism</i>	Adult men with established severe hypogonadism should be screened for concomitant OP. "Severe" is not defined; recommendation is made in the context of late-onset hypogonadism. (Level of Evidence 2	No recommendations	No recommendations	4 (no details regarding systematic literature search; no discussion of evidence, only a reference list)

<sup>6</sup>The Endocrine Society guidelines cite evidence supporting the following risk factors for fracture in men: age, black or Hispanic race (versus white), low body mass index (BMI), alcohol use, smoking, corticosteroid use, parental fracture, history of recent fall, hypogonadism, kidney stones, history of stroke, diabetes, asthma, cardiovascular disease, dementia, osteoporosis, and rheumatoid arthritis.

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
	[≥1 well-designed controlled but nonrandomized study]; Recommendation Grade B)			
<b>Institute for Clinical Systems Improvement (ICSI)</b> (Florence et al., 2013)  <i>Health Care Guideline: Diagnosis and Treatment of Osteoporosis</i>	OP screening (DXA) for <u>women ≥65 yrs of age and in younger women whose 10-yr fracture risk is ≥9.3% from FRAX analysis or other indication of increased risk.</u> (Quality of evidence: moderate; strength of recommendation: strong). OP screening should be <u>considered for men &gt;70 yrs or men age 50-69 with risk factors.</u> (Not rated.) BMD using central DXA recommended to predict fracture risk and <u>monitor treatment.</u> (Quality of evidence: moderate; strength of recommendation: strong). BMD screening recommended for patients using <u>glucocorticoids</u> equivalent to >5 mg prednisone/day for ≥3 months. (Quality of evidence: moderate; strength of recommendation: strong)	Bisphosphonates are recommended for reduction of fracture risk in postmenopausal women, men, and in those taking glucocorticoids (quality of evidence: moderate; strength of recommendation: strong). Men with OP who are undergoing ADT for prostate cancer should receive once-yrly intravenous zoledronic acid (quality of evidence: high; strength of recommendation: strong). Anabolic therapy with parathyroid hormone is recommended for patients with a particularly high risk of fracture (quality of evidence: high; strength of recommendation: strong).	No recommendations	7
<b>International Society for Clinical Bone Densitometry (ISCD)</b> (ISCD, 2013)  <i>2013 ISCD Official Positions – Adult: Indications for Bone Mineral Density (BMD) Testing</i>	BMD testing indicated for: <u>Women:</u> ≥65 yrs of age, postmenopausal women <65 yrs of age with risk factors (e.g., low body weight, prior fracture, high risk medication use, disease or condition associated with bone loss), during menopausal transition with risk factors (e.g., low body weight, prior fracture, high-risk medication), fragility fracture	No recommendations	Serial BMD testing recommended <ul style="list-style-type: none"> <li>• For anyone not receiving therapy in whom evidence of bone loss would lead to treatment.</li> <li>• To monitor treatment effect in any patient receiving therapy.</li> <li>• At intervals of 1 yr after initiation of or change in therapy; longer intervals once therapeutic effect is</li> </ul>	<i>Not intended to be an evidence-based practice guideline.</i>

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
	Men: ≥70 yrs of age, <70 yrs of age with risk factors (as in women), fragility fracture		established <ul style="list-style-type: none"> <li>At shorter intervals for conditions associated with rapid bone loss, e.g., glucocorticoids therapy</li> </ul>	
<b>North American Menopause Society (NAMS)</b> (NAMS, 2010)  <i>Management of Osteoporosis in Postmenopausal Women: 2010 Position Statement of the North American Menopause Society</i>	BMD measurement recommended in all <u>women ≥65 yrs of age, postmenopausal women with medical causes</u> of bone loss, postmenopausal <u>women ≥50 yrs of age with risk factors,</u> <sup>7</sup> and <u>postmenopausal women with a fragility fracture.</u>  BMD measurement should be <i>considered</i> in <u>postmenopausal women ≥50 yrs of age with:</u> certain fractures, BMI <21 kg/m <sup>2</sup> , history of hip fracture in parent, current smoker, rheumatoid arthritis, and >2 units of alcohol intake every day.	Drug therapy is recommended for the following groups: <ul style="list-style-type: none"> <li>All postmenopausal women with an osteoporotic vertebral or hip fracture.</li> <li>All postmenopausal women with T-scores ≤ -2.5 at the lumbar spine, femoral neck, or total hip region.</li> <li>All postmenopausal women with T-scores from -1 to -2.5 and a 10-yr risk of major osteoporotic fracture of ≥20% or hip fracture of ≥3%.</li> </ul> Bisphosphonates are the first-line drugs for treating OP in postmenopausal women.	Repeat BMD measurement after 1-2 yrs of treatment is appropriate; for patients who are untreated; the retesting interval should be 2-5 yrs.	4 (criteria for selecting evidence not described, strengths and limitations of evidence not described, methods for formulating recommendations not described)
<b>National Osteoporosis Foundation</b> (NOF, 2014)  <i>Clinician's Guide to Prevention and Treatment of Osteoporosis</i>	BMD testing should be performed in: <u>Women:</u> ≥65 yrs; also, postmenopausal or in menopausal transition with risk factors or an adult age fracture.	Drug therapy should be considered for postmenopausal women and men ≥50 yrs of age with the following: <ul style="list-style-type: none"> <li>Hip or vertebral (clinical or radiographic) fracture</li> </ul>	No specific recommendation regarding BMD testing after an initial normal scan. Serial DXA testing, generally every 2 yrs, is an important component of OP management.	3 (methods were not described)

<sup>7</sup> Osteoporotic risk factors considered by NAMS include: fracture after menopause (other than fracture of skull, ankle, finger, toe, or facial bone), thinness (BMI <21 kg/m<sup>2</sup>), history of hip fracture in parent, current smoker, rheumatoid arthritis, or >2 units of alcohol intake per day.

Sponsor, Title	Relevant Recommendations			
	Screening	Treatment	Monitoring	Quality/Main Limitations
Also endorsed by the American College of Rheumatology (American College of Rheumatology, 2008)	<p><u>Men</u>: ≥70 yrs of age or between 50-69 yrs of age with risk factors or an adult age fracture</p> <p>Vertebral imaging (x-ray or lateral vertebral imaging assessment) may be considered if BMD testing is not available.</p>	<ul style="list-style-type: none"> <li>T-score ≤ -2.5 at the femoral neck, total hip, or lumbar spine</li> <li>T-scores from -1 to -2.5 and a 10-yr risk of major osteoporotic fracture of ≥20% or hip fracture of ≥3%</li> </ul>		
<p><b>U.S. Preventive Services Task Force (USPSTF) (USPSTF, 2011b)</b></p> <p><i>Screening for OP</i></p>	<p><u>Women</u>: ≥65 yrs of age; younger with 10-yr fracture risk ≥ that of a 65-yr-old white woman without additional risk factors.</p> <p><u>Men</u>: Current evidence is insufficient to assess the balance of benefits and harms in screening men but considerations for screening are described.</p> <p>FRAX is the preferred tool for assessment of fracture risk.</p>	No recommendations	There is a lack of evidence regarding appropriate intervals for repeated BMD measurement.	6